Lecture Notes in Networks and Systems 191

Amit Joshi
Mufti Mahmud
Roshan G. Ragel
Nileshsingh V. Thakur *Editors*

Information and Communication Technology for Competitive Strategies (ICTCS 2020)

ICT: Applications and Social Interfaces

Volume 1

Springer

Lecture Notes in Networks and Systems

Volume 191

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

Amit Joshi · Mufti Mahmud · Roshan G. Ragel ·
Nileshsingh V. Thakur
Editors

Information and Communication Technology for Competitive Strategies (ICTCS 2020)

ICT: Applications and Social Interfaces

Springer

Editors
Amit Joshi
Global Knowledge Research Foundation
Ahmedabad, Gujarat, India

Roshan G. Ragel
University of Peradeniya
Kandy, Sri Lanka

Mufti Mahmud
Computing and Technology
Nottingham Trent University
Nottingham, Nottinghamshire, UK

Nileshsingh V. Thakur
Prof Ram Meghe College of Engineering and Management
Amravati, India

ISSN 2367-3370 ISSN 2367-3389 (electronic)
Lecture Notes in Networks and Systems
ISBN 978-981-16-0738-7 ISBN 978-981-16-0739-4 (eBook)
https://doi.org/10.1007/978-981-16-0739-4

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

Preface ICTCS-2020

The Fifth International Conference on Information and Communication Technology for Competitive Strategies (ICTCS-2020) targets state of the art as well as emerging topics pertaining to information and communication technologies (ICTs) and effective strategies for its implementation for engineering and intelligent applications.

The conference is anticipated to attract a large number of high-quality submissions, stimulate the cutting-edge research discussions among many academic pioneering researchers, scientists, industrial engineers, students from all around the world and provide a forum to researcher; propose new technologies, share their experiences and discuss future solutions for design infrastructure for ICT; provide a common platform for academic pioneering researchers, scientists, engineers and students to share their views and achievements; enrich technocrats and academicians by presenting their innovative and constructive ideas; focus on innovative issues at international level by bringing together the experts from different countries.

The conference was held during 11–12 December 2020 digitally on Zoom and organized by Global Knowledge Research Foundation.

Research submissions in various advanced technology areas were received, and after a rigorous peer review process with the help of programme committee members and external reviewer, 211 papers were accepted with an acceptance rate of 21%. All 211 papers of the conference are accommodated in 2 volumes, and also, papers in the book comprise authors from 15 countries.

This event success was possible only with the help and support of our team and organizations. With immense pleasure and honour, we would like to express our sincere thanks to the authors for their remarkable contributions, all the technical program committee members for their time and expertise in reviewing the papers within a very tight schedule and the publisher Springer for their professional help.

We are overwhelmed by our distinguished scholars and appreciate them for accepting our invitation to join us through the virtual platform and deliver keynote speeches and technical session chairs for analysing the research work presented by the researchers. Most importantly, we are also grateful to our local support team for

their hard work for the conference. This series has already been made a continuous series which will be hosted at different locations every year.

Gujarat, India — Amit Joshi
Nottingham, UK — Mufti Mahmud
Kandy, Sri Lanka — Roshan G. Ragel
Amravati, India — Nileshsingh V. Thakur

Contents

Editors and Contributors

About the Editors

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

Dr. Mufti Mahmud received his Ph.D. degree in Information Engineering (specialized in neuroengineering) from the University of Padova—Italy in 2011. A recipient of the Marie Curie postdoctoral fellowship, Dr. Mahmud has served at various positions in the industry and academia in India, Bangladesh, Italy, Belgium and the UK during the last 17 years. Dr. Mahmud aims to leave behind a healthy, secure and smart world to live in. As Expert of neuroengineering, computational intelligence and data science, his research aims to build predictive, secure and adaptive systems for personalized services to improve quality of life through advancement of healthcare access in low-resource settings. Senior Member of IEEE and ACM, and Professional Member of BCS, Dr. Mahmud holds leadership roles at many technical committees, such as Vice Chair of the Intelligent System Application Technical Committee of IEEE Computational Intelligence Society,

Member of the IEEE CIS Task Force on Intelligence Systems for Health, Co-chair of the IEEE P2733 Standard on Intelligent systems design, with AI/ML, Member of the IEEE R8 Humanitarian Activities Subcommittee and Project Liaison Officer of the IEEE UK and Ireland Special Interest Group on Humanitarian Technology. Dr. Mahmud serves as Associate Editor of the Cognitive Computation, IEEE Access, Brain Informatics and Big Data Analytics journals. He also serves at many technical, programme and organization committees of high-rank conferences including the Local Organizing Chair of IEEE-WCCI2020; General Chair of BI2020; and Programme Chair of IEEE-CICARE2020.

Dr. Roshan G. Ragel is Full Professor in Computer Engineering at the University of Peradeniya, Sri Lanka. He is Skilled University Lecturer and Researcher with a doctorate in computer science and engineering and a passion for teaching, research and administration. His research interest is in the areas of fault-tolerant and dependable computing, accelerated and high-performance computing, wearable computing and synthesis of secure and reliable embedded processors. Prof. Ragel received his Ph.D. in Computer Science and Engineering from the University of New South Wales, Australia, in 2007, and his B.Sc. in Engineering from the University of Peradeniya, Sri Lanka, in 2001. He was appointed Senior Lecturer in Computer Engineering at the University of Peradeniya in December 2007 and Full Professor since October 2017. Prof. Ragel has co-authored about 150 peer-reviewed articles in topics including Micro-Architectural Support for Reliability and Security in Embedded Processors (SoC), Internet of Things (IoT), Side-Channel Attacks and Countermeasures, Application Specific Processor Design, High-Performance Computing such as Hardware/Software Acceleration of Bioinformatics Algorithms, Wearable Computing, Format-Preserving Digitisation and OCR of Local Language Text, and Completely Automated Public Turing test to tell Computers and Humans Apart (CAPTCHAs) in the Context of Mobile Applications. He has won several best paper awards in conference papers, the most prolific Sri Lankan author award from Elsevier Research in 2017 and the Presidential award for scientific publication in Sri Lanka in 2018. He was a visiting research scholar at the UNSW, Sydney, a few times and UIUC, Illinois, once. He is Senior Member of the IEEE and the IEEE Computer Society.

Dr. Nileshsingh V. Thakur received Ph.D. in Computer Science Engineering from Visvesvaraya National Institute of Technology, Nagpur, India, on 1st February 2010. He is having over 28 years of teaching and research experience. Presently, he is Professor in the Department of Computer Science and Engineering at Prof Ram Meghe College of Engineering and Management, Badnera, Amravati, India. He is the author or the co-author of more than 80 scientific publications in international journals, international conferences and workshops. He is Editor of the books published by Tata McGraw-Hill and Springer. Under his guidance, 6 research scholars have completed Ph.D. and 28 scholars have completed M.Tech. He is an editorial board member of 6 international journals. He acted as Programme Committee Member of various international conferences. He is a reviewer for IEEE

and IET Image processing Journals. He delivered more than 60 keynote/expert talks at international and national conferences, and training programmes. He is certified ISO Auditor and worked as External Expert in the selection board for recruitment of Scientist-B for DRDO, India. His research interest includes image and video processing; robotics and vision; pattern recognition and neural network; smart camera and sensor networks; big data analysis; nature-based computing algorithms and evolutionary approaches.

Contributors

Ali Abbas Rizvi Department of Information Technology, Dwarkadas J. Sanghvi College of Engineering, Mumbai, India

Prashant Abbi Department of ISE, RV College of Engineering®, Bengaluru, India

Jayashree Agarkhed Department of Computer Science and Engineering, Poojya Doddappa Appa College of Engineering, Kalburgi, India

Alpana Agarwal Thapar Institute of Engineering and Technology, Patiala, India

Priyanka Agarwal PES University, Bengaluru, India

Prasit G. Agnihotri Civil Engineering Department, Sardar Vallabhbhai National Institute of Technology, Surat, India

Shalini Agrawal DSPM International Institute of Information Technology (IIIT), Naya Raipur, India

Saad Salman Ahmed Faculty of Mathematics, University of Technology and Applied Sciences, Muscat, Sultanate of Oman

Sheeba Uruj Ahmad Department of Computer Science, Birla Institute of Technology & Science, Pilani, India;
Dubai International Academic City, Dubai, UAE

Syed Rafeek Ahmed Faculty of Engineering, University of Technology and Applied Sciences, Muscat, Sultanate of Oman

S. Ajitha M. S. Ramaiah Institute of Technology, Bengaluru, India

B. N. Akash PES University, Bengaluru, India

Hasina Al Harthy Department of Research, Royal Hospital, Muscat, Sultanate of Oman

Huda Salim Al Shuaily Faculty of IT, University of Technology and Applied Sciences, Muscat, Sultanate of Oman

Asha Ambhaikar Kalinga University, Raipur, India

K. M. Amogh PES University, Bengaluru, India

P. P. Amritha TIFAC-CORE in Cyber Security, Amrita School of Engineering, Amrita Vishwa Vidyapeetham, Coimbatore, India

Maleka Anjum Department of Computer Science & Engg, PDA College of Engineering, Kalaburgi, India

N. B. Arunekumar Pondicherry University, Puducherry, India

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

Himanshu Ashar Department of Information Technology, Dwarkadas J. Sanghvi College of Engineering, Mumbai, India

Nishant Bagmar MIT Institute of Design, MIT-ADT University, Pune, Maharashtra, India

Kailash Chandra Bandhu Shivajirao Kadam Institute of Technology and Management—Technical Campus, Indore, India

Shabnam Banerjee Tata Medical Centre, Kolkata, India

Manu Bansal Thapar Institute of Engineering and Technology, Patiala, India

N. Freddy Patricio Baño Carrera de Software, Instituto Superior Tecnológico Mayor Pedro Traversari, Quito, Ecuador

Urjita Bedekar Vivekanand Education Society's Institute of Technology, Mumbai, India

Jorge A. Benitez Departamento de Jurisprudencia, Universidad Técnica Particular de Loja, Loja, Ecuador

Parag Bhalchandra School of Computational Sciences, Swami Ramanand Teerth Marathwada University, Nanded, India

P. K. Bharti Shri Venkateshwara University, Gajraula, UP, India

Janardan Bharvad Devang Patel Institute of Advance Technology and Research (DEPSTAR), Charotar University of Science and Technology (CHARUSAT), Changa, Gujarat, India

Anand Bhaskar SPSU, Udaipur, Rajasthan, India

Pranav A. Bhat Department of ISE, RV College of Engineering®, Bengaluru, India

Gresha Bhatia Vivekanand Education Society's Institute of Technology, Mumbai, India

Manmohan Singh Bhatia Terna College of Engineering, Nerul, India

Nabarun Bhattacharyya Centre for Development of Advanced Computing, Kolkata, India

Pushkar Bhuse Department of Information Technology, Dwarkadas J. Sanghvi College of Engineering, Mumbai, India

Milon Biswas Department of Computer Science and Engineering, Bangladesh University of Business and Technology, Dhaka, Bangladesh

Navneet Kaur Brar Thapar Institute of Engineering and Technology, Patiala, India

Aishwariya Budhrani Devang Patel Institute of Advanced Technology and Research (DEPSTAR), Faculty of Technology & Engineering (FTE), Charotar University of Science and Technology (CHARUSAT), Anand, India

Suneeta V. Budihal School of Electronics and Communication, KLE Technological University, Hubballi, India

N. K. Cauvery Department of ISE, RV College of Engineering®, Bengaluru, India

Kavita Chachadi KLE Technological University, Hubballi, Karnataka, India

Arti Chandani SIMS, Symbiosis International (Deemed University), Pune, MH, India

Sandeep S. Chapalkar Dayananda Sagar College of Engineering, Bengaluru, India

Sachin V. Chaudhari Sanjivani College of Engineering, Kopargaon, SPPU, Pune, Maharashtra, India

Kanika Chauhan Indira Gandhi Delhi Technical University for Women, New Delhi, India

K. Chidananda Department of CSE, JNTUA, Anantapur, Andhra Pradesh, India

Prashant Chintal Marathwada Institute of Technology, Dr. B.A.M University, Aurangabad, Maharashtra, India

Pradnya Chitrao SIMS, Khadki, Pune, India

John Colaco Department of Electronics & Telecommunication Engineering, Goa Engineering College, Ponda, Goa, India

Nitin Darkunde School of Mathematical Sciences, Swami Ramanand Teerth Marathwada University, Nanded, India

R. Deepa Department of Computer Science and Engineering, SRM Institute of Engineering and Technology, Vadapalani, Chennai, India

Rahil Desai Department of Information Technology, Dwarkadas J. Sanghvi College of Engineering, Mumbai, India

Nilesh Deshmukh School of Computational Sciences, Swami Ramanand Teerth Marathwada University, Nanded, India

Anupama Deshpande Department of Electrical Engineering, Shri J.J.T. University, Jhanjhunu, Rajasthan, India

Snehal Dongare G H Raisoni College of Engineering, Nagpur, India

Priyank D. Doshi Atmiya University, Rajkot, Gujarat, India

R. Durga Department of Computer Science and Engineering, SRM Institute of Science and Technology, Kattankulathur, Tamil Nadu, India

Divya Susanna Ebin MS Ramaiah Institute of Technology, Mathikere, Bangalore, India

Lozada T. Edwin Fabricio Universidad Regional Autónoma de Los Andes–UNIANDES, Carrera de Software, Ambato, Ecuador;
Facultad de Informática, Universidad Nacional de La Plata, LIFIA, La Plata, Argentina

Prutha Gaherwar Department of Computer Technology, Pune Institute of Computer Technology, Pune, Maharashtra, India

Pankaj J. Gandhi Civil Engineering Department, Sardar Vallabhbhai National Institute of Technology, Surat, India;
School of Engineering, P. P. Savani University, Kosamba, Surat, India

Rishi Kumar Gandhi Osmania University (OU), Hyderabad, India

Avani Gangrade Shivajirao Kadam Institute of Technology and Management—Technical Campus, Indore, India

Juan Manuel Garcia-Samaniego Departamento de Ciencias Naturales, Universidad Técnica Particular de Loja, Loja, Ecuador

Dweepna Garg Devang Patel Institute of Advance Technology and Research (DEPSTAR), Charotar University of Science and Technology (CHARUSAT), Changa, Gujarat, India

S. K. Gautham BMS Institute of Technology and Management, Bangalore, India

Haftom Gebregziabher Federal Technical and Vocational Education and Training Institute, Addis Ababa, Ethiopia

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

Alokesh Ghosh Centre for Development of Advanced Computing, Kolkata, India

Devdulal Ghosh Centre for Development of Advanced Computing, Kolkata, India

Tarun Kanti Ghosh Centre for Development of Advanced Computing, Kolkata, India

T. Gireesh Kumar TIFAC-CORE in Cyber Security, Amrita School of Engineering, Amrita Vishwa Vidyapeetham, Coimbatore, India

Hector F. Gomez A. Centro de Posgrados, Universidad Técnica de Ambato, Ambato, Ecuador

Chetan Gondaliya AMPICS, Ganpat University, Mehsana, Gujarat, India

S. Gowri Shankar Department of Computer Science and Engineering, BMS College of Engineering, Bangalore, India

Nahid Guard School of Electronics and Communication, KLE Technological University, Hubballi, India

I Gede Aris Gunadi Universitas Pendidikan Ganesha, Singaraja Bali, Indonesia

Amar Nath Gupta Symbiosis International University, Lavale, Pune, India

Kanika Gupta Indira Gandhi Delhi Technical University for Women, New Delhi, India

Nidhi Gupta I. K. Gujral Punjab Technical University, Kapurthala, Punjab, India

Shailendra Gupta Globus Engineering College, Bhopal, India

Md. Habibur Rahman Department of Computer Science and Engineering, Daffodil International University, Dhaka, Bangladesh

K. S. Haritha Assistant Professor, Govt. Engineering college, Kannur, Kozhikkode, Kerala, India

Pradeep Kumar Hazari Birla Institute of Technology, Mesra, Lalpur Ext Ranchi, Jharkhand, India

Sushma Hiremath KLE Technological University, Hubballi, India

Vaishali Vaibhav Hirlekar Sir Padmapat Singhania University, Udaipur, Rajasthan, India

Sanjana S. Hurkadli KLE Technological University, Hubballi, India

Sindhoori Ichalakaranji KLE Technological University, Hubli, Karnataka, India

Gede Indrawan Universitas Pendidikan Ganesha, Singaraja Bali, Indonesia

Mohamed Ismail Department of Mathematics, Sathyabama Institute of Science and Technology, Chennai, India

Nalini C. Iyer KLE Technological University, Hubballi, India;
JSS Dharwad, Dharwad, India

Prashant Jadhav Ramrao Adik Institute of Technology, Nerul, India

Dhiraj Jain Symbiosis Centre of Management Studies, Pune, India

Swati Jain Government J. Yoganandam Chhattisgarh College, Raipur, India

Nilesh K. Jaiswal SVIT Vasad, Vasad, India

Tushar Hrishikesh Jaware R C Patel Institute of Technology, Near Nimzari Naka, Shirpur Dhule, India

K. P. Jayalakshmi St Joseph Engineering College, Mangaluru, Karnataka, India

S. Jeyanthi Department of Computer Science and Engineering, PSNA College of Engineering and Technology, Dindigul, Tamil Nadu, India

K. Suresh Joseph Pondicherry University, Puducherry, India

Solley Joseph CHRIST (Deemed to be University), Bengaluru, India;
Carmel College of Arts, Science and Commerce for Women, Goa, India

Khushali A. Joshi SVIT Vasad, Vasad, India

Raviraj Joshi Department of Computer Science and Engineering, Indian Institute of Technology Madras, Chennai, Tamil Nadu, India

Shraddha Joshi Department of Computer Technology, Pune Institute of Computer Technology, Pune, Maharashtra, India

T. Sheeba Justin Muscat College, Muscat, Sultanate of Oman

Jyoti ICFAI University, Jharkhand, Ranchi, India

Gauri Kalnoor Department of Computer Science and Engineering, BMS College of Engineering, Bangalore, India

Priyadarshini D. Kalwad KLE Technological University, Hubli, Karnataka, India

Vijendra Kamble School of Computational Sciences, Swami Ramanand Teerth Marathwada University, Nanded, India

Suvarna G. Kanakaraddi KLE Technological University, Hubli, Karnataka, India

Anita Kanavalli M S Ramaiah Institute of Technology, Bengaluru, India

Chinky Karda DSPM International Institute of Information Technology (IIIT), Naya Raipur, India

K. Karibasappa Dayananda Sagar Academy of Technology and Management, Bengaluru, India

Sunidhi Kashyap Department of Computer Science, Birla Institute of Technology & Science, Pilani, India;
Dubai International Academic City, Dubai, UAE

Amanpreet Kaur Jamia Millia Islamia, New Delhi, Delhi, India

B. C. Kavitha Department of ECE, Hindustan Institute of Technology and Science, Chennai, India

Rakshil Vijaybhai Kevadiya Devang Patel Institute of Advanced Technology and Research (DEPSTAR), Faculty of Technology & Engineering (FTE), Charotar University of Science and Technology (CHARUSAT), Anand, India

Rahul Khengare Department of Computer Technology, Pune Institute of Computer Technology, Pune, Maharashtra, India

Satyendra Kishore ICFAI University, Jharkhand, Ranchi, India

Shajidmon Kolamban Faculty of Mathematics, University of Technology and Applied Sciences, Muscat, Sultanate of Oman

K. Komathy Academy of Maritime Education and Training, Chennai, India

Shweta Kore JSS College, Dharwad, India;
JSS Dharwad, Dharwad, India

Anjan K. Koundinya BMS Institute of Technology and Management, Bangalore, India

Reshmy Krishnan Muscat College, Muscat, Sultanate of Oman

Govind Kulkarni Department of Computer Science, Parli-Vaijjnath, India

Vinuta Kulkarni KLE Technological University, Hubballi, India

Ajay Kumar Thapar Institute of Engineering and Technology, Patiala, India

Arun Kumar Ir Padmapat Singhania University, Udaipur, Rajasthan, India

Pranab Kumar Birla Institute of Technology, Mesra, Lalpur Ext Ranchi, Jharkhand, India

Eduardo M. Lacap Jr Faculty of Statistics, University of Technology and Applied Sciences, Muscat, Sultanate of Oman

R. B. Lohani Department of Electronics & Telecommunication Engineering, Goa Engineering College, Ponda, Goa, India

Anand Lokapure CSIR, National Institute of Oceanography, Panaji, Goa, India

Narendra Lalchand Lokhande R C Patel Institute of Technology, Near Nimzari Naka, Shirpur Dhule, India

Edwin Fabricio Lozada T. Carrera de Software, Universidad Regional Autónoma de los Andes–UNIANDES, Ambato, Ecuador

Md. Kawsher Mahbub Department of Computer Science and Engineering, Bangladesh University of Business and Technology, Dhaka, Bangladesh

Y. Mahesha St Joseph Engineering College, Mangaluru, Karnataka, India

Ruth Patricia Maldonado Rivera Departamento de Psicología, Universidad Técnica Particular de Loja, Ambato, Ecuador

Kurra Mallaiah Osmania University (OU), Hyderabad, India

P. Manasa Research Scholar, Department of Electronics and Communication Engineering, VTU-RC, Vivekananda Institute of Technology, Bengaluru, India

Manju Mandot Janardan Rai Nagar Rajasthan Vidyapeeth, Udaipur, Rajasthan, India

Venkatesh Mane ECE Department, KLE Tech University, Hubli, India

S. Maria Celestin Vigila Department of IT, Noorul Islam Center for Higher Education, Kumaracoil, Kanyakumari, India

Priscilla Massa-Sánchez Departamento de Economía, Universidad Técnica Particular de Loja, Loja, Ecuador

Sumit Mathur SPSU, Udaipur, Rajasthan, India

Gabriela Elizabeth Armijos Maurad Departamento de Economía, Universidad Técnica Particular de Loja, Loja, Ecuador

Mita Mehta SIMS, Symbiosis International (Deemed University), Pune, MH, India

Wricha Mishra MIT Institute of Design, MIT-ADT University, Pune, Maharashtra, India

V. Mohan Department of ECE, Saranathan College of Engineering, Trichy, Tamilnadu, India

S. A. More Department of Electronics and Telecommunication Engineering, R. C. Patel Institute of Technology, Shirpur, Dhule, India

Nandini Mukherjee Jadavpur University, Kolkata, India

Aniket Muley School of Mathematical Sciences, Swami Ramanand Teerth Marathwada University, Nanded, India

Mahendra Kumar Murmu Department of Computer Engineering, National Institute of Technology, Kurukshetra, India

Khurram Mustafa Jamia Millia Islamia, New Delhi, Delhi, India

S. Muthurajkumar Department of Computer Technology, Anna University, Chennai, India

K. Muthuvel Research Scholar - Assistant Professor, Dept. of EEE, Noorul Islam Centre for Higher Education, Kumaracoil, Tamil Nadu, India

Arun T. Nair Research Scholar - Assistant Professor, Dept. of EEE, Noorul Islam Centre for Higher Education, Kumaracoil, Tamil Nadu, India

Sarachandran Nair Muscat College, Muscat, Sultanate of Oman

C. I. Nalini ECE Department, KLE Tech University, Hubli, India

R. S. Nandhini Vellore Institute of Technology, Vellore, Tamil Nadu, India

Suvarna Nandyal Department of Computer Science and Engineering, PDA College of Engineering, Kalaburgi, Karnataka, India

Nandkishor Narkhede SPSU, Udaipur, Rajasthan, India

Shubha Nayak KLE Technological University, Hubli, Karnataka, India

Sizolwethu Ndlovu University of Zululand, Richards Bay, South Africa

Thi Xuan Binh Ngo Banking University Ho Chi Minh City, Ho Chi Minh City, Viet Nam

Thi Huyen Trang Nguyen Banking University Ho Chi Minh City, Ho Chi Minh City, Viet Nam

Van Thuy Nguyen Banking University Ho Chi Minh City, Ho Chi Minh City, Viet Nam

Tajim Md. Niamat Ullah Akhund Department of Computer Science and Engineering, Daffodil International University, Dhaka, Bangladesh

S. R. Nirmala KLE Technological University, Hubballi, Karnataka, India

Francisca Nonyelum Ogwueleka Computer Science Department, Nigerian Defence Academy, Kaduna, Nigeria

Olukayode Oki Walter Sisulu University, East London, South Africa

Oluwaseyi Ezekiel Olorunshola Computer Science Department, Nigerian Defence Academy, Kaduna, Nigeria

Md. Omar Faroque Department of Computer Science and Engineering, Move Fitness, Frisco, USA

Richard Eduardo Ruiz Ordoñez Instituto Superior Tecnologico Loja, Loja, Ecuador

Dinesh Padole G H Raisoni College of Engineering, Nagpur, India

Anuj Pal Sagar Institute of Science, Technology & Research (SISTec-R), Bhopal, India

Prateek Pande Osmania University (OU), Hyderabad, India

Dipesh B. Pardeshi Department of Electrical Engineering, Shri J.J.T. University, Jhanjhunu, Rajasthan, India;
Sanjivani College of Engineering, Kopargaon, SPPU, Pune, Maharashtra, India

Ajay Patel AMPICS, Ganpat University, Mehsana, Gujarat, India

Atul Patel Charotar University of Science and Technology, Changa, Gujarat, India

Hiral Patel DCS, Ganpat University, Mehsana, Gujarat, India

M. N. Patel Parul University, Vadodara, Gujarat, India

Nirav M. Patel Civil Engineering Department, PIET, Parul University, Vadodara, Gujarat, India

Prashant Patil Department of Electronics and Telecommunication Engineering, R. C. Patel Institute of Technology, Shirpur, Dhule, India

Pushpa Patil Department of Computer Science and Engineering, B.L.D.E.A's, V.P. Dr.P.G.Halakatti, College of Engineering and Technology (Affiliated to Visvesvaraya Technological University, Belagavi 590018), Vijayapur, Karnataka, India

Siddarama Patil Dept of E &CE, Poojya Doddappa Appa College of Engineering, Kalburgi, India

Vijayalaxmi R. Patil Ambedkar Institute of Technology, Bengaluru, India

Vishal B. Pattanashetty KLE Technological University, Hubballi, India

Pullas T. Paul Facultad de Ciencias Humanas y de La Educación, Universidad Técnica de Ambato, Ambato, Ecuador

Soumya Peddi Department of Computer Science and Engineering, Poojya Doddappa Appa College of Engineering, Kalburgi, India

Tanasanee Phienthrakul Department of Computer Engineering, Faculty of Engineering, Mahidol University, Salaya, Nakhon Pathom, Thailand

Preeti S. Pillai ECE Department, KLE Tech University, Hubli, India

E. Poovammal Department of Computer Science and Engineering, SRM Institute of Science and Technology, Kattankulathur, Tamil Nadu, India

Om Prakash SICSR, Symbiosis International (Deemed University), Pune, MH, India

Prakrit Prakash MIT, Manipal Academy of Higher Education, Manipal, KA, India

B. Pranava Department of ISE, RV College of Engineering®, Bengaluru, India

Arpana Prasad Department of Computer Science, Punjabi University, Patiala, Punjab, India

T. Preeti KLE Technological University, Hubli, Karnataka, India

Seema Miranda Priya St Joseph Engineering College, Mangaluru, Karnataka, India

G. Priyanga Department of Computer Technology, Anna University, Chennai, India

V. Priyanka TIFAC-CORE in Cyber Security, Amrita School of Engineering, Amrita Vishwa Vidyapeetham, Coimbatore, India

Sahana Punagin K.L.E. Technological University, Hubli, India

Abhishek Puri ECE Department, KLE Tech University, Hubli, India

Sini Rahuman Sathyabama Institute of Science and Technology, Chennai, India

Meghana K. Raj MS Ramaiah Institute of Technology, Mathikere, Bangalore, India

Annie Rajan Department of Computer Science, University of Mumbai, Mumbai, India;
Department of Computer Science, DCT's Dhempe College of Arts and Science, Panaji, Goa, India

T. Rajasenbagam Department of Computer Science and Engineering, Government College of Technology, Coimbatore, Tamil Nadu, India

K. Rajeshwar NIRD&PR, Hyderabad, India

M. Rajkumar ICFAI University Jharkhand, Ranchi, India

S. Ramachandram Osmania University (OU), Hyderabad, India

L. Ramanathan Vellore Institute of Technology, Vellore, Tamil Nadu, India

Hemant Rathore Department of CS&IS, BITS Pilani, Goa, India

Manorama Singh Rathore DSPM International Institute of Information Technology (IIIT), Naya Raipur, India

Purva Raut Department of Information Technology, Dwarkadas J. Sanghvi College of Engineering, Mumbai, India

Chintal Upendra Raval Devang Patel Institute of Advanced Technology and Research (DEPSTAR), Faculty of Technology & Engineering (FTE), Charotar University of Science and Technology (CHARUSAT), Anand, India

M. Ravikumar DOS in Computer Science, Kuvempu University, Shivamogga, Karnataka, India

Hena Ray Centre for Development of Advanced Computing, Kolkata, India

Md. Hashmathur Rehman ICFAI University Jharkhand, Ranchi, India

S. Reshma Department of Computer Science and Engineering, Visvesvaraya Technological University—Research Resource Center, Belagavi, Karnataka, India

Shivam Ribadiya Devang Patel Institute of Advance Technology and Research (DEPSTAR), Charotar University of Science and Technology (CHARUSAT), Changa, Gujarat, India

Priyadarshini Roy DSPM International Institute of Information Technology (IIIT), Naya Raipur, India

Sonal Mobar Roy NIRD&PR, Hyderabad, India

S. Safiyyah Sanovar Department of Computer Technology, Anna University, Chennai, India

B. M. Sagar Department of ISE, RV College of Engineering®, Bengaluru, India

Sangit Saha Centre for Development of Advanced Computing, Kolkata, India

Sanjay K. Sahay Department of CS&IS, BITS Pilani, Goa, India

Sikder Md. Saiful Islam Department of Computer Science and Engineering, Bangladesh University of Business and Technology, Dhaka, Bangladesh

Rajashekhar Salagar Department of Computer Science and Engineering, B.L.D.E.A's, V.P. Dr.P.G.Halakatti, College of Engineering and Technology (Affiliated to Visvesvaraya Technological University, Belagavi 590018), Vijayapur, Karnataka, India

Ambuja Salgaonkar Department of Computer Science, University of Mumbai, Mumbai, India

Shreyas Salimath KLE Technological University, Hubli, Karnataka, India

N. Samanth Department of Electronics and Communication Engineering, B.N.M Institute of Technology, Bengaluru, Karnataka, India

S. Sampathkumar DOS in Computer Science, Kuvempu University, Shivamogga, Karnataka, India

S. Sampreeth Department of Electronics and Communication Engineering, B.N.M Institute of Technology, Bengaluru, Karnataka, India

Baby Sam Samuel InterTech.LLC, Muscat, Sultanate of Oman

I Made Subrata Sandhiyasa Universitas Pendidikan Ganesha, Singaraja Bali, Indonesia

Ravi Sankar Centre for Development of Advanced Computing, Kolkata, India

M. Sethumadhavan TIFAC-CORE in Cyber Security, Amrita School of Engineering, Amrita Vishwa Vidyapeetham, Coimbatore, India

Mohit Sewak Department of CS&IS, BITS Pilani, Goa, India

Vikrant Shaga Department of Information Communication Technology, Federal Technical and Vocational Education and Training Institute, Addis Ababa, Ethiopia

Happy L. Shah SVIT Vasad, Vasad, India

K. Shaila Department of Electronics and Communication Engineering, Vivekananda Institute of Technology, Bengaluru, Karnataka, India

M. Shamim Kaiser Institute of Information Technology, Jahangirnagar University, Dhaka, Bangladesh

P. Shanmugapriya Department of ECE, Saranathan College of Engineering, Trichy, Tamilnadu, India

A. Sharan Jasmine Department of ECE, Saranathan College of Engineering, Trichy, Tamilnadu, India

Molak R. Sharma Symbiosis International University, Pune, India

Neeraj Sharma Department of Computer Science, Punjabi University, Patiala, Punjab, India;
Gian Jyoti Institute of Management and Technology, Mohali, Punjab, India

Vijayeta Sharma East London NHS Foundation Trust, London, UK

B. S. Sheethal MS Ramaiah Institute of Technology, Mathikere, Bangalore, India

Vinu Sherimon Faculty of IT, University of Technology and Applied Sciences, Muscat, Sultanate of Oman

Raghavendra Shet K.L.E. Technological University, Hubli, India

Sujala D. Shetty Department of Computer Science, Birla Institute of Technology & Science, Pilani, India;
Dubai International Academic City, Dubai, UAE

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

Saroja V. Siddamal KLE Technological University, Hubballi, India

Poly Sil Sen Techno Main Salt Lake, Kolkata, India

Bhavesh Singh Department of Information Technology, Dwarkadas J. Sanghvi College of Engineering, Mumbai, India

Upasana Sinha JK Institute of Engineering Bilaspur, Bilaspur, India

A. P. Siva Kumar Department of CSE, JNTUA, Anantapur, Andhra Pradesh, India

Nithya Rekha Sivakumar Department of Computer Sciences, College of Computer and Information Sciences, Princess Nourah Bint Abdulrahman University, Riyadh, Kingdom of Saudi Arabia

S. Sobin Soniya Department of CSE, Noorul Islam Center for Higher Education, Kumaracoil, Kanyakumari, India

Nirali A. Solanki SVIT Vasad, Vasad, India

Neeru Sood Department of Computer Science, Birla Institute of Technology & Science, Pilani, India;
Dubai International Academic City, Dubai, UAE

Sanjay Sood C-DAC Mohali, Punjab, India

Luis Alberto Soria Nuñez Facultad de Arquitectura Diseño y Artes, Pontificia Universidad Católica del Ecuador, Quito, Ecuador

Sadia Sorna Department of Computer Science and Engineering, Bangladesh University of Business and Technology, Dhaka, Bangladesh

H. N. Sowmya Nitte Meenakshi Institute of Technology, Bengaluru, India

Jagath Prasad Sreedhar Faculty of Statistics, University of Technology and Applied Sciences, Muscat, Sultanate of Oman

G. Sreenivasulu Department of CSE, ACE Engineering College, Hyderabad, India

M. Sreerag TIFAC-CORE in Cyber Security, Amrita School of Engineering, Amrita Vishwa Vidyapeetham, Coimbatore, India

G. Sridhar Department of Computer Technology, Anna University, Chennai, India

Sindhu Sridhar PES University, Bengaluru, India

M. Srikar Department of Electronics and Communication Engineering, B.N.M Institute of Technology, Bengaluru, Karnataka, India

Roopam Srivastava Shri Venkateshwara University, Gajraula, UP, India

Vandita Srivastava Centre for Converging Technologies, University of Rajasthan, Jaipur, India

B. Sujatha Department of CSE, University College of Engineering, Osmania University, Hyderabad, India

Suman Kumar Swarnkar Kalinga University, Raipur, India

Virendra Kumar Swarnkar Bharti College of Engineering and Technology, Durg, India

P. S. Swathi Department of Computer Science and Engineering, PDA College of Engineering, Kalaburgi, Karnataka, India

Abdoulaye Séré Laboratoire d'Algèbre, de Mathématiques Discrètes et d'Informatique (L.A.M.D.I), Université Nazi BONI (U.N.B), Bobo-Dioulasso, Burkina Faso

Vinod Talikoti KLE Technological University, Hubballi, India

Pritam Tamsekar School of Computational Sciences, Swami Ramanand Teerth Marathwada University, Nanded, India

N. Thulasi Chitra MLR Institute of Technology, Dundigal, Hyderabad, India

Farhana Sarmin Tithi Department of Computer Science and Engineering, United International University, Dhaka, Bangladesh

Aditi Tiwari Shivajirao Kadam Institute of Technology and Management—Technical Campus, Indore, India

Ayushi Tiwari Shivajirao Kadam Institute of Technology and Management—Technical Campus, Indore, India

Vivek Tiwari DSPM International Institute of Information Technology (IIIT), Naya Raipur, India

Cheick Amed Diloma Gabriel Traoré Laboratoire d'Algèbre, de Mathématiques Discrètes et d'Informatique (L.A.M.D.I), Université Nazi BONI (U.N.B), Bobo-Dioulasso, Burkina Faso

N. Uma Maheswari Department of Computer Science and Engineering, PSNA College of Engineering and Technology, Dindigul, Tamil Nadu, India

Preetham Umrani KLE Technological University, Hubballi, India

Jaimin N. Undavia Charotar University of Science and Technology, Changa, Gujarat, India

Shambhavi S. Uppin KLE Technological University, Hubballi, India

Priyalba Vaghela Devang Patel Institute of Advanced Technology and Research (DEPSTAR), Faculty of Technology & Engineering (FTE), Charotar University of Science and Technology (CHARUSAT), Anand, India

Nilay M. Vaidya Charotar University of Science and Technology, Changa, Gujarat, India

R. Vallikannu Department of ECE, Hindustan Institute of Technology and Science, Chennai, India

Pratik A. Vanjara Shree M. & N. Virani Science College, Rajkot, Gujarat, India

Deepesh Vanwani Department of Computer Engineering, National Institute of Technology, Kurukshetra, India

Abraham Varghese Faculty of Mathematics, University of Technology and Applied Sciences, Muscat, Sultanate of Oman

Shyla Manavalan Varghese Department of Mathematics, Houston Baptist University, Houston, TX, USA

Ramesh Vasappanavara Vasavi College of Engineering, Ibrahimbagh, Hyderabad, India

Prashant Vats Indira Gandhi Delhi Technical University for Women, New Delhi, India

K. Venu Madhav Department of CS&IS, School of Sciences, BIUST, Palapye, Botswana, South Africa

K. R. Venugopal Vice-Chancellor, Bangalore University, Bengaluru, Karnataka, India

Parul Verma Amity University, Lucknow, UP, India

N. Vismita Department of ISE, RV College of Engineering®, Bengaluru, India

Shreya A. Vyas SVIT Vasad, Vasad, India

Yashraj Wadalkar MIT Institute of Design, MIT-ADT University, Pune, Maharashtra, India

Smita Wagholikar SIT, Symbiosis International (Deemed University), Pune, MH, India

Pavit Wattanakul Department of Computer Engineering, Faculty of Engineering, Mahidol University, Salaya, Nakhon Pathom, Thailand

A Bio-Inspired-Based Degree Constrained MST Algorithm for Cognitive Radio Networks

Deepesh Vanwani and Mahendra Kumar Murmu

Abstract Cognitive radio network (CRN) has been observed as one of the most emerging technologies since last one decade and identified as a natural extension of wireless networks. The CRN consists of secondary user nodes (SUs) interconnected using spectrum holes in environment. In wireless networks, especially in distributed systems, constructing minimum spanning tree (MST) has been experienced as a classical problem. In this, finding degree constraint spanning tree is a well-known solution approach to construct MST. We intend to propose similar works to create MST for CRN using bio-inspired method. In our problem, we restrict every SU node to have a maximum of degree "d." We used the genetic algorithm along with ant colony optimization to solve the problem. It is the first algorithm of its kind to the best of our knowledge.

Keywords Cognitive radio networks · Ant colony algorithm · Secondary users · Primary user · Genetic Algorithm · Bio-inspired · Degree constraint · Minimum spanning tree

1 Introduction

Cognitive radio network (CRN) is frequently termed as cognitive radio ad hoc networks (CRAHNs). The research community has observed and stated that the CRAHN is a natural extension of traditional ad hoc networks. The CRN includes the following components; primary user (PU), secondary user (SU) and spectrum holes. The CRN consumes the unused bandwidth of licensed user (aka PU) in radio environment. The SU or cognitive radio is programmable radio, which dynamically configures to use the best wireless channel in its range to avoid user interference and congestion [1]. Presently, most of the devices use ISM bands within specified range of 900 MHz to 2 GHz. In results, it has been observed more congested. Therefore,

D. Vanwani (✉) · M. K. Murmu
Department of Computer Engineering, National Institute of Technology, Kurukshetra, India

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_1

CRN has been identified as the most appealing technology to resolve this problem which we can use to exploit the unused license bands in radio environment [2].

In cognitive radio networks, primary user is licensed user, who has the license to access the bandwidth [3]. When primary user is not using the bandwidth, it can be sensed and available for SU to use on temporary basis. The secondary user continuously learns the environment and tries to use the unused channels. The SU nodes access the available local channel set (LCS), opportunistically. On PU detection, SU nodes shift to another unused channel, if available, or neighbor nodes might be disconnected if there is no free channel [4]. In nutshell, the SU node continuously accounts the PU activity while creating CRN. In large networks, when source and destination nodes are situated far away, the data transmission becomes hazy. In other words, if data has to reach through multiple intermediate nodes to the destination node from any source node, the connections may face higher interruptions from PU. This can be compromised by creating short route between any pairs. Therefore, creation of a minimum spanning tree (MST) has been observed as a significant problem of cognitive radio networks [5, 6].

Recently, bio-inspired approach has been successfully applied to construct MST in cognitive radio networks [7]. We aim to use genetic algorithm (GA) and edge replacement method, jointly, while computing the MST in CRN. The genetic algorithm is used to sense the network and for pheromone distribution [8]. Then, ant colony optimization is used to construct the spanning tree. Once the spanning tree is created, the edge replacement method is used to optimize the solution or transform spanning tree into minimum spanning tree [9].

In general, MST is a kind of spanning tree with the total lowest possible weight and it is solvable in polynomial time for simple wireless networks [10]. However, if any constraint is added, while constructing MST, like adding a degree constraint on the nodes in the spanning tree, the problem becomes NP-Hard. In general, finding MST is the NP-Hard problem [11]. The problem is NP-complete if we successfully compute a minimum spanning tree in a given path whose maximum degree is "d." To accomplish this, heuristic approach is used. Using degree constraint, we implied that the degree of every node is less than or equals to "d." While constructing spanning tree, if the degree of any node becomes greater than "d," that edge will be rejected. In this paper, we have used two of the popular bio-inspired methods: genetic algorithm and ant colony optimization to construct the spanning tree. Also, we are using prim's-based algorithm to construct the spanning tree. After that, edge replacement method is used to optimize the solution [12, 13].

In this paper, our work has been organized as follows. In Sect. 2, the formal definition of the problem is given. In next Sect. 3, the system model and used data structures are defined. In Sect. 4, the proposed algorithm has been included. Also, a running example is shown for better understanding. Finally, we concluded our works in Sect. 5.

2 Problem Definition

The spanning tree "T" of an undirected graph "G" is a tree that includes all the vertices of "G," with a minimum no. of edges [7]. Our problem can be formally defined as: For a degree constraint undirected graph "G," with no self-loops and non-negative weights on the edges, find the spanning tree with minimum cost and maximum "d" degree for any node.

3 System Model

In this paper, we presume the graph model G = (V, E, W), with V vertices and E edges, and W represents the weight attached to every edge. Here, all the vertices denote the SU nodes in the system. We construct a minimum spanning tree for a graph for SU nodes, with the degree construction that no node can have degree greater than "d." We assume the following assumptions regarding the paper.

- *Distinct weight*: Every link between two SU nodes has given a unique pheromone value.
- *Genetic algorithm*: Using GA, we calculate the pheromone values between any pair of SU nodes. The GA uses to refine the solution until either the desired output has been achieved or the number of iterations is exceeded the predefined limit [14].
- *Ant colony optimization*: Ant colony optimization is used to construct the path to create the minimum spanning tree.
- *Communication channel*: These are used to create a graph between SU nodes. We assume that these channels are reliable during the construction of the MST.
- *PU interference*: If PU interferes during the construction of MST, the algorithm needs to be reinitiated.
- *Neighbor discovery*: Initially, the SU node will be in an idle state. In the idle state, it will generate a "hello" message. These "hello" messages are short messages, which SU transmits after every "t" seconds. In our paper, the interval t = 1 s has been assumed. When a SU node receives the "hello" message first time from a node "i," it will register "i" as the new destination in its routing table. After then, it expects a hello message from the node after every "t" seconds. If the node does not receive the "hello" message after t duration, it waits for a fixed number of t durations, before deleting the node from their routing table [15].

3.1 Data Variables

We have defined the different states in which a "SU" node in Table 1 and different states of edges in Table 2.

Table 1 States of SU nodes

Idle	Inactive SU node
Look	SU node starts looking for an optimal path to its neighbors
Found	SU node has a path to all its neighbors
Joined	SU node added to the MST

Table 2 States of an edge

Basic	Edge of SU node in connected graph
Deny	Rejected Edge
Accepted	The edge is added in the ST

Table 3 Data structures used in the algorithm

SU_i	i^{th} SU nodes
IN_i	i^{th} initiator node
Ph	Pheromone values
Nb_i	Neighbor list of i^{th} SU node
ST	List storing visited nodes
Cut	List storing minimum weight edges
T_m, T_{final}	List storing visited nodes
i_{max}	Maximum iterations (value = 10,000)
i_{stop}	Maximum iterations without improvement

3.2 Data Structures

The data structures used in our algorithm are mentioned in Table 3.

3.3 Message Types

In this model, we used different types of messages for the communication between Su nodes. The pheromone message will contain the source id and the pheromone value and the rejection message will be sent if the edge has been rejected (Table 4) [16].

Table 4 Types of messages used

Packet types	Format
Pheromone message	$< SU_i, Ph >$
Rejection message	$< SU_i$, "cycle form," $Nb_i - 1 >$

4 Proposed Algorithm

The proposed algorithm is used to find the best path between any pair of SU nodes using the GA. By using genetic algorithm, we try to minimize the cost. Different processes like selection, crossover and mutation, are performed to generate results. It is generally better than using techniques like the broadcasting. Once pheromone is distributed, ant colony optimization is used to construct the spanning tree. Ants follow the pheromone trails to trace the path and construct the spanning tree. After then, One-Edge replacement () method is used to construct the MST. Initially, all SU nodes are assumed to be idle. A new SU node can be detected by using "hello" messages. Each SU node maintains a dynamic list of its neighbors [17, 18].

We consider three parameters to define our cost. We assume that:

Encoded power: 3 bits
Encoded bit error rate: 3 bits
Data rate: 3 bits
Total bits: 9 bits.

Every path is a combination of 9 bits. We calculate the fitness value for each path and perform "rank"-based selection to pair chromosomes with higher fitness values as parents (Fig. 1).

0	0	1	1	1	0	1	0	1

1	1	1	1	0	0	0	0	0

Now, we perform "one-point crossover" to generate new generations [19]. (Performed after fourth bit.)

1	1	1	1	0	0	0	0	0
0	0	1	1	1	0	1	0	1

After crossover, we get the following results.

0	0	1	1	0	0	0	0	0

1	1	1	1	1	0	1	0	1

Now, we perform "bit flip mutation" over new chromosomes to avoid local minima. Here, in the first chromosome, we flip the first bit, and in the second chromosome, we flip the last bit.

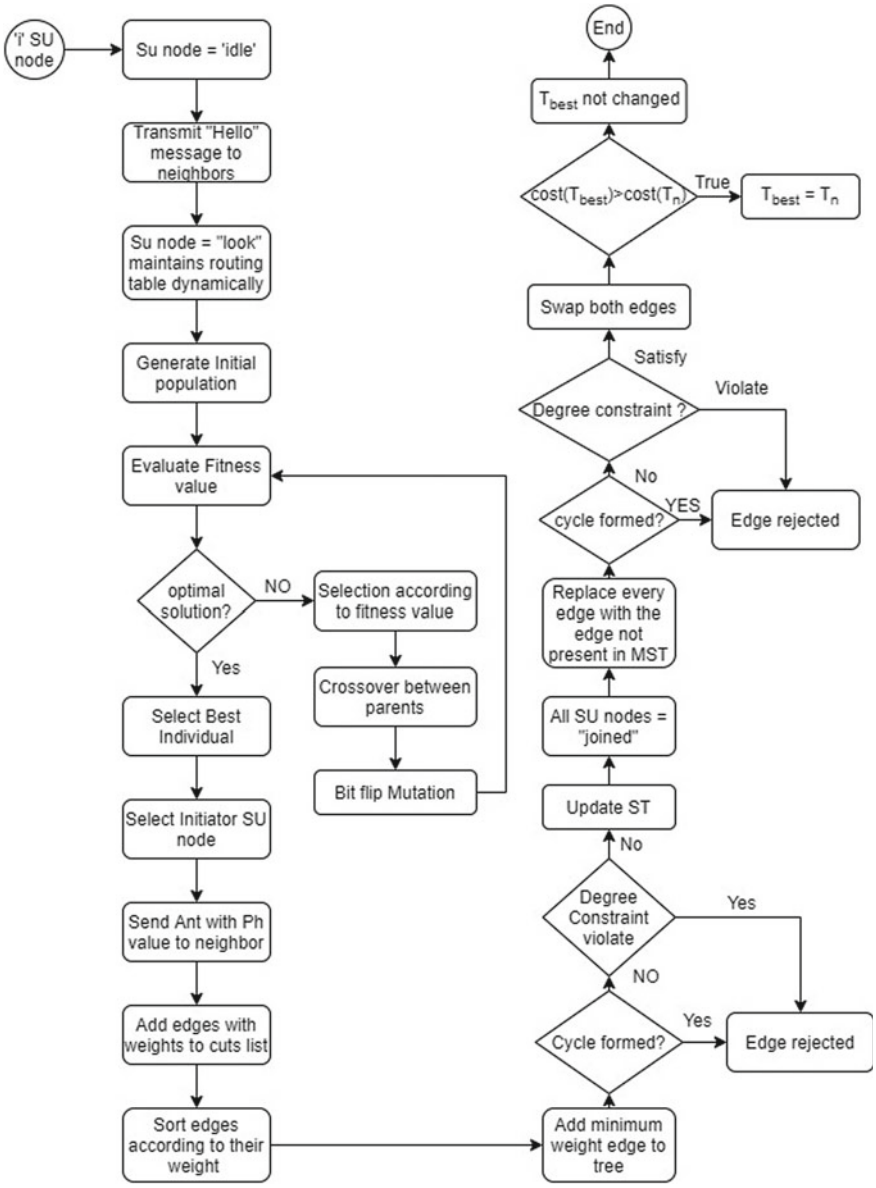


Fig. 1 Flow chart of the algorithm

1	0	1	1	0	0	0	0	0

1	1	1	1	1	0	1	0	0

After mutation, we add these new chromosomes to the previous population to generate a new population. Then calculate the overall fitness value of the new generation. If the overall fitness value of the new generation is better than the old generation, we keep the new population, or else we discard it. Go back to selection step and iterate the whole process for predefined number of times [20].

Check for the termination condition; if satisfies, then terminate and select the chromosome with minimum value.

1	1	1	1	1	0	1	0	0

Send an ant with pheromone value = (511 − cost) to destination. [Max possible value 511 for 9 bits number]. Here, the cost is 500. So, the pheromone value will be = 511 − 500 = 11.

To compute the values, the following formulas are used:

Structure of chromosome:

- *Data rate*
- *Bit error rate encoded*
- *Encoded power*.

Some important formulas [21]:

Fitness function:

$$F = w_{dr}.f_{dr} + w_{pwr}.f_{pwr} + w_{ber}.f_{ber} \tag{1}$$

Here,

$$f_{dr} = f(x) = (d_r - d_c)/d_r \tag{2}$$

$$f_{pwr} = 1-(P/P_{max}) \tag{3}$$

$$f_{ber} = \begin{cases} \log10(\overline{Pbe})/\log10(\overline{\text{Pminber}}), & \text{if } P_{be} \text{ in } 10-3 \\ 1, & \text{Otherwise} \end{cases}$$

Here, d_r is the data rate, and dc is the data rate available on the channel.

The transmission power on a given channel is P, and P_{max} is the max available power to the CR user. P_{be} is the average BER, and P_{minber} is the minimum value of the BER (Fig. 2).

```
Initialize
Su node = idle
For all Suj node == idle
    Suj node = look
    Discover neighbors list Nbi
    While Nbi.size! =0 && SUj node==found
        While I < number of iterations:
            N = Total number of chromosomes
            Calculate fitness value for each chromosome
            Perform Rank based Selection
            Perform One-point Crossover
            Perform Bit Mutation
            Append new species to old generation to generate new generation
            I+1
        End while
        Select the chromosome with the least cost
        Calculate Pheromone value
        Send an ant with Pheromone message to SUj
    End While
    Set INi as the initial SU node having SU node=found
        INi Su node = joined
        Update ST with INi
    Check Edge weight for all Nbi
    While NBi.size! =0
    If (edge weight SUj > edge weight SUj+1 and SUj == Found)
        If (degree constraint not violate)
                SUj edge = accepted
                Update cuts list with SUj edge
                Update ST with SUj
               SUj = joined
                Else if SUj == joined
            SUj edge = deny
    End while
End For
Tfinal = ST

i←1, ibest = 0;
while i<imax and i-ibest<istop
    Tm = One-Edge_Replacement();
    if (W( Tfinal ) > W( Tm ) ))
        Tfinal = Tm;
End while
Return Tfinal
End
```

Fig. 2 Proposed Algorithm

Select a SU node with status = Found and set it as the root node. There can be at most of d = 3 degrees of any node in this example. Here "0" is the initiator node. Change the status of the node to "joined." It compares the cost values of all its neighbors and selects the one with the lowest among them. After adding the node in the ST, we change the status of the node to "joined," so we do not add that node again and can avoid cycles. Also, we need to check for the degree constraint condition to be satisfied. When the destination node is reached, or all nodes are covered, terminate the process (Fig. 3, 4 and 5) [22].

Here, edge (0,5) has the lowest weight among the remaining edges but, if the edge gets added in the tree, it will violate the degree property of node 0. That is why this edge is not included in the tree (Figs. 6 and 7).

The total weight of the spanning tree is $(1 + 1 + 1 + 5 + 6 + 8 + 9 + 10) = 41$.

The generated spanning tree will be sent as input to the "One-Edge_Replacement ()" method. In this method, we are trying to find out if any other spanning tree exists

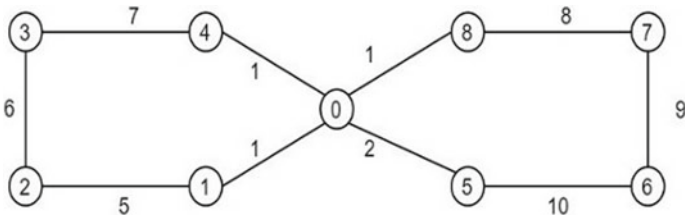


Fig. 3 Logical graph structure in CRN, b selecting path with the lower cost value, c update cuts list

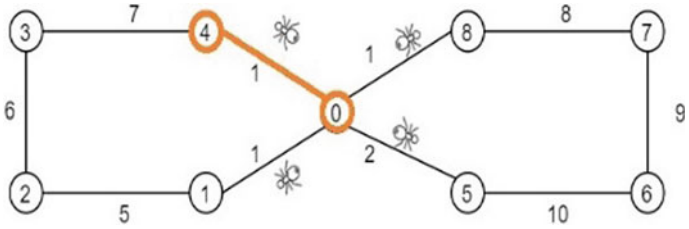


Fig. 4 Selecting path with the lower cost value

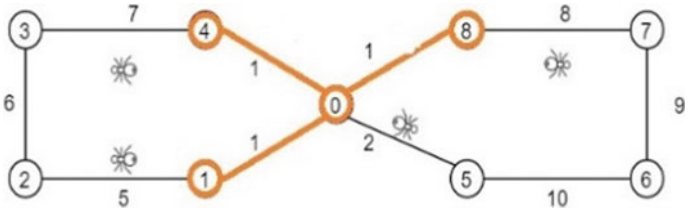


Fig. 5 Update cuts list

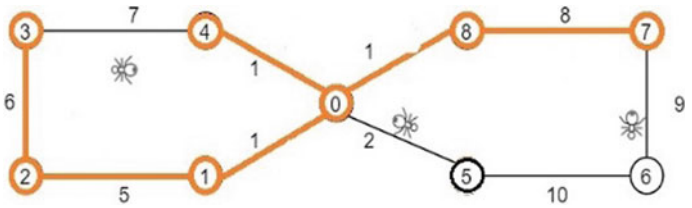


Fig. 6 Update ST

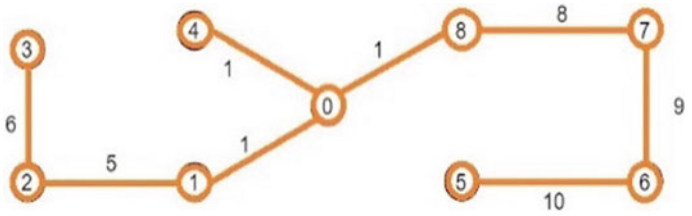


Fig. 7 ST formation

with a total lower cost than the input spanning tree. If it does, the T_{final} will be replaced by the output of the "One-Edge_Replacement () method" [23, 11].

To find this out, every edge in the spanning tree is swapped by the edge in the cuts list, which is not included in the spanning tree. If such swapping does not create a cycle or does not violate the degree constraint property, the flag "change" becomes true. The process continues until all edges combinations are considered. If any of the trees with swapped edges have lesser total weight than input spanning tree, then we replace the T_{final} with the new tree (Fig. 8).

To find this out, every edge in the spanning tree is swapped by the edge in the cuts list, which is not included in the spanning tree. If such swapping does not create a cycle or does not violate the degree constraint property, the flag "change" becomes true. The process continues until all edges combinations are considered. If any of the trees with swapped edges have lesser total weight than input spanning tree, then we replace the T_{final} with the new tree (Figs. 9 and 10).

Here, while executing "One-Edge_Replacement ()," we replace the edge (0,5) with the edge (0,8). While doing so, we check for cycle formation and degree constraint violation. Since both conditions are satisfied, we can swap these two edges. Since a change has occurred, set the "change" flag true and break the loop. Now, the method will be executed again on the new tree.

Similar to the above description, we swap two edges edge (0,4) with (0,8) and (5,6) with (3,4). Now, the total cost of the spanning tree gets computed.

Here, the total cost of the tree is $(1 + 1 + 2 + 5 + 6 + 7 + 8 + 9) = 39$, which is less than the cost of input spanning tree (41). So, the T_{best} will be replaced by the newly generated tree.

```
One-Edge_Replacement (T, E, V, W)
{
        T_final = T;
    T_m = T;
        Sort Edges in decreasing order of their pheromone values
    Repeat
          Change = False
                Sort T_m in increasing order of their edge costs
                j= v-1
       while j >= 0
        m = 0
       while (E)
                if swapping T_m[j] and E[m] does no cycle creation and no
                Degree constraint violation
            E← E U T_m[j]
            T_m = T_m U {E[m]} – T_m[j]
            Change = true
            break;
            m++;
           j = j-1
    Until not changed;
    Return T_m;
}
```

Fig. 8 One-Edge_Replacement()

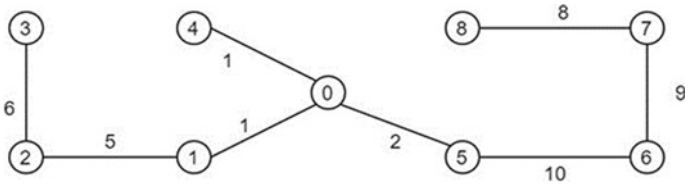


Fig. 9 Swapping edges

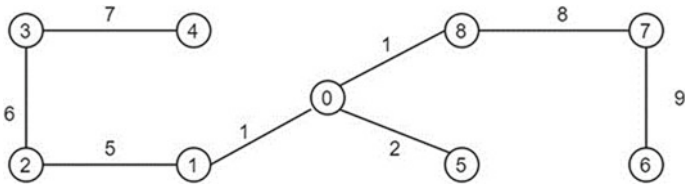


Fig. 10 MST formation

5 Conclusion

In this paper, we have proposed to construct the degree constraint minimum spanning tree algorithm from CRN. It is a bio-inspired algorithm and running example shows its simplicity and effectiveness of the protocol. We have identified AntHocNet patch to implement our work and detail simulations have not been included in this paper. To the best of our knowledge, it is the first protocol in this category.

References

1. Gao, X., Jia, L.: Degree-constrained minimum spanning tree problem with uncertain edge weights. In: School of Mathematical Sciences and Physics (2016)
2. Akyildiz, I.F., Lee, W.Y., Chowdhury, K.R.: CRAHNs: cognitive radio Ad Hoc networks. Ad Hoc Netw. **7**(5), 1–27 (2009)
3. Salgueiroa, R., de Almeida, A., Oliveira, O.: New genetic algorithm approach for the min-degree constrained minimum spanning tree. J. Oper. Res. **258**, 877–886 (2017)
4. Shi, K., Song, Q., Lin, S., Xu, G., Cao, Z.: An improved genetic algorithm for degree constrained minimum spanning trees. In: Chinese Control and Decision Conference, pp. 4603–4607 (2016)
5. Murmu, M.K.: A distributed approach to construct minimum spanning tree in cognitive radio networks. In: International Conference on Eco-friendly Computing and Communication Systems (2015)
6. Akyildiz, F., Lee, W.-Y., Vuran, M.C.: Next generation/dynamic spectrum access/cognitive radio wireless networks: a survey. Comput. Netw. **50**(13), 2127–2159 (2006)
7. Shetty, A., Puthusseri, K.S., Shankaramani, D.R.: An improved ant colony optimization algorithm: Minion Ant (Mant) and its application on TSP. In: IEEE Symposium Series on Computational Intelligence (2018)
8. Genetic algorithm, Wikipidea (2020)
9. Caro, G.D., Ducatelle, F., Gambardella, L.M.: AntHocNet: an ant-based hybrid routing algorithm for mobile Ad Hoc networks. In: Istituto Dalle Molle sull'Intelligenza Artificiale (IDSIA)
10. Guo, W., Zhang, B., Chen, G.: A PSO-optimized minimum spanning tree-based topology control scheme for wireless sensor networks. Int. J. Distributed Sensor Netw. (2013)
11. Cognitive radio. Wikipedia (2020)
12. Zhang, Y., Li, L.: MST ant colony optimization with Lin-Kerninghan local search for the traveling salesman problem. In: International Symposium on Computational Intelligence and Design (2008)
13. Ruzika, S., Hamacher, H.W.: A survey on multiple objective minimum spanning tree problems. Algorithmics Large Complex Netw. **5515**, 104–116 (2009)

14. El Morabit, Y., Mrabti, F., Abarkan, E.H.: Spectrum allocation using genetic algorithm in cognitive radio networks. In: RFID And Adaptive Wireless Sensor Networks, pp. 90–93 (2015)
15. Sun, X., Chang, C., Su, H., Rong, C.: Novel degree constrained minimum spanning tree algorithm based on an improved multicolony ant algorithm. In: School of Computer Science and Software (2015)
16. Alam, S., Marcenaro, L., Regazzoni, C.: Opportunistic spectrum sensing and transmissions. In: Cognitive Radio and Interference Management: Technology and Strategy, pp. 1–28 (2012)
17. Stanzin, T., Murmu, M.K.: A Bio-inspired approach to construct minimum spanning tree in cognitive radio networks. In: International Conference on Communication and Signal Processing (2018)
18. Gallager, R.G., Humblet, P.A., Spira, P.M.: A Distributed algorithm for minimum weight spanning trees. ACM Trans. Program. Lang. Syst. **5**(1), 66–77 (1983)
19. Arun, J., Karthikeyan, M.: Optimized Cognitive Radio Network Using Genetic Algorithm, vol. 22, pp. 3801–3810 (2019)
20. Ning, A., Ma, L., Xiong, X.: A new algorithm for degree-constrained minimum spanning tree based on the reduction technique. In: School of Management (2007)
21. Nayyar, A., Singh, R.: Ant colony optimization computational swarm intelligence technique. In: International Conference on Computing for Sustainable Global Development (2016)
22. Coloroni, A., Dorigo, M., di Elettronica, D.: Distributed optimization by ant colonies. In: European Conference on Artificial Life, pp. 134–142 (1992)
23. Bui, T.N., Deng, X., Zrncic, C.M.: An improved ant-based algorithm for the degree-constrained minimum spanning tree problem. IEEE Trans. Evol. Comput. **16**(2), 266–278 (2012)

Faceless, Cashless and Paperless Scheme for Farmer?—Implementation of Mukhya Mantri Krishi Ashirvard Yojana (MMKAY) in the State of Jharkhand Through ICT: A Case Study

Jyoti, Pradeep Kumar Hazari, Satyendra Kishore, and Pranab Kumar

Abstract As per the census 2011, there are 38 lakh farmers in the state of Jharkhand and farmers' welfare is always the top priority of government. Ensuring the reaching of benefits of schemes to farmers is also a challenging issue before government. Challenges become more challenging when desired infrastructure to reach the beneficiaries is not robust or fully developed. This study focused on the use of ICT tool for implementation of public scheme by the government and role of ICT to solve many operational issues. Use of ICT is one of the efficient tools for brining transparency and efficient in implementation of public scheme.

Keywords ICT · Government scheme · DBT

1 Introduction

As per the census 2011, there are 38 lakh farmers in the state of Jharkhand and farmers' welfare is always the top priority of government. Ensuring the reaching of benefits of schemes to farmers is also a challenging issue before government. Challenges become more challenging when desired infrastructure to reach the beneficiaries is not robust or fully developed. The biggest challenge before government was how to deliver service effectively to farmers without face to face contact with government officials, transfer of benefit into the bank account of farmers through online mode and minimum use paperwork. However, clarity on concept, procedure, objective, dedication, fast decision, sticking to timelines, good teamwork, use of Information and communication and technology can overcome the challenges make the service delivery reach to the farmers at the doorstep effectively.

Jyoti (✉) · S. Kishore
ICFAI University, Jharkhand, Ranchi, India

P. K. Hazari · P. Kumar
Birla Institute of Technology, Mesra, Lalpur Ext Ranchi, Jharkhand, India
e-mail: pranabkumar@bitmesra.ac.in

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_2

Mukhya Mantri Krishi Ashirvard Yojana (MMKAY) is one such example which has been recently implemented in the State of Jharkhand by Department of Agriculture, Animal Husbandry & Cooperative. It is a cash-based incentive schemes for small and marginal farmers. This is one of the flagship scheme of Government of Jharkhand which was designed to implement through online one so that objective of transparency in government service to public and effective use of public money can be achieved. Eventually, this scheme has been well accepted among farmers and helped the state to build more trust and confidence among rural people and also improving the government system credibility and impact rural economy positively (Medani L). The online technology played a very significant role in the implementation of this scheme to reach the farmers and delivery of services and made this a faceless, cashless and paperless programme [1, 2].

2 Conceptualization of Mukhya Mantri Ashirvard Yojana

At present, the Department of Agriculture is running 78 states funded and 26 centrally sponsored schemes (Total 104) in the state which are farmers centric.

During 2018–2019 financial year, the government came up with an idea to provide a direct cash benefit to small and marginal farmers through DBT mode for speedy and timely delivery of benefits to the farmers. The idea was conceived during December 2018 and a team of 5 IAS officers was formed to design the frame work of scheme and action plan to implement it before March 2019.

The study team has prepared a detailed concept note of the scheme for the state while taking the learning from schemes of other states also. The concept note on the scheme was presented before high-level committee for discussion. A high-level committee after detailed deliberation accepted the concept.

3 What's MMKAY?

There are approximately 28 lakh small and marginal farmers in the state whose holding is up to 5 acres. These small and marginal farmers are eligible under MMKAY. The eligible farmers are given cash benefit of minimum of Rs. 5000 per acre up to maximum of Rs. 25,000 (up to 5 acres) based on holding size through DBT mode in three instalments towards the purchase of input like seed, fertilizers, pesticides and labour charges for cultivation purpose. This scheme aims at reducing distress purchase of inputs and doubling farmers' income. This scheme came into force in the state from 1 April 2019. For this purpose, Rs. 2000 crore was earmarked in budget. This was the highest budget allocation so far for single scheme under agriculture department.

Basis of payment to the farmers in MMKAY			
Benefit eligibility			
Sl. No.	Landholding	Amount payable (in Rs)	Remarks
1	Up to 1 Acre	5000	Flat payment
2	More than 1 Acre and Less than 5 Acre	More than 5000 and less than 25,000	Pro rata basis
3	5 Acre	25,000	Maximum payable
4	More than 5 Acre	Nil	No Payment

Illustration			
Sl. No.	Landholding	Calculation	Amount
1	0.1 Acre	Flat rate	5000
2	0.99 Acre	Flat rate	5000
3	1.00 Acre	1 × 5000	5000
4	1.02 Acre	1.02 × 5000	5100
5	1.5 Acre	1.05 × 5000	7500
6	2.3 Acre	2.3 × 5000	11,500
7	3.69 Acre	4.69 × 5000	18,450
8	4.99 Acre	4.99 × 5000	24,950
9	5.00 Acre	5 × 5000	25,000
10	5.1 Acre	0	0

4 Objective of Scheme

Based on the background of study and discussion, the following objectives were defined for the MMKAY

1. To contribute towards doubling of the farmers income.
2. To save farmers from vicious cycle of debt trap and to use the cash for cultivation purpose.
3. To provide the cash benefit to the farmers through DBT mode.
4. To make whole operation online to minimize the hassle.

5 Operational Challenges for the Scheme

To implement this scheme, the biggest challenge before the department was to identify the real landowner. Modernization and digitalisation of land record under

NLRMP is not up to the desired level. Further, tenancy acts namely Chhotanagpur Tenancy Act (CNT Act) and Santhal Pargana Tenancy Act (SPT Act) are stringent in terms of land transfer from one class of landowner to another class. This has resulted in fewer changes on land records in terms of number of mutation and more succession/transfer in real sense which is off the land record. Land revenue matters fall under Revenue and Land Reforms Department.

The second challenge was to obtain the database and maintain the database of farmers along with their details like identity, bank details and other records which are instrumental for online transfer payment.

The third major challenge was to transfer fund directly to eligible farmer's bank account seamlessly.

The fourth challenge was coordination and monitoring with revenue and agricultural officials of all 24 districts on real-time basis.

Lastly, the timeline to implement this scheme was very limited, i.e. 30 June 2019, i.e. before onset of monsoon season and cropping season.

Other peripheral challenges were to choose technology partner, development of software and integration with banks and Public Finance Management System (PFMS); training and capacity building of staff and officials.

6 Action Plan for Overcoming of Operational Challenges

To overcome the above challenges, it was decided to develop online system and it should be run through a single web portal. After several rounds of discussions with Revenue Dept and other departments such as IT Dept and Finance Dept, an action plan was proposed by the Agri Dept before government for perusal and approval. National Informatics Centre (NIC) was the technology partner in execution of this scheme.

All department Agriculture Department, Revenue Department, National Informatics Centre (NIC), District Administration, Banks and PFMS (Public Finance Management System, GOI), UIDAI(regional office) all worked together to implement this scheme under central control room set up at Agri Dept, Ranchi. Director, Agriculture was designated as State Nodal Officer (SNO).

The online system was developed by NIC. Dry run was conducted in one of the circles (tehsil) before it was launched. Software was under continuous modifications as data entry, verification, transfer of data from one level to another level progressed and even at payment level.

This scheme heavily relied on efficiency of district administration in terms of collection/verification/uploading of data, transfer of fund and follow-up.

Social media and technology allowed us to work as a team. Whatsapp group was created where more than 150 officials/technical staffs right from secretary to Collectors to Agriculture officers, NIC officials and other department officials, date entry operators and bankers were connected through Whatsapp. It helped to identify problem at the earliest while running software. Swift response was the key. It enabled

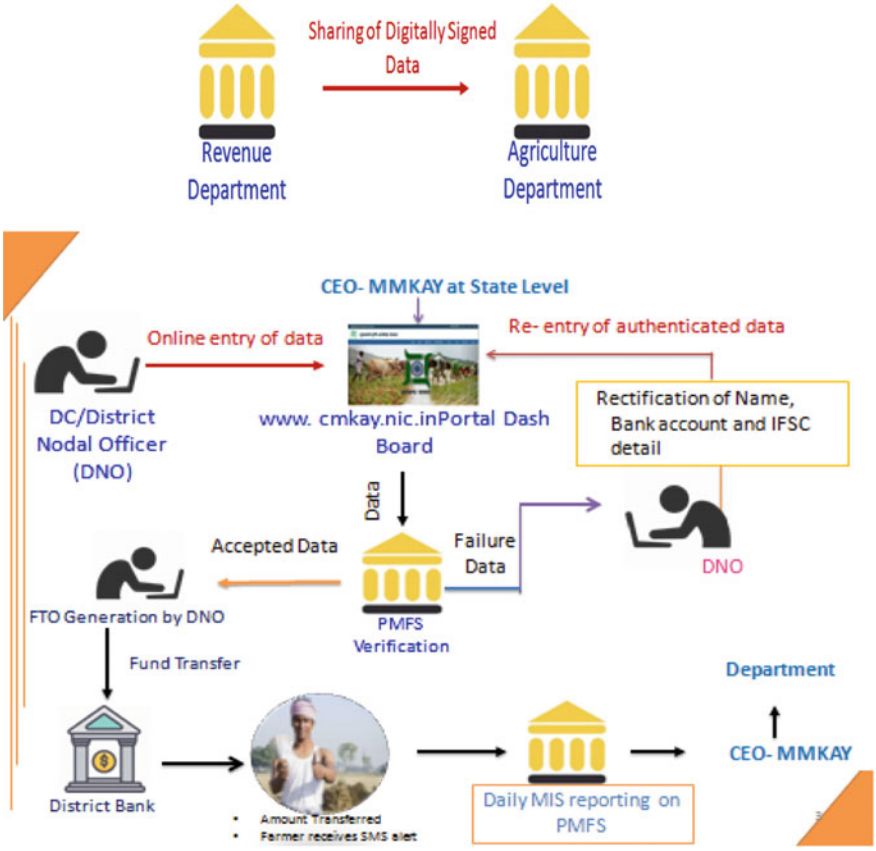
us to work as team and learn from others understandings. At state head quarter, a strong control room created under control of Secretary, Agriculture and round o clock feedback system through Whatsapp was working to support the whole operation [3].

A programme management unit (PMU) was set up with resource persons having diversed background provided by Ernest and Young (E&Y).

Grievance Redressal System was set up at district level which was headed by Additional Collector or District Agriculture Officers level.

7 Process Flow of Whole Operation

The scheme could take off only when farmer's details received from revenue department in digital form.



The verified farmers details were entered directly into official web portal (www.cmkay.gov.in) directly at district level round the clock so that maximum number of

beneficiaries could be covered under the stipulated timeline. The web portal was managed by NIC and data processing was also the responsibility of NIC.

The entered data of farmers were sent to PFMS for processing to make arrangements for release of payment. The PFMS validated data were sent for release of payment through Adhaar-seeded bank account to ensure the transfer of benefits to right farmer. Payment was integrated with mobile-based SMS so that farmer could receive instant SMS on transfer of online payment.

The physical and online verification and authentication of farmer's data were done at 7 levels including gram sabha level. Village-level revenue committee was formed which included village headman, elected representatives, woman and other village members. This committee was mandated with the task of verification of real landowner in case of any dispute. Hence, its role was very critical in the verification. If the beneficiary had any grievances, they could call grievances cell of their block for this scheme. So, the scheme was faceless—no need to contact government official for record verification of entry of his/her name [4].

8 Biggest Challenge and Decision of Government

The strategic decision in view of poor state of updation/modernization of land records of state government played the biggest role in maximum coverage of number of farmers in the state. Succession, mutation and updation of land records were a cumbersome task. It was difficult to achieve the target of covering 25 lakh small and marginal farmer in given timeline. The state took deliberated decision to allow village-level revenue committee to verify the farmers in absence of any legal mutation/succession or in case of any dispute. This decision made the operation easier covering more number of farmers.

9 Output and Status of Scheme

1. So far, more than 25 lakhs farmers have registered on MMKAY portal.
2. Whole process is being done through online hence office is free from paper.
3. 16.15 lakh farmers have received payments. Amount is transferred through DBT which is cashless.

10 Lessons Learnt and Way Forward

1. Inter-departmental and inter-agency coordination are important in sorting out issues.
2. Regular and close monitoring of each steps at each level and of all officials.

3. Technology partner should be fully aware of objectives/timeline/level of verification.
4. Dry run of online operation is essential before rolling out.
5. Leadership role of district collector and suggestion from them helped in improving the online system.
6. Random verification by senior district official is also important to keep revenue official in check.
7. Full integration of online system with PFMS will make payment system robust.

11 Conclusion

Therefore, it can be concluded that successful implementation of beneficiary oriented schemes of state government where large number of beneficiaries to be benefitted, the role and use of ICT is critical and MMKAY portal is one of the best example. Use of ICT in this case not only helped in reducing the number of operational transactions but also helped to bring high-level transparency in terms of the use of public money. Moreover, it helped the government machinery to reduce the public grievances to the maximum level.

Photos and Pictures on Mukhya Mantri Krishi Aashirvad Yojana

References

1. Memdani, L.: An overview of digitalization of rural india and its impact on the rural economy an overview of digitalization of rural india and its impact on the rural economy. https://www.researchgate.net/
2. Bhatia, A., Kiran, C.: Rural development through E- Governance Initiatives in India. IOSR J. Bus. Manage. (ISSN 2278-487X p- ISSN:2319-7668) https://pdfs.semanticscholar.org/911b/28fd574c18d07a7adea1132fa1e907405965.pdf
3. Kumar, T., Prdhi, S., Bisen, A.: Role of digital india in rural areas" International seminar on non-conventional energy sources for sustainable development of rural areas. Int. J. Adv. Eng. Res. Dev. (ISSN:2348–4470, p-ISSN:2348-6406) http://www.ijaerd.com/papers/special_papers/ISNCE08.pdf
4. Mishar, D.C., Mittal, P.K.: e- Governance and Digitization of Indian rural development. Conference paper. https://www.researchgate.net/

Extraction of Clean Speech Along with Emphasis on Essential Noise

V. Mohan, P. Shanmugapriya, and A. Sharan Jasmine

Abstract Recently, the extraction of clean speech from the noisy speech has gained significance in designing the smart hearing aid by using deep neural network-based speech enhancement techniques. In the conventional denoising process, irrespective of the type of the noise, noise was totally removed from the noisy speech signal. But some of the noises such as alert sounds, siren, fire alarm, baby cry are essential and also known as desired noise. Due to this problem, people with hearing issues were troubled to hear these essential noises. The proposed work provides a solution for this problem by classifying the noises into desired and undesired noise using deep learning technique with convolutional neural network. The desired noise is fused with the clean speech and provides the desired noisy speech signal. Speech enhancement is carried out using a suitable Deep Neural network while classification of noise is carried out using Convolutional Neural Network. Experimental setup for extraction of clean speech segments from noisy version, classification of noise based on its characteristics and weighted fusion desired noise with the enhanced speech have been presented in this proposed work. By using deep learning techniques, the better accuracy has been achieved.

Keywords Denoising with essential noise · Classification of noise · Deep neural network · Convolutional Neural Network and Smart Hearing Aid.

1 Introduction

Speech is the basic way for humans to convey the information to one another. It can be defined as the response of the vocal tract to one or more excitation signals [1]. Mostly speech signals are highly non-stationary process with multiple component

V. Mohan (✉) · P. Shanmugapriya · A. Sharan Jasmine
Department of ECE, Saranathan College of Engineering, Trichy, Tamilnadu, India
e-mail: mohan-ece@saranathan.ac.in

P. Shanmugapriya
e-mail: shanmugapriya-ece@saranathan.ac.in

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_3

that may vary in time and frequency. It consists of sequence of sounds, and the transition occurs between the sounds that can carry the information that needs to convey.

In the recent years, speech enhancement process plays vital role on removing the noise in the noisy environment. Most of the researches were done in the single channel speech enhancement technique. The goal of the speech enhancement process is to improve the intelligibility and quality of the speech signal. This is the growing challenge in the real-time applications such as mobile communications, hearing aid and robust speech recognition. At first, the enhancement process done by using Wiener filtering, but some of the musical noise were maintained in the enhanced signal in [2].

After that, the minimum mean square error (MMSE)-based estimation was introduced in [3] for the noise reduction and for speech enhancement process. They focused on the statistical difference between the noise and the speech. In [3], some of the methods like estimation of the power spectral density (PSD) of speech and noise have been used for the denoising process. Then, to learn the higher order statistical information for the enhancement of speech, the deep learning techniques were introduced in [2–8]. In [9] and [10], the audio encoder and the audio denoiser with machine learning algorithms were proposed to enhance the speech.

Audio scene classification is performed through deep belief network [DBN] in [11]. The pre-training has to be done by using restricted Boltzmann machine (RBM) that learns the higher-level features from the datasets in an unsupervised manner. Then, fine-tuning has been done through feed forward neural network (FNN). Initialization of weights for the higher-level features is done through FNN to find the better parameters. These processes were done in the training stage. Then, the regression-based approaches were introduced for the speech enhancement process. MMSE spectral magnitude estimation-based parameters are used for the enhancement of speech in [4]. Raw waveform-based approaches for speech enhancement was implemented through convolutional neural network (CNN) in [11] and [6]. There were many different machine learning techniques introduced for speech enhancement process in [1, 4, 5] and [7]. In all these techniques, the noise was completely removed from the speech signal which is undesired one for smart hearing aid. The main problem is to retain the essential noise without damaging the speech signal. The problem can be solved by analyzing the noise that is to be eliminated from the signal and classifying them under essential and non-essential noise. The essential noise has to be extracted from the speech, and speech signal is enhanced with the essential noise.

In this proposed work, clean speech is obtained through denoising and enhancement using DNN. Analysis and classification of noise are performed through CNN. Based on the classification result, the parameters of essential noises are extracted and speech is enhanced. The deep neural network is used for the enhancement of speech and provides the better performance.

1.1 Importance of Essential Noise

The human–machine interaction system may suffer from the background noises. Especially, the sensorineural system of the hearing-impaired person could be affected in the noisy environment. Still the researches were done in the removal of whole noise and provided clean speech signal by using deep learning algorithms [5]. The mapping of the clean speech and noisy speech done by the supervised learning on regularization techniques. But the hearing – impaired persons are facing the difficulty on hearing the clean speech in the noisy environment. They have the trouble on hearing the alert sound while speaking in the noisy environment.

The classification of sounds still occurs predominantly by applying general classifiers (hidden Markov model, Gaussian mixture model, super vector machine, etc.) that could manually extract the features such as Mel-frequency cepstral coefficients [12]. Some of the applications were presented at the mobile phones for providing alert sounds for the deaf people. That application issued the alert by enhancing the activity or movement of the user at the time of issuing the alert [13].

1.2 Convolutional Neural Network

Later, the sounds were classified by using the deep learning architecture such as convolutional neural network (CNN) [14]. CNN performs classification of sounds, traffic signals, microscopic images, etc., and CNN consists of an input layer, group of convolutional and pooling layer and a limited number of fully connected hidden layer. The input layer accepts the three-dimensional input and passed to the feature extraction layer. The convolutional layer extracts the feature, the pooling layer reduces the dimension of the features, and then, the output layer carries the estimation of that features.

Some of the hearing aids were introduced to remove the babble noise in the lower SNR environment. In that, CNN performed better than the recurrent neural network or feed forward neural network (FNN)with smaller-sized network. The redundant convolutional encoder decoder (R-CED) was presented for the denoising process in [15]. Still the CNN performed the classification of the sounds generally. But here, we used the CNN for the classification of noises.

In this proposed work, the noises can be classified into desired and undesired noise by using CNN. The desired noise is the alerting sounds such as fire alarm, sirens, baby cry and car horn, and the undesired noise is such as crowd noises, babble noise, pink noise and brown noise. The desired noise is considered as the essential noises. Here, the desired noise got fused with the extracted clean speech and presented the clean speech with essential noise. This work is more efficient and useful for the crime sectors, hearing aids, etc., and the detailed procedure of this work is explained as follows.

2 Proposed System

2.1 Speech Enhancement Process

Employing deep neural network for mapping the clean speech from noisy speech for speech enhancement application is the objective of the proposed system. In Fig. 1, the processes such as frame blocking using windowing, spectral estimation of noisy speech samples using STFT, extraction of clean speech spectrum and synthesis of clean speech are illustrated. The mathematical model [16] for the noisy spectrum in presented in Eq. (1) (Fig. 2).

$$Y_{k,l} = S_{k,l} + N_{k,l} \tag{1}$$

where $S_{k,l}$ represents the spectral coefficients of the clean speech component, and $N_{k,l}$ denotes the noise spectral coefficients.

The spectral coefficients are obtained by applying the short time Fourier transform (STFT) on the frames of speech signals where the speech segments are generated with 50% overlapping. In the laboratory environment for experimentation purpose, the noisy speech signals are generated with the control over the amount of noise added by using the additive corruption model. The selection of noise type for the generation of noisy speech signals is also under the control of the user.

Apply square hann window for synthesis and use overlap–add method to reconstruct the time domain signal. The power spectral density (PSD) for noisy and speech signal as in [3, 17].

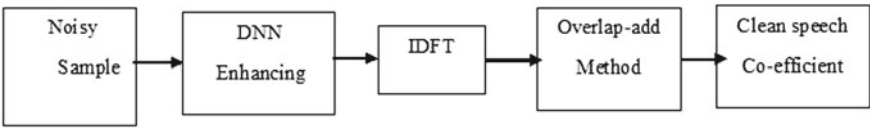


Fig. 1 Pre-processing of speech

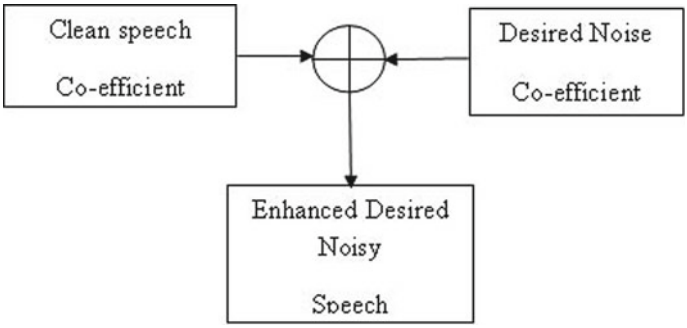


Fig. 2 Estimation of enhanced desired noisy speech

$$P_n(K) = E\{|N_K|^2\} \tag{2}$$

$$P_S(K) = E\{A_K^2\} \tag{3}$$

$P_n(K)$–PSD of noise signal.
$P_S(K)$–PSD of speech signal.

$A_K = |S_K|$ is the spectral amplitude derived from MMS estimation technique. Calculate the signal-to-noise ratio of the signal.

$$\varepsilon_k = P_S(K)/P_n(K) \tag{4}$$

The values are normalized to zero mean and unit variance. There are two process have to be performed in this process. They are pre-training using the RBM and fine-tuning using FNN [18]. The normalized values enter into the restricted Boltzmann machine (RBM) where the higher-level features are extracted on performing the pre-training process by attaining unsupervised learning. The extracted features enter into the feed forward neural network (FNN) for the mapping function done by supervised learning. The targeted outputs are speech and noise of interest.

The dropout regularization techniques are applied to drop out the random neurons during the training process which is useful to generalize the network. This information was fed into the DNN to perform the noise aware training process by adding the noise estimator to predict the desired signal. At the end of the enhancement stage, the desired signal $\hat{X}(d)$ is obtained through mapping using the DNN model. The architecture of DNN used in this work is illustrated in Fig. 3. The desired speech segment is further leveled using post-processing process.

Then, the speech segment is post-processed to remove the effects of normalization by multiplying by an equalization factor η.

$$\hat{X}'(d) = \hat{X}(d) \times \eta \times v(d) + m(d) \tag{5}$$

A global variance parameter which is a variance estimation between the reference and the estimated features can be calculated. In order to enhance the clean speech estimate, the inverse DFT is applied to reconstruct the enhanced clean speech frame.

$$\hat{X}^f(d) = \exp\{\hat{X}'(d)/2\} \exp\{j < Y^f(d)\} \tag{6}$$

Finally, to fully reconstruct the speech signal from the frames, the overlap–add method for the convolution long sequences is used. The proposed method also ensures the smooth merging of amplitude $\hat{A}_{k,l}$ and the desired noise phase $\Phi_{k,l}^y$ along with the controlled amplitude of the desired signal.

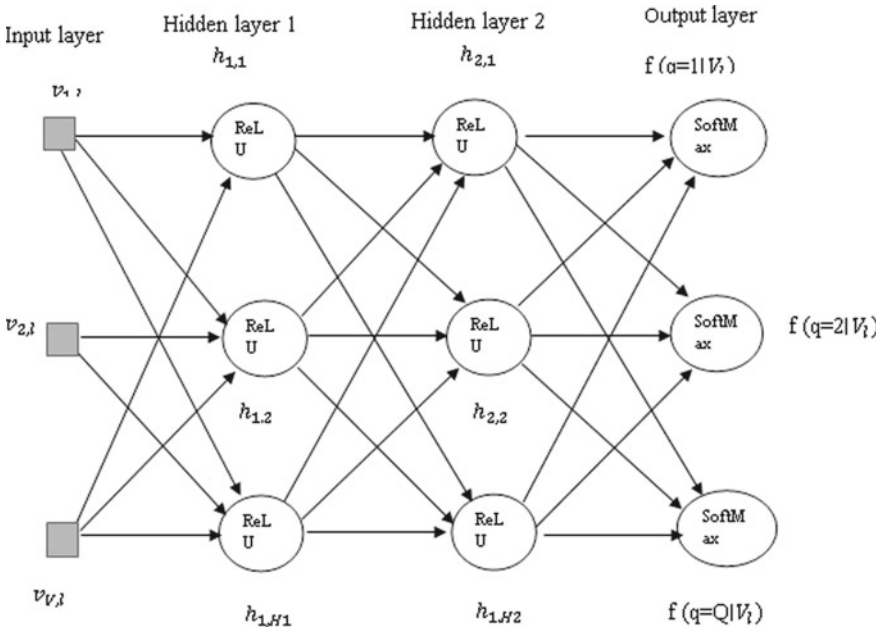


Fig. 3 Architectural design of DNN

2.2 Classification of Noise

Classification of noise is performed through CNN. Initially, the spectrograms obtained from clean speech as well as noisy speech are used for training the network. The input spectrograms are passed to the two convolutional layers followed by the pooling layers. The CNN network architecture used for the classification of noise is illustrated in Fig. 6.

In Fig. 4, the input noisy speech samples fed into the CNN classifier that trained the network to classify the environmental noise as given in [18]. The detailed architecture of convolutional neural network is as shown in Fig. 7. It provides the output of probability based on the threshold values. The threshold values were multiplied with the weights and making several trails, the average values will be calculated, and if it is larger than the threshold, it is identified as noise of interest, and the appropriate weight value makes it audible. Else, the weight decreases the intensity of the noise signals and makes it not to be audible and considered as undesired noise.

The system repeats the process in an iterative manner with the noisy speech utterances. Finally, the clean speech segment and the desired speech spectrums are fused to provide the desired noisy speech.

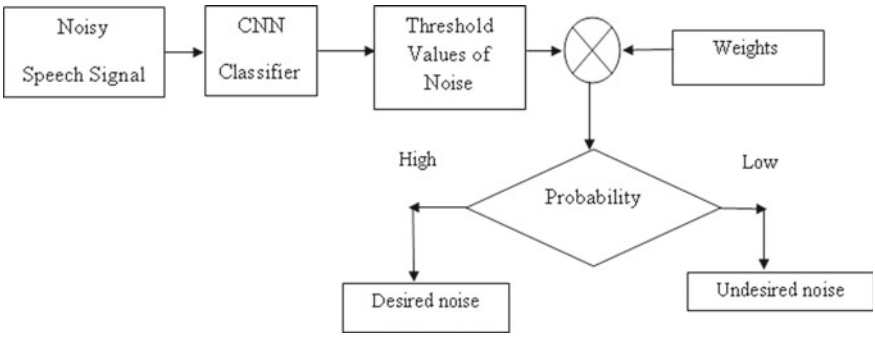


Fig. 4 Classification of noise

3 Methodology

The machine learning techniques such as deep neural network (DNN) and the convolutional neural network (CNN) are used in the proposed work as follows.

3.1 Deep Neural Network (DNN)

The DNN is the class of machine learning algorithms that uses multilayer to extract the higher-level features from the raw input. DNN is an artificial neural network. It finds the correct mathematical manipulation to turn the input into output. DNN architecture generates compositional model where the object expressed as the layered composition of primitives. DNN are typically feed forward network in which the data flows from the input layer to the output layer. It creates a map of virtual neurons and assigns the random numerical values or weights to connect between them. The weights and the inputs got multiplied and provide the output between 0 and 1. The DNN is used for the speech enhancement process in the proposed system. The input is characterized by 13 Mel-frequency cepstral coefficients (MFCC), and the Δ and $\Delta\Delta$ accelerations extracted for each frame. The three previous features and the three future segments are also included by adding the context which resulted in the feature vector $V_l = [v_{1,l}, \ldots, v_{V,l}]^T$ with dimensionality 1×273. The cepstral coefficients are also normalized and mean centered. The normalization is required to make the features to be of robust to the noise added from the environment from which the speech was acquired.

The normalized features were passed through the two hidden layers, and the score $f(q/V_l)$ is obtained for each vector $q \in \{1, \ldots, Q\}$. The H_1 and H_2 were the output of the two hidden layers. The rectified linear units (ReLU) actively work as the transfer function for these layers. The SoftMax transfer function applied in the output layer generates the posterior probability $f(q \mid V_l)$ in [18]. The enhancement process

as illustrated in Fig. 3, the enhancement of speech vectors are obtained from the training of DNN [12, 19]. The pre-training is done using the unsupervised learning by restricted Boltzmann machine (RBM) to estimate the higher-level features, and also, the weights are initialized for the supervised learning-based feed forward neural networks. The targeted outputs are both speech and noise of interest. The noise aware training in the DNN helps to predict the desired signal [4].

The output of the DNN is the estimated desired signal that undergoes post-process to remove the normalization effect. It is multiplied with the normalization factor η to mitigate the residual error. The inverse discrete fourier transform applied to the log spectrum of the desired signal is the output of the post process which is used to reconstruct the enhanced clean speech frame. Then, the overlap method is used to reconstruct the utterance of the enhanced clean speech signal. Finally, the desired clean speech coefficients were obtained.

The spectra of noise and speech could be obtained from the speech enhancement process. Using time varying frequency-dependent smoothing, the parameters are adjusted according to the average over several frames of noisy speech power spectra. It is also possible to estimate the noise using OM-LSA approach [20]. The estimation of clean speech spectra is difficult, and the relationship between the clean speech and the noisy speech is nonlinear. The DNN fed with more noisy samples to get the information of noise for the noise aware training [21].

3.2 Convolutional Neural Network (CNN)

It consists of input layer, output layer and multiple hidden layers. The hidden layer introduced to present the local structure in two-dimensional input data. It limited the layers to process the tiny part of the whole input space called receptive field.

The weight of the hidden layer creates the convolutional kernel to be applied to the input space resulted in a feature map. The pooling layer reduces the dimension of the features, and the output layer carried out the main estimation task of the network [22]. MFCCs of noisy speech signal is shown in Fig. 5.

3.3 Noise Classification

The CNN classifier consists of the convolution layer with max pooling to extract the features and perform the classification. The output layer provided the predictions. The data samples were collected from The noisy speech samples, almost for 5 s, are pre-processed and extracted features of 60 bands of log spectra [13]. The spectrum of each audio sample divided into 41 frames and formed into the segment of about 60rows/bands and 41 columns/bands. The segments passed through the first convolutional layer of size (57×6) with 80 filters and entered the max pooling layer with filter size of about (4×3) and strides of (1×3). The obtained output of the layer fed

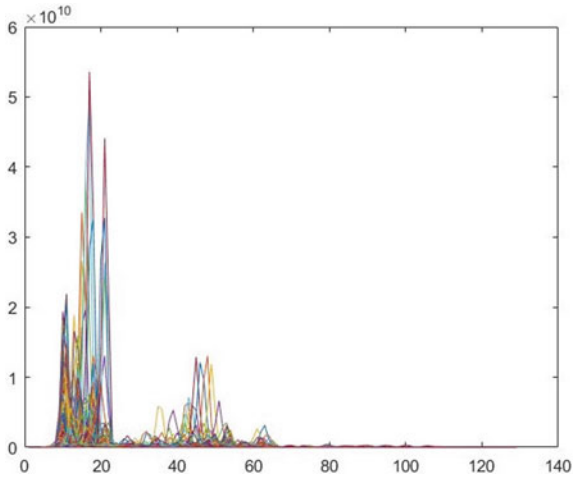


Fig. 5 MFCCs of noisy speech signal

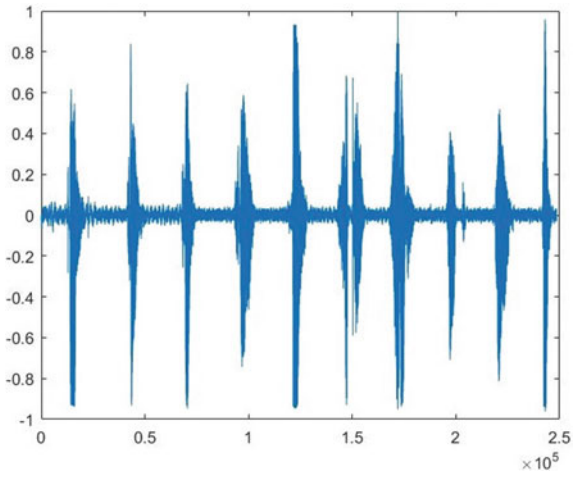


Fig. 6 Voice activity detection

into the next convolutional layer with 80 filters of the size (1 × 3) entered the next max pooling layer with filter size (1 × 3) and strides (1 × 3). The obtained output entered the fully connected layers of about 5000 units with ReLU activation function. Then, the dropout regularization applied and dropped out of about 20 units [23].

4 Experimental Setup

For this proposed work, the corrupted speech is taken from the TIMIT datasets with SNR of 5db [15]. It consists of pink noise, babble noise, restaurant noise, street noise, train noise, airport noise, etc. The 100 environmental sounds and the four noises such as AWGN, babble, restaurant noise and street noise were collected from the Aurora2 database [24]. The speech datasets were collected from the voice bank corpus [25], and also, some of the speech data derived from the TIMIT database [26]. It has around 400 records, 28 English speakers, including 14 female speaker and 14 male speakers.

The speech signal is corrupted with all these types of noise with six different levels of SNR, i.e., 20 dB, 15 dB, 10 dB, 5 dB, 0 dB, and −5 dB to build a noisy speech training set. These are samples used for the enhancement process. 2500 noise data were used for the training process. The 15 unseen noises were collected from the Auroara2 database [24] for testing. The clean and noise waveforms are down sampled to 8 kHz. The frame length used in this work is 256. DNN toolbox is used to plot the performance of the training process.

The samples were applied to the RBM for pre-training; the MFCC features were extracted by computing DFT, and as the result, the log power spectra obtained in [12]. The features are normalized to zero mean, unit variance. The learning rate is set to be 0.0005. Then, it undergoes the fine-tuning process, and the features passed through the FNN having the supervised learning. The learning rate is set for about 0.1 for the first 10 epochs. After each subsequent epoch, it decreased by 10%.

Regarding the classification of noise, the environmental sounds are taken from the UrbanSound8K database [27]. This dataset contains 8732 labeled sounds. It consists of 10 folds of sounds from the environment such as air conditioner, dog barking, car horn, children playing, drilling, engine idling, gun shoot, jackhammer and street music. These sounds are used for the classification of noises [28]. The datasets are divided for both the training and testing; 80% is used for training process, and 20% is used for testing procedure at the ratio of desired noise to speech is 30 : 70. The unseen noise is collected from the NOISEX-92 corpus used for testing purpose.

The log-scaled Mel-spectrogram extracted from all audio samples with window size of 1024, hop size of 512 and 60 Mel bands. The spectrograms are split into 50% overlapping segments of 41 frames (short variant, 950 ms) or 90% overlapping segments of 101 frames. These segments were fed as input into the network. They enter the two fully connected hidden layers of 5000 ReLUs each and a SoftMax output layer.

The training performed using MATLAB. 0.5 dropout probability occurred for fully connected. The training procedure was stopped after 5000 epochs for the short segment variant and 1500 epochs for the long segment variant. The final predictions were achieved by average probabilities of the outputs of segments.

The obtained output is the coefficients of the desired noise signal. It gets fused with the clean speech coefficients and providing the output of the desired noisy speech signal.

5 Results and Discussion

The obtained results on the extraction of clean speech along with emphasis on essential noise are described as follows. The noisy speech signal is given as the input. It consists of N=22529, derived from the short time Fourier transform. The signal is extracted by applying short time Fourier transform on computing the discrete Fourier transform (DFT) on the overlapping windowed frame. The log power spectrum is obtained and denoted as $y^{l}(k)$ [15]. The obtained speech power is −22.6837, and the noise power is −20.9297. The sliding window duration is 256 samples [3]. Sample clean audio signal and sample noisy audio signals are shown in Fig. 7.

It consists of sampling frequency 48,000 Hz. The noise segments are extracted from the noisy speech signal by using speech power and noise power. The obtained speech power is 804.4891, and the noise power is 6.9502. Figure 8 represents the confusion matrix evaluated on the datasets with desired and undesired noises. Figure 9 represents the performance of classification process in terms of receiver operating characteristics.

We have used the datasets at 5 dB and trained the samples. The noise removed from the noisy speech signal using DNN. Figures 10 and 11 represent the spectrograms of clean speech and noise, respectively. The previous systems eliminated any noise from the signal. But our system can eliminate only the undesired noise.

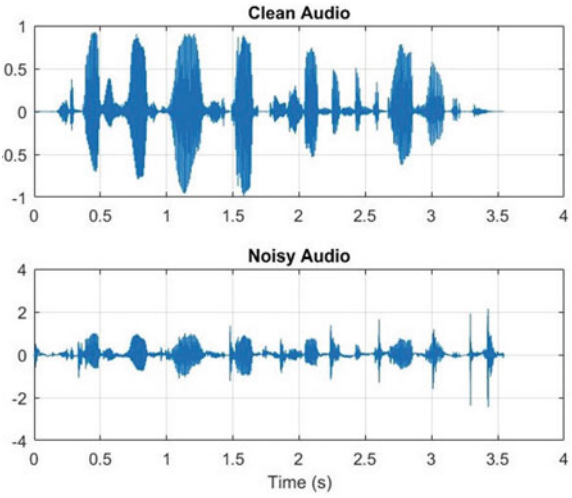


Fig. 7 Sample clean and noisy audio

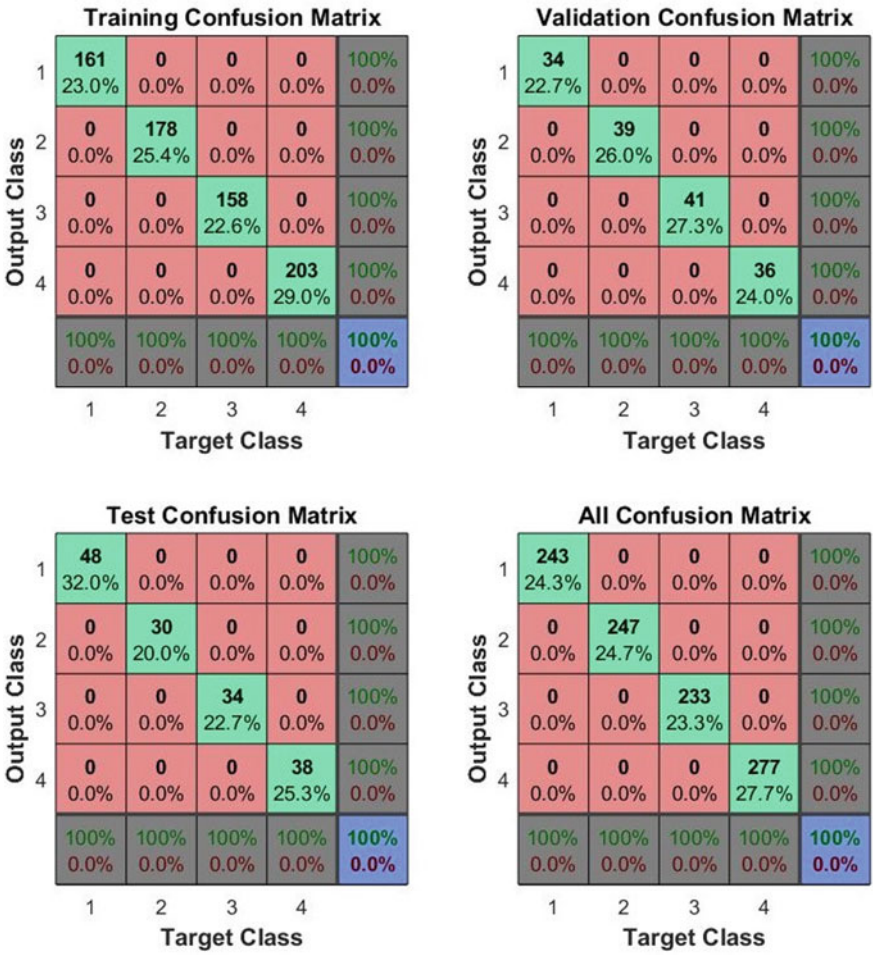


Fig. 8 Confusion matrix

Our system performs the classification process. The CNN classifier classifies the noise as desired noise and undesired noise. The spectrograms of the desired noise signals are given in Figs. 12 and 13, and the spectrograms of undesired noise are given in Figs. 14 and 15. The DNN is trained with a set of noise samples with an objective of emphasizing the noise of interest. The noise aware training performance with the DNN is illustrated in Fig. 16. Our proposed work is applicable to design the smart hearing aid. It can automatically adjust the hearing settings based on detecting the environment such as home, restaurant and traveling.

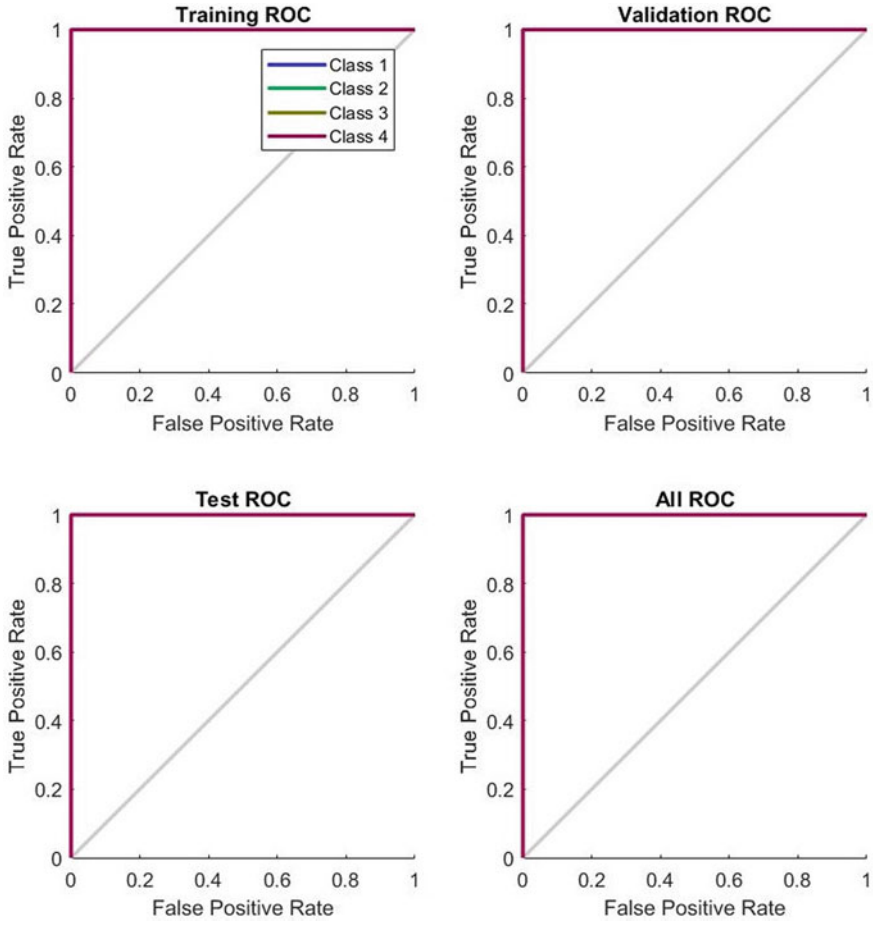


Fig. 9 Receiver operating characteristics

6 Conclusion

It is concluded that the desired noisy speech signal is the most important one for the hearing-impaired person and also useful for the cyber-crime and some security purpose. Here, there were three processes performed to obtain the desired noisy speech. Our system helps to avoid the catastrophic events that were mostly happened for the hearing-impaired persons when the desired noise such as the alert sounds is not heard.

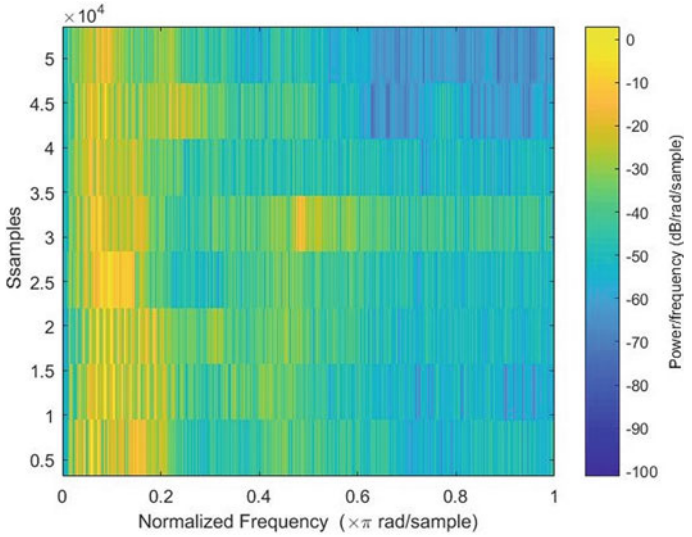


Fig. 10 Spectrogram of clean speech

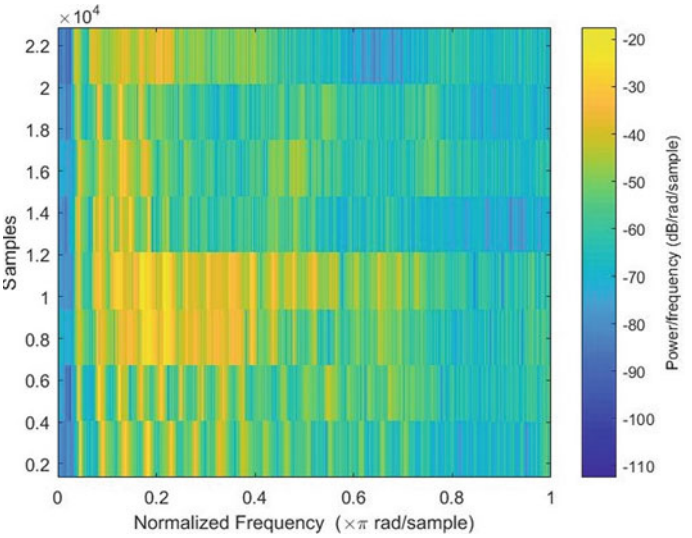


Fig. 11 Spectrogram of noise

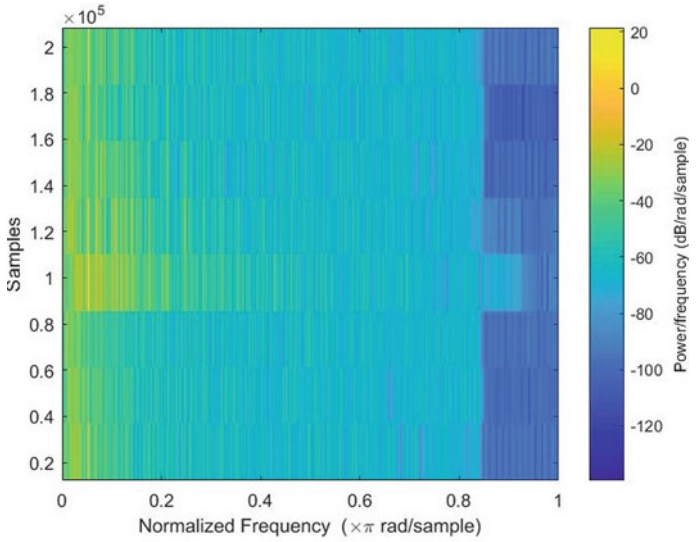


Fig. 12 Spectrogram of church bell noise

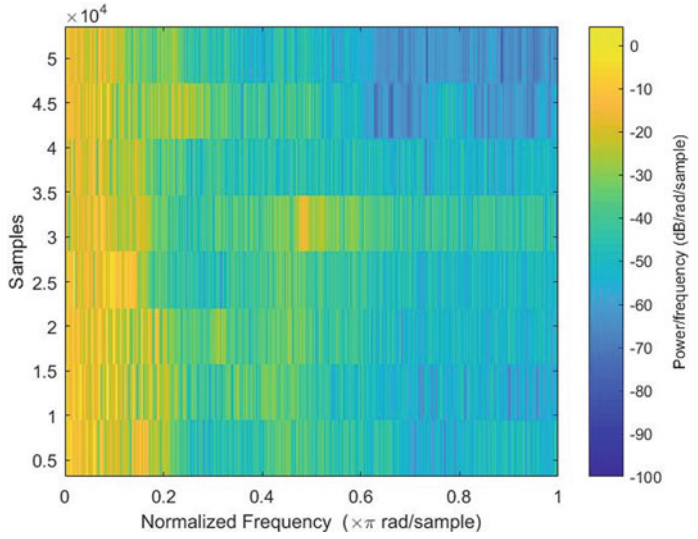


Fig. 13 Spectrogram of barking dog noise

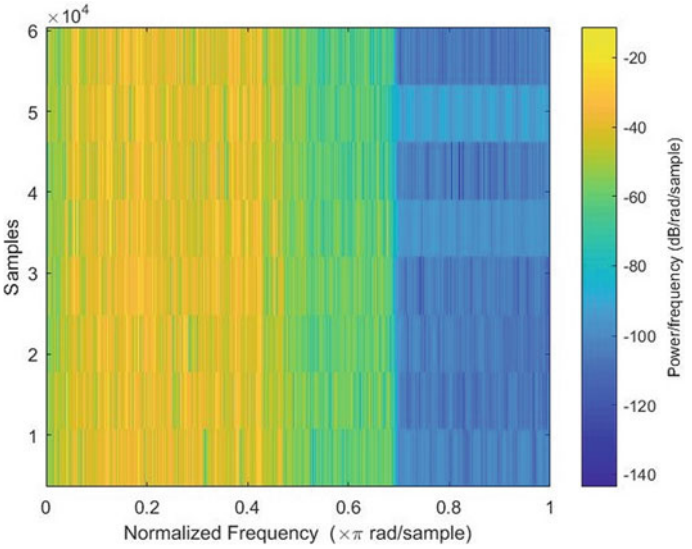


Fig. 14 Spectrogram of undesired restaurant noise

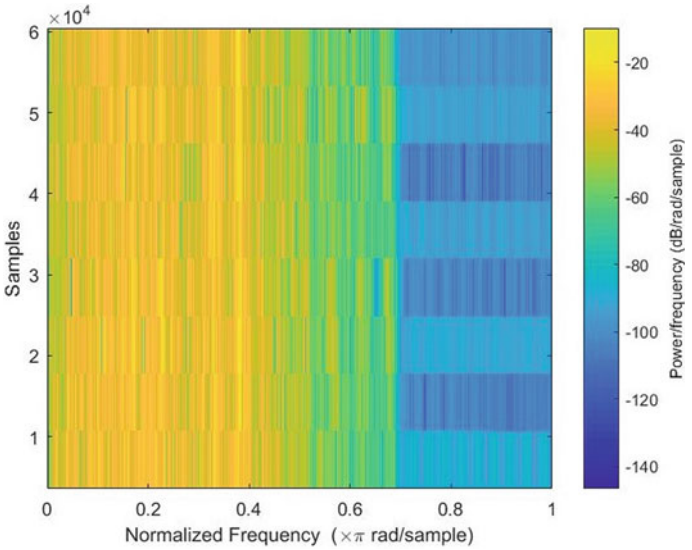


Fig. 15 Spectrogram of undesired street noise

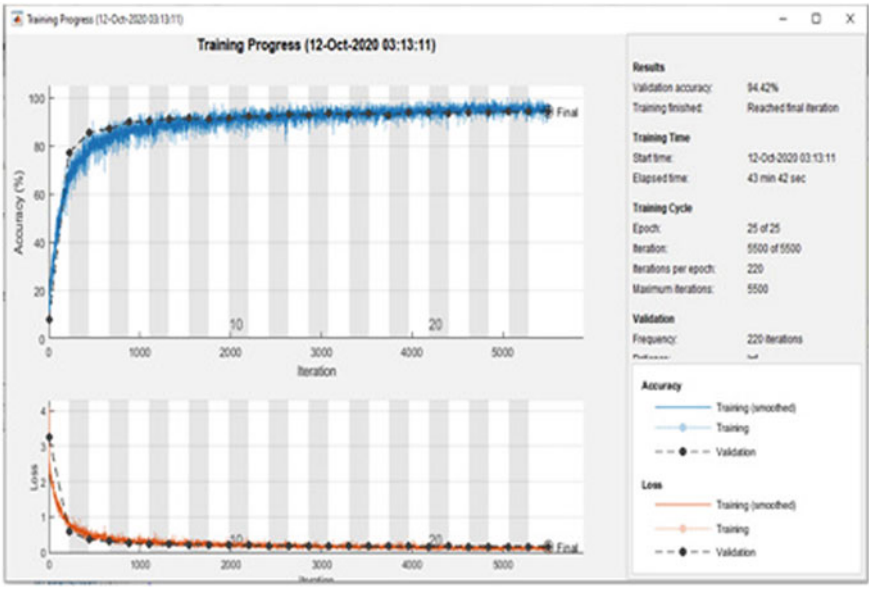


Fig. 16 Training performance of DNN

References

1. fxprosound audio design, Traffic Roadsounds, Jul. 2009. [Online]. Available: https://www.freesound.org/s/75375/
2. Raghuvanshi, S., Goyal, S.: Development of digital signal processing platform for digital hearing aid. Int. J. Adv. Res. Electrical Electronics Instrumentation Eng. (An ISO 3297: 2007 Certified Organization) **3**(2), February 2014
3. Gerkmann, T., Hendriks, R.C.: Noise power estimation based on the probability of speech presence. In: Proc IEEE Workshop Appl. Signal Process. Audio Acoust, pp. 145–148. New Paltz, NY, USA (2011)
4. Breithaupt, C., Krawczyk, M., Martin, R.: Parameterized MMSE spectral magnitude estimation for the enhancement of noisy speech. In :Proc Speech Signal Process, pp. 4037–4040. , Las Vegas, NV, USA, Apr, IEEE Int. Conf. Acoust. (2008)
5. Barchiesi, D., Giannoulis, D., Stowell, D., Plumbley, M.D.: Acoustic scene classification: classifying environments from the sounds they produce. IEEE Signal Process. Mag. **32**(3), 16–34 (2015). https://doi.org/10.1109/MSP.2014.2326181
6. Hinton, G.E., Srivastava, N., Krizhevsky, A., Sutskever, I., Salakhutdinov, R.R.: Improving neural networks by preventing co-adaptation of feature detectors. Arxiv (2012) (Online). Available: http://arxiv.org/abs/1207.0580
7. Hu, G.: 100 nonspeech environmental sounds (2004) (Online). Available: http://www.cse.ohiostate.edu/pnl/corpus/HuCorpus. html
8. Piczak. K.J.: Dataset for environmental sound classification. In: 23rd ACM International Conference on Multimedia, Brisbane, Australia, Oct. 2015, pp. 1015–1018
9. Garofolo, J.S.: Getting started with the DARPA TIMIT CD-ROM: an acoustic phonetic continuous speech database, NIST Tech. Rep (1988)
10. Khandelwal, R., Narayananm S., Lim L.: Emergency alert service (2006) (Online). Available: https://patents.google.com/patent/US7119675B2/en

11. Fu, S.W., Tsao, Y., Lu, X., Kawaim H.: Raw waveform-based speech enhancement by fully convolutional networks. ArXiv (2017) (Online). Available: https://arxiv.org/abs/1703.02205
12. Park, S.R., Lee, J.: A fully convolutional neural network for speech enhancement. Arxiv (2016) (Online). Available: https://arxiv.org/abs/1609.07132
13. Abdel-Hamid, O., Jiang, H.: Fast speaker adaptation of hybrid NN/HMM model for speech recognition based on discriminative learning of speaker code. In: Proc. ICASSP, pp. 7942–7946 (2013)
14. Erhan, D., Bengio, Y., Courville, A., Manzagol, P.-A., Vincent, P.: Why does unsupervised pre-training help deep learning? J. Mach. Learn. Res. **11**, 625–660 (2010)
15. Lim, J.S., Oppenheim, A.V.: Enhancement and bandwidth compression of noisy speech. Proc. IEEE **67**(12), 1586–1604 (1979)
16. Rehr, R., Gerkmann, T.: On the importance of Super-Gaussian speech priors for machine-learning based speech enhancement. IEEE/ACM Trans. Audio Speech Language Process. **26**(2), Feb. 2018
17. Ephraim, Y., Malah, D.: Speech enhancement using a minimum mean-square error log-spectral amplitude estimator. IEEE Trans. Acoust. Speech Signal Process ASSP **33**(2), 443–445 (1985)
18. Erhan, D., Courville, A., Bengio, Y., Vincent, P.: Why does unsupervised pre-training help deep learning. J. Mach. Learn. Res. **11**, 625–660 (2010)
19. Cohen, I., Berdugo, B.: Speech enhancement for non-stationary noise environments. Signal Process **81**(11), 2403–2418 (2001)
20. Xu, Y., Du, J., Dai, L.-R., Lee, C.-H.: Fellow, IEEE. A Regression approach to speech enhancement based on deep neural networks. IEEE/ACM Trans. Audio Speech Language Process. **23**(1), January 2015
21. Wang, Y., Wang, D.: Boosting classification based speech separation using temporal dynamics. In: INTERSPEECH 2012 ISCA's 13th Annual Conference Portland, OR, USA, September 9–13, 2012
22. Steeneken, H.J.M., Geurtsen, F.W.M.: Description of the RSG.10 noise database, TNO Institute for Perception, Soesterberg, The Netherlands, Tech. Rep. IZF 1988–3 (1988)
23. Piczak, K.J.: Environmental sound classification with convolutional neural networks. In: MLSP, 1–6 (2015)
24. Ephraim, Y., Malah, D.: Speech enhancement using a minimum-m square error short-time spectral amplitude estimator. IEEE Trans. Acoust. Speech Signal Process **32**(6), 1109–1121 (1984)
25. Dieleman, S., Brakel, P., Schrauwen, B.: Audiobased music classification with a pretrained convolutional network. In: Proceedings of the 12th International Society for Music Information Retrieval (ISMIR) conference, pp. 669–674 (2011)
26. Xu, Y., Du, J., Dai, L.-R., Lee, C.-H.: An experimental study on speech enhancement based on deep neural networks. IEEE Signal Process. Lett. **21**(1), 65–68 (2014)
27. Hirsch, H.G., Pearce, D.: The AURORA experimental framework for the performance evaluations of speech recognition systems under noisy conditions. In: Proc. ISCA ITRW ASR, pp. 181–188 (2000)
28. Lee, H., Grosse, R., Ranganath, R., Ng, A.Y.: Unsupervised learning of hierarchical representations with convolutional deep belief networks. Commun. ACM **54**(10), 95–103 (2011)

A Model-Based System for Intrusion Detection Using Novel Technique-Hidden Markov Bayesian in Wireless Sensor Network

Gauri Kalnoor and S. Gowri Shankar

Abstract Intrusion detection is one of the major challenges faced in wireless communications. The wireless network, where tiny sensors are comprised with deployment in remote areas and in few critical applications, is called wireless sensor network (WSN). Since WSN is computationally intensive, the framework is designed using the novel technique of machine learning known as hidden Markov Bayesian model. The model is also termed as naïve Bayesian hidden (NBH) as a decision function and knowledge-based Bayesian hidden system (KBHS) for detecting an intruder. Our approach mainly aims to detect the intruder based on the trained data using a novel algorithm applied for intrusion detection system framework. In phase 1, the dataset is preprocessed and trained. Then, the trained data is forwarded for detection where the decision function is applied by using hidden Markov Bayesian approach. The detection is performed in this stage and then updated for testing the detected data in the second phase. The simulation results are tabulated, and experiments obtained determine better performance with high detection rate and accuracy. The results also prove high throughput with minimum delay in transmission. Also, the obtained results are compared with that of weighted support vector machine (WSVM). The comparison shows better performance rate of WSN with high security.

Keywords Naive Bayes' Hidden model · Bayesian network · Hidden Markov model (HMM) · Training · Dataset · Knowledge-based Bayesian hidden system · Intrusion detection system (IDS) · WSN

G. Kalnoor (✉) · S. Gowri Shankar
Department of Computer Science and Engineering, BMS College of Engineering, Bangalore, India
e-mail: kalnoor.gauri@ieee.org

S. Gowri Shankar
e-mail: gowrishankar.cse@bmsce.ac.in

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_4

1 Introduction

The network applications comprised of sensors often have some of the stringent requirements for parameters like delay and throughput. These parameters are based on QoS-based network parameters. The sensor networks are battery operated, and such QoS-based requirements for high performance must be met especially when least amount of energy is consumed. Due to the high percentage of energy that is depleted for data communication, reliable and efficient communication are required to be supported which is one of the critical issues in WSN. However, the factors like mobility, fading and interference based on "time-varying multiuser" cause the quality of channel highly variable. These factors also cause difficulty in achieving the objectives. Certainly, the authors in [1, 2] have observed the variability in a significant manner for link quality in WSN in frequency band of 433 MHz, with radios. The authors point out the probability of packet error varies instantaneously by around 30% approximately with its mean value. Also, it has been observed that packet delivery error [3] exhibits process stochastic with significantly long-term dependence.

The objective of the proposed work is to provide the detection of an intruder based on hidden Markov model (HMM) and Bayesian network such that decision function is updated based on detection of trained data. Thus, the detection methodology is presented based on combination of HMM and Bayes' network called as NBH. The information scattered from multiple nodes is modeled with NBH using the large number of states that is considered as infinite, out of which, a posterior full distribution based on the number of states associated with the targeted nodes. The detection technique is a useful tool which is applied to model the information, and it is the optimal modeling approach to minimize error in false positive while the detection rate is maximized.

The work is organized in different sections. Section 2 explains the extensive research work in the related area. The framework and model are designed and described in Sects. 3 and 4, respectively. Results and analysis are presented in Sect. 5 with concluded remarks in Sect. 6.

2 Related Research

The techniques applied to gain the better performance of the network are reviewed, and extensive literature survey is performed such that our proposed work approaches to the specified aim for which the network is modeled.

In [4], an algorithm called as negative selection has been proposed with spatial partition in hierarchical WSN. The distribution of self-set values in the real-valued space is analyzed initially by the algorithm designed and also divided into several subspaces. These different subspaces are filled by selves, and the algorithm is implemented in the subspace. The selves need to be generated randomly, and then, the

Table 1 Comparison of IDS-based techniques

Methodology for detection	Strengths	Weaknesses
Fuzzy C-means clustering [6]	Results of clustering are less sensitive clusters	Communication head is increased
A hybrid detection using back-propagation neural network [7]	Increased detection accuracy Reduced energy consumption	Central point of failure at cluster heads
Detection using support vector machine [8]	No central points of failure since all nodes have the same capability of detection	Learning method for misuse detection is not applied

candidate detector is tolerated with divided selves in the subspace where the detection of subspace is made, and nit the Selves.

In [5], the authors proposed the novel scheme using variance and reduction method base on the authentic Markov process in WSN such that there was increase in accuracy. An algorithm based on "Markov chain Monte Carlo" has been proposed by the author such that computation of matrix for the pair of dominant eigenvalues was obtained which shows increase in energy formed. The localization accuracy was also analyzed using parameter at high level in WSN. The authors have also proposed the clustering process in a decentralized manner and are carried out autonomously by the sensor nodes. The process was carried to save the energy of node, without any communication of radio signal to the sink. The author describes rotation of cluster head (CH) that depends on the CH's residual energy. The frequency/timing of rotation was considered based on the consumption of energy the sensor nodes to perform various tasks by the nodes during the lifetime of network. Thus, the sensor's energy consumption was ensured in a balanced way. The proposed work in this article estimates the presence of node in a cluster, resulting in extended network lifetime.

Table 1 explains the various techniques based on IDS that are proposed in the previous literature of the area.

3 Framework Designed

In the proposed work, the traffic in the network is discretized while dividing the time into number of time slices denoted by Δt. The model proposed is described in four steps below as shown in Fig. 1. This technique is performed at the node level of base station.

1. Preprocessing of dataset: The data that has to be trained is normalized by the method of normalization called as min-max-based normalization.
2. Training the dataset: The training data which is normalized are grouped and classified as the certain clusters with large number of datasets in each cluster. The clustering is done based on the improved mean shift clustering algorithm. The clusters are then merged eventually into two different sub-clusters depending

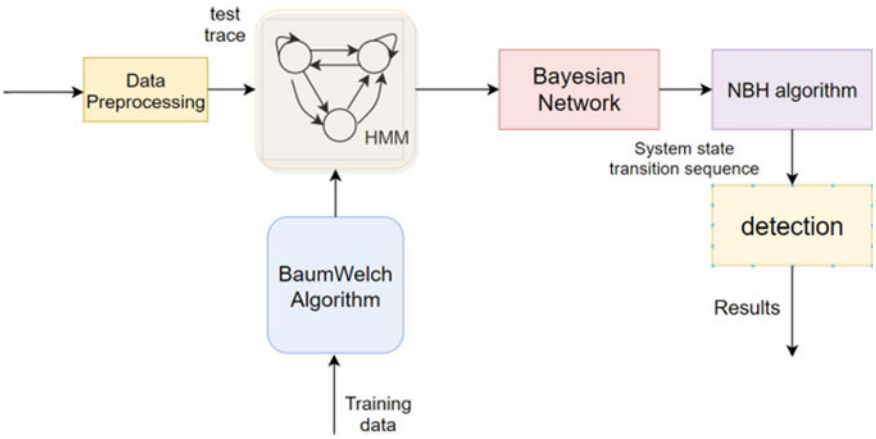


Fig. 1 Proposed system model

on the distance between themselves and the center point of another cluster. The normal data is compared with each feature vector among the training data and also tagged as anomaly or normal with the clustering result. Furthermore, a weight is assigned to every feature vector that represents the distance between itself and the centroid of its cluster group. The weighted labels for all the training data are the inputs to NBH such that a decision function is established.

3. Detection of trained data: The data tested is flagged as anomaly or normal by the decision function.
4. Updating: In the phase of testing, the processed feature vectors are combined with the training data, and the decision function is rebuilt at a specific interval $t = kt$, $kN+$. The algorithm using intrusion detection is utilized at the base station of sensor network, whereas the other nodes only capture and transmit their respective network status.

4 Hidden Markov Bayesian Model in WSN

Discrete values are generated from categorical distribution, whereas continuous values are generated from the Gaussian distribution. The parameters of HMM are of two types: emission probabilities (also called as output probabilities) and transition probabilities. Given the values of hidden state at time t-1, the values of hidden state at time t are chosen and controlled by the transition probabilities.

The data that is preprocessed is used for guidance and also testing. The five attributes that are chosen are prototype_type, src_bytes, flag, dst_bytes and count. The incessant variables are discretized and the obtained discrete values are represented using symbols.

4.1 Build the Bayes Network

The simplest representation of dynamic Bayesian network (DBN) is known as hidden Markov model (HMM). A DBN is a network which includes the sequence order of the state variables.

The state variables in the network have the graphical representation of a node. Thus, the ordered sequence more often is either symbolic sequences or time series. The edges in the graph represent the relationship and dependency among the nodes of the network, which are the directed edges. Every relationship (edge-node pair) has an associated table called conditional probability table (CPT). The records for normal type and attack type for applying the dataset, a Bayesian networks, are built.

4.2 Initialization of HMM Parameters

The dataset for the randomly generated data like KDD is used to design the IDS, which has 41 features. Among these, five features are chosen known as state hidden variables, named as protocol type, dst_bytes, src_bytes, flag and count. Each variable chosen has the distinct values set that emits as a symbol. Also, each state hidden variable has "K" distinct values, say, V_1, V_2, …, V_k. And then forming $X(t)$ observations.

The state transition diagram describes for both normal and attack records as discussed below.

- No. of state hidden variables $N = 5$, which is equal to no. of variables chosen. Thus, the dimension of "state transition matrix" is 5×5.
- No. of emission symbols those are distinct $M = 18$. So, the size of matrix for emission transition probability is 5×18.
- The probability distribution at initial stage.

$$\pi = \{0.000581, 0.261902, 0.08983, 0.375828, 0.271858\}.$$

The matrices of state transition A_{attack} and A_{norm} are initialized along with the random variables based on the Bayesian model (Fig. 2).

B_{norm} and B_{attack} were initialized based on the state variables of emission probability as shown in Tables 2 and 3, respectively.

4.3 Estimating the Probabilities: Transition and Emission

The rough guess has to be made about matrix of state modification and emission probabilities initially, such that the constraints of HMM can be determined. Once the

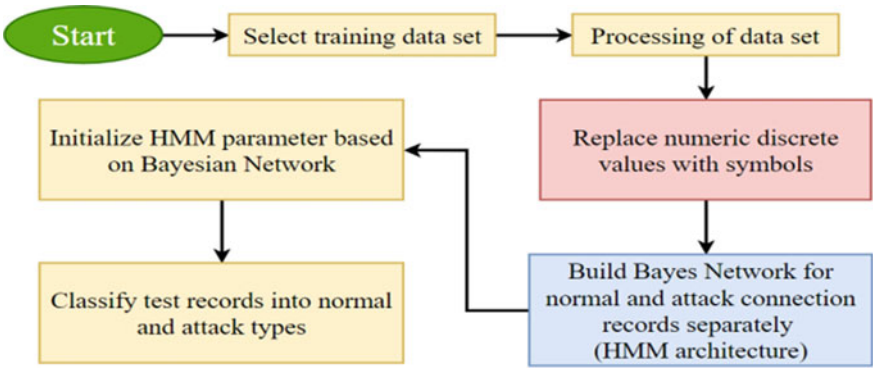


Fig. 2 IDS framework based on HMM and Bayes network model

Table 2 Matrix of state transition of A_{norm} of normal records

Protocol type	Flag	Scr_Byte		Dst_Bytes	Count
Protocol type	0	0.33	0	0.33	0.33
Flag	0	0	0	0	0
Scr_Bytes	0	0	0	0	0
Dst_Bytes	0	0	1	0	0
Count	0	0	0	0	0

Table 3 Matrix of state transition of A_{attack} of attack records

Protocol type	Flag	Scr_Byte		Dst_Bytes	Count
Protocol type	0	0.4	0.3	0.2	0.1
Flag	0	0	0	0	0
Scr_Bytes	0	0	0	0	0
Dst_Bytes	0	0	1	0	0
Count	0	0	0	0	0

values are predicted initially, the most accurate (i.e., maximum likelihood) parameters are discovered by applying the formulae known as "Baum-Welchre-estimation formulae (BW re-estimation)."

4.4 Evaluation of the Model: The Forward Algorithm and Evaluation Problem

When the HMM parameters (π, A, B) are known, the probability of an observed sequence should be found. To find the probability of an observed sequence, consider the problem that corresponds to the connected records consisting of observations from the dataset. The observations considered are as follows: dos, u2r, r2l, normal and probe, whereas in each of the observations, the records taken are as follows: DoS, U2R, R2L, normal and probe. Among the records, the normal records are only considered to build the model. Using such data, hidden states are shown as trellis. Thus, the forward algorithm is used to calculate the probability of observation sequences.

Once the model (π, A, B) is built, the observation sequences O_1, O_2, …, O_T and the probability $P\{O|\lambda\}$ must be found. Using the simple probabilistic arguments, this quantity can be calculated. But the calculation is very large, even if the sequence length "T" is moderate, as the calculation includes the number of operations in the orders of N_T. Because of the complexity in calculation, the other method of calculation can be used which uses an "auxiliary variable," called an "auxiliary variable" $t(i)$, called as "forward variable." The probability of "partial observation sequence" O_1, O_2, O_T, when at state "i," it terminates, then the variable is called forward variable. Thus, mathematically,

$$\alpha_t(i) = p(O_1, O_2, \ldots, O_t, q_t = i|\lambda\} \tag{1}$$

The forward algorithm, which represents the complexity of the above method, is directly proportional to N^2T, i.e., linear with respect to T, whereas the earlier method of direct calculation, and it has an exponential complexity.

Thus, we can also define the backward variable $\beta_t(i)$ as the probability of "partial observation sequence":

$O_{t+1}, O_{t+2}, \ldots O_T$ given that the present state is i. The mathematical representation for backward variable is as follows:

$$\beta_t(i) = p(O_{t+1}, O_{t+2}, \ldots, O_T|q_t = i, \lambda) \tag{2}$$

Thus, the decision function is often updated to improve the accuracy of detection. The classification of the normal and attack types is done based on the decision function using all the steps from preprocessing till the detection phase of the network. HMM and Bayes model are used at the decision function unit. The intrusion detection strategy is used with knowledge-based hierarchical system to improve performance of the network.

5 Results and Analysis

To assess the performance of the proposed methodology, the experiments have been conducted. This experiment scenario was simulated by using NS2 on a PC, with Inter (R) Core, 3.54 GHz and 8 GB RAM. Two types of network structure were considered: flat network structure and hierarchical network structure. The sensor nodes were deployed and scattered in the region with the dimension of 1000 (m) × 1000 (m) based on the type of network structures. The nodes are also deployed at the base station in the center of the deployed region. Among all the nodes, ten percent of the nodes were detected as malevolent nodes which performed different types of attacks. Comparing the flat network, the hierarchical network type of WSN is more appropriate for networks with large-scale nodes which reduces the energy consumption at a node and also bandwidth of the communication medium. Thus, the total number of nodes considered was 100, and the total nodes with attack type was 10.

For the analysis of proposed and present the NS2 setup is used for practical simulation of WSN network, as shown in the figure, the base station is connected with 100 nodes and simulation using both present and proposed algorithm is done to measure the various results as shown in the graphs obtained in Figs. 3, 4 and 5.

The simulation in NS2 for two kinds of network structure in WSN includes flat network type and hierarchical network type. In flat network type, the nodes convey data in a multihop way, and in hierarchical network, the transmission of data at the base station is in a hierarchical manner. For every network type, the routing protocol and MAC layer for all devices were “ad hoc on-demand distance vector” routing (AODV) and IEEE 802.11, respectively. For each experiment scenario type, the simulation time was set to 10,000 s. The traffic flow of network was simulated

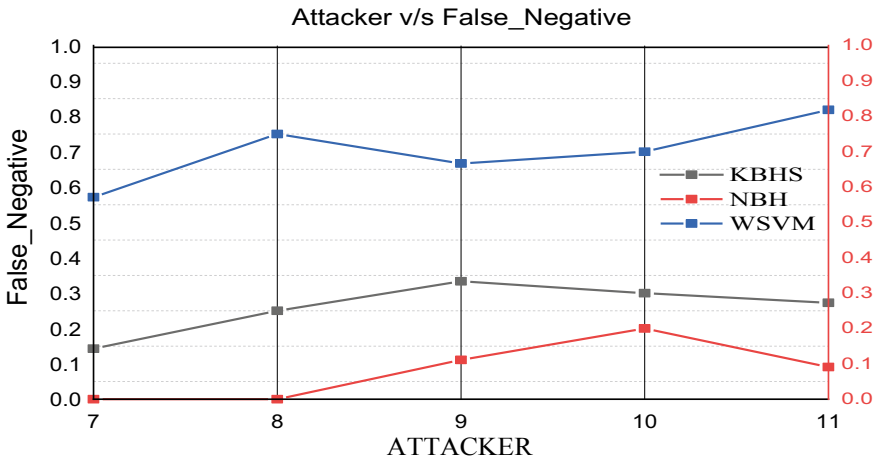


Fig. 3 Attacker versus false negative alarm

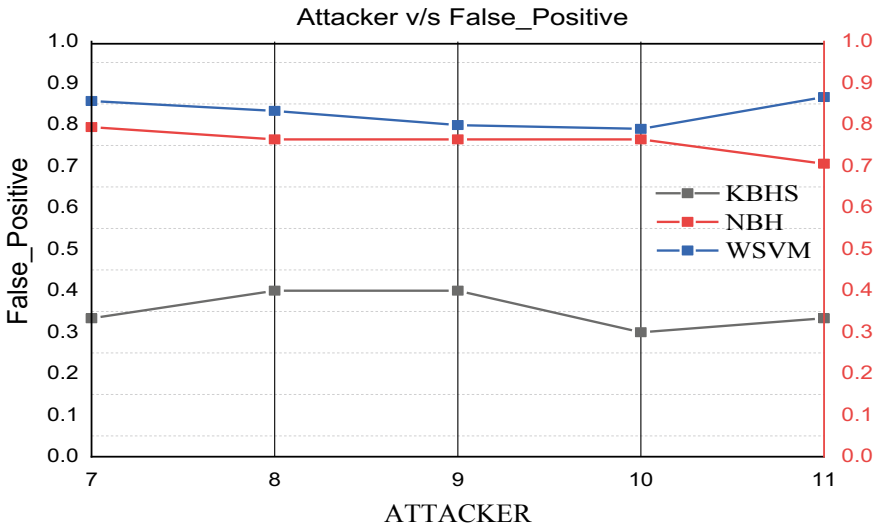


Fig. 4 Attacker versus false positive alarm

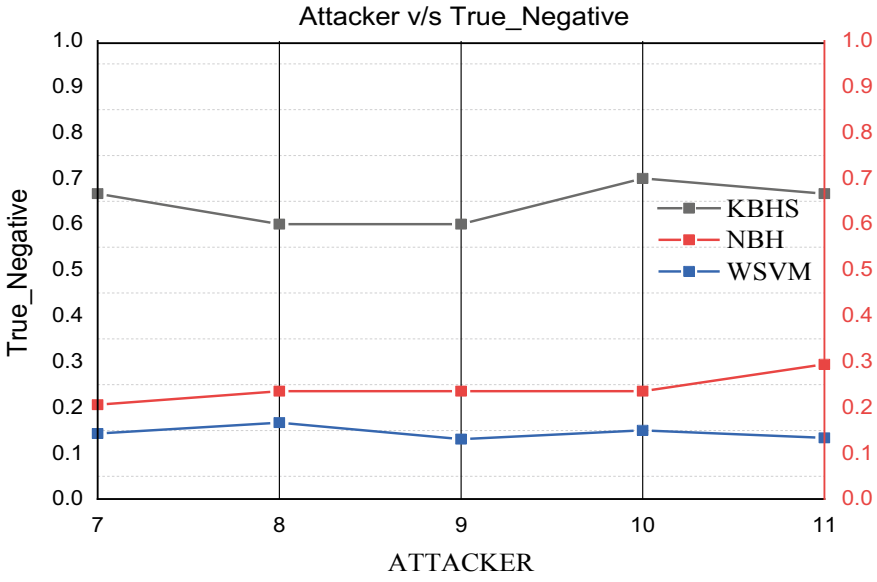


Fig. 5 Attacker versus true negative alarm

using the constant bit rate (CBR) with the packet size of 512 bytes. A mobility model, known as random waypoint (RWP) model, was simulated for the nodes, having a sleep time of 5 s and 5 m/s of maximum speed. The statistical data was simulated for counting the energy consumption by using radio energy model. The key parameters for experiments of simulation were obtained. In every scenario of singular attack, among the three types, only one attack type was launched (i.e., flooding attack, rushing attack and black hole attack) during the interval of 4000–7000 s. The multiple attacks scenario includes all the three attack types launching simultaneously, at the same time interval. Every time the different initial positions of the sensor nodes were set, where each scenario was repeated five times.

The basic unit used to process data in KBHS is known as the feature vector, which is obtained by capturing 13 feature types to represent status of nodes. The efficiency of KBHS is tested by estimating the false alarm rate (FAR) and the detection rate which is used as assessment index. The detection rate was estimated by calculating the percentage of the number of successful detected malicious attack with the total number of malicious nodes. Also, the false alarm rate is estimated as the percentage of the number of false alarms triggered by the system to the total number of nodes with normal data.

The simulation results of proposed work were analyzed and compared with the methods of other mainstream intrusion detection, such as K-means, decision tress (DT), PCA-based centralized intrusion detection (PCACID), mean shift and also logistic regression (LR). The comparison is made to evaluate the performance of network based on the detection efficiency, and it is adaptively to network structures. The simulated results showed that among all the experiment-type scenarios, the false alarm rate and average detection rate (DR) of proposed work obtained were 1.76% and 97.54%, respectively. However, the small deviation obtained was 1.054% and 0.920% for both FAR and DR. Thus, it was demonstrated that the proposed system is more efficient than the other IDS mainstream methods. Though the DT, LR and K-means obtained average DR more than 92% and average FAR less than 3%, the average FAR had slightly large fluctuations in the results at the same situation, i.e., 4.29% and 4.83%, 1.33% and 1.349%, also, 3.9222% and 2.551%. It showed the disadvantage of the existing methods when dealt with various attack forms. In few cases, mean shift and PCACID also achieved low FAR than proposed method. But they failed in detecting the malicious activity in nodes in all the scenarios, at constant rate. Thus, it is showed that our proposed method has advantages over the mainstream methods, mainly in terms of detection rate and FAR. However, our method has achieved the stable performance in all the experiment-type scenarios, mainly in different network structures.

6 Conclusion

In WSN, security is one of the major issues that need to be taken into consideration. The model is designed and implemented using for detection of an intruder

with maximum accuracy rate. Bayesian network and hidden Markov are the novel model designed by applying algorithm during the detection phase, and a decision function is updated every time when the trained data is tested and detected. The normal and abnormal behavior of the trained data are observed, and simulation results are noted which gives better performance in detection rate and throughput. The results obtained are compared with that applied by other mainstream techniques like weighted SVM (WSVM). Minimum transmission delay with minimum energy consumption is obtained in our experimental results.

References

1. Tan, X., Su, S., Huang, Z., Guo, X., Zuo, Z., Sun, X., Li, L.: Wireless sensor networks intrusion detection based on SMOTE and the random forest algorithm. Sensors **19**(1), 203 (2019)
2. Yang, G., Yu, X., Xu, L., Xin, Y., Fang, X.: An intrusion detection algorithm for sensor network based on normalized cut spectral clustering. PLoS ONE **14**(10), e0221920 (2019). https://doi.org/10.1371/journal.pone.0221920
3. Alqahtani, M., Gumaei, A., Mathkour, H., Maher Ben Ismail, M.: A genetic-based extreme gradient boosting model for detecting intrusions in wireless sensor networks. Sensors **19**(20), 4383 (2019)
4. Nema, A.: Innovative approach for improving intrusion detection using genetic algorithm with layered approach. In: Securing the Internet of Things: Concepts, Methodologies, Tools, and Applications, pp. 273–298. IGI Global (2020)
5. Vijayanand, R., Devaraj, D., Kannapiran, B.: Intrusion detection system for wireless mesh network using multiple support vector machine classifiers with genetic-algorithm-based feature selection. Comput. Secur. **77**, 304–314 (2018)
6. Zhang, Y., Li, P., Wang, X.: Intrusion detection for IoT based on improved genetic algorithm and deep belief network. IEEE Access **7**, 31711–31722 (2019)
7. Hamamoto, A.H., Carvalho, L.F., Sampaio, L.D.H., Abrão, T., Proença, M.L., Jr.: Network anomaly detection system using genetic algorithm and fuzzy logic. Expert Syst. Appl. **92**, 390–402 (2018)
8. Reddy, S.S.S., Chatterjee, P., Mamatha, C.: Intrusion detection in wireless network using fuzzy logic implemented with genetic algorithm. In: Computing and Network Sustainability, pp. 425–432. Springer, Singapore (2019)

Analysis of Academic Performance Based on Hierarchical Clusters: First Notes

Lozada T. Edwin Fabricio, Ruth Patricia Maldonado Rivera, Pullas T. Paul, and Luis Alberto Soria Nuñez

Abstract The student evaluation processes are very important in order to propose improvement strategies. In this work, we propose a methodology that verifies the formation of groups related to the academic performance of students, which allows to visualize strategies in front of the anticipated preparation or not to take an exam. We consider as a main element that the evaluation cannot be by individual subjects, but that academic performance must be verified in a group or general way in order to obtain a correct response from the teaching–learning process applied to the student. The work is still in the investigative process; however, the conclusions show that exist groups between different subjects, including that there is no clear relationship between preparation or not for an exam and academic performance.

Keywords Academic performance · Teaching process learning · Clusters

L. T. Edwin Fabricio (✉)
Universidad Regional Autónoma de Los Andes–UNIANDES, Carrera de Software, Ambato, Ecuador
e-mail: ua.edwinlozada@uniandes.edu.ec; edwin.lozadat@info.unlp.edu.ar

Facultad de Informática, Universidad Nacional de La Plata, LIFIA, La Plata, Argentina

R. P. Maldonado Rivera
Departamento de Psicología, Universidad Técnica Particular de Loja, Ambato, Ecuador
e-mail: rpmaldonado@utpl.edu.ec

P. T. Paul
Facultad de Ciencias Humanas y de La Educación, Universidad Técnica de Ambato, Ambato, Ecuador
e-mail: paulspullas@uta.edu.ec

L. A. Soria Nuñez
Facultad de Arquitectura Diseño y Artes, Pontificia Universidad Católica del Ecuador, Quito, Ecuador
e-mail: soria089@puce.edu.ec

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_5

1 Introduction

The evaluation during the teaching–learning process is a necessary process in order to verify the progress of the students. However, this process can become a stress factor, causing academic performance problems, leading to student dropouts. There are different factors that can affect this performance, such as the student who works outside her study hours, whose work does not allow her to prepare adequately for her exams [1]. Even social influence can cause serious inconveniences related to academic performance, this being a factor not necessarily of studies but rather of culture or social affectation [2]. It is therefore very important to identify the causes of the problem of the decrease in academic performance, which may be related to psychological, economic, sociological, organizational and interaction dimensions. Even the rigid organizational structures of universities can cause student disinterest, which together with the parents' disinterest in the student can form sufficient grounds for dropping out [3]. When a student's academic record is analyzed, it is usually done individually subject by subject. This makes it possible to lose the vision that a person cannot be a specialist in all subjects and that there is a valid hypothesis that is good for one or a couple of aspects. For this reason, it is very important to obtain valid datasets from the teaching–learning process that allow identifying the problem of academic performance for strategic purposes of improving student achievement. The use of educational algorithms based on data mining, clustering is still in an early stage, but they show results of academic performance clustering when applied to an academic context dataset [4]. There is the possibility of identifying homogeneous groups in the student dropout population. The cluster analysis allows the identification of groups of school refusal, the same ones that can be grouped by not in school, mixed reinforcement, tangible reinforcement and negative reinforcement, each with its own group score, which is directly related to social influence [5]. Learning strategies can be temporarily analyzed and appropriate feedback can be generated to form new strategies. Through groupings, these learning tactics can be approached in order to identify their relationship with academic performance and new learning methods, this type of process is generally related to personalized feedback and the generation of effective strategies that maintain behavior suitable for improving academic performance [6].

The quality of teaching is also related to academic performance. The evaluation processes can be grouped in order to identify the evaluation quality, in order to simplify the number of indicators and the weights assigned to each of them and thereby obtain a better understanding of the evaluation processes and with this improves the teaching–learning process [7]. It is even possible to group the learning processes themselves such as physical activity in order to verify if there is a direct relationship with academic performance that is generally considered positive, since the improvement of physical health promotes an increase in the academic performance of the students [8]. In this work, we propose the use of clusters in order to identify if there is a relationship in the academic performance of students for different subjects, since we consider that academic performance should be analyzed in a grouped way

and not individually. To do this, we verify whether academic performance groups are formed, and we analyze the number of groups that are formed. Experimental data[1] show that if students' performance can be assessed in a group and that there is not even a direct relationship between preparing or not for an exam. To better describe our proposal, the following sections describe the steps of the methodology and the experiments and conclusions resulting from its application.

2 Methodology

In order to verify whether there is a group relationship between the grades obtained by students in different subjects and with their preparation, it is proposed to follow the following steps:

- Select dataset
- Check if the groups appear graphically
- Apply hierarchical clusters
- Verify the number of clusters obtained from the maximum distance
- Relate the students with the clusters obtained.

In the following section, the proposed methodology is applied, and from the results, the conclusions and future works are obtained as a result of the experimental application.

3 Experimentation

The datasets_74977_169835_StudentsPerformance file was used, and the appearance of the groups was graphically verified.

In Fig. 1, it is shown that if groups are formed between the selected subjects, whose color in the figure ranges from blue to orange. Based on that, the hierarchical clusters were applied.

Figure 2 shows the hierarchical cluster in which three clusters are formed from the maximum distance (blue line), with which it can be verified if there is a relationship between the subjects and the preparation for their exams.

Figure 3 shows the grouping by three clusters, which implies that academic performance can be related between the subjects considered. Based on this, each student (400) is related to the identified cluster, an example is shown in Table 1.

The result after processing the data was that cluster 0 has the highest number of students whose grades vary between 27 and 97 points, with a correlation coefficient between exam preparation and math grade of 0.19, and with a reading of 0.24.

[1] https://www.kaggle.com/spscientist/students-performance-in-exams?select=StudentsPerformance.csv.

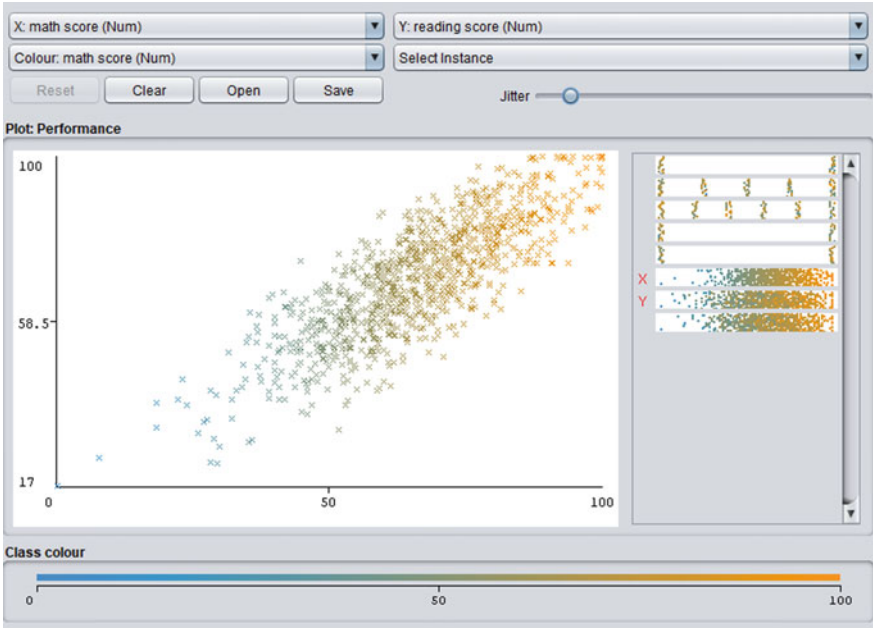


Fig. 1 Groups between the marks obtained in mathematics and in reading

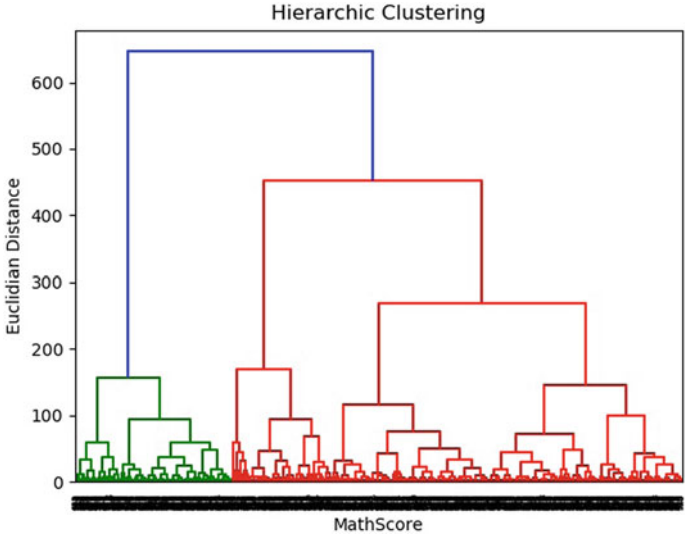


Fig. 2 Hierarchical academic performance cluster

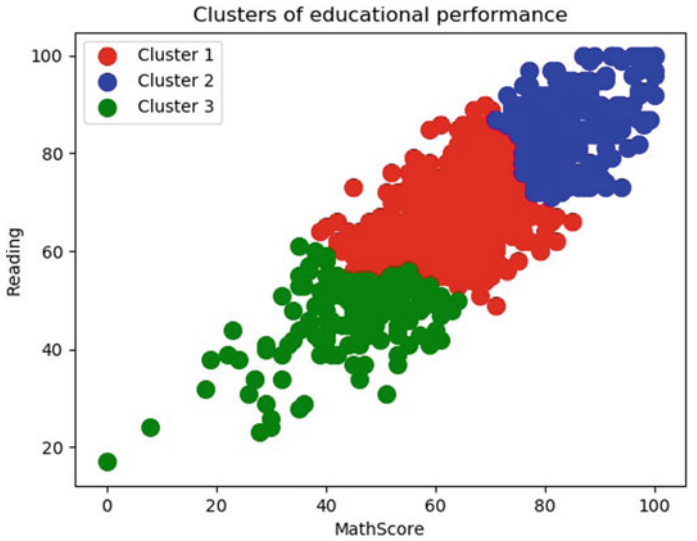


Fig. 3 Grouping by academic performance and test preparation

Table 1 Hierarchical cluster

Test preparation	ClusterMR	Math score	Reading score
None	0	72	72
Completed	0	69	90
None	0	47	57
None	0	71	83
Completed	0	64	64
None	0	58	54
None	0	65	81
None	0	69	75
None	0	54	58
None	0	66	69
Completed	0	65	75
None	0	69	73
Completed	0	74	71
None	0	73	74
None	0	69	54
None	0	67	69
None	0	70	70
None	0	62	70
None	0	69	74
None	0	63	65

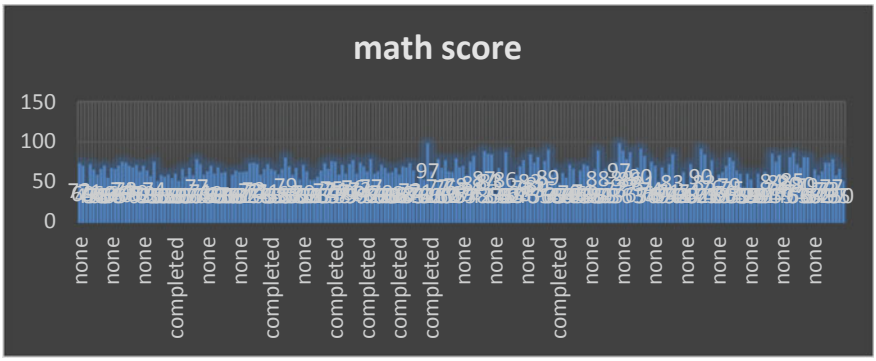


Fig. 4 Preparation for the exam and grade obtained in mathematics

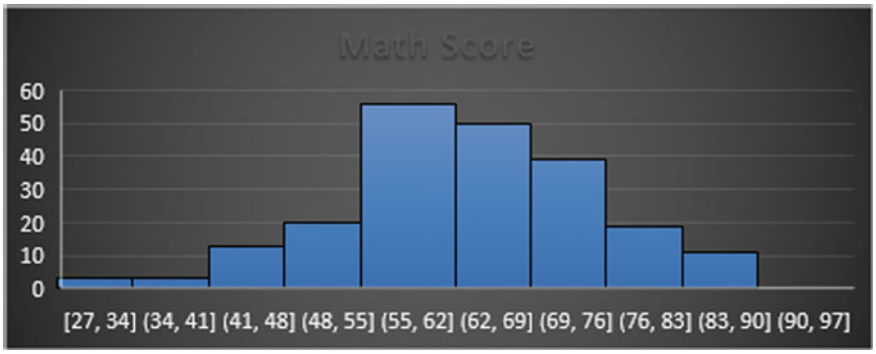


Fig. 5 Preparation for the math exam

In Fig. 4, it is shown that cluster 0 supports both those who prepared and those who did not have any preparation for the exam, while.

Figure 5 shows that the majority of students did not prepare for the exam, but their average is between 55 and 76 points.

4 Conclusions

This work tries to verify whether academic performance groups can be created through hierarchical clusters. The creation of groups was possible, and it is determined that in a cluster there may be students who are preparing and not for an exam. In addition, it is verified with the correlation coefficient that there is no certain relationship between academic performance and preparation for the exam, and that students who do not prepare for the exam can even obtain the highest marks on it.

These conclusions are very interesting even so that educators and parents can mediate the preparation for the exams to reduce stress on them.

References

1. Bartoli, T., Polanec, S.: Does work harm academic performance of students? Evidence using propensity score matching, Res. High. Educ., pp. 401–429 (2018.
2. Garcia Martínez, L., Charczuk, M., Baldizzoni, E.: Identification of Causes of Desertion and Disengagement of Undergraduate Students in Systems Using Information Exploitation Engineering, from IGLU (2015)
3. García, J., Manuel, G., Zanfrillo, A.: Desgranamiento University: Student Perspective in Engineering Communication. de International Colloquium on Gestao University na America do Sul (2011)
4. Dutt, A., Aghabozrgi, S., Ismail, M., Mahroreian, H.: Clustering algorithms applied in educational data mining. Int. J. Inf. Electron. Eng., p. 112 (2015)
5. Gonzalvez, C., Ingles, C., Kearney, C., Sanmartin, R., Vicent M., García-Fernandez, J.: Relationship between school refusal behavior and social functioning: a cluster analysis approach (2019)
6. Matcha, W., Gasevic, D., Uzir, N., Jovanovic, J., Pardo, A.: Analytics of learning strategies: associations with academic performance and feedback. In: Proceedings of the 9th International Conference on Learning Analytics & Knowledge, 2019.
7. Zhang, M., Wang, J., Zhou, R.: Entropy value-based pursuit projection cluster for the teaching quality evaluation with internal number. Entropy **21**(2), 203 (2019)
8. Lima, R., Pfeiffer, K., Moller, N., Andersen, L., Bugge, A.: Physical activity and sedentary time are positively associated with academic performance: a 3-year longitudinal study. J. Phys. Act. Health **16**(3), 177–183 (2019)

Application of Hierarchical Clusters to Obtain Legal Reference Structures

Priscilla Massa-Sánchez, N. Freddy Patricio Baño, Gabriela Elizabeth Armijos Maurad, and Richard Eduardo Ruiz Ordoñez

Abstract In the fields of Legal Law, it is vitally important to know the greatest number of legal documents to make an impartial decision. In this work, we propose to identify the grouping of decision making based on the case, jurisdiction and type of applied law. For this, we apply hierarchical clusters, and the methodological application has shown that in effect the applied laws are grouped, forming a reference structure for similar cases. The work is in progress, but the initial conclusions show that it is possible to obtain group cases for decision making in legal sentences.

Keywords Sentences · Legal cases · Cluster

1 Introduction

In legal fields, it is necessary to verify the largest amount of bibliography or legal theories that allow improvement in decision making. Knowing what the elements of judgment have been in similar cases or what decisions were made individually is essential for a judge's decision. The difficulty appears because in the science of law, there is a lot of information, legal domains that must be studied and classified in order to process impartially. The processes of drafting the legal sentences participate in the veracity, the direct and definitive language and the acceptance of the

P. Massa-Sánchez (✉) · G. E. A. Maurad
Departamento de Economía, Universidad Técnica Particular de Loja, Loja, Ecuador
e-mail: pmassa@utpl.edu.ec

G. E. A. Maurad
e-mail: gearmijos1@utpl.edu.ec

N. F. P. Baño
Carrera de Software, Instituto Superior Tecnológico Mayor Pedro Traversari, Quito, Ecuador
e-mail: freddy.bano@institutotraversari.edu.ec

R. E. R. Ordoñez
Instituto Superior Tecnologico Loja, Loja, Ecuador
e-mail: rruiz@institutos.gob.ec

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_6

judicial verdict, which involves the reasoning and the connectionist and coherent model of the facts, which implies that there must be categories or groups that allow the sentences to be connected in order to obtain adequate legal results [1]. The analysis of legal texts is also examined by semantic structures, providing the description of legal sentences and their corresponding application in various fields of law [2]. There are processes that help the selection and classification of legal texts and even allow to obtain summaries of these, which is already of great help for the legal investigative process [3]. Having basic elements for legal decisions ends in improving administrative efficiency and judicial coherence. Automatic classification of legal texts makes it possible to highlight the search text, and this serves as input to predict decisions, including improving the prediction of human decisions, especially when working on the structure and semantics of the document in search of predictions and their explanatory value [4]. Even the legal arguments with which to make a defense must be properly founded. Correctly identifying the most relevant and discriminatory characteristics is of vital importance in making legal decisions. For this, grouping algorithms have been developed that compare their results with human experts in order to identify which characteristics are appropriate for a correct defense mechanism of a legal case [5]. Especially when working on the structure and semantics of the document in search of predictions and their explanatory value [4]. Even the legal arguments with which to make a defense must be properly founded. Correctly identifying the most relevant and discriminatory characteristics is of vital importance in making legal decisions. For this, grouping algorithms have been developed that compare their results with human experts in order to identify which characteristics are appropriate for a correct defense mechanism of a legal case [5]. Especially when working on the structure and semantics of the document in search of predictions and their explanatory value [4]. Even the legal arguments with which to make a defense must be properly founded. Correctly identifying the most relevant and discriminatory characteristics is of vital importance in making legal decisions. For this, grouping algorithms have been developed that compare their results with human experts in order to identify which characteristics are appropriate for a correct defense mechanism of a legal case [5]. It is of vital importance in legal decision making. For this, grouping algorithms have been developed that compare their results with human experts in order to identify which characteristics are appropriate for a correct defense mechanism of a legal case [5]. It is of vital importance in legal decision making. For this, grouping algorithms have been developed that compare their results with human experts in order to identify which characteristics are appropriate for a correct defense mechanism of a legal case [5].

The study of legal judgments implies the analysis of taxonomies where the empirical methodology appears, which must be improved by the acquisition of restrictions and rewriting rules in order to form adequate searches in legal texts based on response patterns [6]. It is even possible to find the disparity of legal decisions between courts, through text extraction techniques based on the coherence of the sentences and the importance of their location [7]. Legal narratives allow identifying relationships and instances in order to infer knowledge, this can be treated as groups and even individually, especially when legal drafting processes are assumed with linguistic elements

that allow the management of opinion [8]. The objective of this work is to determine if legal decisions given the jurisdictions and the type of law can be grouped in the sense that these decisions can be taken into account among similar cases. The data was obtained from the Supreme Court of the USA.[1] To do this, we prepared the methodology in which hierarchical clusters were applied since these allow us to observe the detail of the formation of groups, and then verify if there are differences in the formation of legal decisions. In the experiment, we worked with the data observing differences in the conformation of groups of the years 2015–2016, and that therefore these data must be analyzed in greater depth in the case of making predictions or grouping of sentences. The following sections describe our research in greater detail, concluding that it is a first-stage study but that it shows that legal decisions can be grouped together and serve as the basis for the descriptions of sentences.

2 Methodology

In order to obtain sufficient data and apply the hierarchical cluster model, it is necessary to follow the following steps:

- Select the files of the data set based on case, jurisdiction and type of applied law.
- Delete the blank data.
- Check if cases can be grouped visually.
- Apply the hierarchical cluster models and visualize the maximum distance, in order to select the number of clusters.
- Identify each case with the cluster.

These steps have been applied in the same experiment as described below.

3 Experimentation

We worked with three datasets OT2015_20160212, OT2015_20160228 and OT2015_20160416 and with the case, jurisdiction and type of law fields. The dataset is visually grouped as shown in Fig. 1.

In Fig. 1, it can be seen that the data is grouped so the methodology can be applied, before the cleaning of the blank data has been carried out.

Figure 2 shows the dendrogram of cases in which the appearance of two groups (green and red) is verified from the maximum distance represented by the orange line. With this procedure, we proceed to verify the groups.

In Fig. 3, the result of the grouping is observed, so that each case corresponds to a type of decision or similar applicable law.

[1] https://github.com/mjbommar/scotus-predict

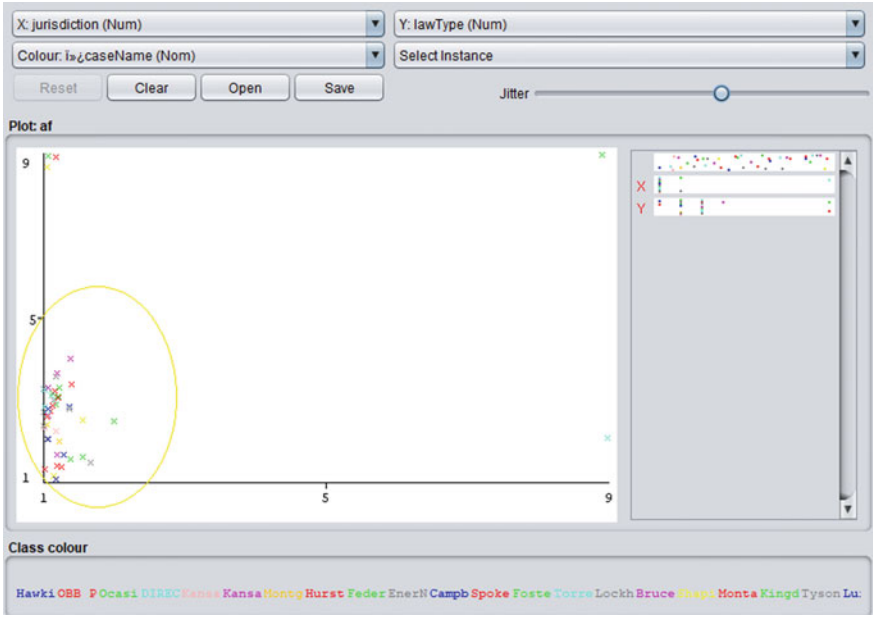


Fig. 1 Data grouping

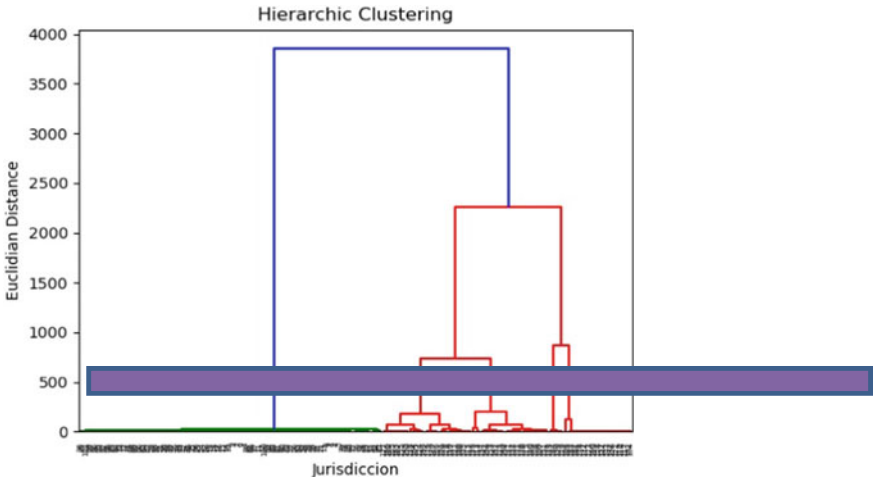


Fig. 2 Dendrogram of cases

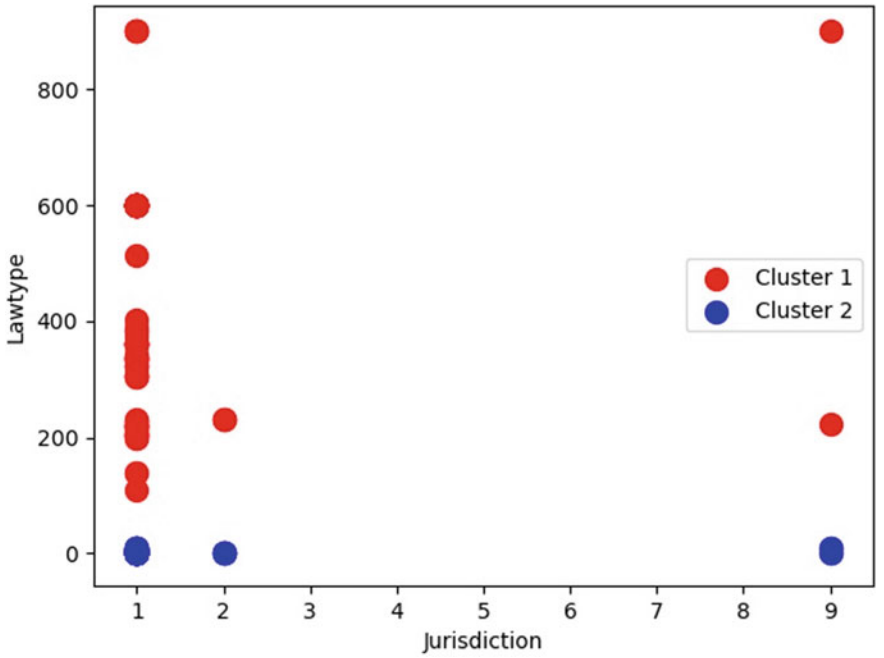


Fig. 3 Case groups

Table 1 shows the assignment of the group of laws to each of the subject cases of the experiment.

4 Conclusions

This work shows a methodology that verifies whether it is possible to apply the same set of laws to different cases and also in different jurisdictions. The result shows the grouping, and therefore, the process can be taken into account for predictions or as a basis of legal application for similar cases. Table 1 shows the assignment of the groups to each case, it is striking that in the cases of 2015–2016 for the file, OT2015_20160416, there is a difference in the application of the rules, there may have been slight variations in the legal contribution to each one of the cases, an element to take into account when it is necessary to apply in similar legal cases. On the other hand, toward 2016 the decisions and the application of laws remain stable, so it is suggested at the time of taking these files as a basis that the bibliographic legal reference be from that year. This investigation is in process, so in future cases we will address the methodological process with a greater number of files; however, it can be concluded at a glance that legal decisions in different jurisdictions can be grouped by types of applied laws.

Table 1 Case versus cluster

caseName	Cluster
Acosta-Febo v. Franklin California Tax-Free Trust	0
Americold Realty Trust v. ConAgra Foods	1
Americold Realty Trust v. ConAgra Foods	1
Americold Realty Trust v. ConAgra Foods	0
Bank Markazi v. Peterson	1
Bank Markazi v. Peterson	1
Bank Markazi v. Peterson	0
Bernard v. Minnesota	0
Betterman v. Montana	0
Beylund v. Levi	0
Birchfield v. North Dakota	0
Bruce v. Samuels	1
Bruce v. Samuels	1
Bruce v. Samuels	0
Campbell-Ewald Company v. Gomez	1
Campbell-Ewald Company v. Gomez	1
Campbell-Ewald Company v. Gomez	0
CPV Maryland LLC v. PPL EngergyPlus LLC	1
CPV Maryland LLC v. PPL EngergyPlus LLC	1
CPV Maryland LLC v. PPL EngergyPlus LLC	0

References

1. Galgani, F., Compton, P., Hoffmann, A.: Knowledge Acquisition for Categorization of Legal, pp. 118–132. Springer, Berlin (2012)
2. Casanovas, P.: Semantic Web Regulatory Models: Why Ethics Matter. Lecture Notes, vol. 28 (2014)
3. Kanapala, A., Pal, S., Pamula, R.: Summarization from Legal Documents a Survey. Artif. Intell. Rev. **51**, 371–402 (2019)
4. Branting, K., Weiss, B., Brown, B., Pfeifer, C., Chakraborty, A., Ferro, L., Ptaff, M., Yeh, A.: Semi-supervised methods for explainable legal prediction. In: Seventeenth International (2019)
5. Poudyal, P., Gonçalves, T., Quaresma, P.: Using Clustering Techniques to Identify Arguments in Legal Documents. ASAIL@ ICAIL (2019)
6. Levelt, W.: On empirical methodology, constraints, and hierarchy in artificial grammar learning. Top. Cogn. Sci. (2019)

7. Pina-Sánchez, J., Grech, D., Brunton-Smith, I., Sferopoulos, D.: Exploring the origin of sentencing disparities in the crown court: using text mining techniques to differentiate disparities. Soc. Sci. Res. **84** (2019)
8. Gómez, A.H.F., Benitez, J., Galarza, L.R.J., del Salto, V.H., Guerrero, D.S., Torres, G.G.: Semantic analysis of judicial sentences based on text polarity. In: Gran Canaria: CISTI 2016 (2016)

COVID-19 Cases—Deaths: First Approach to the Ecuadorian Instance

Juan Manuel Garcia-Samaniego, Hector F. Gomez A., Jorge A. Benitez, and Edwin Fabricio Lozada T.

Abstract The global pandemic caused by COVID-19 is of global concern; therefore, it is vitally important to identify the effects of the virus. In the present work, we work with neural networks to predict the effect of cases versus deaths in Ecuador. The experimental data were collected from official reports from March to May 2020, which served to predict the relationship between the parameters considered. This work is still under development, but it is very important when proposing algorithms based on machine learning and artificial intelligence as decision-making mechanisms.

Keywords COVID-19 · Ecuador · Machine learning

1 Introduction

The severe acute respiratory syndrome Coronavirus2 is responsible for the current critical health situation. The World Health Organization assigned it as the International Classification of Diseases (ICD) code, "COVID-19." The coronavirus can be transferred from animals to humans and then spread between people, which has been increasing exponentially with a growing number of deaths; hence, the goal of

J. M. Garcia-Samaniego (✉)
Departamento de Ciencias Naturales, Universidad Técnica Particular de Loja, Loja, Ecuador
e-mail: mgarcia@utpl.edu.ec

H. F. Gomez A.
Centro de Posgrados, Universidad Técnica de Ambato, Ambato, Ecuador
e-mail: Hf.gomez@uta.edu.ec

J. A. Benitez
Departamento de Jurisprudencia, Universidad Técnica Particular de Loja, Loja, Ecuador
e-mail: jabenitezxx@utpl.edu.ec

E. F. Lozada T.
Carrera de Software, Universidad Regional Autónoma de los Andes–UNIANDES, Ambato, Ecuador
e-mail: ua.edwinlozada@uniandes.edu.ec

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_7

governments worldwide is to reduce the rate of transmission [1]. The regional human mobility criteria represent the connectivity of regions in time, which allows determining the spread and evolution of COVID-19, which is very important in order to verify the spread routes of the virus [2]. Neural networks can be used to predict the effects of the virus when working with time series and in combination with fuzzy logic [3]. In fact, the deep neural network can be used in order to predict deaths from COVID-19 and compare them with the real data of the course of the disease, being consistent with the data predicted by clinical methods, and it is concluded that the outbreak has exponential growth, with a high prediction rate [4]. Machine learning methods are efficient in predicting risks and effects of COVID-19, which makes it possible to prevent the spread of the virus; hence, it is even possible to categorize countries according to their risk, using the Bayesian method, which helps to plan measures preventive in order to counteract the disease [5]. Computational learning allows knowing the parameters of the spread of the virus, such as its level of reproduction, fatality rate and duration of the infection period, and these; they can be used to predict the results of the behavior of the virus, since knowing the real number of infected people is very complex, and therefore, it is necessary that decisions are based on models that allow making predictions, even reaching estimates of the number of daily deaths in the course of a time series [6]. In fact, software such as CoronaTracker were developed in order to provide reliable data and news for efficient decision making, with the aim of predicting and forecasting cases, deaths and recoveries of the virus, as well as evaluating policies and economic decisions related to the spread of the disease [7]. The analysis of the incidence of the virus worldwide can be approached by convolutional neural network mechanisms to determine confirmations and deaths from COVID-19 with just a 7-day time window, with which measures can be significantly expedited. Be taken in cities that would be considered the most affected by the disease [8]. The proposal of this work consists of applying deep neural networks in conjunction with Keras-TensorFlow to predict and verify the number of infections vs deaths in Ecuador from March to May 2020.[1] The methodology consists of improving the predictive model, and the experimental data comes from the official bulletins of the Ecuador. To better explain our proposal, the following section describes the methodology, the experiment and the conclusions resulting from the research.

2 Methodology

To obtain a prediction model that makes it possible to relate COVID-19 cases and deaths, we propose to follow the following steps:

- Obtain the dataset from official bulletins.
- Verify the behavior of the data using histograms.

[1] https://www.gestionderiesgos.gob.ec/informes-de-situacion-covid-19-desde-el-13-de-marzo-del-2020/.

- Improve the predictive model through the adaptability of the curve.
- Train the neural network and obtain the prediction results.

In order to carry out the previous steps, it is necessary to experiment with the neural network with "Keras and Tensor FlowPrograms by Software," "A User-Friendly and Extendable Data Distribution System," "Multi-flip Networks: Parallelizing GenSAT," "Self-determinations of Man."

3 Experimentation

There are two time series that are provided with this data: the time series of COVID-19 cases (officially diagnosed cases, a number generally far from the actual number of cases and difficult to correlate with actual statistics) and the series of time of deaths in which COVID-19 are the proven causes.

Figure 1 shows cases and deaths from COVID-19 in Ecuador until May 2020. We will use 80% of our sample as training samples and the remaining data as test samples, and we will concentrate on predicting deaths.

Figure 2 shows the adaptability of the predictive model, so it is possible to train the neural network and verify the correlational results between cases and deaths.

Figure 3, the histograms tend to look more and more like normal distributions. Finally, let's review our network with these new data and parameters to train the network, with a sigmoid function between −1 and 1.

Prediction model shown in Fig. 4 obtains temporal differences, which must be analyzed due to the proportion between cases and deaths.

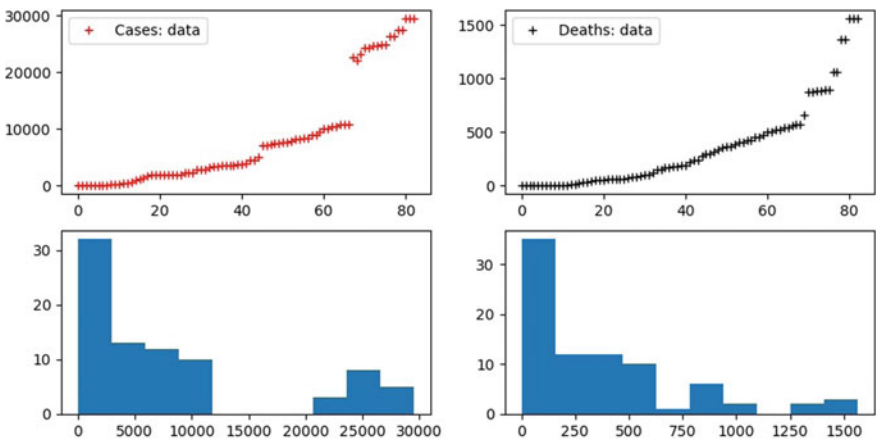


Fig. 1 Comparison of cases versus deaths

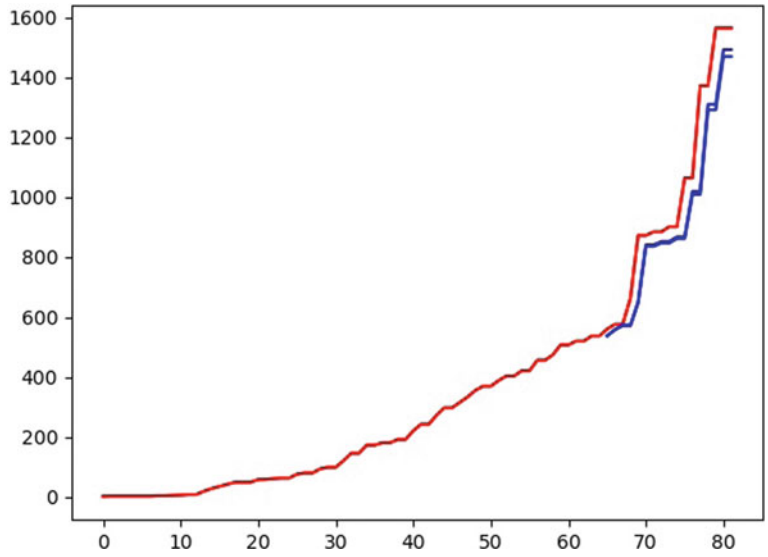


Fig. 2 Prediction model

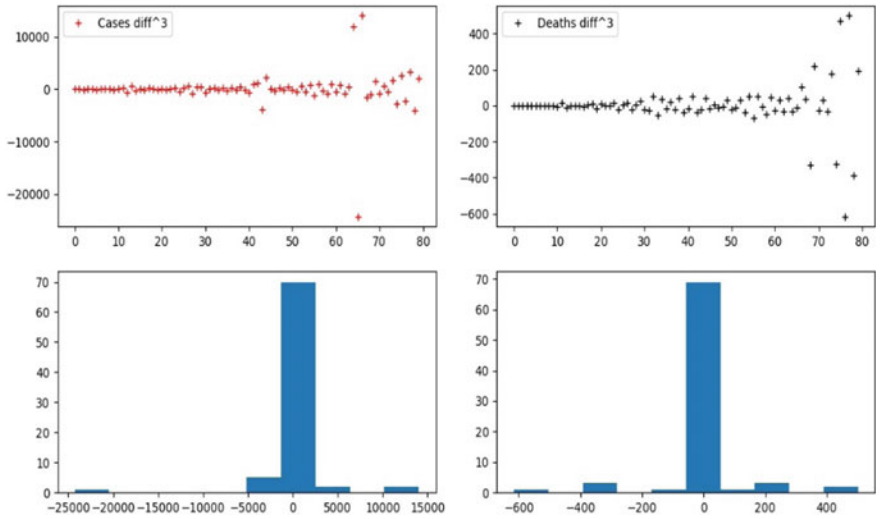


Fig. 3 Comparison of cases versus deaths

4 Conclusions

The goal of this work is to design a prediction model in order to verify the relationship between confirmed cases and deaths from COVID-19, in models which the hard

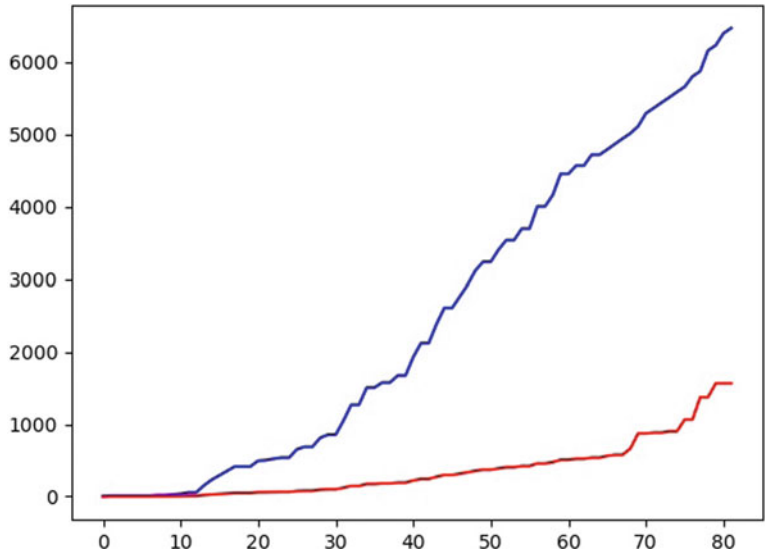


Fig. 4 Adapted prediction model with sigmoid function

adjustment due to the lack of sufficient data to strengthen any prediction model. The model based on neural networks and in Keras with Tensorflow showed that it is possible to have a prediction model, although not with the necessary robustness but that serves as a reference for decision making. If Figs. 2 and 3 are observed, the model can be normalized and even adapted to the functioning of a linear model of a neural network, taking the first steps for a basic prediction model,[2] very close to what can be seen in the final predictive model. The model can be considered reliable, since the data collected is from March to May 2020; however, in the long run, the model can be mistrustful since it is only based on an approximation.

References

1. Chatterjee, A., Gerdes M., Martinez, S.: Statistical explorations and univariate timeseries analyis on COVID-19 datasets to understand the trend of disease spreading and death. Sensors **11**(3089) (2020)
2. Kappor, A., Ben, X., Liu, L., Perozzi, B., Barnes, M., Blais M., O'Banion, S.: Examining COVID-19 forecasting using spatio-temporal graph neural networks. arXiv (2020)
3. Melin, P., Monica, J., Sanchez, D., Castillo, O.: Multiple ensemble neural network models with fuzzy response aggregation for predicting COVID-19 time series: the case of Mexico. Healthcare **8**(2), 181 (2020)
4. Bandyopadhyay, S., Dutta, S.: Machine learning approach for confirmation of covid-19 cases: positive, negative, deatch and release. medRxiv (2020)

[2] https://www.google.com/search?q=cases+covid+ecuador&rlz=1C1CHBD_esEC899EC899&oq=cases+covid+ecuador&aqs=chrome.0.0l7j69i60.2412j0j7&sourceid=chrome&ie=UTF-8.

5. Pal, R., Sekh, A., Kar, S., Prasad, D.K.: Neural network based country wise risk prediction of COVID-19. arXiv
6. Stochitoiu, R.D., Rebedea, T., Popescu I., Leordeanu, M.: A self-supervised neural-analytic method to predict the evolution of COVID-19 in Romania. arXiv (2020)
7. Hamzah, F.B., Lau, C., Nazri, H., Ligot, D.V., Lee G., Tan, C.L.: CoronaTracker: worldwide COVID-19 outbreak data analysis and prediction. Bull. World Health Organ. **1**(32) (2020)
8. Alazab, M., Awajan, A., Mesleh, A., Abraham, A., Jatana V., Alhyari, S.: COVID-19 prediction and detection using deep learning. Int. J. Comput. Inf. Syst. Ind. Manag. Appl. **12**, 168–181 (2020)

Traffic Rules Violation Detection System

Vishal B. Pattanashetty, Venkatesh Mane, Nalini C. Iyer, and Shweta Kore

Abstract In the world of new evolving technologies, violation of traffic rules has become a critical issue for the majority of the developing countries. With growing population, the number of vehicles on road is also increasing rapidly, and also, violations in traffic rules are increasing exponentially. Hence, managing the traffic violations has become a tedious task. Even though there are several automated technologies to manage the traffic violations, due to the nonuniform illumination conditions, diversity of license plate formats, it is very challenging to handle those conditions. Hence, the solution to this problem is to develop a system which is coupled with several parameters like traffic signal detection, speed estimation to determine how often a driver violates the rules of the road. This system detects the violation of traffic signals which can be combined with speed estimation to check for the over speed, and this information is sent to the database where the necessary actions can be taken against those who are violating the rules by the concerned authorities.

Keywords Red light violation · Speed estimation · Open ALPR · Traffic violation

1 Introduction

Over the time period of three decades, there has been a great surge in national road networks in urban and semi-urban areas. This escalation in road networks has given rise to the new problem of increase traffic. Hence, there is a requirement of

V. B. Pattanashetty (✉) · V. Mane · N. C. Iyer
KLE Technological University, Hubballi, India
e-mail: vishalbps@kletech.ac.in

V. Mane
e-mail: mane@kletech.ac.in

N. C. Iyer
e-mail: nalini.c@kletech.ac.in

S. Kore
JSS, Dharwad, India

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_8

effective management and monitoring of traffic. Some basic techniques for traffic measurements like inductive loops, sensors, EM wave detectors suffer from serious shortcomings and are also very expensive to install on the road side as they are very bulky and require a lot of manpower to install them [1]. On the contrary, primarily the video-based structures are very clean to put in as they use existing visitors surveillance infrastructure. Moreover, these systems also can be without problems upgraded with the emerging technology, and they are flexible to redesign. These structures can be altered simply by converting the algorithms [2]. Those systems are used to manage the traffic congestion, count the number of vehicles and identification of traffic incidents. Different video and image processing techniques can be deployed to detect the vehicles and objects. In order to solve this real-time problem of traffic and to stay away from these unexplored consequences, traffic rules violation detection systems are very much needed. These traffic violation detection systems enforce the proper and well-regulated identification and controlling of traffic violations [3]. The violation of rules should be realized in real time so that the traffic police department or the authorities can easily track the road all the time. Hence, it will not be easier for the traffic enforcers to implement the safe roads efficiently [4] These automatic traffic rules violation detection systems can detect traffic light violations and over speeding of vehicles in real time. And the user-friendly graphical user interface which is associated with the system helps to easily operate the system and monitor the traffic. Then, necessary actions can be taken against the violation of traffic rules [5].

2 System Design

In this section, we list out the interfaces. Functional block diagram helps in giving the brief explanation about the process. The alternative parameters are explained. As shown in Fig. 1, video is taken as input from the traffic surveillance camera, and using deep learning model, vehicle detection and tracking is done to check if any violations.

Check for violation is carried out, and details of violators license plate number are extracted sent to the traffic department to take further actions accordingly.

Fig. 1 Block diagram

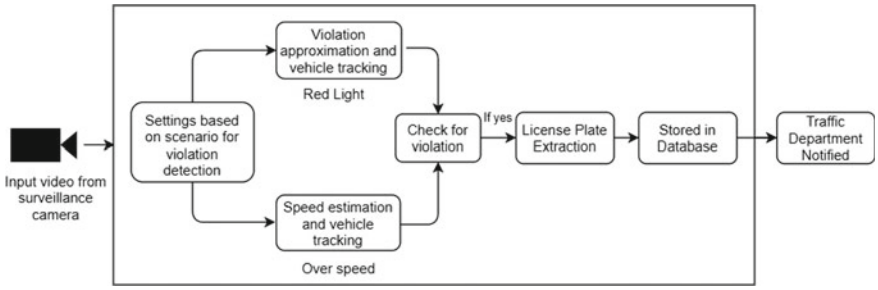


Fig. 2 System architecture

3 Methodology

3.1 System Architecture

As shown in Fig. 2, the system architecture represents the complete view of the traffic rule violation detection system. The system concentrates on detection of two violations, i.e., red light violation or over speed estimation based on the settings done by the user on the basis of scenario.

3.2 Red Light Signal Violation

As shown in Fig. 3, the system is divided into components and mainly consists of violation line approximation, vehicle detection and vehicle tracking.

Violation line approximation

The image is converted to gray scale so that it can be used as an input to binary level thresholding. As shown in Fig. 4, erosion and dilation process are carried out in order to improvize the components representing the crosswalk and reduce the noise of the image. Contours are detected from the binary image. Hence, it is possible to represent the closed boundaries. The OpenCV polygon approximation is done to accept only the rectangular-shaped contours. For better visualization, contours near the traffic signals are represented by bounding boxes. And finally, the horizontal line is drawn in order to represent the violation line.

Fig. 3 Red light signal violation

Fig. 4 Pre-processing techniques

Vehicle detection and tracking: For vehicle detection and tracking, the system uses tensor flow with SSD object detection model. From these detections each vehicle can be tracked across the video. Hence, we can check whether the vehicle has crossed the red light signal.

Violation detection: The vehicle is considered to have violated the rule if crosses the line of approximation. If the vehicle's center y value crosses the approximated violation line its tracker is moved to the violated tracking list.

3.3 Speed Estimation

As shown in Fig. 5 the input video is analyzed. For vehicle detection and tracking, tensor flow with SSD object detection model is used. From this detection, each vehicle can be tracked across the video. Hence, we can check whether the vehicle has exceeded the threshold speed.

For finding the speed of the vehicle, the system uses the simple formula. For this, the input video frame is taken, and width of the road is calculated digitally in pixels. To map these distances in pixels and real world, the pixels of the image are divided by width of the road. Pixels (dpixels) are then converted to meters (dmeters) and then multiplied with the frames per second to get the speed of the vehicle. The formula for calculating the speed of the vehicle is as shown.

$$\text{Speed} = \text{dmeters} * \text{fps} * 3.6 \tag{1}$$

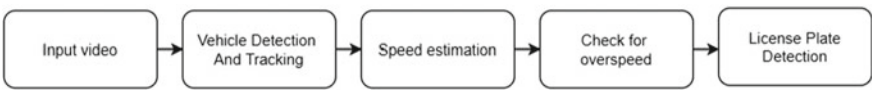


Fig. 5 Speed estimation

As shown in Eq. 1, the dmeters is multiplied with frames per second (fps) and 3.6 to obtain speed value in km/h.

3.4 Vehicle Detection and Tracking

For vehicle detection and trackin, tensor flow object detection API was used with SSD Mobile Net V2 deep learning model (Fig. 6).

The SSD architecture is a single convolution network that learns to predict bounding box locations and classify these locations in one pass. Hence, SSD can be trained end to end. The SSD network consists of base architecture (Mobile Net in this case) followed by several convolution layers. The Mobile Net model is based on depth-wise separable convolutions which are a form of factorized convolutions. These factorize a standard convolution into a depth-wise convolution and a 1*1 convolution called a point-wise convolution.

1. **Vehicle Detection**
   At each five frames, a detection is done. The detection output is formulated by several steps, from filtering the bounding boxes with low confidence rate and filtering any bounding box that is not a vehicle to finally doing non-maximum suppression to the detected boxes, so that each vehicle has only one bounding box.
2. **Vehicle tracking**
   Keeping an accurate localization on each vehicle can be very tedious task, and tracking the first frame vehicles is not enough, due to incoming and outgoing vehicles of the scope of the images. To resolve this, a merging between tracking and detection tasks is needed. The tracking of the vehicles is done at every frame, but every five frames, a new detection occurs, then for each detected bounding box, a measure of IoU (intersection over union) is done with the current tracking bounding boxes. If a detected box matches a tracking box with a relatively good percentage, then it is the same box that in the tracking list, but the new detected

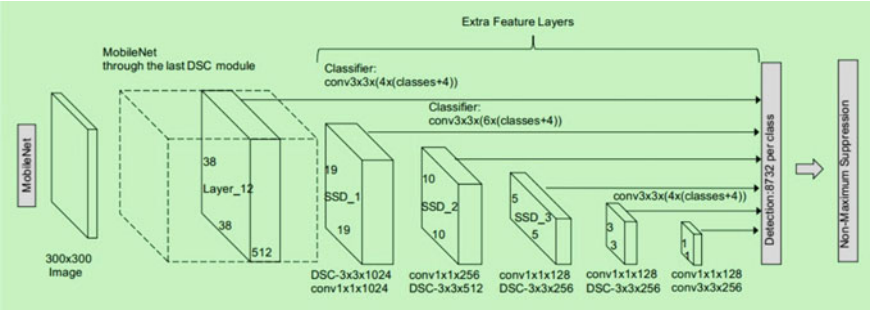


Fig. 6 Mobile Net-SSD model

bounding box has to be more accurate than the tracking one, so an adjustment operation occurs to the tracking object bounding box.

3.5 License Plate Detection

To extract the license plate number of the violator vehicle, OpenALPR was used. OpenALPR is an open source automatic license plate recognition library. The library analyzes images and video streams to identify license plates. It makes use of OpenCV and Tesseract OCR. OpenCV uses edge detection and feature detection techniques for locating the plate. The Tesseract OCR engine is used to identify the detected characters on the plate. The output is the text representation of any license plate characters.

3.6 GUI

The graphical user interface is developed using Tkinter in order to maintain the details of the vehicle owner who has violated the rule. The license plate of the vehicle extracted, and details like vehicle ID, vehicle type, registration number, type of rule broken, date, time, fine to be paid and the status are recorded in Google sheet (Figs. 7 and 8).

The images of the violated vehicles as shown in Fig. 9 are stored in the Google drive. The GUI is shared to all the traffic police stations. The vehicle owner can go to the nearest police station, then, the authorized user can log in with the details, and license plate number is entered and searched if the particular vehicle has violated the rule. If the vehicle owner's details are found in the database, the amount of fine to

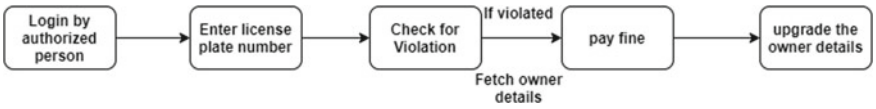


Fig. 7 Block diagram of GUI

Fig. 8 Search box

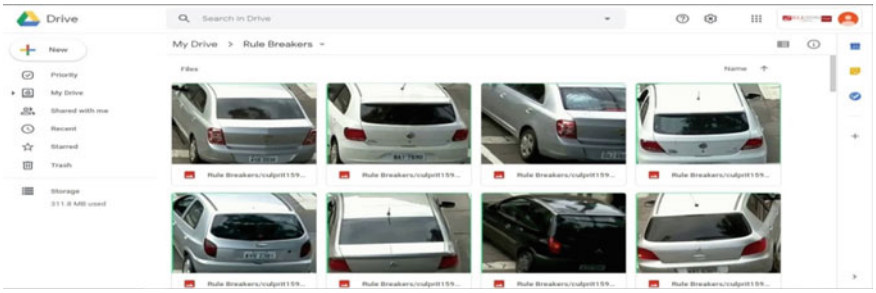


Fig. 9 Search box

be paid is mentioned. If the vehicle owner pays the fine, then the status is updated in the database.

3.7 Flowchart

See Fig. 10.

4 Results

In this section, we represent the results of detection of red signal violations, speed estimation and overview of the GUI.

1. Speed Estimation (Fig. 11)
2. Red Signal Violation Detection
   The detection of tracking box changes its color from green to red as shown in Fig. 13 signifying the crossing of approximated violation line by the vehicle (Fig. 12).
3. Database
   A violated vehicle image is captured and can be seen in top right corner of Fig. 14. License plate number of this violated vehicle is extracted. The Google sheet as shown in Fig. 15 consists of the details of the violated vehicles. Once the fine is paid by the violator, status is updated and reflected in the Google sheet.

5 Conclusion

The system does violation detection of red light and over speed. The number plate is successfully extracted of the violator. The GUI can be used to search the details of

Fig. 10 Flowchart

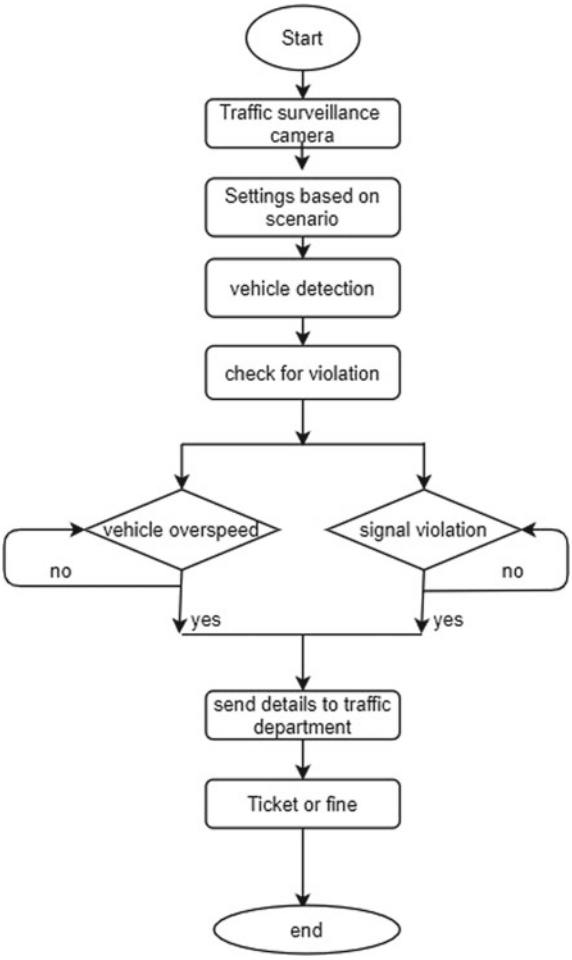


Fig. 11 Example 1 of speed estimation

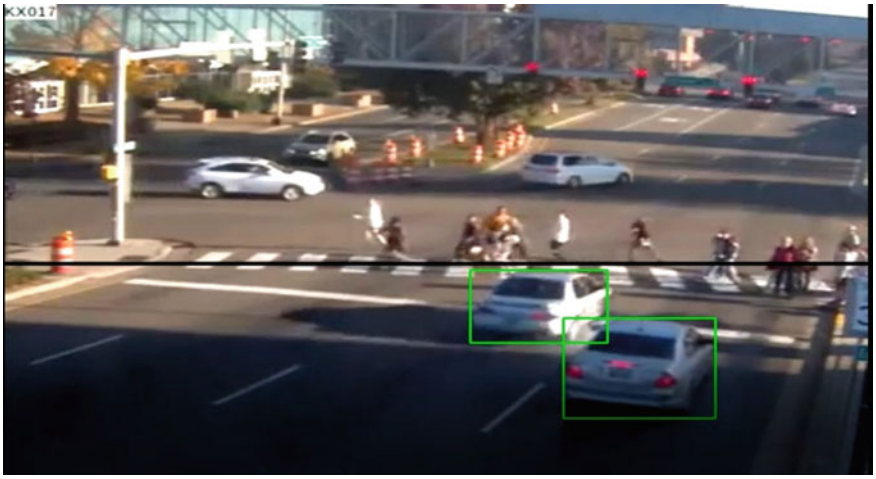
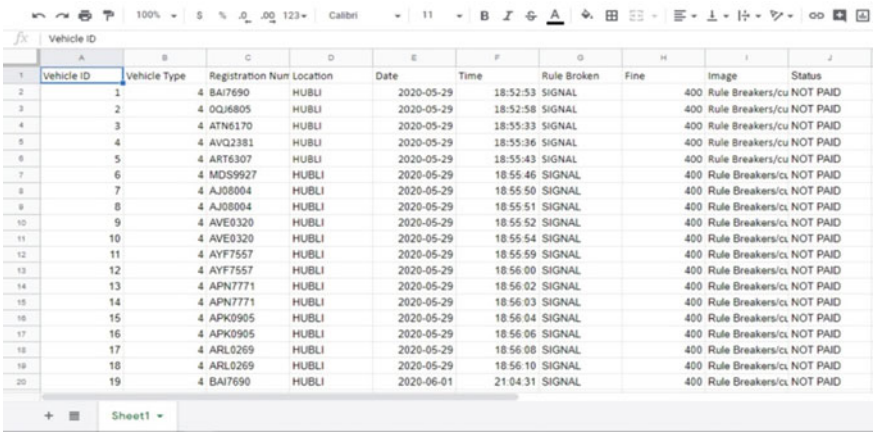


Fig. 12 Before violation

Fig. 13 After violation

the violator and collect the fine. Hence, an integrated traffic rules violation detection system was designed. In future, we would like to detect violations in night traffic vision. Also, we would like to detect other types of traffic rule violations and make the system more robust.

Fig. 14 License plate detection



	Vehicle ID	Vehicle Type	Registration Num	Location	Date	Time	Rule Broken	Fine	Image	Status
2	1	4	BAI7690	HUBLI	2020-05-29	18:52:53	SIGNAL	400	Rule Breakers/cu	NOT PAID
3	2	4	0QJ6805	HUBLI	2020-05-29	18:52:58	SIGNAL	400	Rule Breakers/cu	NOT PAID
4	3	4	ATN6170	HUBLI	2020-05-29	18:55:33	SIGNAL	400	Rule Breakers/cu	NOT PAID
5	4	4	AVQ2381	HUBLI	2020-05-29	18:55:36	SIGNAL	400	Rule Breakers/cu	NOT PAID
6	5	4	ART6307	HUBLI	2020-05-29	18:55:43	SIGNAL	400	Rule Breakers/cu	NOT PAID
7	6	4	MDS9927	HUBLI	2020-05-29	18:55:46	SIGNAL	400	Rule Breakers/cu	NOT PAID
8	7	4	AJ08004	HUBLI	2020-05-29	18:55:50	SIGNAL	400	Rule Breakers/cu	NOT PAID
9	8	4	AJ08004	HUBLI	2020-05-29	18:55:51	SIGNAL	400	Rule Breakers/cu	NOT PAID
10	9	4	AVE0320	HUBLI	2020-05-29	18:55:52	SIGNAL	400	Rule Breakers/cu	NOT PAID
11	10	4	AVE0320	HUBLI	2020-05-29	18:55:54	SIGNAL	400	Rule Breakers/cu	NOT PAID
12	11	4	AYF7557	HUBLI	2020-05-29	18:55:59	SIGNAL	400	Rule Breakers/cu	NOT PAID
13	12	4	AYF7557	HUBLI	2020-05-29	18:56:00	SIGNAL	400	Rule Breakers/cu	NOT PAID
14	13	4	APN7771	HUBLI	2020-05-29	18:56:02	SIGNAL	400	Rule Breakers/cu	NOT PAID
15	14	4	APN7771	HUBLI	2020-05-29	18:56:03	SIGNAL	400	Rule Breakers/cu	NOT PAID
16	15	4	APK0905	HUBLI	2020-05-29	18:56:04	SIGNAL	400	Rule Breakers/cu	NOT PAID
17	16	4	APK0905	HUBLI	2020-05-29	18:56:06	SIGNAL	400	Rule Breakers/cu	NOT PAID
18	17	4	ARL0269	HUBLI	2020-05-29	18:56:08	SIGNAL	400	Rule Breakers/cu	NOT PAID
19	18	4	ARL0269	HUBLI	2020-05-29	18:56:10	SIGNAL	400	Rule Breakers/cu	NOT PAID
20	19	4	BAI7690	HUBLI	2020-06-01	21:04:31	SIGNAL	400	Rule Breakers/cu	NOT PAID

Fig. 15 Details of violated vehicles

References

1. Mampilayil, H.R., Rahamathullah, K.: Deep learning based detection of one way traffic rule-violation of three wheeler vehicles. In: 2019 International Conference on Intelligent Computing and Control Systems (ICCS), Madurai, India, pp. 1453–1457 (2019). https://doi.org/10.1109/ICCS45141.2019.9065638
2. Huang, T.: Traffic Speed Estimation from Surveillance Video Data (2018). https://doi.org/10.1109/CVPRW.2018.00029
3. Mallissery, S., Manohara Pai, M.M., Ajam, N., Pai, R.M., Mouzna, J.: Transport and traffic rule violation monitoring service in ITS: a secured VANET cloud application. In: 2015 12th Annual IEEE Consumer Communications and Networking Conference (CCNC), Las Vegas,

NV, pp. 213–218 (2015). https://doi.org/10.1109/CCNC.2015.7157979
4. Wang, X., Meng, L.-M., Zhang, B., Lu, J., Du, K.-L.: A video-based traffic violation detection system. In: Proceedings—2013 International Conference on Mechatronic Sciences, Electric Engineering and Computer, MEC 2013, pp. 1191–1194 (2013). https://doi.org/10.1109/MEC.2013.6885246
5. Younis, A., Shixin, L., Jn, S., Hai, Z.: Real-time object detection using pre-trained deep learning models MobileNet-SSD CCS concepts

Effectiveness of Online Shopping Advantages of Healthy Food Products on Consumer Buying Behaviour

Amar Nath Gupta and Pradnya Chitrao

Abstract *Purpose* The reason for this exploration is to examine the connection between consumer buying behaviour and online shopping advantages for packaged healthy food products. The research aims to provide knowledge and insights on effective online marketing strategies for marketers engaged with the online creation and appropriation of healthy food products. *Design/methodology/approach* The Cronbach's Alpha Test was applied to test the unwavering quality of the scale, and the outcomes demonstrated that the Likert scale of the survey is dependable. The standard deviation was calculated to know the variation in the opinions of the respondents. The researchers have considered only packaged healthy food products under the study. The SPSS-22 version was utilized for statistical analysis of the data. Primary data was collected through randomly distributing the questionnaire among 392 consumers particularly those who buy healthy food products online from August 2019 to March 2020. The answers of the respondents were estimated on a five-point Likert scale. ANOVA and F-tests were applied to test the null hypothesis. The correlations test was applied to analyse the connection between the factors under investigation. The regression model was created dependent on the examination of information. *Findings* The discoveries of the exploration uncovered that online shopping advantage has a critical relationship on consumers buying behaviour. The study further revealed that the consumer would buy more healthy food products online if the marketers provide online shopping advantages like convenience, reward point system, cashback offers and comparison to purchasing healthy food products. *Research Implications and Limitations* Practical implications—The results of the study highlight the significance of online marketing strategies in influencing consumer buying behaviour. The investigation recommends that marketers must connect with buyers utilizing the manner in which they wish to be locked in by structuring more fitting promotional

A. N. Gupta (✉)
Symbiosis International University, Lavale, Pune, India

P. Chitrao
SIMS, Khadki, Pune, India
e-mail: pradnyac@sims.edu

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_9

techniques. The restriction of the study is that it has been led on healthy food products. Future examinations could be stretched out to different classes of fast moving consumer goods to sum up the discoveries of the investigation.

Keywords Online shopping advantage · Consumer buying behaviour · Healthy food products · Health conscious · Online purchasing · Convenience · Reward

1 Introduction

Technological advancements have strikingly changed human life in a noteworthy way. The infiltration of wireless network among buyers has made new open doors for the producers and marketers for the brand positioning and creating brand image of their offers adequately and effectively [31, 37, 40]. The e-commerce industry in the country has experienced immense growth in the last few years due to an increase in the total number smartphone users, penetration of Internet services through the introduction of 4G networks by the cellular service providers in the urban and rural parts of the country, the advancements in the field of information technology especially data analytics, artificial intelligence and search engine optimization (SEO). The growing incomes of middle-class consumers have further fuelled the growth of the e-commerce industry in the country. Morgan Stanley [38] sees the Indian e-commerce industry of $ 15 billion in 2016 to $ 200 billion by 2026.

Moreover, as web-based shopping in India is developing quickly, more exact examination is expected to comprehend shopper conduct and perspectives in this advancing digital environment. Online shopping offers many benefits that physical stores cannot offer, i.e. ease and convenience of purchase, promoting user-friendly websites, offering a wide range of products, reducing transaction costs and saving time of the customers. Customers have the chance to purchase anything from any piece of the world without really setting off to the outlets whenever. It permits the customers to put in a request for the merchandise/administrations at their straightforwardness and comfort additionally diminishes exchange costs. In addition, it gives customers the opportunity to make 24-h business bargains. Web-based shopping sets up customer and supplier relationships and improves customer steadfastness and maintenance. Accordingly, it assists with pulling in countless purchasers to shop on the web [28, 34, 41].

The activities on the Internet are included in online marketing strategies [5]. Therefore, online marketing strategies can be used as Internet marketing strategies [5]. For marketing along with achieving the marketing goals, the Internet has become a tremendous vehicle and has become the forefront application of the consumer purchase applications [27]. It has been observed that the quality of the website plays a key role in online impulse buying behaviour (Akram et al. 2018).

As per the Indian tradition, the consumers have mostly demanded the food products which are fresh and unprocessed as compared to processed and packaged food; nevertheless, the consumers have demonstrated huge changes in their consumption

patterns which have given a ton of probability to the makers and advertisers of healthy food products in the nation [9, 12, 24]. Health food marketers are trying to explore every single opportunity to engage their consumers in the way they wish to be further engaged. The consumers have become more health conscious in recent times as they have more access to such information related to their nutrition needs. Subsequently, they are making healthier choices for food products as per their lifestyle and dietary requirements. The growing awareness among consumers has led to an increasing demand for natural and organic food products. Many previous research studies have avowed that while analysing the nutrition label of the food products, the consumers are very keen to know about the total fats and calorie counts [11, 13, 15, 18, 17, 21, 25]. The role of the web and web-based media for getting the hang of, instructing and sharing data about eating habits has expanded. Consumers are presented to an assortment of messages from numerous Internet users, including bloggers, YouTubers, peers and numerous other social media players. Healthy food marketers are progressively holding onto social-media as a device to speak with their customers.

Couched in this background, this analysis investigates the relationship between consumer buying behaviour and online shopping advantages on packaged healthy food products. To have a competitive edge for superior market attention, it is critical for the producers and marketers to have a sound understanding of the consumers' buying behaviour. According to Perry [30], marketers using personalized online communication use online marketing strategies to build brand awareness and brand loyalty. The research aims to provide knowledge and insights on effective online marketing strategies for marketers involved in the online production and distribution of healthy food brands. The online shopping advantages like convenience, reward point system, cashback offers and comparison to purchasing healthy food products play a significant role in influencing consumer buying behaviour.

2 Literature Review

2.1 Understanding Online Shopping Advantages and Its Impact on Consumer Buying Decision

The quantity of cell phones and other web-empowered gadgets and their applications and administrations makes getting to the Internet simpler and more beneficial. This has led to a tremendous increase in online shopping, which in turn has led to increased competition in the online marketplace [23, 39]. In the conventional physical shopping framework where customers need to go to the shops to purchase merchandise, the contemporary web-based shopping makes the correct situation for different organizations to arrive at buyers legitimately around the globe [4, 22]. The investments in new product development of the food industry in healthy food segment include remarkable innovations in the product such as health claims for food and the nutrition, sales distribution, brand differentiation, improved nutrition profiles, "hyper-personalized"

products, digitization, nutritional genomics, material diversification and packaging [3, 8, 10, 14, 29, 32, 33].

Online shopping offers many benefits that physical stores cannot offer, i.e. ease and convenience of purchase, promoting user-friendly websites, offering a wide range of products, reducing transaction costs and saving time of the customers. Clients have the chance to purchase anything from any piece of the world without really heading off to the outlets whenever. It permits the clients to submit a request for the merchandise/administrations at their simplicity and accommodation additionally lessens exchange costs. Additionally, it furnishes clients with the office to make 24-h business bargains. Web-based shopping sets up client and provider connections and improves client devotion and maintenance. Consequently, it assists with pulling in countless consumers to shop on the web [28, 34, 41].

2.2 Constructs for Online Shopping Advantage

2.2.1 Consumer Awareness

The awareness has been portrayed as human perception and consistent response to the condition of what they devour or use it regularly. Awareness is the first stages in the buying decision process where purchasers who from the outset are not very much familiar with the item gotten familiar with it [7]. Awareness is the initial phase in the consumer's buying behaviour that is concerned about anything which could prompt consumer's advantage trailed by different stages in the purchasing process [6]. Numerous studies underscored that awareness positively affects customer purchase intention [5–7, 19, 20, 26]. The online marketing strategies help in increasing the market share and profitability of the organizations in the competitive marketplace [36].

2.2.2 Convenience on Online Shopping

Convenience is viewed as one of the most significant components of influencing web-based shopping. Not at all like offline-based shopping, the striking highlights of Internet shopping are its comfort, which has been depicted as the primary reason for customers to make online purchases. [28]. The online sales promotional offers play a key role in influencing consumer buying behaviour. According to Singh et al. [35], "The gigantic inflow of online deals and limited-time offers in various web-based business sites which are delighting consumers to go for online shopping". It has been seen that the design of the site assumes a crucial role in online motivation purchasing conduct. The review of the literature and opinions of the experts indicate that online shopping advantage is important for the marketing of healthy food products as it influences consumer buying behaviour. According to Perry [30], marketers using personalized online communication use online marketing strategies to build brand

awareness and brand loyalty. The research aims to provide knowledge and insights on effective online marketing strategies for marketers involved in the online production and distribution of healthy food brands.

3 Defining the Research Gap and Formulation of Research Questions and Objectives

3.1 Research Gap

From the literature survey, it has been distinguished that there is an absence of enough examination on inspecting the effect of Internet shopping advantage on the buyer purchasing conduct and identifying the linkage between online shopping risks and consumer buying behaviour towards packaged healthy food products.

3.2 The Following Research Questions Were Identified for the Study

1. Is there any shopping advantage for the customers while buying healthy food products online?
2. Is there any relationship between consumer buying behaviour and online shopping advantage on healthy food products for the consumers?

3.3 The Research Objectives Were

(a) To study the relationship between consumers buying behaviour and online shopping advantages on packaged healthy food products.
(b) To explore whether the purchase intention of consumers is influenced by online shopping advantages or not.

4 Design/Methodology/Approach

The Cronbach's Alpha Test was applied to test the unwavering quality of the scale, and the outcomes demonstrated that the Likert scale of the survey is dependable. The standard deviation was calculated to know the variation in the opinions of the respondents. The researchers have considered only packaged healthy food products under the study. The SPSS-22 version was utilized for statistical analysis of the data. Primary data was collected through randomly distributing the questionnaire

among 392 consumers particularly those who buy healthy food products online from August 2019 to March 2020. The reactions of the respondents were estimated on a five-point Likert scale. ANOVA and F-tests were applied to test the null hypothesis. The correlations test was applied to analyse the connection between the factors under investigation. The regression model was created dependent on the examination of information. **Primary data was analysed using percentages, arithmetic mean and standard deviation.** To compare the mean value of the variables the **percentages** were used. **The arithmetic mean** values were used to analyse the impact of variables and to test the hypothesis of the research. **The standard deviation (SD)** was calculated to analyse the variation in the responses. The convenience sampling (nonprobability sampling) method was used to collect the primary data.

5 Results and Discussion

It was seen that the larger part (very nearly 60%) of the respondents was in the age group of 25–35 years followed by 23% in the age group of 36–45 years, and the rest of the rate was of over 45 years age group.

Test of Reliability of Scale: To approve the unwavering quality of the Likert scale utilized in the survey, Cronbach's alpha test is applied for inquiries of every factor. The after-effects of the Cronbach's alpha test are introduced in Table 1 (Table 2).

The consequences of the study showed that most of the respondents concurred that the online purchase of packaged healthy food products is a convenient procedure. Further, it was found that the majority of respondents strongly agreed that they can easily place the order of purchasing the packaged healthy food products online without any difficulties. Most of the respondents concurred that they get ease of comparison while placing the order for the purchase of healthy food products online. The outcomes have affirmed that the vast majority of the customers search for online offers, rebate coupons, reward focuses and other such motivating forces while making

Table 1 Result of Cronbach's alpha, α Test

Construct	Cronbach's alpha value	Result of test
Advantages of online shopping	0.888	Accepted

Table 2 Descriptive statistics

Descriptive statistics					
	N	Minimum	Maximum	Mean	Standard deviation
Mean score shopping advantage	392	21	100	82.26	8.386
Valid N (listwise)	392				

the purchase decision. It was confirmed that the easy accessibility of the website also plays a crucial role in influencing the purchase intention of the consumer. It was found that consumers prefer to buy branded food products as they trust the brands. The respondents agreed that online shopping allows them to return the product to the company if it does not meet their expectations. The results of the study have further confirmed that the purchase intention is positively influenced by online shopping advantages. The researchers have found that there is always a risk factor associated with the online transaction, and it negatively affects the consumer buying decision process. The research findings indicated towards minimization of the risk factors associated with online shopping to influence the consumers to place the order of purchase. Overall, the study proves that online marketing strategies play a significant role in creating a brand image and brand loyalty.

Testing of Hypothesis

Null Hypothesis H0: There is no relationship between consumer buying behaviour and online shopping advantages on healthy food products.

Alternative Hypothesis H1: There is some relationship between consumer buying behaviour and online shopping advantages on healthy food products.

To test the above null hypothesis, Pearson correlation is applied. The outcomes are appeared in Table 3.

Interpretation: The table shows that the calculated Pearson correlation coefficient between buying behaviour and online shopping advantages on the packaged health food products is 0.462. The calculated *p*-value is 0.000. This is under 0.05. In this way, the test is rejected. Consequently, the null hypothesis is rejected, and the alternate hypothesis is acknowledged. Subsequently, it is reasoned that there is a huge positive connection between shopper purchasing conduct and online shopping preferences on packaged health food products. It implies that the web-based shopping advantage decidedly impacts the consumer's purchase intention towards healthy food products.

Table 3 Correlations test results

Correlations			
		Mean score of buying behaviour	Mean score shopping advantage
Mean score of buying behaviour	Pearson correlation	1	0.462**
	p-value (2-tailed)		0.000
	N	392	392
Mean score shopping advantage	Pearson correlation	0.462**	1
	p-value (2-tailed)	0.000	
	N	392	392

**Correlation is significant at the 0.01 level (2-tailed)

6 Findings and Conclusion of the Study

The research has proved that respondents below 25 years and those who are aged between 25 and 35 years have a significantly higher positive attitude towards buying health products as compared to respondents aged 36–45 years and above 45 years. The study has further indicated that price and promotional offers play a significant role in influencing the buying behaviour of the respondents. The findings revealed that online shopping advantage has a significant correlation on consumers buying behaviour. The study further revealed that the consumer would buy more healthy food products online if the marketers provide online shopping advantages like convenience, reward point system, cashback offers and comparison to purchasing healthy food products. It was also found that online shopping offers many benefits that physical stores cannot offer, i.e. ease and convenience of purchase, promoting user-friendly websites, offering a wide range of products, reducing transaction costs and saving time of the customers. Consumers have the chance to purchase anything from any piece of the world without really setting off to the outlets whenever. It permits the consumers to submit a request for the products/administrations at their straightforwardness and comfort additionally decreases exchange costs. Additionally, it furnishes consumers with the opportunity to make 24-h business bargains. Web-based shopping builds up consumer and marketers relationships and improves consumer satisfaction and retention. Therefore, the outcome of the study is in line with the findings of Omotayo and Omotope [28], Sareen and Jain [34], Zhou et al. [41].

It is concluded that there is a significant positive relationship between consumer buying behaviour and online shopping advantages on packaged health food products. It implies that the web-based shopping advantage emphatically impacts the shopper's buy goal towards healthy food products. The outcomes have affirmed that the vast majority of the customers search for online offers, rebate coupons, reward focuses and other such motivating forces while making the purchase decision. It has been further concluded that the convenience is viewed as one of the most significant components of influencing web-based shopping for the healthy food products. The study further concluded that online shopping provides for the ease of comparison to the consumers while placing the order for the purchase of healthy food products online. The study would also help in creating brand awareness of the packaged health food brands online, maintaining or achieving profitability through the sustainability of business by creating strong brand positioning in the market.

7 Research Implications and Limitations

The results of the study feature the noteworthiness of web-based marketing strategies in impacting buyer purchasing conduct. The examination proposes that marketers must draw in consumers through the manner in which they wish to be engaged by planning more suitable online marketing strategies. The consequences of this study

may give healthy food product marketers an understanding of compelling web-based marketing techniques and tools to increase the overall profits by improving the brand image of the products and contributing in the brand positioning. Finally, the study confirms the findings of Halabi and Lussier [16], which stated that implementing organizations' marketing efforts and online marketing strategies can increase overall profitability and reduce the percentage of business failures. The primary confinement of this exploration study is that it was essentially constrained to students and employees working in private and government sector organizations in the Pune and Mumbai region and did not include other respondents from other parts of the country as the researchers were having the easy access to such respondents. Therefore, there is a possibility that the outcome of this research does not represent the population of the other parts of the country in terms of consumer buying behaviour. The second limitation of the study is that there is a possibility that some of the respondents might have provided biased information that might have affected the outcome of the study. Future investigations could be reached out to different classes of fast moving consumer goods to sum up the discoveries of the research.

References

1. Abubakar, F.M., Ahmad, H.B.: The moderating effect of technology awareness on the relationship between UTAUT constructs and behavioural intention to use technology: a conceptual paper. Austr. J. Bus. Manage. Res. **3**(2), 14–23 (2013)
2. Agwu, E., Murray, P.J.: Drivers and inhibitors to e-commerce adoption among SMEs in Nigeria. J. Emerg. Trends Comput. Inf. Sci. **1**(3) (2014)
3. Akkerman, R., Farahani, P., Grunow, M.: Quality, safety and sustainability in food distribution: a review of quantitative operations management approaches and challenges. Oper. Res. Spectrum **32**(4), 863–904. https://doi.org/10.1007/s00291-010-0223-2
4. Al-Maghrabi, T., Dennis, C., Halliday, S.: Antecedents of continuance intentions toward e-shopping: the case of Saudi Arabia. J. Enterp. Inf. Manag. **24**(1), 85–111 (2011)
5. Aleksandar, G., Olja, M.: The Many Faces of Internet Marketing. IRENET—Society for Advancing Innovation and Research in Economy, Dubrovnik, pp. 412–417
6. Ansari, Z.A., Alhazemi, A.: An empirical study of the consumer awareness and acceptance of online shopping in Saudi Arabia. Int. J. Dev. Res. **6**(2), 6918–6925 (2016)
7. Bashir, A.M., Bayat, A., Olutuase, S.O., Abdul Latiff, Z.A.: Factors affecting consumers' intention towards purchasing halal food in South Africa: a structural equation modelling. J. Food Prod. Market. **25**(1), 26–48 (2018)
8. Battini, D., Calvazara, M., Persona, A., Sgarbossa, F.: Sustainable packaging development for fresh food supply chains. Packag. Technol. Sci. **29**(1), 25–43. https://doi.org/10.1002/pts
9. Chengappa, P.G., Achoth, L., Rashmi, P., Dega, V., Reddy, B.M.R., Joshi, P.K.: Emergence of organised retail chains in India during post liberalization era. Paper presented at the South Asia Regional Conference of the International Association of Agricultural Economists, Globalisation of Agriculture in South Asia, Hyderabad, The World Bank, Washington, DC (2005)
10. Collins, O., Bogue, J.: Designing health-promoting foods for the ageing population: a qualitative approach. Br. Food J. **117**(12), 3003–3023. https://doi.org/10.1108/09564230910978511
11. Cowburn, G., Stockley, L.: Consumer understanding and use of nutrition labelling: a systematic review. Public Health Nutr. **8**(1), 21–28 (2004)

12. Deininger, D.U., Sur, M.: Food Safety in a Globalizing World: Opportunities and Challenges for India. South Asia Sustainable Development Department, World Bank, Washington, DC (2007)
13. Dumanovsky, T., Huang, C.Y., Basset, M.T., Silver, L.D.: Consumer awareness of fast-food calorie information in New York City after implementation of a menu labelling regulation. Res. Pract. **100**(12), 2520–2525 (2010)
14. Foroudi, P., Gupta, S., Sivarajah, U., Broderick, A.: Investigating the effects of smart technology on customer dynamics and customer experience. Comput. Hum. Behav. **80**, 271–282. https://doi.org/10.1016/j.chb.2017.11.014
15. Grunert, K.G., Wills, J.M.: A review of European research on consumer response to nutrition information on food labels. J. Public Health **15**, 385–399 (2007)
16. Halabi, C.E., Lussier, R.N.: A model for predicting small firm performance: increasing the probability of entrepreneurial success in Chile. J. Small Bus. Enterp. Dev. **21**, 4–25 (2014). https://doi.org/10.1108/jsbed-10-2013-0141
17. Harnack, L.J., French, S.A., Oakes, J.M., Story, M.T., Jeffery, R.W., Rydell, S.A.: Effects of Calorie labelling and value size pricing on fast-food meal choices: results from an experimental trial. Int. J. Behav. Nutrit. Phys. Activity **63**(5), 1186–1479 (2008)
18. Higginson, C.S., Rayner, M.J., Draper, S., Kirk, T.R.: How do consumers use nutrition label information? Nutrit. Food Sci. **32**(4), 145–52 (2002)
19. Kabango, C.M., Asa, A.R.: Factors influencing e-commerce development: implications for the developing countries. Int. J. Innov. Econ. Dev. **1**(1), 64–72 (2015)
20. Kiwanuka, A.: Acceptance process: the missing link between UTAUT and diffusion of innovation theory. Am. J. Inf. Syst. **3**(2), 40–44 (2015)
21. Kozup, J.C., Creyer, E.H., Burton, S.: Making healthful food choices: the influence of health claims and nutrition information on consumers' evaluations of packaged food products and restaurant menu items. J. Mark. **67**(2), 19–34 (2003)
22. Laohapensang, O.: Factors influencing internet shopping behaviour: a survey of consumers in Thailand. J. Fash. Mark. Manag. **13**(4), 501–513 (2009)
23. Limbu, Y.B., Wolf, M., Lunsford, D.: Perceived ethics of online retailers and consumer behavioral intentions: the mediating roles of trust and attitude. J. Res. Interact. Mark. **6**(2), 133–154 (2012)
24. Ling, S., Pysarchik, D.T., Choo, H.J.: Adopters of new food products in India. Mark. Intell. Plan. **22**(4), 371–391 (2004)
25. Mesias, F.J., Escribano, M., de Ledesma, A.R., Pulido, F.: Consumer preferences for beef in the Spanish region of Extremadura: a study using conjoint analysis. J. Sci. Food Agric. **85**, 2487–2494 (2005)
26. Mohamad, S.A., Kassim, S.: Examining the relationship between UTAUT construct, technology awareness, financial cost and E-payment adoption among microfinance clients in Malaysia. In: The 1st Aceh Global Conference (AGC 2018). Atlantis Press. https://doi.org/10.2991/agc-18.2019.56
27. Omar, M., Bathgate, I., Nwankwo, S.: Internet marketing and customer satisfaction in emerging markets: the case of Chinese online shoppers. Int. Bus. J. **21**(2), 224–237 (2011)
28. Omotayo, F.O., Omotope, A.R.: Determinants of continuance intention to use online shops in Nigeria. J. Internet Bank. Commer. **23**(2), 1–48 (2018)
29. Peeroo, S., Samy, M., Jones, B.: Facebook: a blessing or a curse for grocery stores? Int. J. Retail Distrib. Manage. **45**(12), 1242–1259. https://doi.org/10.1108/IJRDM-12-2016-0234
30. Perry, M.: Market orientation in small businesses: creative or lacking? Mark. Manag. J. **24**, 96–107 (2014)
31. Persaud, A., Azhar, I.: Innovative mobile marketing via smartphones: are consumers ready? Mark. Intell. Plan. **30**(4), 418–443 (2012)
32. Rosenbaum, M.S., Ramírez, G.C., Edwards, K., Kim, J., Campbell, J.M., Bickle, M.C.: The digitization of health care retailing. J. Res. Interact. Market. **11**(4), 432–446 (2017). https://doi.org/10.1108/JRIM-07-2017-0058

33. Samoggia, A., Arvola, A., Bertazzoli, A., Gurinovic, M., Hendrixson, V., Rivaroli, S., Ruggeri, A.: Offering low-cost healthy food: an exploration of food manufacturers' and retailers' perspectives. Int. Food Agribusiness Manage. Rev. **17**(4), 27–58 (2014)
34. Sareen, M., Jain, A.: The role of social influence and consumers' effort expectancy on online shopping: an empirical study of India. Int. J. Manage. Res. Bus. Strat. **3**(1), 138–158 (2014)
35. Singh, P., Keswani, S., Singh, S., Sharma, S.: a study of adoption behavior for online shopping: an extension of Tam model. Int. J. Adv. Soc. Sci. Hum. **4**(7) (2018)
36. Smith, P.R., Chaffey, D.: E-Marketing Excellence (EMarketing Essentials). Routledge Taylor and Francis Group, New York (2008)
37. Smutkupt, P., Krairit, D., Esichaikul, V.: Mobile marketing: implications for marketing strategies. Int. J. Mobile Market. **5**(2), 126–139 (2010)
38. Stanley, M.: Morgan Stanley explains why India's e-commerce market is a hot investment opportunity (2017). Available at: www.qz.com/1089559/morgan-stanley-explains-why-indias-ecommerce-market-is-a-hot-investment-opportunity/. Accessed 17 April 2017
39. Vazquez, D., Xu, X.: Investigating linkages between online purchase behaviour variables. Int. J. Retail Distrib. Manage. **37**(5), 408–419 (2009)
40. Wei, J., Liu, L.C., Koong, K.S.: An onion ring framework for developing and assessing mobile commerce security. Int. J. Mobile Commun. **4**(2), 128–142 (2006)
41. Zhou, L., Dai, L., Zhang, D.: Online shopping acceptance model—a critical survey of consumer factors in online shopping. J. Electron. Commerce Res. **8**(1) (2007)

Robots and Artificial Intelligence: An Aid to Dependent People

Arti Chandani, Om Prakash, Prakrit Prakash, Smita Wagholikar, and Mita Mehta

Abstract In the world of digitalization and automation, there were astonishing myths coexistent that were not only surprising, but also incredible. Whilst the move is on to usher into the world of automation, where even the manufacturing units are devising the methods to automate their functions, it is only ironic to note that people have also developed several myths on using a robot. For one, robots have been associated with a device that would take away one job. For another, a robot is considered risky as if it will start ruling the very person that made it. Apparently, a research gap exists that needs to be plugged along with the understanding gaps to demystify the use of robots. COVID-19 has brought the world to such a stage where companies are now rethinking about the use of robots. This paper is written to explain the use of robots, clarify the need of robots in the society and present the growing demand of this technology. A research was also conducted to present the availability and state-of-art robot technology. It was also shown in the paper as to how some domains like health care are in dire need of robots, as the pandemic continues to wreak havoc on society.

Keywords Robots · Artificial intelligence · Dependent people · TIAGo

A. Chandani (✉) · M. Mehta
SIMS, Symbiosis International (Deemed University), Pune, MH 411028, India
e-mail: arti.chandani@sims.edu

O. Prakash
SICSR, Symbiosis International (Deemed University), Pune, MH 411028, India

P. Prakash
MIT, Manipal Academy of Higher Education, Manipal, KA 411028, India

S. Wagholikar
SIT, Symbiosis International (Deemed University), Pune, MH 411028, India

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_10

1 Introduction

Robots are machines with immense potential and application. Robots have ventured into domestic, mobility, guard, rescue, astronaut, industrial and social sectors of human life. The humanoid robots display several features that enable it to play games, dance and more interestingly perform different tasks to help people. Robots known as Androids are able to demonstrate and give training like human beings. Social robots are very useful and are even able to cater to some needs of small children. Industrial robots are widely used in industry to carry out work. In the medical field, robots are known to perform operations. As the robots are able to cater to a variety of applications, we often consider robots and artificial intelligence systems as intruders who can steal jobs or make decisions against humans. However, the truth is that these devices are already helping to improve people's lives. For example, these have been shown to be effective in support tasks for people with disabilities. These robots are already playing a vital role in the services to the elderly people who have developed some disabilities or the rehabilitation of patients with injuries. Robots are also being deployed for patients with the chronic pathologies. Robots are the most "human" side of artificial intelligence. These abilities of robots make them novel machines which can take instructions and implement them in order to get the desired work done. The motivation behind using robots is that they contribute in making human life easy and comfortable. The different type of robots are in prevelence to take care of varied human tasks (Table 1).

Table 1 Types of robots in prevalence

S. No.	Type of robots
1	Humanoid
2	Animal (four legged)
3	Androids
4	Guard
5	Social
6	Mobility
7	Domestic
8	Rescue
9	Astronaut
10	Industrial
11	Medical field

1.1 Dependent People

Dependent people are those who have developed a health conditions or exist with a condition that makes them in the need of others help. For example, old people or people with special needs or people who have developed a health condition, which makes them dependent on help and support of caretakers. One special example on today's context is people who have contracted the pandemic disease of COVID-19. The diseases are such infectious that if someone contacts this disease, the person is immediately quarantined in a hospital. In such case, the subject is completely dependent on all day-to-day need on the caretakers. In effect, the caretaker is able to provide only limited support as the caretakers are also at grave risk of contracting such disease. This means that the scenario is already set to take the technology in the healthcare domain to the next level, where a device like a robot may give such risky care. Today, as the statistics reveal, almost 20% of the worldwide population may be living in a stage where they might heavily depend on others. Such dependency has to be taken up by a robot, and once again the society needs a dire and immediate help from technology.

1.2 Need of Robots and AI

Robots are the devices, which automate some human action with the help of some intelligence built-in for their operations. This intelligence simulates human intelligence, at least to some level. Such intelligence, which has been artificially built by developers, is termed as artificial intelligence. Robots, powered by such artificial intelligence, can help people who are severely dependent on others. Imagine an old person, who is left with no one to help and lives in a condition that makes him almost impossible to get up from the bed to fulfill some daily needs. Arguably, there is no solution better than placing a robot to help such a person who cannot even get up to fill water for himself. In other cases, many people who have contracted COVID-19 disease are quarantined, and caretakers are very cautious in dealing with them. Pandemic has almost certainly proved the immediate need of a large number of robots to be acquired by any healthcare center.

2 Goal of Robotics to Help Growing Population of Elderly People

Statistics indicated that the elderly population is growing gradually. Health care has improved a lot, as there is cure for several diseases that were incurable a few decades ago. There is a cure for cancer and other deadly diseases, which is responsible for increasing the average life expectancy. Life expectancy is increasing as a result of

healthcare improvement, which indicates that the elderly population will grow a lot in the coming decades. However, it will be increasingly difficult to reach all the people who need some human means or intervention. Companies like PAL Robotics are now increasingly aware of this horizon. "There are many studies that say that there will not be enough resources in long-term care to provide assistance to everyone in a dignified and quality way. That is why robotics is entering this field," says Viladomat [5]. The robotics companies also do not believe that robots will conflict with caregivers, as there are a lot of things that a human being contributes that a robot cannot supply. Their goal is not to make robot friends that have intelligence and simulate the emotions or feelings, just like a human being. Their object is to make tools with the ability to interact with the environment and help people who have limiting conditions using advanced concepts of artificial intelligence (AI) and machine learning (ML). With this development, robots are going to fill jobs that may involve physical risk for nurses or caregivers [4]. As we are in the age of COVID-19, one can imagine the potential of robots. They can save countless lives of caregivers, as for sure the COVID-19 virus cannot affect a robot [5].

3 Ways in Which Robot and AI Are Helping Dependent People

3.1 TIAGo Robot Helping People with Condition

In the age of digitalization and automation, as the industries are preparing to enter into *industry 5.0—automation*, there are devices like TIAGo robot, to help those who have a medical condition or certain disability. TIAGo is now available to lend a hand in whatever it takes. The device has one arm and no feet but a round base that turns and rolls on the ground. It has a face, somewhat similar to what one saw in the sci-fi movies Wall. E's and ET's, where a smile is always drawn. TIAGo robot is an acronym that means "*Take It And Go!"* and is able to help with household chores to an elderly person, or a severe disability, who lives alone at home. And not only that—through its optical sensor, this robot can know how that person is and communicates with them to ask what they need at all times [8]. A nurse can monitor the person from a distance to provide simple care for that person and thus avoid having to travel to the health center. Sometimes, it is only necessary for someone to be aware that no accident happens to the subject when that person goes to a place like a bathroom. Some patients would not mind even if a device accompanies them to certain places, whereas they would have discomfort if a person like a nurse accompanies then to the washroom, especially if they have severe disabilities. This is how the TIAGo robot can be programmed to help and is even capable of placing a blanket well on the bed at night if the subject happens to be uncovered [7].

TIAGo is very much there in the market, and it is great news for many. The models are manufactured by PAL Robotics, a leading company in robotics applied

to assistance to people with various medical conditions. Robots can support subjects suffering from sight, hearing impairment or disability in ambulation as well as other medical conditions like Alzheimer's disease. The latter psychological conditions develop in many old people, who appear to be normal but actually are severely limited. They need someone to keep reminding them of simple things as they get increasingly forgetful.

3.2 Robotics Development

Several robotics companies now predict that in a matter of five years they will see "TIAGos" in private homes. In a longer period of time, humanoids will also enter the scene. Humanoids are also termed sometimes as bipedal robots, with anatomy and stature similar to those of the human beings. With the help of AI/ML-based programming, these devices will be able to adapt one hundred percent to the external environment where they would work. These would be designed according to human anatomy and biology. Currently, the development of humanoids is being slower than that of other models due to navigation and displacement problems.

3.3 Handmade

The main burden for integrating robots into people's daily lives is their high price. Many robotics models are handcrafted in-house, and very few units are made commercially. This brings the price of the most basic TIAGo model to 30,000 Euros and the most complete ones to 60,000 Euros. In the case of humanoids, the cost may be between 150,000 and 900,000 Euros. Currently, there is no demand for this kind of technology, but once people try it and see that it is useful, they will be marketed to end users and not just to university R&D departments, etc.

3.4 Rehabilitation Robots

At the National Hospital for Paraplegics in Toledo, there are two robots that carry out their work every day normally. These are the Locomat, which adhere to the legs of the spinal cord injured to help them in their rehabilitation march. "They have been involved in the rehabilitation of people for 10 years now. It is one more treatment, as is physiotherapy, occupational therapy, etc." explains Antonio del Ama, Head of the Biomechanics and Technical Aids Unit at the hospital [2]. In addition, in the center, there are three units of another type of ambulatory robot for gait compensation [2].

Fig. 1 Locomat for rehabilitation. *Source* Marchal-Crespo et al. [6]

3.5 Exoskeletons Help in the Rehabilitation of Spinal Cord Injuries

Before the appearance of the Locomat (as shown in Fig. 1), the intensive repetition of the gait movement, which the patient must do as a fundamental therapy for his or her rehabilitation, was an immense effort for physiotherapists. The patients were placed on a treadmill, supported from above with a harness [3]. Two physiotherapists moved his legs manually while another held his body. After 10–15 min, the physiotherapists were physically exhausted [4]. Now thanks to these robots, the movement can be repeated for hours. There is false fear that their jobs will be taken away, but it is a paradigm shift [10].

3.6 AI as Neurological Therapy

JosepLluís Arcos, from the Artificial Intelligence Research Institute, is in charge of three of the most advanced machine learning projects applied to cognitive stimulation and rehabilitation. The Innobrain and Cognitio projects benefit people with degenerative diseases, while Amate helps people with autism spectrum disorders (ASD) to show their emotions [1]. Arcos assures that a large number of AI and machine learning projects are emerging in the area of neurology that, although they are in an incipient phase, are already beginning to show great potential. "There are projects that, by exploiting GPS technology, allow people with neurodegenerative diseases to be given more autonomy by monitoring their movements and helping in the event of disorientation or loss," [9].

3.7 Robots for COVID-19 and Other Pandemics

COVID-19 has made the collaboration between humans and robots in a natural form, and this is very true where one of the good walls to contagion is robots for social contact as shown in Fig. 2.

"The pandemic opens the eyes of organizations to realize that robots are required to avoid one-on-one contact," said Aldo Luévano, CEO of Roomie, the Mexican engineering firm that developed the robot. He also mentioned that large technology firms who are engaged in the development and implementation of robotics in Mexico could develop a robot which can help in the diagnosis of COVID-19 within a short span of two weeks [11].

The company worked with two infectologists and relied on technology from Intel, which provides AI and processing solutions to analyze large amounts of data, and from Amazon Web Services, which provides the cloud, machine learning and the analysis and identification of people and places through Amazon recognition [11].

4 Discussion

Need of robots has been increasingly felt, not only by companies, but also society. As the robots have gradually entered the market, companies, homes and even commercial places, the myths related to robots are being shattered. The myths that the robots are going to take jobs, or that robots would be security threats, are now being rejected as a myth. Many science fiction movies are responsible for establishing the myths, and some of them are so weird that they claim that ultimately robots may rule society. The state of art of development of robots is still in the nascent stage. There are only specific areas where robots are proving helpful like providing a helping hand to people who are dependent on others for their daily chores.

COVID-19 has sealed the issue in favor of robots. Many caretakers have contracted the deadly disease while treating COVID patients. Caretakers like nurses have to give all-round service for extended hours. Some nurses contracted the disease and risked

Fig. 2 Robot for COVID-19. *Source* The Hindu [12]

their lives especially as a robot can easily carry such activities. For example, a robot can be present at the hospital and provide services like taking temperature, providing medicine doses, providing helping hand to patients in their miscellaneous need during their quarantine states and most importantly relieve nurses with some of their utmost busiest schedules. Arguably, robots are the best way to serve people who are living with some physical or mental condition that makes them constantly dependent on others. This is especially true if a person also has no alternative but to live alone and in that case, it is also being regarded as best for the safety and security of the person living alone. Robot is programmed to do just the specific activity and will not have enough "brains" to plot some mischief. Moreover, the company who supplies robots can also remotely monitor the robot functioning and would work ethically and in a responsible manner.

4.1 A Pandemic to Develop Robots

The Roomiebot COVID-19 is equipped with technologies to detect possible cases of Coronavirus and that is why it was developed in conjunction with two infectologista. Luévano explained that this tool comes equipped with a thermometer and an oximeter that allows the detection of dyspnea or shortness of breath, a common condition among patients with the disease. The idea, he said, is not to replace medical personnel, but to streamline the process of caring for potential patients. With this tool, he explained that what is sought is to avoid the collapse of medical institutions due to lack of personnel by letting the first line of contact be a robot that, in case of detecting symptoms in a patient, can channel it effectively. The robot is capable of identifying symptoms such as cough, headache, arthralgia, myalgia, odynophagia, rhinorrhea, conjunctivitis and chest pain and can also obtain demographic or clinical data to identify patients at increased risk for pre-existing diseases.

5 Conclusion

The healthcare experts have not stressed that at this time some medical firms do have the capacity to deploy hundreds of robots that can be used in hospitals, companies and airports, and it is expected that in about a few weeks they will have a presence in at least some hospitals and other facilities. The experts also explained that in the case of robots like TIAGo robot and Roomiebot COVID-19, the firms are willing to operate in the joint venture model, in which the medical institutions cover some expenses for the robot's raw material to allow its use and impact to expand rapidly. The investment cost for a robot is high. However, it must be remembered that the service received by mankind is of great importance and is worthwhile as human life is invaluable. Robots are into use not to replace humans but to provide service in this

type of pandemic and help people. Robots are just undertaking the riskier tasks to reduce the risks of doctors or nurses catching the disease.

5.1 Future Directions

With the advent of technology in machines and the advent of artificial intelligence, the future seems to be promising for the comfort and convenience of mankind. Robots are known for carrying out tasks effectively. The use of technology and robots will gain more popularity owing to its advantages and applications in varied sectors. Use of robots in service to patients, dependents and people with autism is a boon to mankind.

References

1. Arcos, J.L., De Mantaras, R.L., Serra, X.: Saxex: A case-based reasoning system for generating expressive musical performances. J. New Music Res. **27**(3), 194–210 (1998)
2. del-Ama, A.J., Gil-Agudo, Á., Pons, J.L., Moreno, J.C.: Hybrid gait training with an overground robot for people with incomplete spinal cord injury: a pilot study. Front. Hum. Neurosci. **8**, 298 (2014)
3. Frizera, A., Elias, A., Del-Ama, A. J., Ceres, R., Bastos, T.F.: Characterization of spatio-temporal parameters of human gait assisted by a robotic walker. In: 2012 4th IEEE RAS & EMBS International Conference on Biomedical Robotics and Biomechatronics (BioRob), pp. 1087–1091. IEEE (2012)
4. Grama, L., Rusu, C.: Extending assisted audio capabilities of TIAGo service robot. In: 2019 International Conference on Speech Technology and Human-Computer Dialogue (SpeD), pp. 1–8. IEEE (2019)
5. Lafrenz, R.: Archive 2017
6. Marchal-Crespo, L., Tsangaridis, P., Obwegeser, D., Maggioni, S., Riener, R.: Haptic error modulation outperforms visual error amplification when learning a modified gait pattern. Front. Neurosci. **13**, 61 (2019)
7. Mocanu, I., Axinte, D., Cramariuc, O., Cramariuc, B.: Human activity recognition with convolution neural network using TIAGo robot. In: 2018 41st International Conference on Telecommunications and Signal Processing (TSP), pp. 1–4. IEEE (2018)
8. Pages, J., Marchionni, L., Ferro, F.: Tiago: the modular robot that adapts to different research needs. In: International workshop on robot modularity, IROS (2016)
9. Plaza, E., Arcos, J. L., Martin, F.: Cooperative case-based reasoning. In: Distributed Artificial Intelligence Meets Machine Learning Learning in Multi-agent Environments, pp. 180–201. Springer, Berlin, Heidelberg (1996)
10. Torricelli, D., Del Ama, A.J., Gonzalez, J., Moreno, J., Gil, A., Pons, J.L.: Benchmarking lower limb wearable robots: emerging approaches and technologies. In: Proceedings of the 8th ACM International Conference on PErvasive Technologies Related to Assistive Environments, pp. 1–4

11. The Economic Times News: Mexican robot takes care of COVID patients, 31 Aug 2020. Retrieved from https://economictimes.indiatimes.com/news/international/world-news/mexican-robot-takes-care-of-covid-patients/stressbuster/slideshow/77847975.cms
12. The Hindu: Coimbatore start-up builds robot for COVID-19. Coimbatore start-up builds robot for COVID-19. 29 April 2020. Retrieved from https://www.thehindu.com/news/cities/Coimbatore/coimbatore-start-up-builds-robot-for-covid-19/article31466385.ece

Predicting Performance of Students Considering Individual Feedback at Online Learning Using Logistic Regression Model

Vikrant Shaga, Haftom Gebregziabher, and Prashant Chintal

Abstract Online learning apps or Web sites can provide education-beneficial interventions. There are different variables affecting students' performance. Nowadays, researchers working on a smart education system using which students can perceive the concepts for lifelong learning and not just for grades or for clearing the examinations. It is difficult to measure the student's understanding or performance based on the feedback they were given after completion of every course. This paper is predicting the student's performance concerning feedback of 137 postgraduate students using the logistic regression model of machine learning. The purpose of using logistic regression is the response variable from the dataset which is restricted to 0 or 1, and the probability range would be between 0 and 1. Determining the accuracy rate of the model is the key result in this research, and our case, logistic regression model has achieved the prediction accuracy of 72%. Teachers using online educational utilities for enhancing teaching–learning would be motivated by this research.

Keywords Online learning · Smart education · Machine learning · Logistic regression model

1 Introduction

Massive open online courses (MOOCs) and learning management systems (LMS) are free open-source web-based learning platforms. In most universities, online

V. Shaga (✉)
Department of Information Communication Technology, Federal Technical and Vocational Education and Training Institute, Addis Ababa, Ethiopia

H. Gebregziabher
Federal Technical and Vocational Education and Training Institute, Addis Ababa, Ethiopia
e-mail: haftom.gebregziabher@ethernet.edu.et

P. Chintal
Marathwada Institute of Technology, Dr. B.A.M University, Aurangabad, Maharashtra, India

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_11

learning through these open-source free platforms has a good impact on the learners as well on the teachers. Using these platforms, teachers can predict the performance of the students based on various in-classes and out-class activities [1]. The activities, feedback, and assessment together play a crucial role in preparing instructional resources. The use of technology in education facilitates learning, improving performance, and managing technological processes and resources. The appropriate or disciplined use of pedagogical approaches, instructions, tools, and technologies enhances the learning of individuals as well as their performances. Also, learning activities between learners can be appreciated with specific tools and resources to achieve specific outcomes. The teacher's role is to engage learners in learning activities to achieve outcomes [2]. According to Zhu et al. [3], 'Smart education' in education initiates intelligent, personalized, and adaptative education using the most advanced technologies, which has been carried out worldwide for personalizing the teaching and learning processes. The student-centric model is designed by teachers by incorporating the contents and activities for promoting online education in teaching–learning. This model was specifically designed for enhancing the practical approach of learners. The gender-wise understanding and opinion from postgraduate students were highlighted in the research [4]. In this research, various learning activities have been undertaken by learners (postgraduate students) as a part of e-learning. For pedagogic decisions, few in-classes and out-class segments have been designed. The in-class segment consists of quizzes, assignments, and discussion forums, whereas the out-class segment consists of problem statements with expected output and videos. All the documents and resources are licensed under creative commons. A feedback survey involves a predefined set of questions of Moodle (Gnomio). To collect the information about the teaching–learning mechanism, course contents, and instructor's interaction, survey is divided into six sections. Questions in these sections are categorized based on relevance, reflective thinking, interactivity, tutor support, peer support, and interpretation [5]. The use of online learning platforms and tools enhances student learning. But it is difficult to predict performances and progress of the students especially after their feedback about the courses they have learned through an online e-learning platform. It seems to be very contradictory. As per our dataset, average time taken for submitting feedback by failed students is 2.9 min, and passed students is 4.64 min (Fig. 1).

2 Statistical Overview

Students from the reputed higher education institutions are involved in the study. All students are pursuing their post-graduation and belong to the technology field. In the present paper, a structured questionnaire has been used as the entire course designed in Gnomio or Moodle (e-learning community) with a predefined set of questions available with feedback survey facilities. To gather the information, we used Gnomio's predefined question set. Figure 2 shows the feedback survey report of one of our online courses with attributes.

Fig. 1 Average time taken by students for feedback submission

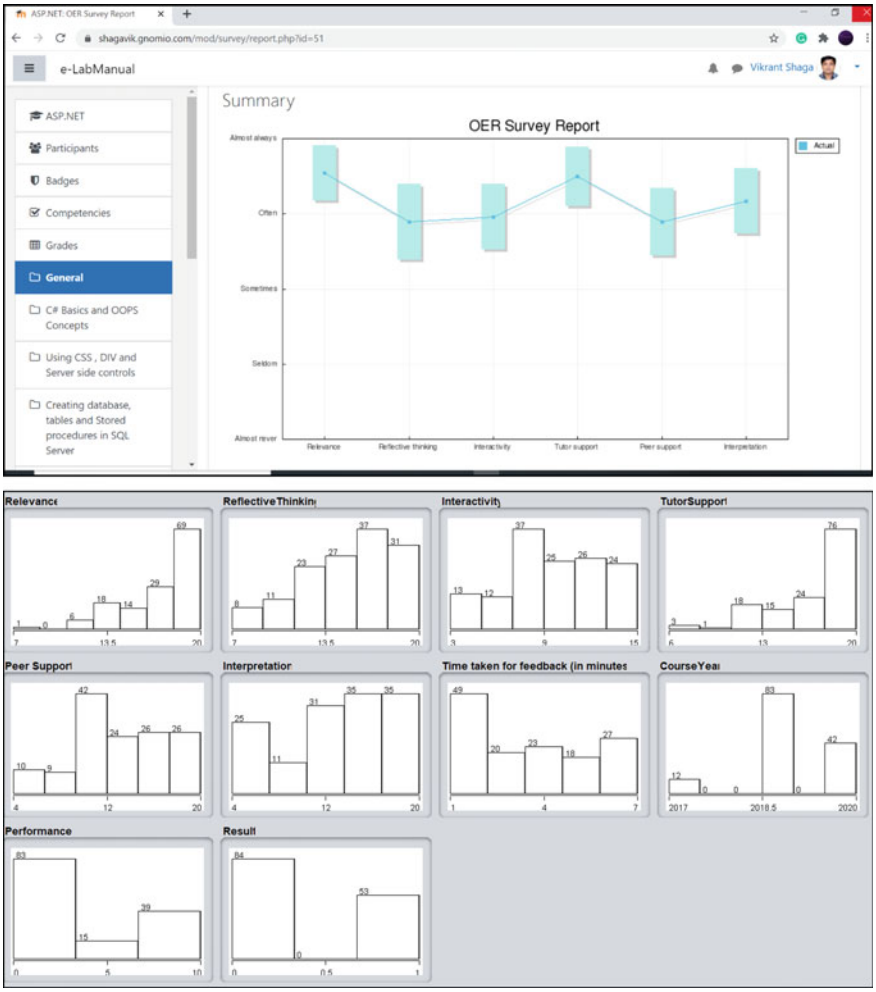


Fig. 2 Student feedback score for each attribute

Table 1 Important attributes for performance measurement

S. No.	Attributes	Mean	Standard deviation
1	Relevance	17.606	2.702
2	Reflective thinking	14.985	3.308
3	Interactivity	10.204	3.066
4	Tutor support	17.073	3.205
5	Peer support	13.511	3.976
6	Interpretation	13.197	4.783
7	Time taken for feedback (in minutes)	3.577	1.793
8	Performance	3.618	4.314

The result of the student is either pass or fail based on our predefined criteria. If student score more than 50% means pass and failed if less than 50% marks in the final online quiz (Table 1).

3 Machine Learning

Learning is not just gaining knowledge out of different courses. It comprises of knowledge acquisition, knowledge management, and knowledge augmentation. Machine learning simply means making machine intelligence by computer algorithms, statistical methods, and some data. It means machine learns based on statistical methods, algorithms, and more importantly existing data. Programming computers to optimize a performance criterion using example data or experience [6]. A field of study that gives computers the ability to learn without being explicitly programmed [7]. Machine learning is also a technical tool of data science that creates logic from data by converting existing data into knowledge. Many prevailing algorithms from the area of machine learning are developed for pattern recognition, forecasting from the previous activities, and obtaining insights. Appropriate classification of characteristics read by the machine instead of the brain means machine learning algorithms are more efficient and appropriate than human beings for classification. But it has also been observed that no model is perfect, and results may be accurate but still prone to errors. The crucial part is to develop a model that provides an accurate result [8] (Fig. 3).

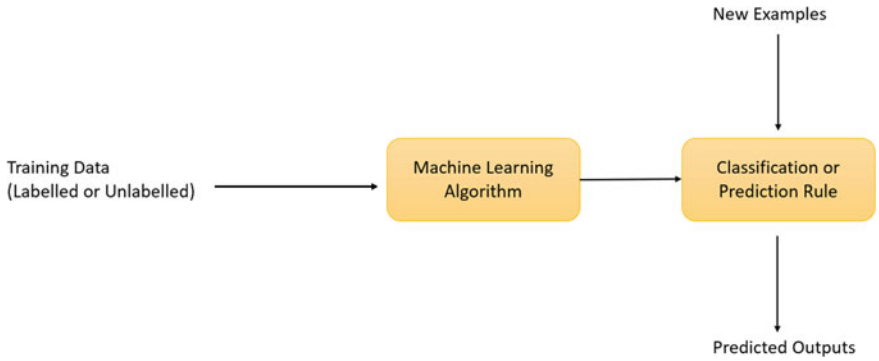


Fig. 3 Machine learning approach [8]

4 Logistic Regression Model

The statistical analysis helped in solving various problems in many sectors and domains. The logistic regression is the baseline supervised machine learning algorithm for classification and estimates the effects of independent variables on the result variables as a probability. The logistic regression ensuring the determination of the risk factors as the probability is a method that investigates the relationship of the result variables with independent variables in binary or multiple phases. The logistic regression analysis is a method that helps in the categorization and appointment process. Contrasting regression analysis, there is no precondition such as a normal distribution variance or continuous variance. The effects of expressive variables on the dependent variables are estimated as probability thereby enabling the risk factors to be determined as a probability. The logistic regression is a statistical method that allows for categorization as appropriate to the rules of probability by estimating the values for the dependent variable. It analyzes tabulated or pure datasets [9, 10].

Using the logistic regression model of machine learning, categorical variables can be predicted. This algorithm generates as output the probability of a categorical dependent variable. This variable produces the data coded in the domain [0, 1], where 1 usually means positive (yes, true, success, etc.) and 0 means negative (no, false, failure, etc.), though the result is not mathematically restricted to this interpretation. Therefore, the logistic regression model generates a probability $P(Y = 1)$ as a function of the input(s) X. This type of regression analysis is also known as 'binary logistic regression analysis.' It means if there is only one independent variable, we can use binary logistic model; but if there are multiple independent variables, then multiple binary logistic regression model can be used as below [11]:

$$P = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots \beta_n X_n)}} \tag{1}$$

P Observing the probability of an analyzed event.

β_0	Another statement for the value of the dependent variable when independent variables get a value of zero.
β_0, β_1	Regression coefficients of independent variables.
$X_1, X_2 \ldots X_n$	Independent variables.
n	The number of independent variables.
e	Error term.

5 Dataset

The dataset consists of records of 137 postgraduate students from two different countries—India and Ethiopia. Students from the reputed higher education institutes are involved in the study. All students are postgraduates and belong to the computer application and technology field. The examination results and feedback comprised two postgraduate programs, i.e., MCA and MSc IT within departments of the institute for the student intakes from the year 2017–2020. Out of 137 students, 60 (44%) are female and 77 (56%) are male. The online courses for which students have given the feedback are ASP.NET, advance Java, Web technologies and E-learning design and development, and software engineering. The feedback questions were a predefined set of questions from Gnomio (e-learning platform) which includes six sections. The first section is in relation with *relevance* with variables from 1 to 4, second section is in relation with *reflective thinking* with variables from 5 to 8, the third section is in relation with *interactivity* with variables from 9 to 12, the fourth section is in relation with *tutor support* with variables from 13 to 16, the fifth section is in relation with *peer support* with variables from 17 to 20, and sixth section is in relation with *interpretation* with variables from 21 to 24.

The feedback given by the students for the online courses (in numbers of minutes) is appropriate or not based on their performances in the exam; this can be identified using a logistic regression model.

6 Experimental Interventions

We have used python and its class libraries to implement a logistic regression model. We are classifying success in exams based on course 'feedback time (in minutes)' of 137 postgraduate students which are our independent variable. We have the 'Result' column is our response or dependent variable which tells whether the result is pass or fail. Here, we have assumed X as 'feedback time (in minutes)' which will be our input vector. The machine learning system for logistic regression implementation steps is as follows [11–13]:

(a) *Algorithm*: Implementation of a logistic regression model with improved performance requires the 'sigmoid' equation. It has a very important advantage of taking real-valued numbers and maps it to the range [0, 1] which is required

for the probability. 'Sigmoid' is also known as 'linear regression equation' which is as follows:

$$y = g(z)\frac{1}{1+e^{-z}} \tag{2}$$

where $z = \theta^T x$ (matrix multiplication of the parameters θ and input vector X).

(b) *Loss function*: It is the summation of all training data samples which is used to determine the best θ parameters of the logistic regression model. The objective of this function is to minimize the error. It is also called as log loss.

$$J(\theta) = -\frac{1}{m}\left[\sum_{i=1}^{m} y^{(i)} \log h_\theta\left(x^{(i)}\right) + \left(1 - y^{(i)}\right) \log(1 - h_\theta\left(x^{(i)}\right)\right] \tag{3}$$

(c) *Gradient Descent*: It is an iterative optimization algorithm to minimize the above-given loss function. For the loss function J, the gradient at the point j is defined as:

$$\frac{\partial J(\theta)}{\partial \theta_j} = \sum_{i=1}^{m} (h_\theta\left(x^{(i)}\right) - y^i) x_j^{(i)} \tag{4}$$

Repeat until convergence {

$$\theta = \theta - \alpha \frac{\partial J(\theta)}{\partial \theta_j}$$

}

where

θ training parameter.

α learning rate for controlling the size during the optimization process (assumed to be 0.1 in our case).

(d) *Prediction*: As our main focus is on the prediction of student's performance based on their feedback and our logistic regression is also trained, we are ready for predictions on the data by assuming the threshold value 0.51. Probabilities are higher than the threshold value considered to be a success/positive and less than the threshold value considered to be failure/negative.

7 Results

In this section, we calculate the accuracy of the logistic regression model to predict whether feedbacks given by students contradict the performance. We implemented

this machine learning predictive logistic regression model using python programming. We used loss function to determine the appropriate value of θ and then used gradient descent to minimize the error between actual and predictive values (Figs. 4 and 5).

('loss:', 0.6918714971844803)
('loss:', 0.6111908592507376)
('loss:', 0.5780858804459077)
('loss:', 0.5631700538428338)
('loss:', 0.555905421387823)
('loss:', 0.5521612337974974)
('loss:', 0.5501518960173133)
('loss:', 0.5490417287183086)
('loss:', 0.5484151896594208)
('loss:', 0.5480559999154101)

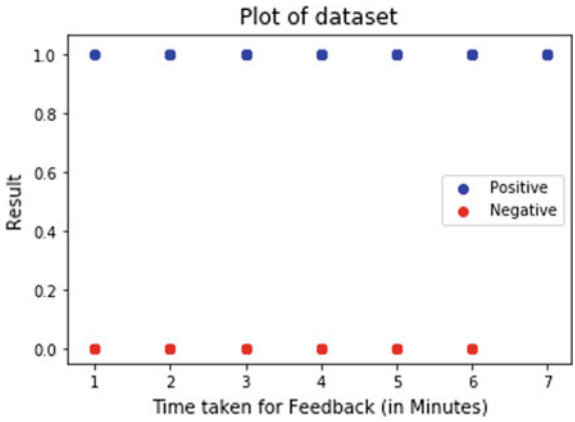


Fig. 4 Dataset points for feedback time and result

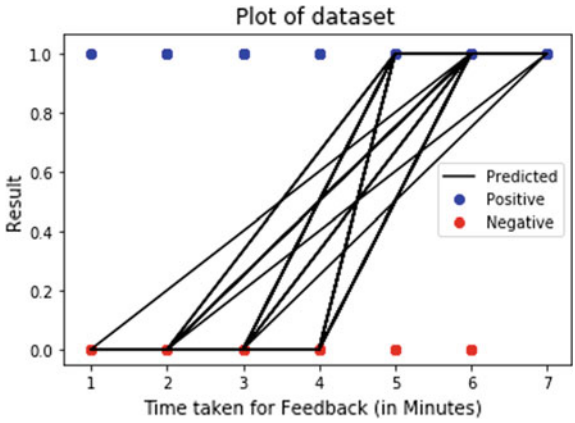


Fig. 5 Dataset points and prediction curve for feedback time and result

time taken = 0.30 s
('(preds = = y).mean() = ', 0.7226277372262774)
('theta = ', [−2.68440542, 0.59767034])

The time taken by this model is 0.30 s, and it is observed that 0.69 is the initial loss. The loss is then reduced to approximately 0.54 after 1000 iterations. The accuracy of the model is found out to be 72% means out of 137 students, 98 results are correctly predicted, and 39 are wrongly predicted. This result of the logistic regression model is achieved using a learning rate value $\alpha = 0.1$ and 1000 iterations. The final parameter θ is found out to be [−2.68440542, 0.59767034].

8 Conclusion

In this research, the logistic regression model is used to predict whether student responses in terms of feedback and exam results for the courses are contradictory or not. The results of the experimental program clearly show that the accuracy of the model is very significant to motivate teachers to use this type of e-learning mechanism. It also signifies that teachers should use these types of online e-learning platforms for improving the teaching–learning of students. More than 50% positive predictive values for students' feedback and their results show that students who gave positive feedback about the different parameters of the courses, their performance are also good. It is expected in teaching–learning process that the above experiment should be applied for different set of students too, so that the above experiment can be rectified applying more types of algorithms to get better results. As overall teaching learning mechanism with student feedback has been applied for various postgraduate students, it can be applied for under graduates and diploma students also. Then, after applying different machine learning algorithms on new datasets can reveal useful predictions for our research.

References

1. Hussain, M., Zhu, W., Zhang, W., et al.: Using machine learning to predict student difficulties from learning session data. Artif. Intell. Rev. **52**, 381–407 (2019). https://doi.org/10.1007/s10462-018-9620-8
2. Huang, R., Michael Spector, J., Yang, J.:. Educational Technology: A Primer for the 21st Century. Springer (2019). ISBN 9811366438, 9789811366437
3. Zhu, Z.T., Yu, M.H., Riezebos, P.: A research framework of smart education. Smart Learn. Environ. **3**, 1–17 (2016). https://doi.org/10.1186/s40561-016-0026-2
4. Shaga, V., Samee, S., Vengatesan, K., Abhishek, K.: Fact findings of exploring ICT model in teaching learning. Int. J. Sci. Technol. Res. **8**(12) (2019). ISSN 2277-8616
5. Shaga, V., Samee, S., Mohammed, A., Vengatesan, K.: Enhancing Empirical approach in teaching-learning using ICT. Int. J. Pure Appl. Math. **118**(20), 2727–2734 (2018). ISSN: 1314-3395 (on-line version)

6. Alpaydin, E.: Introduction to Machine Learning, 3rd edn. Cambridge, MIT Press (2014)
7. Samuel, A.L.: Some studies in machine learning using the game of checkers. IBM J. Res. Dev. **3**, 210–290 (1959)
8. Liakos, K.G., Busato, P., Moshou, D., Pearson, S., Bochtis, D.: Machine learning in agriculture: a review (2018). Received: 27 June 2018; Accepted: 7 Aug 2018; Published: 14 Aug 2018
9. Korkmaz, M., et al.: The importance of logistic regression implementations in The Turkish Livestock Sector and logistic regression implementations/fields. J. Agric. Fac. HR.U **16**(2), 25–36 (2012)
10. Verma, C., Stoffova, V., Illés, Z., Dahiya, S.: Binary logistic regression classifying the gender of student towards Computer Learning in European schools (2018)
11. Russano, E., Avelino, E.F.: Fundamentals of Machine Learning using Python (e-book Edition 2020). Arcler Press, Oakville, ON L6H 5R7, Canada. ISBN: 978-1-77407-427-5
12. Jurafsky, D., Martin, J.H.: Speech and Language Processing. Copyright @2019. All rights reserved. Draft of 2 Oct 2019
13. Chintal, P., Gaikwad, R.J., Deshmukh, R.R.: Cyber crime analysis of Maharashtra state using gradient descent approach with linear regression. Int. J. Pure Appl. Math. **119**(16), 3537–3542 (2018). ISSN: 1314-3395

Survey of NLP Resources in Low-Resource Languages Nepali, Sindhi and Konkani

Annie Rajan and Ambuja Salgaonkar

Abstract Research results in different subdomains of NLP are enumerated and compared for Nepali, Sindhi, and Konkani languages. These languages are spoken by relatively small populations, they have low NLP resources, and the numbers of their native speakers are falling. Hence, the list of results points to areas where more research is required. As far as we are aware, this is the first comprehensive survey that brings together NLP resources in Nepali, Sindhi, and Konkani.

Keywords Nepali language · Sindhi language · Konkani language · Natural language processing

1 Introduction

Information retrieval, speech recognition, and machine-assisted multilingual conversations are a few among the natural language processing (NLP) specializations that call for interdisciplinary research in linguistics and artificial intelligence. Digitization of linguistic resources, developing NLP tools, and applications like translation are areas in which such research has taken place. The NLP tools in any language are in the domains of text-to-speech (TTS) conversion, parts-of-speech tagging (PoS), sentiment analysis, morphological analysis, named entity recognition (NER), word segmentation, text segmentation, word sense disambiguation (WSD), WordNet, text data mining, spell checker, translation, and transliteration.

The development of digitized resources for NLP in small population Indian regional languages is comparable with the research in low-resource European languages [1]. NLP tools for Indian languages are not keeping up with the ever-increasing demand for them, especially for languages like Nepali, Sindhi, and

A. Rajan (✉) · A. Salgaonkar
Department of Computer Science, University of Mumbai, Mumbai, India

A. Rajan
Department of Computer Science, DCT's Dhempe College of Arts and Science, Panaji, Goa, India

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_12

Konkani which lack standardization of script and usage due to various regional, historical, and cultural factors.

Nepali, Sindhi, and Konkani are Indo-European languages. Nepali is the official language of Nepal, and it has 2.8 million native speakers in India [2]. Sindhi, one of the official languages of Pakistan, has 2.5 million speakers in India [3]. Konkani, the official language of the Indian state of Goa, is spoken by a population of 2.3 million spread across neighboring states of Goa, Maharashtra, Karnataka, and Kerala. Konkani has reached Portugal, Kenya, Uganda, Pakistan, and the Persian Gulf as a language of people of Indian origin [4–6]. According to the 2011 census of India, the populations of native speakers of Nepali, Sindhi, and Konkani languages each amount to about 0.25% of the total population of India [7].

The present paper aims at a comprehensive survey of NLP research done so far in Nepali, Sindhi, and Konkani and attempts to identify the potential for further research. The common factors among these languages are that they have low NLP resources and approximately the same number of native speakers. Since the native speakers form a small fraction of the population of India, their languages are prone to getting a low priority in terms of digitization and NLP. This is the relevance of the content of the present paper. Section 2 below is a brief description of NLP resources available in Nepali, Sect. 3 is on resources in Sindhi, and Sect. 4 on resources in Konkani. Section 5 has a comparison of NLP resources in the three languages, and Sect. 6 has the conclusions and suggestions for work in the future.

2 NLP Research in Nepali

NLP research results for Nepali have appeared in the literature since about 2004 [22], and the bulk of the results is from the last 5–10 years [8–32]. A summary of this work is given below in Sect. 5.

PoS tagging of Nepali text of 5373 words using the Viterbi algorithm gave an accuracy of 97% [8]. For a data set of 10,775 words, the accuracy obtained using support vector machines (SVM) is 93% [9]. PoS tagging using the hidden Markov model (HMM) for 15,005 words resulted in an accuracy of 76%, and with the Viterbi algorithm, it was 95% [10]. HMM with a rule-based method obtained an accuracy of 93% on a data set of 15,720 words [11]. The general regression neural network (GRNN) has proved to be exceptionally good, with 100% accuracy of PoS tagging on a data set of 48,100 words [12]. Another experiment with HMM on a data set of 150,839 words gave an accuracy of 96% [13]. Based on these results, one may conclude that research in PoS tagging in Nepali has reached maturity.

Research on sentiment analysis of Nepali text has employed a Naïve Bayes classifier. For a data set of 384 documents, 500 movie reviews, and 20,000 sentences, the accuracy obtained is 73%, 78%, and 73%, respectively [14–16]. The relatively low accuracy of about 75% is generally due to the fact that some words are used with positive polarity at certain times and with negative polarity at other times. It is difficult to disambiguate such words. NER for Nepali has performed very well

irrespective of the size of the data employed in the training phase. For a data set of 344 words, an accuracy of 93% is obtained with HMM [17]. Using support vector machines (SVM), the accuracy for a data set of 29,298 words is 92% [18]. Work in word sense disambiguation (WSD) in Nepali appears to have been initiated only recently. Using a semantic graph-based approach, it has demonstrated a 15% increase in accuracy for disambiguating nouns and 39% in the case of adjectives [19].

Researchers have given a thought to text data mining in Nepali. Clustering and particle swarm optimization have been employed for 50 documents [20]. For a document size of 45 key words and 45 multi words, accuracy of, respectively, 90% and 91% has been obtained using the classical vector space approach [21]. Design and implementation issues and linguistic aspects of a morphological analyzer and stemmer for Nepali language are discussed in [22]. A morphological analyzer tool for Nepali nouns is described in [23]. A Nepali morph analyzer built using Apertium's LT toolbox has covered 20,000 words with 219 paradigms [24]. Another morph analyzer has shown 70% accuracy [25]. A proposed Nepali text-to-speech (TSS) synthesis system based on concatenation employs the epoch synchronous non-overlap add (ESNOLA) method [26].

The details of building a Nepali national corpus of 14 million PoS words are described in [27]. An overview of major primary resources and NLP applications is explained in [28]. A Nepali spell checker based on a lexicon of 500 words is described in [29]. Version 1.0 of a Nepali spellchecker was released in the year 2005 with 300,000 words, while release 1.1 in 2007 has 930,000 words [30]. A Nepali WordNet has been developed under the technology development of Indian languages (TDIL) programme [62]. It is a machine-readable lexical database based on expansion from the Hindi and English WordNets [31]. A model has been proposed for Nepali to English translation that produces the equivalent feature structures by using morphological analysis, a bilingual lexicon, and syntactic, semantic and pragmatic procedures [32]. This seems to be the beginning of machine translation in Nepali. The references given in this section are organized by category in Sect. 5.

3 NLP Research in Sindhi

Sindhi is spoken in the Sindh province of Pakistan, the Sindhi community dispersed over much of India, and the diaspora around the world. Though Sindhi is not an official language of any Indian state, it nevertheless is a language officially recognized by the government. One potential obstacle for Sindhi NLP is the multiplicity of scripts: Perso-Arabic in Pakistan and mostly Devanagari in India. Even so, Sindhi NLP has made rapid strides. The Sindhi community will be able to use NLP tools in any script of its choice.

A supervised approach has been used for the development of a PoS tagger for Sindhi, with a training data set of 26,366 words and a test corpus of 6783 words [33]. The accuracy of the PoS tagger without WordNet is 96%, and with WordNet it is 97%. An SVM-based PoS tagger is proposed for Sindhi with a data set of 28,000

words, and the accuracy obtained is 97% [34]. Using SVM and K-NN techniques, a supervised machine learning model performs sentiment analysis on a corpus of 9779 records. The f-score with SVM for positive, negative, mix, and neutral sentiments is 79%, 61%, 75%, and 35%, respectively, and with K-NN, it is 77%, 62%, 55%, and 28%, respectively [35]. The conclusion is that overall SVN outperforms K-NN. A rule-based approach is used for developing an NER system for Sindhi with a data set of 936 words, and the accuracy obtained is 98% [36]. A model has been proposed for word segmentation of Sindhi text. Passages were taken from the Internet and from publications like general articles, news articles, book chapters, and digital dictionaries. Out of a total of 157,509 words that were collected, 146,132 words were marked as valid Sindhi word tokens. The cumulative accuracy is 92% [37]. The word segmentation problem was addressed using a lexicon-driven approach for Sindhi text on a data set of 16,601 words, and the segmentation error rate observed was 9.54% [38].

Based on the inflection rules for nouns, an equivalent computational rule is defined for a finite state transducer [39]. Apertium's LT toolbox has been used to develop a morphological analyzer with a data set of 384 words, for which the f-score is 83% [40]. To resolve the problem of segmenting complex Sindhi words and possible morphemes for those words, an algorithm has been designed and implemented. The resulting system was tested on a collection of 1681 words, and the cumulative segmentation error rate calculated to be 5.02% [41].

A tool has been developed for stemming and lemmatization of Sindhi words, on a corpus of 17,061 words [42]. Sindhi morphology with addition, subtraction, and replacement methods as well as inflection of nouns with number, gender, and case is discussed in detail [43]. A brief introduction is given to Sindhi morphology to understand the internal structure of words, along with descriptions of word formation with reference to word types, prefixes and suffixes, and their outcome on the words [44]. This paper concludes that the morphological construction of Sindhi is either derivational or inflectional. There is an analysis and review of morphemes, including the function, structure, and nature of compound words, prefix words, suffix words, and prefix-suffix words. Moreover, a comparative analysis is carried out to explicate the formation of Sindhi morphology.

The WordNet approach is used for the development of a Sindhi TSS system on test data of 1500 sentences. Its accuracy without WordNet is 83%, and with WordNet, it is 90% [45]. Letter frequency distribution was calculated in a Sindhi corpus containing almost 14 million characters. The frequency distribution in a file with over 50,000 words was found to be identical to the distribution in the entire corpus. Further, a total of 4.1 million words were analyzed, and 70,576 distinct word forms were found [46]. A hybrid architecture is given, consisting of an algorithm that computes edit distances to find a similarity metric between two strings, the phonetics-based SoundEx algorithm to index words by sounds, and the ShapeEx algorithm to develop pattern matching for a Sindhi spellchecker [47]. Another Spell checker is developed for Sindhi, also using the SoundEx and ShapeEx algorithms. A combinational approach is used to generate a misspelled similar word list [48].

Various spelling errors and their types are studied for Sindhi. It is found that substitution errors occur due to two reasons: similarity of the letters in the alphabet and similar pronunciation of different letters. The other type of error is the omission or deletion of the space character at word boundaries [49]. A Sindhi spell checker with error detection and correction techniques was given a collection of 1744 Sindhi words. It recognized all 482 mistakes and suggested 2449 words in place of these incorrect words [50]. A Sindhi text tokenization model has been put forward using simple, compound, and complex words [51]. An algorithm to transliterate between Roman and Sindhi scripts has been proposed. Transliteration is done by analyzing the letters that come before and after letters without diacritical marks. Complexities and problems of Sindhi transliteration are discussed in detail [52].

4 NLP Resources in Konkani

The small size of Goa state belies its importance on the national stage. This is the relevance of NLP tools in Konkani, Goa's official language.

Konkani numerals from 1 to 100 were voiced by eSpeak, a rule-based speech synthesis tool [53]. Further, a set of five Konkani sentences was produced by the same software, by concatenating the recorded voices of phonemes, diaphones, commonly used words, and numbers. The natural feel and clarity of the voicing were rated by 20 subjects as having an average of 1.8 on a five-point scale [54]. Analysis of these findings has been provided [55]. A letter-concatenation-based TTS model was trained with utterances of a corpus of 3000 characters, vowels, diacritics, and other distinctive symbols. A voicing of a Konkani text written in Devanagari and Roman has been demonstrated [56]. An optical character reader (OCR) for Roman and Devanagari has been used for pre-processing sentences [57]. Phonetization accuracy of 64% was obtained for text-to-speech conversion of Devanagari-scripted Konkani after the model was trained with the festival synthesizer and Festvox sound constructor framework [58]. The natural feel and intelligibility of the sounds were improved to 3.5 and 2.5, respectively.

For purposes of PoS tagging, annotation standards for Konkani have been published by the Bureau of Indian Standards BIS [59]. There are eleven parts of speech: in the tag set: noun, pronoun, demonstrative, verb, adjective, adverb, post-position, conjunction, quantifiers, particles, and residuals. A PoS tagger for Konkani has been described that employs HMM as its basis and recommends the Viterbi algorithm to find a suitable tag for a given word [60]. An automatic PoS tagger has been trained with a data set of 5088 code-mixed Konkani-English sentences and tested with 4,000 Twitter and Facebook posts. Classifiers based on support vector machines (SVM) and conditional random fields (CRF) demonstrated accuracy of 92% and 91%, respectively, for the Twitter posts and 91% and 90% for the Facebook posts. Thus, SVM slightly outperformed CRF [61].

An HMM-based PoS tagger has been designed on a corpus of 50,000 Konkani sentences, using the NLTK natural language tool kit written in Python. Its accuracy

is 73% [63]. A PoS tagger has been built by employing 25,000 sentences from the healthcare domain and 20,000 sentences from the tourism domain. 5000 sentences from health and 1000 sentences of tourism domains were tested. It demonstrated accuracy of 89% and 87%, respectively [64].

A finite state machine for paradigm selection has been built with a corpus of 268,000 words, and this system yields an f-score of 95% [65]. The AutoParSe parsing system automatically assigns paradigms to nouns using morphophonemic rules, and for the nouns that are not resolved in this manner, a rule-based method is used [66]. Linguistic data for Konkani has been uploaded on the GitHub repository of Apertium, a machine translation platform [67]. KOP, a model of an opinion mining system that works on Konkani text, has been proposed [68]. A model trained by employing SVM for 2910 manually tagged synsets of Konkani words has 48% accuracy [69]. A Naïve Bayes classifier was tested with 50 Konkani poems [70]. Its accuracy is 82% for known data and 70% for unseen data.

The Indian languages corpora initiative (ILCI) has created corpora in the domains of health, tourism, agriculture, and entertainment. In the first stage, collections of 25,000 sentences each were developed for health and tourism, and in the second phase, 30,982 sentences were gathered for agriculture and entertainment. Also in the second stage, a monolingual general domain corpus was created with 30,595 sentences. Other resources created under the ILCI project are Konkani Sakal Bharati OTF fonts, English–Konkani translated strings with over 7,500 words, and automated paradigm selection for an FSA-based verb morphological analyzer.

In the second stage of the project, a WordNet named Indradhanush created a Konkani synset with 32,372 words [62]. The linguistic data consortium for Indian languages (LDC-IL) has a Konkani text corpus of 3,995,616 words, a speech corpus of over 136 h from 441 speakers of Konkani language, and there is a speech recognition database gathered from 450 different native speakers, each of whom recorded a news text [71]. For the analysis of code-mixed data, a methodology has been described for the creation of a normalized data set for Konkani-English code-mixed social media text [72]. A paper describes the creation of a WordNet using expansion. In this method, the Konkani WordNet is created from Hindi as the source language [73].

A word-level identification system for Konkani-English has been developed, with 5088 sentences of Facebook posts adding up to 60,118 words in Konkani and English. Its accuracy is 90% for random field (RF), 89% for SVM, and 97% for CRF [74, 75]. Konkanverter, a transliteration engine based on a finite state transducer, performs machine transliteration between Devanagari, Kannada, and Roman scripts. The accuracy is the highest at 97.87% for Romi-Kannada script [76]. A recent survey of NLP resources in Konkani has enumerated the references cited above [81].

5 Comparison between Nepali, Sindhi, and Konkani

In Table 1, a list of references of publications relevant to NLP in Nepali, Sindhi, and Konkani is given, classified according to their core topics.

Table 1 List of references of Nepali, Sindhi, and Konkani according to core topics

Core Topics of NLP	Nepali	Sindhi	Konkani
PoS	[8–13]	[33, 34]	[59–61, 63, 64]
Morphological analyzer	[22–25]	[39–44]	[65–67]
Speech synthesis	[26]	[45]	[53–58]
Sentiment analysis	[14–16]	[35]	[68–70]
Resource generation	[27–31]	[46]	[62, 71, 72]
NER	[17, 18]	[36]	–
Word segmentation	[77, 78]	[37, 38]	–
Spell checker	[30]	[47–50]	–
Translation	[32]	–	–
Transliteration	–	[52]	[76]
Text segmentation	–	[51]	–
WordNet	[31, 79]	–	[73]
WSD	[19]	–	–
Text data mining	[20, 21]	–	–

6 Conclusion and Future Work

We compared NLP research in three languages for two reasons: one, they are spoken by about the same number of people, and two, they share their origin to the extent that they are derived from the medieval language Paishachi [80]. Development of these low-resource languages needs urgent attention. From Table 1, we can conclude that there has been some progress in NLP work for Nepali. For Sindhi, there has been some work in translation, WordNet, word sense disambiguation, and text data mining, while for Konkani there is little or no research work in NER, word segmentation, spell checkers, translation, test segmentation, text mining, and word sense disambiguation. It is a fact that the population of people speaking these languages is reducing, and moreover, many among the younger generation are migrating to countries where they are losing touch with their native languages. To prevent these languages from being wiped out or eliminated, more efforts need to be undertaken in areas where research has been initiated.

References

1. Quality Translation: https://www.qt21.eu/?target=Introduction. Last accessed 19 Aug 2020
2. Languages of Nepal: https://en.wikipedia.org/wiki/Languages_of_Nepal. Last accessed 19 Aug 2020
3. Sindhi Language: https://en.wikipedia.org/wiki/Sindhi_language. Last accessed 19 Aug 2020
4. Almeida, S.M.: A Description of Konkani. Thomas Stephens Konkani Centre, Miramar, Goa (1989)
5. Konkani language: https://en.wikipedia.org/wiki/Konkani_language. Last accessed 19 Aug 2020
6. Borkar, S.: Lets Learn Konkani. Rajhauns Publisher, Panjim, Goa (2006)
7. List of Languages by Number of Native Speakers in India. https://en.wikipedia.org/wiki/List_of_languages_by_number_of_native_speakers_in_India. Last accessed 19 Aug 2020
8. Yajnik, A.: General regression neural network based PoS tagging for Nepali text. In: *4th International Conference on Natural Language Computing*, pp. 35–40 (2018). https://doi.org/10.5121/csit.2018.80603
9. Shahi, T.B., Dhamala, T.N., Balami, B.: Support vector machines based part of speech tagging for Nepali text. Int. J. Comput. Appl., 38–42 (2013). https://doi.org/10.5120/12217-8374
10. Yajnik, A.: Part of speech tagging using statistical approach for Nepali text. World Acad. Sci. Eng. Technol. Int. J. Cogn. Lang. Sci., 76–79 (2017). https://www.researchgate.net/publication/325440900
11. Sinha, P., Veyie, N.M., Purkayastha, B.S.: Enhancing the performance of part of speech tagging of Nepali Language through hybrid approach. Int. J. Emerg. Technol. Adv. Eng., 354–359 (2015). https://www.semanticscholar.org/paper/Enhancing-the-Performance-of-Part-of-Speech-tagging-Sinha-Veyie/415e970bee7291b7fe5679a3c7f845f446f899e7
12. Yajnik, A.: ANN based PoS tagging For Nepali text (2018). https://doi.org/10.5121/ijnlc.2018.7302
13. Paul, A., Purkayastha, B.S., Sarkar, S.: Hidden Markov model based part of speech tagging for Nepali language. In: International Symposium on Advanced Computing and Communication, Silchar, India, pp. 149–156 (2015). https://ieeexplore.ieee.org/document/7377332
14. Thapa, L.B.R., Bal, B.K., Classifying sentiments in Nepali subjective texts. In: 7th International Conference on Information, Intelligence, Systems & Applications (IISA), Chalkidiki, Greece, pp. 1–6 (2016). https://ieeexplore.ieee.org/document/7785374
15. Pant, A.K., Yadav, A.: Sentiment analysis on Nepali movie reviews using machine learning, researcher CAB. J. Res. Dev. (2004). https://www.academia.edu/12075554
16. Gupta, C.P., Bal, B.K.: Detecting sentiment in Nepali texts: a bootstrap approach for sentiment analysis of texts in the Nepali language. In: International Conference on Cognitive Computing and Information Processing, pp. 1–4 (2015). https://ieeexplore.ieee.org/document/7100739
17. Dey, A., Paul, A., Purkayastha, B.S.: Named entity recognition for Nepali language: a semi hybrid approach. Int. J. Eng. Innov. Technol., 21–25 (2014). https://www.ijeit.com/Vol%203/Issue%208/IJEIT1412201402_04.pdf
18. Bam, S.B., Shahi, T.B., Named entity recognition for Nepali text using support vector machines. Intell. Inf. Manage., 21–29 (2014). https://doi.org/10.4236/iim.2014.62004
19. Roy, A., Sarkar, S., Purkayastha, B.S. Knowledge based approaches to Nepali word sense disambiguation. Int. J. Nat. Lang. Comput., 51–63 (2014). https://doi.org/10.5121/ijnlc.2014.3305
20. Sarkar, S., Roy, A., Purkayastha, B.S.: A Comparative analysis of particle swarm optimization and K-means algorithm for text clustering using Nepali Wordnet. Int. J. Nat. Lang. Comput. (2014). https://doi.org/10.5121/ijnlc.2014.3308
21. Sitaula, C.: Semantic text clustering using enhanced vector space model using Nepali language. Comput. Sci. Telecommun., 41–46 (2012). https://www.scirp.org/(S(i43dyn45teexjx455qlt3d2q))/reference/ReferencesPapers.aspx?ReferenceID=877603

22. Bal, B.K., Shrestha, P.: A morphological analyzer and a stemmer for Nepali. PAN Localization, Working Papers, 2007, pp. 324–31 (2004). https://ww.panl10n.net/english/final%20reports/pdf%20files/Nepal/NEP06.pdf
23. Chhetri, I., Dey, G., Das, S.K., Borah, S.: Development of a Morph Analyser for Nepali Noun Token. In: International Conference on Advances in Computer Engineering and Applications, pp. 984–987 (2015). https://doi.org/10.1109/ICACEA.2015.7164849
24. Bhat, S.M., Rai, R.: Linguistic data consortium for Indian Languages Central Institute of Indian Languages, building morphological analyzer for Nepali. J. Mod. Lang., 45–58 (2012). https://www.academia.edu/5484848
25. Sitaula, C.: A hybrid algorithm for stemming of Nepali text. Intell. Inf. Manage., 136–139 (2013). https://doi.org/10.4236/iim.2013.54014
26. Chettri, B., Shah, K.B.: Nepali text to speech synthesis system using ESNOLA method of concatenation. Int. J. Comput. Appl., 24–28 (2013). https://doi.org/10.5120/10053-4909
27. Yadava, Y. P., Hardie, A., Lohani, R.R., Hall, P., Allwood, J., McEnery, T., Gurung, A., Gurung, S., Regmi, B.N.: Construction and Annotation of a Corpus of Contemporary Nepali, pp. 213–225. https://doi.org/10.3366/E1749503208000166
28. Bal, B.K.: Towards building advanced natural language applications—an overview of the existing primary resources and applications in Nepali. In: Proceedings of the 7th Workshop on Asian Language Resources, Singapore, Association for Computational Linguistics Stroudsburg, PA, USA, pp. 165–170 (2019). https://www.aclweb.org/anthology/W09-3424/
29. Bal, B.K., Rupakheti, P., Khatiwada, L., Pandey, B.: Nepali spell checker, Research Report on the Nepali Spell Checker (2008). https://www.panl10n.net/english/Outputs%20Phase%202/CCs/Nepal/MPP/Papers/2007/0702/mpp_reports_pdf/report_nepali_spell_checker.pdf
30. Bal, B.K., Karki, B., Khatiwada, L.P., Shretha, P.: Nepali Spellchecker 1.1 and the Thesaurus, Research and Development (2007). https://www.semanticscholar.org/paper/Nepali-Spellchecker-1-.-1-and-the-Thesaurus-%2C-and-Bal-Karki/a48160f9e1035b0a5804ccf344bd44887be77d90?p2df
31. Chakrabarty, A., Purkayastha, B.S., Roy, A.: Experiences in building the Nepali WordNet insights and challenges. In: 5th WordNet Conference, IIIT Mumbai (2010). https://www.cfilt.iitb.ac.in/gwc2010/pdfs/35_napali_wordnet__Chakrabarty.pdf
32. Shrestha, H.K.: Rule based machine translation system in the context of Nepali text to English text. Department of Computer Science, University of Oklahoma (2005)
33. Mahar, J.A., Memon, G.Q.: Sindhi part of speech tagging system using WordNet. Int. J. Comput. Theory Eng. (2010). https://doi.org/10.7763/IJCTE.2010.V2.198
34. Surahio, F.A., Mahar, J.A.: Prediction system for Sindhi parts of speech tags by using support vector machine. In: International Conference on Computing, Mathematics and Engineering Technologies, pp. 1–6 (2018). https://doi.org/10.1109/ICOMET.2018.8346331
35. Ali, M., Wagan, A.I.: Sentiment summerization and analysis of Sindhi text. Int. J. Adv. Comput. Sci. Appl., 296–300 (2017). https://doi.org/10.14569/IJACSA.2017.081038
36. Jumani, A.K., Memon, M.A., Khoso, F.H., Sanjrani, A.A., Soomro, S.: Named entity recognition system for Sindhi language. In: International Conference for Emerging Technologies in Computing, pp. 237–246. Springer (2018). https://doi.org/10.1007/978-3-319-95450-9_20
37. Bhatti, Z., Ismaili, I.A., Soomro, W.J., Hakro, D.N.: Word segmentation model for Sindhi text. Am. J. Comput. Res. Repository, 1–7 (2014). https://doi.org/10.12691/ajcrr-2-1-1
38. Mahar, J.A., Memon, G.Q., Danwar, S.H.: Algorithms for Sindhi word segmentation using lexicon driven approach. Int. J. Acad. Res. (2011). https://www.researchgate.net/publication/285754096
39. Arain, M.U.R., Bhatti, M.I.: Finite state morphology and Sindhi noun inflections. In: Proceedings of the 24th Pacific Asia Conference on Language, Information and Computation, Japan, pp. 669–676 (2010). https://www.researchgate.net/publication/265883958
40. Motlani, R., Tyers, F.M., Sharma, D.M., A finite-state morphological analyser for Sindhi. In: Proceedings of the Tenth International Conference on Language Resources and Evaluation, pp. 2572–2577 (2016). https://www.aclweb.org/anthology/L16-1409/

41. Narejo, W.A., Mahar, J.A., Mahar, S.A., Surahio, F.A., Jumani, A.K.: Sindhi morphological analysis: an algorithm for sindhi word segmentation into morphemes. Int. J. Comput. Sci. Inf. Secur., 293 (2016). https://www.academia.edu/27075171
42. Dootio, M.A., Wagan, A.I.: Automatic stemming and lemmatization process for Sindhi text. J. Soc. Sci. Interdiscip. Res., NED University Engineering Technology, Karachi Sindh Pakistan, 19–28 (2017). https://www.researchgate.net/publication/328202210
43. Arain, M.U.R., Sindhi morphology and noun inflections. In: Proceeding of the Conference on Language and Technology, Lahore, Pakistan, pp. 74–81 (2009). https://www.researchgate.net/publication/242510188
44. Narejo, W.A., Mahar, J.A.: Morphology: Sindhi morphological analysis for natural language processing applications. In: International Conference on Computing, Electronic and Electrical Engineering, pp. 27–31 (2016). https://ieeexplore.ieee.org/document/7495248
45. Mahar, J.A., Memon, G.Q., Shah, S.H.A.: WordNet based Sindhi text to speech synthesis system. In: Second International Conference on Computer Research and Development, pp. 20–24 (2010). https://doi.org/10.1109/ICCRD.2010.31
46. Rahman, M.U.: Towards Sindhi corpus construction. Linguist. Lit. Rev., 39–48 (2015). https://doi.org/10.32350/llr/11/04
47. Bhatti, Z., Ali Ismaili, I., Nawaz Hakro, D., Javid Soomro, W.: Phonetic-based Sindhi spellchecker system using a hybrid model. Digit. Scholarship Hum., 264–282 (2016). https://doi.org/10.1093/llc/fqv005
48. Bhatti, Z., Waqas, A., Ismaili, I. A., Hakro, D.N., Soomro, W.J.: Phonetic based Soundex & Shapeex Algorithm for Sindhi Spell Checker System, pp. 1147–115 (2014). https://www.researchgate.net/publication/262302047
49. Bhatti, Z., Ismaili, I.A., Shaikh, A.A., Javaid, W.: Spelling Error Trends and Patterns in Sindhi (2014). https://arxiv.org/abs/1403.4759
50. Dahar, I.A., Abbas, F., Rajput, U., Hussain, A., Azhar, F.: An efficient Sindhi spelling checker for Microsoft Word. Int. J. Comput. Sci. Netw. Secur., pp. 144–150 (2018). https://www.researchgate.net/publication/325966088
51. Mahar, J.A., Shaikh, H., Memon, G.Q.: A model for Sindhi text segmentation into word tokens. Sindh Univ. Res. J., 43–48 (2012). https://www.researchgate.net/publication/273449362
52. Leghari, M., Rahman, M.U.: Towards Transliteration between Sindhi Scripts Using Roman Script. Linguist. Lit. Rev., 95–104 (2015). https://doi.org/10.32350/llr.12.03
53. Dessai, N.F., Naik, S., Salkar, S., Mohanan, S.: Text to speech for Konkani language. In: International Conference on Computing and Control Engineering (2012).https://doi.org/10.1109/ICCMC.2018.8487620
54. Dessai, N.F., Naik, G., Pawar. J.: Development of Konkani TTS system using Concatenative Synthesis. In: International Conference on Control, Instrumentation, Communication and Computational Technologies, Kumaracoil, India, p. 344 (2016). https://ieeexplore.ieee.org/document/7987971
55. Mohanan, S., Salkar, S., Naik, G., Dessai, N. F., Naik, S.: Text reader for Konkani language. CIIT Int. J. Autom. Autonom. Syst. (2012). https://www.ciitresearch.org/dl/index.php/aa/article/view/AA072012029
56. Borkar, S.: Text to speech System for Konkani (GOAN) language, M.E. (Electronics) Dissertation, Rajarambapu Institute of Technology, Islampur, Maharashtra, India (2006). https://www.w3.org/2006/10/SSML/papers/paper.pdf
57. Colaco, J., Borkar, S.: Design and implementation of Konkani text to speech generation system using OCR technique. Imper. J. Interdiscip. Res. (2016). https://www.academia.edu/29776714/
58. Dessai, N.B.F., Naik, G.A., Pawar, J.D.: Implementation of a TTS system for Devanagari Konkani Language using Festival, pp. 386–391 (2017). https://www.ijarcs.info/index.php/Ijarcs/article/view/3309
59. Sardesai, M., Pawar, J., Vaz, E., Walawalikar, S.: BIS annotation standards with reference to Konkani Language. In: Proceedings of the 3rd Workshop on South and Southeast Asian Natural Language Processing, pp. 145–152 (2012). https://www.aclweb.org/anthology/W12-5012/

60. Khorjuvenkar, D.N.P., Ainapurkar, M., Chagas, S.: Part of speech tagging for Konkani language. Int. J. Eng. Res. Comput. Sci. Eng. (2018). https://doi.org/10.1109/ICCMC.2018.8487620
61. Phadte, A., Arsekar, R.: Part-of-speech tagger for Konkani-English code-mixed social media text. In: International Conference on Applications of Natural Language to Information Systems, pp. 303–307 (2018). https://doi.org/10.1007/978-3-319-91947-8_31
62. Technology Development for Indian Languages. https://tdil.meity.gov.in/. Last accessed 19 Aug 2020
63. Rajan, A.: Design and implementation of a PoS tagger for Konkani using NLTK (2016), MPhil, Goa University (unpublished)
64. Kane, M.M.P.: Part of speech tagging for Konkani Corpus. Int. J. Eng. Res. Comput. Sci. Eng. (2017). https://ijercse.com/abstract.php?id=10298
65. Pawar, J., Desai, S., Bhattacharya, P.: Automated Paradigm Selection for FSA based Konkani Verb Morphological Analyzer, pp. 103–110 (2012). https://www.aclweb.org/anthology/C12-3013/
66. Desai, S., Desai, N., Pawar, J., Bhattacharyya, P.: AutoParSe: An Automatic Paradigm Selector for Nouns in Konkani (2014). https://www.aclweb.org/anthology/W14-5136/
67. Apertium Linguistic Data for Konkani. https://github.com/apertium/apertium-kok. Last assessed 22 Aug 2020
68. Miranda, D.T., Mascarenhas, M.: KOP: an opinion mining system in Konkani. In: IEEE International Conference on Recent trends In Electronics Information Communication Technology, Bangalore, India (2016). https://ieeexplore.ieee.org/document/7807914/
69. Karmali, R., Pawar, J., Fondekar, A.: Konkani SentiWordNet: Resource for Sentiment Analysis using Supervised Learning Approach, pp. 55–59 (2016). https://irgu.unigoa.ac.in/drs/handle/unigoa/4462
70. Rajan, A., Salgaonkar, A.: Sentiment analysis for Konkani language: Konkani poetry a case study. In: ICT Systems and Sustainability, pp. 321–329. Springer, Singapore (2020). https://doi.org/10.1007/978-981-15-0936-0_32
71. Linguistic Data Consortium for Indian Languages. https://www.ldcil.org. last accessed 19 Aug 2020
72. Phadte, A.: Resource creation for training and testing of normalization systems for Konkani-English code-mixed social media text. In: International Conference on Applications of Natural Language to Information System, Paris, France, pp. 264–271 (2018). https://doi.org/10.1007/978-3-319-91947-8_26
73. Walawalikar, S., Desai, S., Karmali, R., Naik, S., Ghanekar, D., D'Souza, C., Pawar, J.D.: Experiences in Building the Konkani WordNet using the Expansion Approach (2010). https://www.researchgate.net/publication/228726698
74. Phadte, A., Thakkar, G., Towards normalising Konkani-English code-MIXED social media text. In: Proceedings of the 14th International Conference on Natural Language Processing, pp. 85–94 (2017). https://www.aclweb.org/anthology/W17-7511/
75. Wagh, R., Phadte, A.: Word level language identification system for Konkani-English code-mixed social media text (CMST). In: Proceedings of the 10th Annual ACM India Compute Conference, pp. 103–107 (2017). https://doi.org/10.1145/3140107.3140132
76. Rajan, V.: Association for Computational Linguistics and Dublin City University, Konkanverter—a finite state transducer based statistical machine transliteration engine for Konkani language. In: Proceedings of the Fifth Workshop on South and Southeast Asian Natural Language Processing, pp.11–19 (2014). https://www.aclweb.org/anthology/W14-5502/
77. Pandey, R.C., Dawadi, B.R., Sharma, S., Basnet, A.: Dictionary based Nepali word recognition using neural network. Int. J. Sci. Eng. Res., 473–479 (2017). https://www.researchgate.net/publication/317304847
78. Neupane, A.: Development of Nepali character database for character recognition based on clustering. Int. J. Comput. Appl., 42–46 (2014). https://doi.org/10.5120/18799-0315

79. Roy, A., Sarkar, S., Purkayastha, B.S.: A proposed Nepali Synset entry and extraction tool. In: 6th International Global WordNet Conference, Matsue, Japan, pp. 312–316 (2012)
80. Wikipedia, Sanskrit (2019). https://en.wikipedia.org/wiki/Sanskrit. Last accessed 19 Aug 2020
81. Rajan, A., Salgaonkar, A., Joshi, R.: A survey of Konkani NLP resources. Comput. Sci. Rev. **38**, 100299 (2020). https://www.sciencedirect.com/science/article/abs/pii/S1574013720303993

Cyber-Physical System—An Architectural Review

L. Ramanathan and R. S. Nandhini

Abstract Cyber-physical systems (CPS) are very complex systems that are integrated with collaboration of communication, computation, and control together termed as 3C technology. Researchers and engineers are moving toward progress in the exploration of CPS. However, the intricacy of computing and physical advancements leads to following challenges in the development of CPS, such as physical abstraction, robustness, scalability, efficiency, and security. This paper first introduces the architectural analysis and design. Then, different CPS architectures are discussed from the perspective of manufacturing and industrial revolution. Finally, it presents the applications of CPS regarding the same.

Keywords Cyber-physical system (CPS) · 3C technology · CPS architecture · Industrial revolution

1 Introduction

Cyber-physical system is not a consolidated concept. Generally, the integration of the cyber world and the dynamic world can be defined as cyber-physical system. It is a physical and engineering system that supervises and controls real-time situations. Cyber-physical systems observe the physical world, process the data by the systems that influence, and alter the physical world. The traditional 3C concept is applied to the cyber-physical system architecture—communication, computation, and control [1]. The collaboration from the design optimization problems from the interaction between the physical environment and the computational elements has become recent research focus in cyber-physical system. These interactions should be achieved on a large scale. CPS is the fourth industrial revolution [2].

L. Ramanathan (✉) · R. S. Nandhini
Vellore Institute of Technology, Vellore, Tamil Nadu 632014, India
e-mail: lramanathan@vit.ac.in

R. S. Nandhini
e-mail: rameshsneka.nandhini@vit.ac.in

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_13

The relevance of CPS is predetermined as the composition of computational, networked resources with multiple-physical systems in the industry. These systems include mechanical, chemical, electrical, etc. Embedded systems, wireless network, and sensors are associated with CPS. But CPSs have certain characteristics of their own such as cyber capability in every physical component, networked at multiple and extreme scales, multiple spatial–temporal constraints, unconventional computing, and physical substrates, high degree of automation, dynamic recognizing [3]. Earlier, CPS had two-tier architecture, which consists of the physical part and the computation part. Sensing the physical environment, collecting data, and executing the decisions made by the computing part are functions of physical part. Making a decision by evaluating and data processing from the physical part are functions of computing part. These two parts are integrated with each other. Later, a three-tier architecture was proposed by La et al. [4], where a new service tier was introduced. Figure 1 shows the three-tier architecture of CPS.

The following are the ways how these tiers are associated with the architecture of CPS:

- Environment tier: constitutes physical devices and real-time environment, which encompasses target users and their associated environment.
- Control-tiers: processes the collected data from the sensors, decision making, finding services for the framework, and let the change invoke in the physical device.
- Service-tiers: a service-oriented computing environment where CC (Cloud Computing) and SOA services are provided.

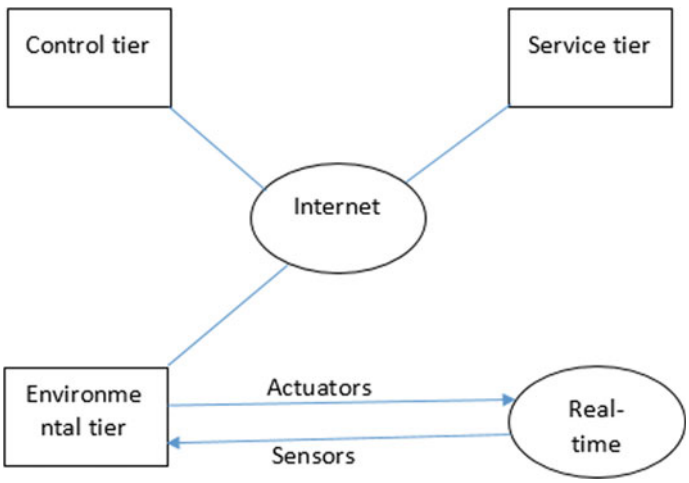


Fig. 1 Three-tier architecture of CPS

CPS can be applied in many fields. Their applications include healthcare and industrial automation, control technology, distributed energy system, aircraft control, etc. [5–7].

2 Architectural Analysis and Design of CPS

The physical world became changing is complicated. The challenges in CPS's development and design are posed from the fundamental distinction to identify between the physical and the cyber worlds. These are major considerable challenges. Safety and reliability requirements are introduced and raised from the physical components of CPS, whereas the cyber components pose security challenges. The present networking and computing technologies are questioned because the physical parts are quantitatively vary from object-oriented software components [8]. The specific CPS challenges are as follows:

- Pattern abstraction: The abstractions provided should also be able to address from the application programmers point of view, which is the complex low-level concepts and ease of the design and implementation, and also should create a logical collection of sensing devices. These abstractions are focused on wireless sensor networks rather than cyber-physical systems [9]. The potentiality of CPS can be realized only if we rebuild the computing and networking abstractions, where the physical dynamics and computations will have to clinch in a combined way.
- Scale and Efficiency: The sensors of large size are densely deployed which can screen on a broad scope of areas, which maintains the quality of monitoring and event detection control and establishes predictions in real time on large-scale integrated physical modules network systems [10].
- Robustness: The influence of robustness in engineering is prevalent, but it is less clear how to design robustness for cyber-physical systems. Robustness is a definitive concept in control system which is a fundamental requirement while designing control theory. Continuous mathematics and continuity concepts are adapted for modeling and evaluating physical systems, which help explain robustness for CPS [11].

Considering these challenges of CPS, the analysis of the following is addressed:

Rare event detection and event analysis are essential requirements for CPS. In cyber-physical systems, there will often be a data stream in a real-time environment. Time will become semantic property at such times. Kang et al. [12] present real-time data services for CPS, which supports the requirements for above challenges in CPSs. It is a new information-centric approach, where the network-enabled real-time embedded databases (nRTEDBs) communicate and control with each other. The database communicates in a secure, timely manner with wireless sensors. The proposed database has four characteristics—Real-time data processing, real-time routing, reliable event detection, security, and robustness.

Bhuiyan et al. [13] present the detection of events through differential pattern mining in CPS. A comprehensive data mining framework called Dpminer is proposed for event detection which functions in a distributed and parallel manner. Here, the actual data for pattern detection is not necessarily modified, works by sensing the actual values, and helps with outputs for event detection. A new behavioural pattern mining technique called Differential Sensor Pattern (DSP), which observes different frequencies and the set of sensors is considered rather than the traditional binary patterns. The extraction of sensor patterns that may have event information is also possible. Wang et al. [14] group the workflow as follows:

- Monitoring: Environment is an elementary function of CPS, where physical processes are monitored. Feedback on past actions is given that are considered by the CPS to make sure correct operations and decisions are made in future.
- Networking: Data aggregation and diffusion are dealt in this module. Very large amounts of data are generated from various sensors, and these data need to be aggregated or spread to analyze for further process.
- Computing: The data collected from the monitoring module is processed and analyzed further. This module also checks if the physical process clearly identifies the criteria that are pre-defined.
- Actuation: The execution of actions during the computing phase. The workflow is illustrated in Fig. 2.

This workflow also introduces the context-aware security framework [14].

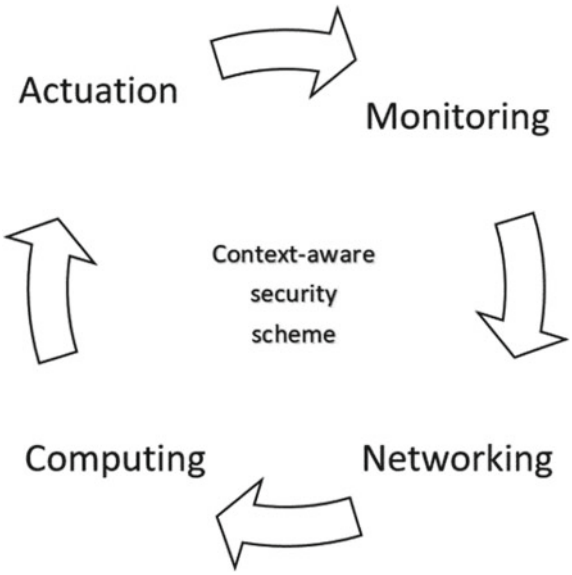


Fig. 2 Workflow of CPS

3 Architectural Survey of CPS

3.1 5 Component Architecture of CPS

The five-component architecture was proposed by Lee et al. where the cyber-physical system consists of five levels—connection level, conversion level, cyber level, cognition level, and configuration level [15].

Connection level. Acquires accurate and reliable data from machines and their parts. This can be considered as initial step in advancing toward CPS in smart factories. This level uses various devices or sensors from the manufacturing systems are used, such as ERP, MES, SCM, and CMM for data collection. A range of data, like the voltage, temperature, rotating speed, vibration, humidity, lightness, atmospheric pressure, and images and videos are collected. Selecting proper sensors must be considered [15].

Conversion level. Data to information conversion is done using tools and methodologies in this level. The mechanisms are specifically developed for advances in indication and medical management applications. This brings conscious knowledge to machines [15].

Cyber level. The processed information from the below level is received and acts as principal information in this architecture. This level runs as a sophisticated fleet-based analytical method. They perform performance analysis among the fleet, which also helps to predict future behavior.

Cognition level. The decision making using analytic information is done at this level. Online monitoring and adaptive learning are necessary for making correct decisions.

Configuration level. This level provides with response from the cyber world to the physical world. For the machine to self-configure and self-adapt, the level always administers. The level acts as the resilience control system (RCS) to monitor the system [15].

3.2 8C Architecture of CPS

Large size, densely deployed sensors can screen a broad range. An evolution in the 5C architecture is proposed in [16] called 8C architecture. This architecture is attained by including three component aspects into the five-component architecture. Coalition, customer, and content are other three components added to the architecture.

Coalition. This aspect focuses integration between different parties in terms of production process. The value chain integration and production integration are considered as factors.

Customer. This aspect focuses on the character which the customers play in the production process.

Content. This facet focuses on extracting, storing, and inquiring the content of product traceability.

3.3 ACPS Architecture

ACPS refers to anthropocentric cyber-physical systems. Highest abstraction level was provided to this model. This model integrates the physical component, the cyber component, and the human component. Service-oriented architectures (SOA), semantic web, and human–machine interaction (HMI) support the ACPS reference model. Adaptive and dynamic vision is achieved with unified integrity [17]. This model considers the human–machine interface and embeds it into the device. The interacting format converts the signals to a specific format, and the relations between elements are made via adaptors that helps in converting but optional.

3.4 SOA Architecture

SOA is service-oriented architecture. Software component technology is used to integrate loosely coupled services into a workflow in real time. A new generic CPS architecture is proposed [18]. There are five tiers in SOA architecture:

Perceive tier. The data collected from the sensors are considered for pre-processing. WSN is one main source for collecting data.

Data tier. Data processing techniques, storage devices, and computations are present in this layer. Normalization, noise reduction, and data storage are some of the data processing techniques used. This acts as the mediator between service and producer.

Service tier. Services to consumers are directly provided by this. They provide typical functions, including APIs to consumers, decision making, task analysis, and task schedule.

Execution tier. The commands from the system are received and executed in the physical environment. Its actions are reflected in the physical device or actuators.

Security assurance. Security is collective in the entire system and applies to every tier. Access security, data security, and device security are provided. The module also guarantees the system security from prohibited access or attacks. The service-oriented architecture (SOA) is the combination of services or components [18].

3.5 3C Architecture

3C architecture strengthens the CPS-enabled smart manufacturing in industries. This architecture is inspired from the ACPS architecture that consists of three components such as human component (HC), cyber component (CC), and physical component (PC). In addition to these components, the architecture adopts sub-elements of ACPS architecture. The interfacing parameters such as connectors and protocols are also adopted [19]. The 3C architecture can be reviewed as a reference for later smart manufacturing systems.

4 Applications

CPS has a vast and extensive range of application fields, which mainly include aircraft control, medical devices, intelligent manufacturing, battlefield surveillance, civil infrastructure monitoring, collaborative robotics, environmental monitoring, energy consumption and regeneration, future defense systems, energy management, smart home, and smart cities, etc., With the steady development and advancement in science and engineering, potential development in cyber-physical systems can be expected in the areas of mediations such as medical surgery using robots and nano-scale manufacturing, data mining, fatal operating environments (fire control, explorations), coordination (air and road traffic control, combats), efficiency (green buildings), etc. [19].

4.1 Agriculture

CPS in agriculture can be applied to develop food production, monitor food distribution, and consumption, precision in agriculture, smart water management. CPS can best be used to monitor the environment and its impact on the crops. Mehdipour et al. [20] propose a solution for rodents and pests that damage crops through the cyber-physical system. Caramihai and Dumitrache [21] offer a control approach that ensures agricultural proactivity versus environmental and market changes based on CPS. They contend increasing the food supply gap can be tackled by applying real-time control services.

4.2 Energy Management

CPSs are distributed systems. In spite of the devices in CPSs needs less energy, the inconvenience in energy demand and supply poses a significant [22]. A smart

grid is the next-generation electrical power grid with the efficiency to adapt and optimal power generation, consumption, and distribution [23]. Proposes a new CPS application for energy management framework (EMF) toward autonomous electric vehicle (AEV) in smart grid and charging stations in the smart grid. Li et al. [24] built a model for the class of power system with data attack, with an assumption of attack detectability which is dynamically distributed.

4.3 Environmental Monitoring

The importance of environmental monitoring plays a vital role. Mois et al. [25] present the advancement of a CPS which monitors the environmental conditions in remote locations and propose a system that presents a direct solution for ecological and environmental monitoring applications. Extreme weather conditions may damage the safety and reliability of the system. CPS is a complex system where a single error may cause abnormalities in surroundings, which results in damage in the entire system. Sierla et al. [26] propose a framework for CPS for interactions between them and its surroundings. This framework is applied to certain boundaries that consists spaces with defined physical boundaries. Sanislav et al. [27] presents an open architecture that is applied in environmental monitoring that consists of different layers. The bottom layer allows the addition of data. This comprises of wireless sensor nodes and moves to the above layers. In the top and middle layers, complex data analysis and decision rules are made. Web services and database management assure proper operation to make decisions in this layer.

4.4 Smart Manufacturing

CPS is next industrial revolution also known as Industry 4.0. The smart manufacturing results in fulfilling the dynamic needs along with the integration of human ingenuity and automation by collecting, processing, and evaluating the data and communicate with the system to act and control the environment. CPS connects the real-time environment with the automation process and takes appropriate decisions. Thoben et al. [28] analyze the applications of the CPS in smart manufacturing and also give its objective on which areas such as product design, production, logistics, maintenance, and exploitation can be applied efficiently.

5 Conclusion

CPSs are very complex and dynamic systems that pose different challenges in analysis and design, such as pattern abstraction, scale and efficiency, robustness,

storage restrictions, resource constraint, and so on. It is important to analyze today's computing and networking technologies and verify if they provide sufficient foundations for CPS. At last, some applications of CPS are discussed—Agriculture, energy management, and environmental monitoring.

6 Future Scope

Researches in CPS are limited to a specific real-time applications, current development and advancement, and theoretical systems. Most of the architectures established are for specific local applications. Therefore, a unified architecture for CPS is necessary independent of applications.

References

1. He, J.: Cyber-physical systems. Commun. CCF **6**(1), 25–29 (2010). https://doi.org/10.1109/ISORCW.2012.15
2. Jacobson, C.: Cyber-physical systems. ERCIM NEWS, broj, 97 (2014)
3. Rajkumar, R., Lee, I., Sha, L., Stankovic, J.: Cyber-physical systems: the next computing revolution. In: Design Automation Conference, pp. 731–736. IEEE (2014). https://doi.org/10.1145/1837274.1837461
4. H.J. La, S.D. Kim.: A service-based approach to designing cyber physical systems. In: 2010 IEEE/ACIS 9th International Conference on Computer and Information Science, pp. 895–900. IEEE (2010). https://doi.org/10.1109/ICIS.2010.73
5. Zhang, Y., Qiu, M., Tsai, C.W., Hassan, M.M., Alamri, A.: Health-CPS: healthcare cyber-physical system assisted by cloud and big data. IEEE Syst. J. **11**(1), 88–95 (2015). https://doi.org/10.1109/JSYST.2015.2460747
6. Mosterman, P.J., Zander, J.: Industry 4.0 as a cyber-physical system study. Softw. Syst. Model. **15**(1), 17–29. https://doi.org/10.1007/s10270-015-0493-x
7. Rajkumar, R.: A cyber–physical future. Proc. IEEE 100(Special Centennial Issue):1309–1312. https://doi.org/10.1109/JPROC.2012.2189915
8. Lee, E.A.: Cyber physical systems: design challenges. In: 2008 11th IEEE International Symposium on Object and Component-Oriented Real-Time Distributed Computing (ISORC), pp. 363–369. IEEE (2008). https://doi.org/10.1109/ISORC.2008.25
9. Sobhrajan, P., Nikam, S.Y., Pimpri, D., Pimpri, P.D.: Comparative study of abstraction in cyber physical system. Int. J. Comput. Sci. Inf. Technol **5**, 466–469 (2014)
10. Abdelzaher, T.: Research challenges in distributed cyber-physical systems. In: 2008 IEEE/IFIP International Conference on Embedded and Ubiquitous Computing, vol. 1, pp. 5–5. IEEE (2008). https://doi.org/10.1109/EUC.2008.190
11. Tabuada, P., Caliskan, S.Y., Rungger, M., Majumdar, R.: Towards robustness for cyber-physical systems. IEEE Trans. Autom. Control **59**(12), 3151–3163 (2014). https://doi.org/10.1109/TAC.2014.2351632
12. Kang, K.D., Son, S.H.: Real-time data services for cyber physical systems. In: 2008 The 28th International Conference on Distributed Computing Systems Workshops, pp. 483–488. IEEE (2008). https://doi.org/10.1109/ICDCS.Workshops.2008.21
13. Bhuiyan, M.Z.A., Wu, J., Weiss, G.M., Hayajneh, T., Wang, T., Wang, G.: Event detection through differential pattern mining in cyber-physical systems. IEEE Trans. Big Data (2017). https://doi.org/10.1109/TBDATA.2017.2731838

14. Wang, E.K., Ye, Y., Xu, X., Yiu, S.M., Hui, L.C.K., Chow, K.P.: Security issues and challenges for cyber physical system. In: 2010 IEEE/ACM Int'l Conference on Green Computing and Communications & Int'l Conference on Cyber, Physical and Social Computing (pp. 733–738). IEEE (2010). https://doi.org/10.1109/GreenCom-CPSCom.2010.36
15. Bagheri, B., Yang, S., Kao, H.A., Lee, J.: Cyber-physical systems architecture for self-aware machines in industry 4.0 environment. IFAC-PapersOnLine **48**(3), 1622–1627 (2015). https://doi.org/10.1016/j.ifacol.2015.06.318
16. Jiang, J.R.: An improved cyber-physical systems architecture for Industry 4.0 smart factories. Adv. Mech. Eng. **10**(6), 1687814018784192 (2016). https://doi.org/10.1177/1687814018784192
17. Pirvu, B.C., Zamfirescu, C.B., Gorecky, D.: Engineering insights from an anthropocentric cyber-physical system: a case study for an assembly station. Mechatronics **34**, 147–159 (2016). https://doi.org/10.1016/j.mechatronics.2015.08.010
18. Hu, L., Xie, N., Kuang, Z., Zhao, K.: Review of cyber-physical system architecture. In: 2012 IEEE 15th International Symposium on Object/Component/Service-Oriented Real-Time Distributed Computing Workshops, pp. 25–30. IEEE (2012). https://doi.org/10.1109/ISORCW.2012.15
19. Wolf, W.: Cyber-physical systems. Computer **3**, 88–89 (2009). https://doi.org/10.1109/MC.2009.81
20. Mehdipour, F., Nunna, K.C., Murakami, K.J.: A smart cyber-physical systems-based solution for pest control (work in progress). In: 2013 IEEE International Conference on Green Computing and Communications and IEEE Internet of Things and IEEE Cyber, Physical and Social Computing, pp. 1248–1253. IEEE. https://doi.org/10.1109/GreenCom-iThings-CPSCom.2013.217
21. Caramihai, S.I., Dumitrache, I.:Agricultural enterprise as a complex system: a cyber physical systems approach. In: 2015 20th International Conference on Control Systems and Computer Science, pp. 659–664. IEEE (2015). https://doi.org/10.1109/CSCS.2015.147
22. Shi, J., Wan, J., Yan, H., Suo, H.: A survey of cyber-physical systems. In: 2011 International Conference on Wireless Communications and Signal Processing (WCSP), pp. 1–6. IEEE (2011). https://doi.org/10.1109/WCSP.2011.6096958
23. Wan, J., Yan, H., Li, D., Zhou, K., Zeng, L.: Cyber-physical systems for optimal energy management scheme of autonomous electric vehicle. Comput. J. **56**(8), 947–956 (2013). https://doi.org/10.1093/comjnl/bxt043
24. Li, Y., Wu, J., Li, S.: State estimation for distributed cyber-physical power systems under data attacks. Int. J. Model. Ident. Control **26**(4), 317–323 (2016). https://doi.org/10.1504/IJMIC.2016.081137
25. Mois, G., Sanislav, T., Folea, S.C.: A cyber-physical system for environmental monitoring. IEEE Trans. Instrum. Meas. **65**(6), 1463–1471 (2016). https://doi.org/10.1109/TIM.2016.2526669
26. Sierla, S., O'Halloran, B.M., Karhela, T., Papakonstantinou, N., Tumer, I.Y.: Common cause failure analysis of cyber–physical systems situated in constructed environments. Res. Eng. Design **24**(4), 375–394 (2013). https://doi.org/10.1007/s00163-013-0156-2
27. Sanislav, T., Mois, G., Folea, S., Miclea, L., Gambardella, G., Prinetto, P.: A cloud-based Cyber-Physical System for environmental monitoring. In: 2014 3rd Mediterranean Conference on Embedded Computing (MECO), pp. 6–9. IEEE (2014). doi:https://doi.org/10.1109/MECO.2014.6862654
28. Thoben, K.D., Wiesner, S., Wuest, T.: "Industrie 4.0" and smart manufacturing-a review of research issues and application examples. Int. J. Autom. Technol. **11**(1), 4–16 (2017). https://doi.org/10.20965/ijat.2017.p0004

ICT-Based Intelligent Control of a Biogas Plant

Manmohan Singh Bhatia, Ramesh Vasappanavara, and Prashant Jadhav

Abstract With green initiatives gaining primacy, biogas plants are gaining popularity and acceptance in large towns/cities because of their potential for waste-to-fuel conversion. They are now mushrooming in many sizes to cater to a household or a large community. While the construction of biogas plants has picked up pace, smooth running of plants remains poor. Many well-designed plants fail, for lack of intelligent management. The control in these plants needs mathematical calculations as well as diagnostics and remedial actions. This paper summarizes our six-year experience on procedures for efficient and trouble-free operation of a mid-sized biogas plant (input 500 kg/day of organic kitchen waste) and generation of 20 m^3/day of biogas output. The science behind these management rules is briefly stated. Our experience shows that this control can be implemented via ICT. With this, a supervising engineer and a semiskilled labour (part time) are enough to manage such a biogas plant.

Keywords Biogas · Kitchen waste · Bacteria · Anerobic digestion · Biogas plant design · Methane · Digester · ICT controlled biogas generation

1 Introduction

Biogas plants are a boon and deserve a place in towns/cities to make our environments clean, green, and healthy [1–4]. Making a straight headway to the problem of efficient and trouble-free running of biogas plants, the focus of all human decisions and actions needs to be related to bacteria, who are the sole reason of breaking down the biomass

M. S. Bhatia (✉)
Terna College of Engineering, Sect 22, Nerul 400706, India
e-mail: manmohansinghbhatia@ternaengg.ac.in

R. Vasappanavara
Vasavi College of Engineering, Ibrahimbagh, Hyderabad 500089, India

P. Jadhav
Ramrao Adik Institute of Technology, Sect 7, Nerul 400706, India
e-mail: prashant.jadhav@rait.ac.in

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_14

to liberate methane gas through the process of anaerobic digestion [5, 6]. If living bacteria are present in the requisite good numbers at any time and they get their required nutrition, then continual bacterial presence through growth by mitosis is assured. The average life span of bacteria is about 12 h after which they die out [7]. However, if good nutrition is provided, the bacteria multiply before their death by dividing its cell into two new living cells by a process known as mitosis. So, unless steps are taken for nutrition and growth of bacteria, the generation of methane gas will not take place after bacteria die out. In a healthy situation also, some bacteria will die out, but if some others grow in numbers, the number of living and active bacteria will be maintained to carry out anaerobic digestion. It thus becomes mandatory for the design and operation of the plant to include all factors on which the presence of living and active bacteria is ensured on a continual basis [5, 6]. Thus, reiterating an obvious conclusion, even if the design of the plant is good, it will not work unless operational aspects promote bacteria growth continually—the prime task of design and operation of biogas plant. We state these factors in the next section.

2 Growth Cycle of Bacteria

Bacteria are single cell organisms that have life processes running inside the cell boundary. These include metabolic activities that synthesize nutrients and lead to growth of offspring(s) by cell division (or mitosis). The average growth time scales of any bacteria type are shown below [taken from Encyclopaedia Britannica] (Fig. 1).

The growth assumes enough supply of nutrients and optimum growth conditions or growth factors.

3 Growth Factors for Bacteria

The essential factors for bacteria growth in biogas plant are listed in Table 1.

Since biogas plants are screened off from oxygen supply, most bacteria that operate are facultative anaerobes although obligate anaerobes are also not ruled out. The process of digestion here is fermentation. Most bacteria are facultative anaerobes, so their initial supply is through their presence in the supplied biomass.

The bacteria that act in biogas plants are mesophiles which thrive in the temperature range 25–45 °C. So, a daily temperature variation and the seasonal one should be within the temperature range specified above. If not, then temperature control may be necessary [8, 9]. The figure below gives temperature zones in which bacteria types survive optimally (Fig. 2).

pH level is important for cell metabolism of bacteria. Since bacteria are single cell organisms, their surrounding medium should be pH compatible. Generally, the bacteria here are ***neutrophiles*** which require a pH range 5 to 8 on a 14-point logarithmic scale. The control of pH is linked to the nature of biomass loaded and the

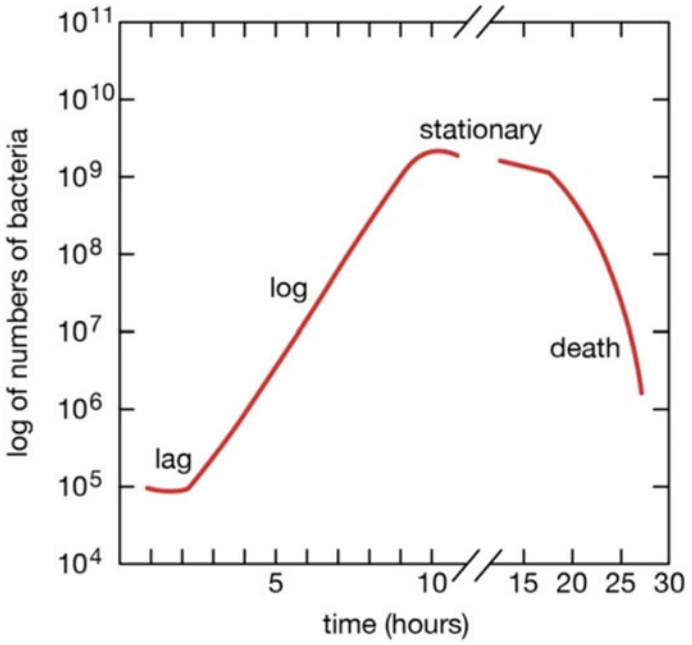


Fig. 1 Depiction of an average growth cycle of bacteria

Table 1 Growth factor and its range

Factor	Range
Oxygen availability	No external supply
Temperature	25–45 °C (Ambient)
pH level	5–8
Osmosis facilitation across cell membrane	Salt concentration < 5–7%
Energy source and nutrition	No external supply

water input. For example, if the biomass loaded is only vegetables and grains, the pH is generally in the range mentioned above with water content equal in weight to biomass. However, if the input is a mixture of animal dung, plant leaves, and vegetable and nonvegetarian leftovers (minus bones), the pH level needs to be monitored in the main digester or output discharge. Should pH level fall below 5 level, some basic salts need to be added in the right proportion to bring pH in the benign range.

The osmotic pressure on the cytoplasmic bacteria cell membrane is critical to the transport of nutrients across the pores in the membrane. The inside and outside of cell environments are critical for metabolic activities inside the cell and are said to be in either of the three states—isotonic, hypotonic, and hypertonic. Isotonic state is

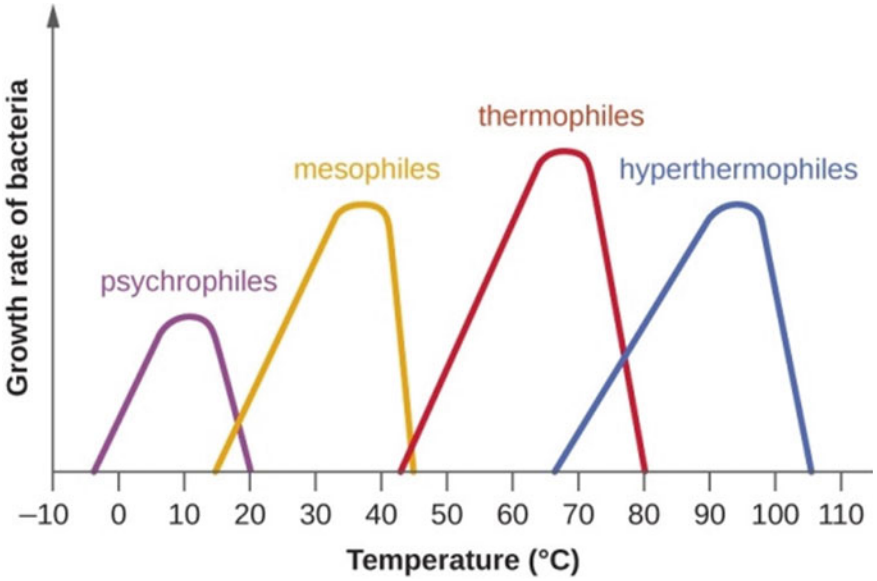


Fig. 2 Bacteria types and their optimal growth temperature ranges

realized when the solute concentration outside the cell membrane and inside is at the same level. This state is beneficial for most bacteria, and thus, adequate biomass to water ratio is critical here. One can begin with a biomass/water ratio of 1 by volume and iterate a little around this, to fix on the exact ratio that registers a maximum yield. The solute concentration by weight anywhere in the plant is kept well below 10% level on weight to weight basis. Our experience with canteen food waste with 80:20 vegetable/non-veg waste was a volume ratio of 0.9–1.0 in the plant feed. This can be different with change in composition of a different biogas mixture input.

The energy requirement and carbon intake of bacteria, to successfully grow and partake in mitosis (cell division) to yield a continuous supply of living and healthy bacteria, are met by the nature of biomass. Our experience showed that for the chosen biomass input, that is kitchen waste, there is no need of external nutrients beyond what is added. The bacteria thus could be classified as ***chemohetrotrophs*** to mean bacteria types that can derive their energy and carbon needs from the organic matter (canteen waste).

4 Design and Construction of the Plant

The plant design [10–16], and construction details with a daily intake of 500 kg of biomass iareshown below. The main parts are (Fig. 3).

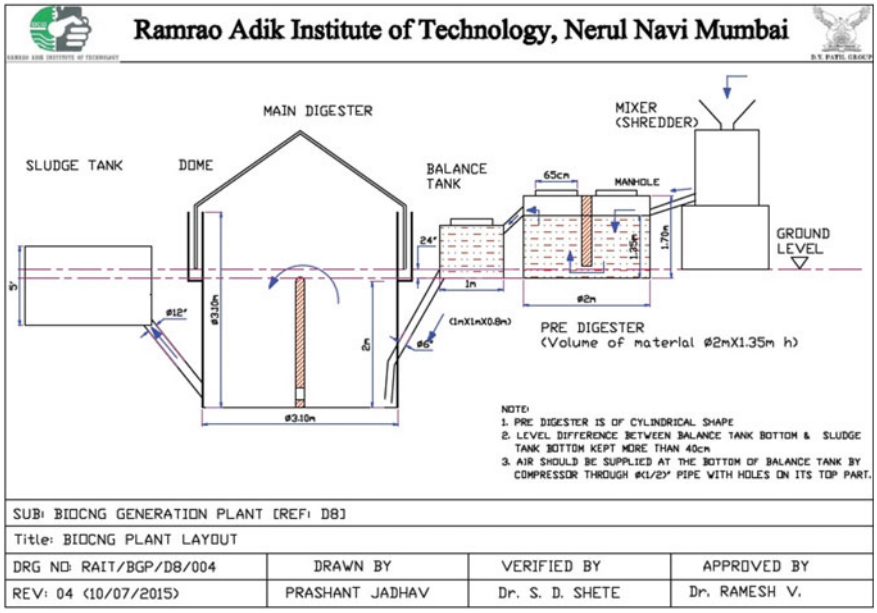


Fig. 3 Proportionate schematic of the biogas plant at RAIT, Nerul, Navi Mumbai

1. Mixer-shredder,
2. Pre-digester,
3. Balance tank,
4. Main digester with dome,
5. Sludge tank.

The dome is constructed from thick iron sheet that is painted on both the sides to prevent corrosion. The dome ends float in water to maintain a seal against atmosphere and weighs more than 90 kg. This dome weight keeps a positive pressure above atmosphere that is enough to pump the methane gas mixture towards the inflatable balloon of capacity 10 m^3 that is placed at an elevation from the plant level. The dome moves upwards against gravity when the methane gas mixture collects under it in the main digester. The connection of digester and storage balloon is via a valve in the pipeline. The balloon area is kept isolated from fire hazard due to inflammable nature of methane gas which has about 10 times more calorific value than LPG gas.

The whole design of plant is based on natural gravity flow right from pre-digester tank to the sludge tank. The flow of the biomass solution is streamline, without eddies, and proceeds from pre-digester tank to the sludge tank. The total residence time inside plant to complete one cycle of flow is about 22–24 days. It is the gravity head of pre-digester tank vis a vis the sludge tank and the pipe diameters that decide the residence time of biomass solution inside the plant. In the pre-digester tank, the solute is in shredded particulate or suspension form, but as it proceeds towards the sludge tank, the whole suspension progressively homogenizes into a fine particulate suspension due to action of bacteria inside the plant. The outflow from the biogas

plant is an excellent manure that can be used as fertilizer for gardens or general vegetation area. The figure below gives the actual picture of the main parts of the plant [17, 18] (Fig. 4).

As shown in Fig. 5, the dome is in a lifted condition with a few m^3 of methane gas mixture inside. The brown, thick marker line, comes down to the lower boundary when no gas is accumulated inside.

The most convenient location of the diagnostic inlet points was the pipe regions that connect the three tanks of the plant. These are designed so as to cause minimal invasion to the flow of solution inside the plant. The removal of small amounts of biomass solution at any stage is included in the design. The testing on , samples can include pH, fluidity, particulate-sizing and bacteria culture. This analysis can help move the plant towards higher productivity.

The output gas mixture can be analysed from time to time to see purity level of methane gas (CH_4 gas). The diagnostic analyzer result on one run of our plant is shown below. These gas analyzers are available commercially and can be suited to the needs of the plant. Depending on the plant factors, the composition of the output mixture is determined. However, we can always add a purification section to get to higher CH_4 concentrations through a multistage column [19, 20].

With better control of parameters of feed and physical conditions inside the plant, the CH_4 content can be increased to above 75% level.

Once the feed is standardized to a good extent [21–23], the diagnostics and experimentation can be just replaced by simple run procedures by the control engineer who sets the stage for optimal operation of the plant. This can be done remotely too with direct supervision of feed and water and simple on/off procedures left in the hands of an unskilled assistant. It is here that AI and ICT procedures become relevant [24–27].

Fig. 4 Shredder above pre-digester

Fig. 5 Dome above main digester

5 ICT-Based Control of Biogas Plant

Having understood the steps that play a role in efficient running of a biogas plant, the next step is to make this knowledge and control methodology widespread so that maximum efficiency is realized. It is here that ICT-based control of biogas plants in an area of a town/city can be adopted. Figure 6 shows the use of ICT as part of the solution, and it is in this sense that we plan to adopt ICT (Figs. 7 and 8).

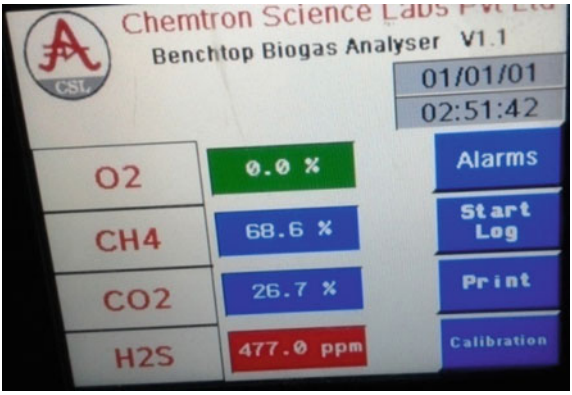


Fig. 6 Analyzer readings of biogas tested online

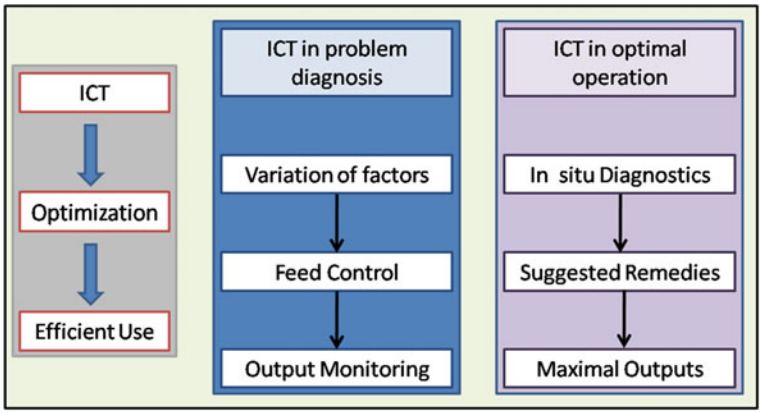


Fig. 7 Use of ICT in biogas plant control

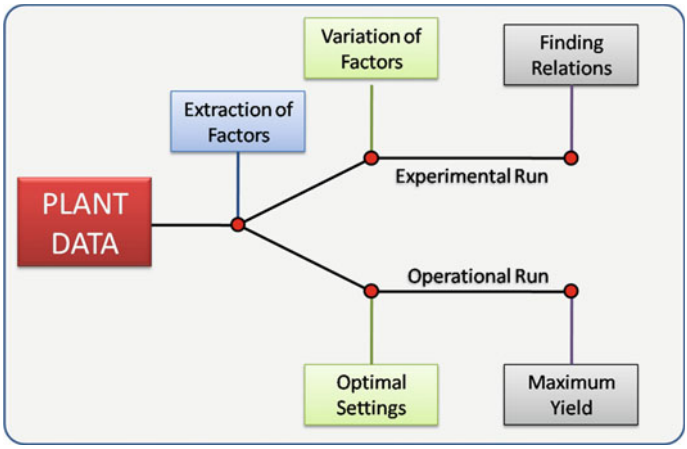


Fig. 8 ICT system for control

While the experimentation to optimize the plant can be done on one plant, the experience gained can be used in running several plants.

Two algorithms for the two modes of the control and running of any plant are delineated as

1. Experimental optimization mode.
2. Running mode.

The first mode will employ variations in feed to ascertain the optimum batch factors that yield maximum methane yield for a class of input feed. This will need data from plant to be shared with the skilled engineer in terms of plant settings and control factors as well as plant feed factors. The second mode can be running of the plant by unskilled person who records and uploads the data to reach the skilled engineer via

ICT technologies, and all decisions can be taken by the engineer. Considerable saving and efficiency in running many biogas plants by a single engineer can result. As stated in the beginning, many biogas plants fail to run for lack of skilled intervention and control.

Lastly, we would like to add that the canteen that supplied the biomass was supplied with the methane gas for use in their kitchen with a suitably designed burner [28, 29] and this resulted in considerable saving on fuel costs besides reducing the need to transport the waste to landfills by municipal trucks.

6 Conclusions

Not only design, but operational aspects have to be right, to derive a trouble-free working life of a biogas plant. Bacteria, the main players for methane gas liberation through anerobic digestion, need their replenishment on a continual basis. The factors that affect growth of bacteria for a given input need to be monitored and regulated to derive maximum gas generation. This is realized in an efficient manner through ICT technology where unskilled workers man the biogas plants in an area and send data to a skilled engineer who takes informed decisions on how to run the plants by sending back instructions to the workers at various plant sites.

References

1. BIOGAS NORD GmbH, Werningshof, 2-4, 33719 Bielefeld, Germany (2007). https://www.biogasnord.com
2. Basic Data on Biogas; Swedish Gas Center (2012). https://www.sgc.se
3. Gustafsson, M., Anderberg, S.: Dimensions and characteristics of biogas policies—modelling the European policy landscape. Renew. Sustain. Energy Rev. **135**, 110200 (2021). ISSN 1364-0321. https://doi.org/10.1016/j.rser.2020.110200
4. Park, Y.S., Szmerekovsky, J., Dybing, A.: Optimal location of biogas plants in supply chains under carbon effects: insight from a case study on animal manure in North Dakota. J. Adv. Transp. **2019**, Article ID 5978753, 13 p. (2019). https://doi.org/https://doi.org/10.1155/2019/5978753
5. Banks, C.: Optimizing anaerobic digestion: evaluating the potential for anaerobic digestion to provide energy and soil amendment. Engineering and the Environment, University of Southampton. 41 p. Buren, K.A.V. (1983)
6. Blumensaat, F., Keller, J.: Modelling of Two-Stage Anaerobic Digestion Using the IWA Anaerobic Digestion Model No. 1 (ADM1). Water Res. **39**(1), 171–183 (2005). ISSN 0043–1354
7. Encyclopaedia Britannica
8. AbundeNeba, F., Tornyeviadzi, M., Asiedu, N.Y., Addo, A., Morken, J., Østerhus, S.W., Seidu, R.: Can the operating limits of biogas plants operated under non-isothermal conditions be defined with certainty? Modeling self-optimizing attainable regions. Comput. Chem. Eng. **141**, 107001 (2021). ISSN 0098-1354. https://doi.org/10.1016/j.compchemeng.2020.107001

9. Ericsson, N., Nordberg, Å., Berglund, M.: Biogas plant management decision support—a temperature and time-dependent dynamic methane emission model for digestate storages. Bioresour. Technol. Rep. **11**, 100454 (2020). ISSN 2589-014X. https://doi.org/10.1016/j.biteb.2020.100454
10. Biogas Plant Design. https://bio-gas-plant.blogspot.com
11. Varsha, B., Soniya, K., Ankita, K., Sandesh, D.: Design of floating biogas digester and production of biogas using kitchen waste. Int. J. Eng. Res. Technol. (IJERT) **8**(6) (2019). ISSN: 2278-0181
12. Damrongsak, D., Chaichana, C., Wongsapai, W.: Small-scale biogas plant from swine farm in Northern Thailand. Energy Procedia **141**, 165–169 (2017). https://doi.org/10.1016/j.egypro.2017.11.031. ISSN 1876–6102
13. Dorella, M., Romagnoli, F., Gruduls, A., Collotta, M., Tomasoni, G.: Design of a biogas plant fed with *Cladophora* Sp. algae and wheat straw. Energy Procedia **147**, 458–466 (2018). https://doi.org/10.1016/j.egypro.2018.07.046. ISSN 1876–6102
14. Rahmat, B., Hodiyah, I., Supriadi, A., Hikmat, M., Purnama, G.: Design of biogas digester with thermophilic pretreatment for reducing fruits wastes. Int. J. Recycl. Organic Waste Agric. **8**(Suppl 1):S291–S297 (2019). https://doi.org/10.1007/s40093-019-00301-y
15. Biogas Plant Design. https://www.nzdl.org/gsdlmod
16. Calculations for biogas plant design. https://www.biogasworld.com/biogas-calculations
17. Planning & Construction of Biogas Plant, Torsten Fischer, Andres Krieg. https://www.kriegfischer.de/
18. Samer, M.: Biogas Plant Constructions, Cairo University, Faculty of Agriculture, Department of Agricultural Engineering, Egypt. https://cdn.intechopen.com/
19. Baccioli, A., Ferrari, L., Marchionni, A., Desideri, U.: Biogas upgrading and liquefaction in an anaerobic digester plant. Energy Procedia **148**, 655–662 (2018). ISSN 1876-6102. https://doi.org/10.1016/j.egypro.2018.08.154
20. Mardani, A., Mahalegi, H.K.M.: Hydrogen enrichment of methane and syngas for MILD combustion. Int. J. Hydrogen Energy **44**(18), 019, 9423–9437. https://doi.org/10.1016/j.ijhydene.2019.02.072. ISSN 0360-3199
21. Sakiewicz, P., Piotrowski, K., Ober, J., Karwot, J.: Innovative artificial neural network approach for integrated biogas—wastewater treatment system modelling: effect of plant operating parameters on process intensification. Renew. Sustain. Energy Rev. **124**, 109784 (2020).https://doi.org/10.1016/j.rser.2020.109784. ISSN 1364-0321
22. Daniyan, I.A., Daniyan, O.L., Abiona, O.H., Mpofu, K.: Development and optimization of a smart system for the production of biogas using poultry and Pig Dung. Procedia Manuf. **35**, 1190–1195 (2019). https://doi.org/10.1016/j.promfg.2019.06.076. ISSN 2351-9789
23. Singh, A.D., Upadhyay, A., Shrivastava, S., Vivekanand, V.: Life-cycle assessment of sewage sludge-based large-scale biogas plant. Bioresour. Technol. **309**, 123373 (2020). https://doi.org/10.1016/j.biortech.2020.123373. ISSN 0960-8524
24. Singh, B., Szamosi, Z., Siménfalvi, Z., Rosas-Casals, M.: Decentralized biomass for biogas production. Evaluation and potential assessment in Punjab (India). Energy Reports **6**, 1702–1714 (2020). https://doi.org/10.1016/j.egyr.2020.06.009. ISSN 2352-4847
25. De Clercq, D., Jalota, D., Shang, R., Ni, K., Zhang, Z., Khan, A., Wen, Z., Caicedo, L., Yuan, K.: Machine learning powered software for accurate prediction of biogas production: a case study on industrial-scale Chinese production data. J. Cleaner Prod. **218**, 390–399 (2019). https://doi.org/10.1016/j.jclepro.2019.01.031. ISSN 0959-6526
26. Logan, M., Safi, M., Lens, P., Visvanathan, C.: Investigating the performance of internet of things based anaerobic digestion of food waste. Process Safety Environ. Prot. **127**, 277–287 (2019). ISSN 0957–5820. https://doi.org/10.1016/j.psep.2019.05.025.
27. Gandiglio, M., Lanzini, A., Santarelli, M., Acri, M., Hakala, T., Rautanen, M.: Results from an industrial size biogas-fed SOFC plant (the DEMOSOFC project). Int. J. Hydrogen Energy **45**(8), 5449–5464 (2020). https://doi.org/10.1016/j.ijhydene.2019.08.022. ISSN 0360-3199

28. Jadhav, P., Sudhakar, D.S.S.: Analysis of burner for biogas by computational fluid dynamics and optimization of design by genetic algorithm. Int. J. Res. Emerg. Sci. Technol. **2**(7) (2015)
29. Decker, T., Baumgardner, M., Prapas, J., Bradley, T.: A mixed computational and experimental approach to improved biogas burner flame port design. Energy Sustain. Dev. **44**, 37–46 (2018). https://doi.org/10.1016/j.esd.2018.02.008. ISSN 0973-0826

Machine Learning Techniques for Music Genre Classification

Nandkishor Narkhede, Sumit Mathur, and Anand Bhaskar

Abstract Music plays a significant role in one's life. The quantity of music released daily is increasing tremendously on Internet platforms like Soundcloud and Spotify. Music brings like-minded people together. Classification has been a challenging task in the field of music retrieval (MIR). Categorization of music can help explain some of the exciting problems such as creating song references, finding related songs, and finding communities that will love that particular song. The aim of our research is to find the best machine learning algorithm that predicts the type of songs using machine learning methods. This paper provides a comparative analysis of the varied machine learning methods used in the field of genre planning.

Keywords Genre · MIR · SVM · KNN · CNN · Accuracy · GTZAN

1 Introduction

Wikipedia says: "Music is a general category that identifies specific parts of music, such as shared traditions or conventions. It should be distinguished from the form and style of music, but in practice these words are sometimes used differently."

Audio is a sound signal method that defines parameters such as frequency, decibels, and bandwidth. The standard audio signal is displayed as amplitude and time function. These audio signals come in a variety of formats, including MP3 format, WMA format (Windows Media Audio), and WAVE format (Waveform Audio File) and are read, analyzed, and computer friendly.

N. Narkhede (✉) · S. Mathur · A. Bhaskar
SPSU, Udaipur, Rajasthan, India
e-mail: narkhede.nandkishor@spsu.ac.in

S. Mathur
e-mail: sumit.mathur@spsu.ac.in

A. Bhaskar
e-mail: anand.bhaskar@spsu.ac.in

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_15

Companies like Spotify, Soundcloud, Apple, etc., use music classification as a recommendation to their customers or simply as a product like Shazam. In performing any of these two functions, music genre recognition is the very first step.

Music is a global language and symbol that we create, understand, and enjoy individually and in groups. It is available in nearly limitless forms–alternatively referred to as genres. There are prevalent capabilities of music that every music form has, regardless of their rhythm, form, scoring, or timbre. Thus, we can summarize a song into a series of values, and it becomes a data factor that may be analyzed.

We postulate that we are able to determine the genre–musical-class of a song by the usage of these features. The excessive dimensionality of the data representing a song makes this a complex task. Machine learning, specifically neural networks, specialize in deciphering patterns from complex datasets, accordingly we thought that the use of machine learning to try to learn how to distinguish between specific genres of song can be effective.

Typically, making use of a genre tag to a piece of song has been a human task. As genres are a human abstraction, it is relatively smooth for us to listen to a piece of song and file what class it belongs to. It is much more tough for a machine, as the entire human listening revel in, is surely represented by a vector of capabilities of the song.

1.1 Problem Statement

Music plays a very important role in human life. Music is the glue that holds like-minded people together and holds groups together. Associations can be determined with the help of songs they compose or focus on. Different associations and organizations focus on different types of songs. An important work that distinguishes one type of music from another is the style of the song. Objective of the research:

a. Creating a machine learning model that classifies music into a relevant category.
b. To compare the accuracy of this machine learning model and previous models and to come to the necessary conclusions.

1.2 Objectives

1. Develop a machine learning model that automatically classifies music based on various features instead of manually entering the category.
2. The model correctly classifies new music into its category to achieve better accuracy.

2 Related Work

Music research has not been completed as regards the collection of musical information and is still an active research subject. This section describes the activities performed in the classification of a music category.

Chillara et al. [1] developed multiple taxonomic models such as logistic regression, ANN, and CNN and trained them through the free music archive (FMA) dataset. They compared the performance of these models and logged their results with estimated accuracy. Trained models of the Mel-spectrograms of songs and their audio features. The CNN model, which uses only spectrograms as input to estimate the music segment with an experimental accuracy of 88.54%, has been observed to excel in terms of accuracy [1].

Bahuleyan [2] compared the performance of different classifiers such as logistic regression, SVM, random forest, and CNN [2]. The feature vectors used are from the frequency domain, time domain, and MFCC properties. The accuracy of the CNN classification is 89.4%.

Elbir et al. [3] explored the fascinating features of music using digital signal processing techniques, followed by KNN, Naive Bayesian, random forest, decision tree, SVM, and CNN machine learning techniques music with recommendation. SVM worked well with 69.90% accuracy [3].

Chavan et al. [4] used machine learning and in-depth learning methods to automatically classify a music category. The authors used traditional neural networks based on an in-depth study to assess the type of music label using a spectrogram [4]. The authors suggested classification by reducing the noise using the bold sound reduction used by Fourier analysis to improve the audio frequency spectrum. Similarly, they applied and monitored machine learning models with features and compared their performance [4].

Bisharad and Laskar [5] proposed an algorithm for genre recognition using a convolutionary, recurrent neural network-based model that is trained on mel spectrogram extracted from the GTZAN dataset 3 s period audio clip. The proposed model delivered 85.36 percent accuracy. The model provided an accuracy of 86.06 percent when applied to the MagnaTagATune dataset [5].

Ramirez and Julia Flores [6] made use of classifiers like decision trees, linear SVMs, Naïve Bayes, DNNs, and RNNs for classification of music genres. With only a few seconds of training, the Naïve Bayes classifier has provided interesting results for some classes with AUC over 0.9 [6].

Bhowmik and Chowdhury [7] suggested a better way of automatically classifying music into different categories using the machine learning method and demonstrated insights and results in the application of a specific scheme for the classification of a large group of bangla music content [7]. Machine learning algorithms J48, SVM, and Naïve Bayes were used. SVM achieved 85% high accuracy.

Duggirala and Moh [8] have tried to use cutting-edge technology such as hierarchical attention networks within the text taxonomy domain to classify different categories of tracks. To do this, the music was first separated into different tracks

(drums, vocals, bass, appendix) and converted into symbolic text data. The obtained accuracy was 90% [8].

Pelchat and Gelovitz [9] reviewed some machine learning methods for music classification. The research used images of spectrograms generated from song time pieces as input to a neural network called CNN to classify songs into respective musical categories [9]. In that they achieved an accuracy of 67%.

Bisharad and Laskar [10] proposed a model classification based on a residual neural network, trained in short clips of just 3 s. The specific model has an error rate of 18%, 9%, and 5.5% when estimating for the music clip in the Top-1, Top-2, and Top-3 categories, respectively [10].

Munkhbat and Ryu [11] proposed a synergistic approach that integrates machine learning algorithms such as KNN, Naïve Bayes, multi-layer perceptron and random forest for music-based stress relief. The obtained accuracy is 78.3%, and the precision and recall values were 80.5% and 54.8%, respectively [11].

Using the residual attention network (RAN), Nguyen et al. [12] proposed a new approach to classify music. The methodology was analyzed based on music data from the Salo AI "Music Genre Classification" challenge. This dataset contains recordings of ten different musical classes in Vietnam. Data reinforcement in the evaluation set was implemented along with error analysis. The results obtained were 71.7% accurate on the test set [12].

Patil and Nemade [13] have proposed learning and network-based learning methods that support audio processing, fragmentation, extraction, categorization, and audio signal retrieval in the database. The authors proposed novel classification and retrieval methods using FPNN, including the fuzzy concept and features of PNN [13]. The authors found that FPNN classifier provided a better accuracy of 93.67%, F1-score of 95.99%, and a Kappa coefficient of 81.04% compared to SVM, kNN, and PNN classifiers.

Leartpantulak and Kitjaidure [14] categorized songs from audio signal to music category using feature extraction. Trimble format, rhythmic content, and pitch content were used as key feature sets. The feature was selected using particle swam optimization (PSO) and sent to the selected feature classifier [14]. The classification effect was less accurate. Therefore, the stacking synchronization method was used to improve the reference. Stacking integration includes basic taxonomy and meta-taxonomy. In the basic classification, classification is done using K-near neighbors (KNN), decision tree (DT), random forest, support vector machines (SVM), and Naive Bayes. This process is generated to generate multiple classifier estimates and sent for meta-classification. Excellent classification with 84.17% accuracy was obtained for SVM.

Dokania and Singh [15] discussed the strong inductive bias of the graph neural networks and showed that the integration of CNN and GNN can achieve the best results on GTZAN and AudioSet (Imbalanced music) datasets [15]. The authors discuss the role of Siamese neural networks, similar to GNN, in studying the edge similarity criteria. Graham (Siamese graph NN) shows signal improvement (99.5%) to establish cutting-edge results.

Table 1 summarizes the details discussed.

Table 1 Literature review of music genre classification

Authors	Technique used	Applications	Remark
Chillara et al. [1]	Logistic regression, ANN, CNN	Free music archive (FMA) dataset	Classification accuracy = 88.54%
Bahuleyan [2]	Logistic regression, SVM, random forest, and CNN	GTZAN dataset for music signals	Classification accuracy = 89.4% for CNN
Elbir et al. [3]	KNN, SVM, and CNN	GTZAN dataset	Classification accuracy = 69.90% for SVM
Chavan et al. [4]	Convolutional neural network, audacity noise reduction	Music signals	Only proposed idea
Bisharad and Laskar [5]	Convolutional recurrent neural network	GTZAN dataset	Accuracy = 85.36%
Ramirez and Julia Flores [6]	Decision trees, Naïve Bayes classifiers, linear SVMs, DNNs	GTZAN Dataset	Accuracy = 90% for Naïve Bayes
Bhowmik and Chowdhury (2019) [7]	J48, SVM, and Naïve Bayes	Bangla music	Classification accuracy = 85% for SVM
Duggirala and Moh (2020) [8]	Hierarchical attention networks	Music signals	Accuracy = 90%
Pelchat and Gelowitz [9]	Convolutional neural network	Music signals	Classification accuracy = 67%
Bisharad and Laskar [10]	Residual neural network	Music signals	Error rate = 18%
Munkhbat and Ryu [11]	Ensemble approach using KNN, Naïve Bayes, and random forest	Music signals	Accuracy = 78.3%
Nguyen et al. [12]	Residual attention network (RAN)	Zalo AI music dataset	Classification accuracy = 71.7%
Patil and Nemade [13]	SVM, KNN, fuzzy PNN (FPNN)	GTZAN dataset	Classification accuracy 93.67% for FPNN
Leartpantulak and Kitjaidure [14]	Particle swarm optimization (PSO), KNN, decision tree, SVM	Music signals	Classification accuracy = 84.17% for SVM
Dokania and Singh [15]	Graph NN with Siamese	GTZAN dataset	Classification accuracy = 99.5%

3 Conclusion

From Table 1, it is clear that many of the researchers have used GTZAN dataset for music genre classification. GTZAN dataset contains 1000 audio tracks of music signals, each of 30 s duration in mono mode with frequency of 22,050 Hz. These 1000 tracks consist of ten different classes of music genres like blues, classical, country, disco, hiphop, jazz, metal, pop, reggae, and rock. Many of the researchers have used SVM and KNN classifier. The minimum classification accuracy obtained in Table 1 is 67% using CNN, and maximum classification accuracy obtained is 99.5% using graph neural network. We propose to compare the classification accuracy of SVM, kNN, PNN, and CNN for music genres.

References

1. Chillara, S., Kavitha, A.S., Neginhal, S.A., Haldia, S., Vidyullatha, K.S.: Music genre classification using machine learning algorithms: a comparison. Int. Res. J. Eng. Technol. (IRJET) **06**(05), 851–858 (2019)
2. Bahuleyan, H.: Music genre classification using machine learning techniques (2018). https://arxiv.org/abs/1804.01149
3. Elbir, A., Çam, H.B., İyican, M.E., Öztürk, B., Aydın, N.: Music Genre Classification and Recommendation by Using Machine Learning Techniques. IEEE (2018)
4. Chavan, O., Kharade, N., Chaudhari, A., Bhalke, N., Nimbalkar, P.: Machine learning and noise reduction techniques for music genre classification. Int. Res. J. Eng. Technol. (IRJET) **06**(12), 225–228 (2019)
5. Bisharad, D., Laskar, R.: Music genre recognition using convolutional recurrent neural network architecture, pp. 1–13. wileyonlinelibrary.com/journal/exsy, April 2019
6. Ramirez, J., Julia Flores, M.: Machine learning for music genre: multifaceted review and experimentation with audioset. J. Intell. Inf. Syst. (2019)
7. Bhowmik, A., Chowdhury, A.E.: Genre of Bangla music: a machine classification learning approach. AIUB J. Sci. Eng. **18**(02), 66–72 (2019)
8. Duggirala, S., Moh, T.-S.: A Novel Approach to Music Genre Classification using Natural Language Processing and Spark. IEEE (2020)
9. Pelchat, N., Gelowitz, C.M.: Neural network music genre classification. In: Proceedings of IEEE Canadian Conference of Electrical and Computer Engineering (CCECE) (2019)
10. Bisharad, D., Laskar, R.: Music genre recognition using residual neural network. IEEE (2019)
11. Munkhbat, K., Ryu, K.H.: Classifying songs to relieve stress using machine learning algorithms. Adv. Intell. Inf. Hiding Multimedia Signal Process. Smart Innov. Syst. Technol. **157**, 411–417 (2020)
12. Nguyen, Q.H., et al.: Music genre classification using residual attention network. In: Proceedings of the IEEE International Conference on System Science and Engineering (ICSSE), 2019, pp. 115–119.
13. Patil, N.M., Nemade, M.U.: Content-based audio classification and retrieval using segmentation, feature extraction and neural network approach. In: Advances in Computer Communication and Computational Sciences, Advances in Intelligent Systems and Computing, pp. 263–281 (2019)

14. Leartpantulak, K., Kitjaidure, Y.: Music genre classification of audio signals using particle swarm optimization and stacking ensemble. In: IEEE Proceedings of the 2019 International Electrical Engineering Congress
15. Dokania, S., Singh, V.: Graph Representation learning for Audio & Music genre Classification (2019)

A Review on Multipolarity in Sentiment Analysis

Roopam Srivastava, P. K. Bharti, and Parul Verma

Abstract Sentiment analysis can be best understood as scientific computation analysis and investigation that explains opinions, attitudes, emotions and sentiments of people. Sentiment analysis has been very popular for the last ten years; therefore, it is still approached by most researchers and companies as a polarity detection problem. Sometimes, if we analyze a given document or sentence or unit of text, result may depict multipolarity, and hence, the net result maybe misleading similar to result of average that does not give information of all its constituent members and their variation. Some sentiment analysis challenges are like sarcasm detection, negation detection, word ambiguity and multipolarity also. This review paper explores the sentiment analysis challenge of multipolarity for understanding and improving this important form of analysis.

Keywords Sentiment analysis · Machine learning · NLP

1 Introduction

Sentiment analysis is a detailed study of opinions, emotions, attitudes and sentiments that people possess for a certain entity which they express as text in written form. It also finds an extensive usage in monitoring social media as it gives general idea of the larger public view of popular issues and subjects. It accelerates public sentiment in areas such as social media, politics, and the stock market etc. Therefore this analysis is also known as opinion mining, for example the sentence:

R. Srivastava (✉) · P. K. Bharti
Shri Venkateshwara University, Gajraula, UP, India
e-mail: roopam.amogh@gmail.com

P. K. Bharti
e-mail: vc@svu.edu.in

P. Verma
Amity University, Lucknow, UP, India
e-mail: Pverma1@lko.amity.edu

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_16

"My vacations been delayed Brilliant!"

If we go with the above sentence, we come to know that the delay of a vacation is not a happy experience for most people. The above sentence clearly indicates a delay of vacation which is not a happy experience for any human being, here the person is speaking sarcastically. But the machine analyses the word "brilliant" and categorises it as positive statement. Natural language processing (NLP) and machine learning (ML) are playing an important role in the sentiment analysis.

1.1 Sentiment Analysis Application

Opinion is very significant for corporate, organization and business as well as for politics. Consumer and people's opinion is always wished by them about their products and services. The area of sentiment analysis application is wide. It is also used in many application like hotels, restaurants, online shopping or stock and share market, government, politics, hospitals, etc. People go through the review and find the best for them. Sentiment analysis converts the unstructured information into the structured information.

Different level of Analysis

Document Level—At **t**his level, it is categorized whether the complete and final. Opinion of a document gives positive or negative sentiment.
Sentence Level—At the next level, analysis is done for each sentence to decide whether each sentence expresses a true, negative or neutral opinion.
Aspect Level—Analysis of document level and sentence level does not discover.

What public like and dislike precisely. To put it another way, they don't say what each Opinion is about; instead, they focus on the target of the opinion.

1.2 Approaches of Sentiment Analysis

To explore sentiment analysis system by using set of algorithm and techniques as following ways

Rule Based: This approach uses a set of human-crafted rules to assist identify subjectivity, polarity or subject opinion. These rules may include various techniques developed in computational linguistics, such as stemming, tokenization, part of approaches to NLP. Regular expressions and context-free grammars are textbook samples of rule-based approaches to NLP. Rule-based tends to target pattern matching or parsing. Finally, this procedure analyzes a test as positive when the presence of number of positive words is greater than the number of negative words vice versa, otherwise return neutral.

Automatic: This method is opposite to rule-based system. It depends on machine learning techniques. For building a machine learning-based system, we first require to collect a training dataset that has expression with a positive, negative and neutral sentiment. Since the algorithms do not directly understand text, the examples need to be transformed into vectors. A simple but successful method for doing this is bag of words.

Hybrid: The combination of rule-based and automatic approach is called hybrid method. The concept of hybrid method is very spontaneous. Accuracy and precision are improved significantly by combining both the above methods.

1.3 Sentiment Analysis Techniques

Machine-Based Approach

As we all know, machine learning is a logical learning based on the use of algorithm process and statistical models by a system to accomplish a definite task without using explicit and clear instruction. It is the largest subset of AI (Fig. 1), which contains the reference [1]. ML algorithms can be classified as—supervised and unsupervised.

Lexicon Approach—This method is guided through the utilization of dictionary, which comprises pre-tagged lexicons. Using seed words, dictionaries can be easily created, either manually or automatically, by increasing the list of words. An input text is converted to a token by the aid of the tokenizer, and all newly arriving tokens are then matched for the lexicons in the dictionary. On encounter of a positive match,

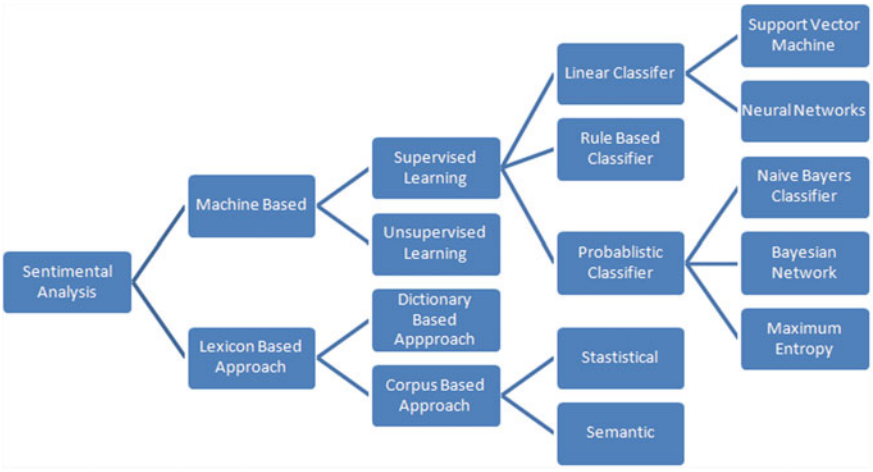


Fig. 1 Sentiment classifier technique [1]

it is further added to the entire pool of scores for the text that has been inputted. For e.g. if the word nice is positively matched in dictionary, it will increment this text's total score else decrement or tag that word as negative. Lexicon algorithms are further classified as—dictionary and corpus approach.

1.4 Challenges of Sentiment Analysis

If we dive deeper, we will locate that there are numerous challenging problems in sentiment analysis. Some sentiment analysis challenges like multipolarity, word ambiguity, sarcasm detection and negation detection, etc. are illustrated here-

Multipolarity

Sometimes, we can observe in a given sentence or document or in any unit of text; we need to examine that some words are having positive polarity and some having negative polarity. This depicts multipolarity in the given sentence, and it leads to misleading final results of the analysis. We can easily understand multipolarity in the given example.

"The speaker quality of my new mobile is very nice but poor battery life."

In this example, the word "speaker" is one aspect and is assigned a positive polarity. Similarly, the word "battery life" is another aspect but it denotes a negative polarity.

Word Ambiguity

It is another challenge of sentiment analysis. The ambiguity words show the different meanings of using a same word in a given sentence. It shows the context polarity in the sentence.

An opinion lexicon consists of the opinion words with their corresponding polarity value. SentiWordNet, General Inquirer are some examples of opinion lexicons. An example of ambiguity is as given below

1. "My son plays well with the **racket**."
2. "The minister was involved in the money laundering **racket**."
3. The above two examples demonstrate how polarity of words influences the sentiment opinion. The polarity of word "racket" is predicted as positive in the first example, and the same word "racket" denotes a negative polarity in the second example.

Sarcasm Detection

Negative sentiments using positive words in a sentence are called sarcastic text. Sentiment analysis models can be cheated easily if sarcasm is allowed except when they are predesigned to take its likelihood of occurrence into consideration. We can easily find sarcasm in the content generated by user from comments of Facebook and tweets from Twitter, etc. Numerical sarcasm example is given below

"This car has a great mileage of 4 km per liter."

Negation Detection

Reversing the polarity of even sentences, phrases and words is termed as negation, in linguistics. With help of various linguistic rules, researchers can detect the occurrence of negation but at same the time, it is also important to determine the word range being influenced by negation words. There is no fixed size for the scope of affected word. In the *example* "the play was not interesting," the range is the next word after the negation word. The actual meaning of word changes whether a positive or negative word falls within the range of negation, opposite polarity will be returned in this case.

Negative opinion can be expressed in many ways, for example,

- When a sentence has a *prefix* ("dis-," "non-") or a *suffix* ("-less"), negation can be morphological there.
- The other form to express a negative in the sentence we come to know that negation maybe implicit. *For example,* "If you do not perform well, it will be your first and last play."
- In the given example, although not a single negative word is used, its sense is negative.
- Sometimes, negation can be explicit.

2 Literature Review

This review paper focuses on multipolarity issue of sentiment analysis. Various researchers have worked in this respect and tried to find out the solution to this issue. This section discusses the work of those researchers.

Morgane Marchand et al. [2] studied multipolarity words into phases. The first phase supervised approach is used while another phase is used opinion classification tools based on machine learning algorithm for evaluating the influence of multiple polarity of word. They also used boosting method and Boos Texter as weak classifiers. Their automatic detection method is well performed to detect multipolarity word.

Doaa Mohey El-Din [3]—In this paper, they focus on the one important sentiment challenge which is called negative polarity level. A new technique was developed by them which improve the accuracy by analyzing negative reviews. This research includes five polarity levels that can affect the words or sentences. This technique applies to two datasets which are training and verified datasets. Their technique enhanced the bag-of-words model.

Antonio Moreno-Ortiz et al. [4]—Their method is primarily intended for analysis of text that is in general language using domain-specific analysis in a lexicon system. They used weirdness ratio (R) and "plug-in" lexical approach to develop specialized and general language corpus. For their dataset, they used economic news segment.

Mohan et al. [5]—A corpus-based method was developed by them to reverse the review by building a pseudo-antonym dictionary. Considering two sides of review by dual prediction classifies test review. An unsupervised learning approach was used by them in which they developed the dictionary-based algorithm to remove the polarity shift problem. Restaurant review used as a dataset. NLTK library, tokenizer and POS used in data preprocessing step.

Kalaivani and Thenmozhi [6]—This paper emphasizes recent studies regarding the execution of DL models like " (i) deep neural networks (DNN), (ii) deep-belief network (DBN), (iii) convolution neural networks (CNN) together with, (iv) recurrent neural network (RNN) model." Those DL models aid in resolving different issues of SA like (a) sentiment classification, (b) the classification methods of (i) rule-based classifiers (RBC), (ii) KNN and (iii) SVM classification methods. They also describe the general structure for a lexicon-based approach.

Xinmiao Li et al. [7]—To enhance sentiment analysis, they proposed a global optimization-based sentiment analysis (PSOGO-Senti) approach that uses IG for feature selection and SVM as the learning engine. To achieve a global optimal combination of feature dimensions and parameters in the SVM, the PSOGO-Senti method uses a particle swarm optimization algorithm.

Sujithra et al. [8]—In this planned system, polarity distribution is enforced with multiple polarities. It will increase the accuracy of the polarity consistency check. They have a tendency to scale back the polarity uniformity downsides to the satiable drawback and utilize two quick SAT solvers to discover instability in an exceedingly sentiment lexicon. Their experiments carry out on five sentiment dictionaries and WordNet to indicate inter- and intra-dictionaries inconsistencies.

Mohsen Ghorbani et al. [9]—A convolutional neural network (CNN) and long short-term memory (LSTM) network to spot the polarity of words on the Google cloud and performing arts computations on Google Colaboratory were designed by them. It supported deep learning algorithms with word embedding techniques and features learn through CNN, and these options are fed directly into a bifacial LSTM layer to capture architecture feature dependencies. Dataset used movie review (MR).

Wei Zhang et al. [10]—Emotion in online reviews depends on a cogitative evaluation theory and sentiment analysis; a novel approach is based on the conjugation of support vector machine, and the latent semantic analysis was proposed. The emotion in online review is divided into four classes, i.e., happiness, hope, disgust and anxiety.

Sandeep Nigam et al. [11]—A lot of algorithms found in sentiment analysis. This paper focuses on the taxonomy of different sentiment analysis methods. Task shows that logistic regression provides high score accuracy as compared to other technique. Sentiment140, dataset used for training data. Their model performed well.

Data Table

Study	Method	Algorithms	Features	Dataset	Limitation
Morgane Marchand et al.	Lexicon based	State of the art and supervised approach	Unigram or bigram	Review collected on Amazon	Classifier will not perform without adaptation until it is trained on domain
Doaa Mohey El-Din	Lexicon technique	NB and SVM	Bag of word, POS, n-gram	SentiWordNet, Hownet	The more the research in a sentiment challenges, the less the average of accuracy rate
Antonio Moreno-Ortiz et al.	"Plug-in" lexicon based	Weirdness ratio(R) Measure	Parts of speech	Specialize language and general language corpus	
Mohan et al.	Affective Lexicon	Synset,SID	POS	Japanese restaurant review	Categorize the opinion of emoticons
Kalaivani A, Thenmozhi D	Deep Learning	SVM, KNN, CNN	Bigrams and trigrams	SentiWord--Net, Twitter	DL approaches have the potentiality to make the SA easier. Due to its best, it will more demanding than traditional method
Hongbin Wang et al.	Machine learning,NLP	TF-IDF , M, HKE BiLSTM		SMP-ECISA2019	The model cannot have sturdy reasoning ability, and it learns hidden sentiment tendencies from the corpus

(continued)

(continued)

Study	Method	Algorithms	Features	Dataset	Limitation
Xinmiao Li et al.	ML,NLP	SVM,KNN,NB	POS	Ctrip dataset And Guahao dataset	PSOGO-Senti approach achieves higher improvement for a more complicated SA task
C. Sujithra et al.	SWDs	WordNet,CNF-SAT		Four sentiment dictionary and WordNet	Set of inconstancy words allows the dictionaries to be improved
Mohsen Ghorbani et al.	NLP,DL	CNN,LSTM		Movie Reviews	
Wei Zhang et al.	Supervised	SVM and LSA	TF-IDF	Amazon Chinese reviews data	
Sandeep Nigam et al.	Supervised	Logistic regression, SVM, L1-based SVM, multinomial Naive Bayes, etc.	TF-IDF, trigram	Sentiment 140, dataset	If the size of the corpora gets large, the selection of vocabulary should get too large to process. So that made a decision to restrict the selection of vocabularies

Analysis

Finally, we analyze in the above review paper that has been described the multipolarity challenges to improve SA. In the area, many researchers have done a magnificent work, especially in multipolarity. Here, we analyze the some work in context to our review.

Morgan et al.—Their paper contributed to SemEval 2013 task 2 for tested opinion on Twitter data. They submitted dual tasks. In massage polarity, they focused on the impact of domain information. For detection of multipolarity word, remove-all version improves the F-measure among all five. Progress is more sensible (+1.46% versus + 0.70%) in domain confident experiment due to much smaller dataset [2].

Doaa et al.—They introduced a new method to improve the accuracy of negative review. They include five negative polarity stages. The accurateness of their approach is in four levels. Average accuracy of all levels is 87.4%. For negative precision,

their technique indicates 10% increase in performance. Their technique enhanced the performance of bag-of-words model [3].

Antonio Moreno-Ortiz et al.—Their technique is designed for general language text. They used "plug-in" lexical approach in their technique. Their algorithm returned a precision of 84.69% and accuracy 84.21% [4].

Mohan et al.—A system called dual sentiment analysis was developed by them performed better when compared to bag of words (BOW). They used unsupervised learning approach. They used restaurant review as a dataset and NLTK library used for preprocessing task. As a result, the overall recommendations of percentage of restaurant reviews are 52.0%, which shows the effectiveness of their system [5].

Hongbin Wang et al.—They introduced a brand new methodology to resolved the matter of "weak features" and "multiconfounding weak features" in implicit sentiment analysis of Chinese text. This paper uses the accuracy; recall rate and F1 score measure the model to comprehensively and accurately evaluate the model as an entire. The accuracy is 82.3% and F1 score is 89.1%. F1 score of their technique is 5.9% higher than best model [12].

Ghorbani et al.—This work is aimed to supply an applicable resolution for analyzing sentiments and classification of the opinions into positive and negative categories. With a small number of epochs and a short amount of time, the model will achieve an accuracy of 89.02%. The best results were obtained compared to previous models [7].

Sandeep Nigam et al.—A series of tests were conducted with the aim of validating their model for aspect identification and sentiment classification at the aspect level. Unigrams were evaluated with and without stopwords, as well as with custom stopwords, while Bigram and Trigram were only evaluated "with stopwords." The results depicted during this paper clearly show that the accuracy is healthier with logistic regression on TF-IDF vectorizer with stop words on Trigrams with an 82.59% [11].

3 Conclusion

After studying various approaches and models related to the multipolarity issue of sentiment analysis, we observed that NLP technique is quite a popular approach used for researchers for the classification of polarity. Sentiment analysis has limitation that it is based upon the predefined lexicons. Sentiment word based upon the lexicon is not vague. There are few works of multipolarity based on lexicon-based approach and it has been observed that they outperformed for large number of domains. However, an interesting future work is being pointed out during this review work, which is classifying sentiment of emoticons.

References

1. Medhat, W., Hassan, A., Korashy, H.: Sentiment analysis algorithms and applications: a survey. Ain Shams Eng. J. **4**, 10931113 (2014). https://doi.org/10.1016/j.asej.2014.04.01
2. Marchand, M., Besancon, R., Mesnard, O., Anne, V.: Domain adaption for opinion mining: a study of multipolarity word (2013)
3. Mohey El-Din, D.: Negative Polarity Levels for Sentiment Analysis, February 2017
4. Antonio, M.-O., Javier, F.-C.: Identifying polarity in financial Texts for sentiment analysis: a corpus-based approach. In: 7th International Conference on Corpus Linguistics: Current Work in Corpus Linguistics: Working with Traditionally-Conceived Corpora and Beyond (CILC 2015). https://doi.org/10.1016/j.sbspro.2015.07.451
5. Mohan, I., Rani, D.G., Pranathi, M.J.A.: Corpus Based Dual Sentiment Analysis, 7 June 2019. Stats and author profiles for this publication at: https://www.researchgate.net/publication/333651351
6. Kalaivani, A., Thenmozhi, D.: Sentimental analysis using deep learning techniques. Int. J. Recent Technol. Eng. (IJRTE) **7**(6S5) (2019). ISSN: 2277-3878, Retrieval Number: F11070476S519/19©BEIESP
7. Li, X., Li, J., Wu, Y.: A global optimization approach to multi-polarity sentiment analysis. PLoSONE **10**(4), e0124672 (2015). https://doi.org/10.1371/journal.pone.0124672
8. Sujithra, C., Arunkumar, A.: Polarity consistency checking for multipolarity based domain independent sentiment dictionaries. Int. J. Sci. Eng. Res. **8**(5) (2017). ISSN 2229-5518
9. Ghorbani, M., Bahaghighat, M., Xin, Q., Özen, F.: ConvLSTMConv network: a deep learning approach for sentiment analysis in cloud computing. J. Cloud Comput. Adv. Syst. Appl. **9**, 16 (2020). https://doi.org/10.1186/s13677-020-00162-1
10. Zhang, W., Kong, S.-x.,·Zhu, Y.-c., Wang, X.-l.: Sentiment classification and computing for online reviews by a hybrid SVM and LSA based approach. Received: 19 Nov 2017/Revised: 20 Dec 2017/Accepted: 29 Dec 2017 © SpringerScience+BusinessMedia, LLC, part of SpringerNature2018. https://doi.org/10.1007/s10586-017-1693-7
11. Nigam S., Das, A.K., Balabantaray, C.R.: Machine learning based approach to sentiment analysis. In: International Conference on Advances in Computing, Communication Control and Networking (ICACCCN2018) ISBN: 978-1-5386-4119-4/18/$31.00 ©2018 IEEE
12. Wang, H., Hou, M., Li, F., Zhang, Y.: Chinese Implicit Sentiment Analysis Based on Hierarchical Knowledge Enhancement and Multi-Pooling. Received 1 June 2020, accepted 7 July 2020. Date of publication 13 July 2020, date of current Version 21 July 2020.https://doi.org/10.1109/ACCESS.2020.3008874

A Deep Learning-Based Approach for Detection of Viral and Bacterial Pneumonia from Chest X-Rays

Himanshu Ashar, Bhavesh Singh, Rahil Desai, Ali Abbas Rizvi, and Purva Raut

Abstract Pneumonia is a disease which has been prevalent for a long period of time, despite active attempts to limit the harm caused due to it. The effect it can have on an individual can vary diversely from person to person, with age being a critical factor. Infants and people of an older age tend to be at maximum risk due to this disease, which affects the lungs. Two categories of pneumonia exist, viral pneumonia and bacterial pneumonia. Over the past few years, the deep learning community has made significant contributions, especially in order to assist medical staff in the diagnosis of pneumonia in suspected patients. Several pretrained models, whose weights have been trained on different datasets, have also been applied for this application by making use of transfer learning. In this paper, various such methods have been analyzed. Pretrained models, customized to suit the current purpose of detecting pneumonia, have been implemented to observe the performance. Along with this, we have proposed our own convolutional neural network architecture in order to detect pneumonia. We compare the results of these models with our proposed system, by performing experiments on a dataset containing three categories: normal, bacterial pneumonia and viral pneumonia.

Keywords Pneumonia · Deep learning · Convolution neural network

H. Ashar (✉) · B. Singh · R. Desai · A. Abbas Rizvi · P. Raut
Department of Information Technology, Dwarkadas J. Sanghvi College of Engineering, Mumbai, India
e-mail: himanshu.ashar@djsce.edu.in

B. Singh
e-mail: bhavesh.singh@djsce.edu.in

R. Desai
e-mail: rahil.desai@djsce.edu.in

A. Abbas Rizvi
e-mail: aliabbas.rizvi@djsce.edu.in

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_17

1 Introduction

Pneumonia is an infection in either or both lungs caused due to fungi, bacteria or viruses. Symptoms of pneumonia can range from mild to severe, including cough, fever and trouble in breathing. Although most cases recover successfully, it can take several weeks to fully recover. However, infants and adults over the age of 65 are at a higher risk due to the disease. Adults over the age of 65 account for a major amount of deaths in the US every year [2]. Among infants in India, pneumonia accounts for a higher mortality rate, with severe cases arising.

Early detection of pneumonia in patients can act as an effective preventive measure for the deaths caused by it all around the world. Chest X-ray (CXR) images have proven to be a prominent source for detecting pneumonia in patients. They can be processed using Python's OpenCV library to locate white spots which could potentially indicate the presence of pneumonia in an individual [1]. Various convolutional neural network architectures have been created in the last few years to identify pneumonia cases from CXRs.

Some neural network architectures such as the CheXNet [3] have displayed encouraging results which are comparable with expert radiologists in the field to accurately detect pneumonia. In this paper, various architectures such ResNet101 and DenseNet are presented by making use of transfer learning approaches to classify CXR images into viral, bacterial pneumonia, as well as the absence of pneumonia. Along with this, an architecture which utilizes convolutional neural networks is proposed for the same classification purpose.

The following sections of the paper are organized as follows. Section 2 summarizes the past research that has been conducted in the field of pneumonia detection with deep learning techniques. Our proposed methodology for a neural network architecture is explained in Sect. 3. Section 4 explains the experimentation and results. The conclusions drawn from the study are presented in Sect. 5.

2 Related Works

Over the last few years, deep learning has made active contributions to extract crucial features from CXR and CT images, in order to assist radiologists in their diagnosis of patients.

In November 2019, Aledhari et al. [4] analyzed the performance of VGG16, Resnet-50 and InceptionV3 models with pretrained weights from ImageNet. The last layers of all the models were modified by adding multiple convolutional layers, as well as global and max pooling layers, with batch normalization. The results obtained were in a binary classification format; normal or pneumonia infected patients. The InceptionV3 and ResNet-50 models showed test accuracies of 53% and 58%, respectively. The proposed fine-tuned VGG-16 model displayed the highest accuracy of

75%, with 76% specificity. This system can help diagnose pulmonary abnormalities with a higher accuracy by using CXR images.

In January 2020, Sharma et al. [5] proposed various deep convolutional neural networks for extracting essential features from CXR images to detect pneumonia. For comparing the regular dataset with a larger dataset, expanded by data augmentation methods, they performed experiments on the CNNs with the original as well as enlarged dataset. Both the proposed CNNs have a similar architecture, one has a dropout layer, while the other does not. Features are extracted using the convolutional and max pooling layers and divided into two parts, the difference being the filter size and pool size of the layers in each part. The extracted features are passed through a dense layer to classify the CXR image. Data augmentation is performed to avoid the model overfitting to the training data. For the model with a dropout layer, augmented data shows an accuracy of 90.68%, while non-augmented data shows an accuracy of 79.80%. The model without a dropout layer displays an accuracy of 89.32% with augmented data and an accuracy of 74.98% with non-augmented data.

In June 2020, Li et al. [6] proposed an improvement to the Lenet-5 network, originally experimented for recognizing handwritten digits, to build a model for detection of pneumonia. Two datasets were used for this purpose, with dataset 1 being smaller than dataset 2. Four convolutional layers are added in addition to the two convolutional layers of Lenet-5, with pooling performed after every convolutional layer. On experimentation, dataset 1 obtained 98.83% training accuracy and 97.26% testing accuracy, while dataset 2 displayed 98.44% training accuracy and 91.41% testing accuracy. Hence, this modified architecture can serve as a satisfactory source of assistance to radiologists for studying CXR images to diagnose cases of pneumonia.

In February 2020, Labhane et al. [7] demonstrated the use of a simple convolutional neural network and various pretrained models using transfer learning to detect pediatric pneumonia. The simple CNN model contained multiple sequences of the convolutional and pooling layers to extract the features, followed by dense layers to classify the images. Using transfer learning, the architecture and weights of the pretrained models InceptionV3, VGG19 and VGG16 were used, pretrained on the ImageNet dataset. The CNNs were followed by a flattening layer and a dense network of layers to perform classification of the CXR images. On evaluation, the following results were obtained: accuracy of 97% on the Basic CNN, 98% on VGG16, 97% on VGG19 and 98% on InceptionV3.

3 Methodology

In this following section, a brief description of the dataset under study is provided which is followed by the various data augmentation techniques incorporated in our research. Lastly, the different architectures that have been studied in this research are described in detail.

3.1 Dataset

The dataset used to conduct the experiments contains 5859 chest X-Ray pictures, put together for public use by Kermany [8]. Dataset contains 1583 normal CXRs and 4273 CXRs of patients infected with pneumonia. The three categories of X-Rays are—normal (no infection), bacterial pneumonia and viral pneumonia. The dataset contains images with a large range of resolutions, from about 700×400 to 2300×2000. From the dataset, 80% is used as a training dataset, approximately 10% is used as a validation set, and 10% is used as a test dataset. An accurate representation of the test set data is specified in the confusion matrices in Sect. 4.

3.2 Data Augmentation

CNN is currently a significant tool for image classification and image analysis purposes. Proper learning algorithms and deep neural networks have provided good results but one in all the key problems is having sufficient data [9, 10]. Deep learning models need huge amounts of data to train properly and provide good results. Certain applications of deep learning often cannot afford to have high losses and incorrect predictions, such as medical applications of deep learning. Medical applications demand high accuracy, and this is possible if the model trains for a large dataset before being tested. To cater this requirement, it is necessary to have a large dataset. For performing data augmentation, zoom range of 0.15 magnitude is used, shear range of 0.15 is used which rotates the image in counter-clockwise direction, horizontal flip is set to true which randomly flips the input images, the wide range shift and height range shift are set to 0.2, and rotation range which performs random rotations on input images is set to 30.

3.3 Transfer Learning

Another method to overcome the computational issues and training over large networks is taken care of by transfer learning. Transfer learning is using pretrained models and weights to train a large network [11]. Three different approaches for transfer learning are fine-tuning, pretrained models and feature extractor [12]. In this study, a pretrained model for training is used. The last layers must be added according to the output requirements. In this examination, we used Resnet101 and Densenet169. The Resnet101 is a 101 layers deep convolutional neural network. It contains 44,707,176 parameters. Densenet169 has only 14,307,880 parameters while it is 169 layers deep convolutional neural network. All these models used the weights of ImageNet which is trained for over 1.2 million images, classified into 1000 categories [13]. Dense layers at the end of the network are classified between three different categories— normal, bacterial pneumonia and viral pneumonia.

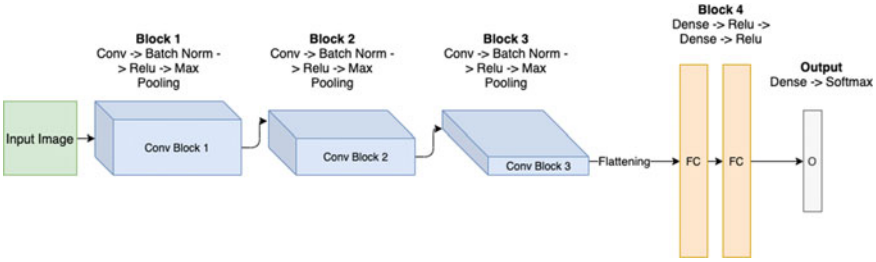


Fig. 1 Proposed Architecture

3.4 Proposed Architecture

The input image is of dimension $256 \times 256 \times 3$, which is passed through the first convolutional block, which contains a 2D convolutional layer, followed by batch normalization, relu activation and a max pooling layer. The convolutional layer has 64 filters of dimension 7×7, and a stride of 2. On applying this filter, the dimension of the image changes to $128 \times 128 \times 64$. Batch normalization and the relu activation function do not affect the input dimension in any of the layers. After applying max pooling, the output from the first convolution layer has a dimension of $127 \times 127 \times 64$. This output is passed through a second convolutional block, which in a similar way converts the features to $32 \times 32 \times 128$.

The next convolutional block receives this as input, processes it and produces features having dimensions $8 \times 8 \times 128$. The output of this block is flattened to give 8192 values. Following this, two dense layers with batch normalization and relu activation are used, which are connected to a final dense layer, which utilizes a softmax function to classify the image into one of the given three categories: normal, viral pneumonia and bacterial pneumonia, according to the class which has the highest probability. Figure 1 shows a complete implementation of the architecture.

4 Experiment and Results

In order to train the model on the given dataset [11], the images are preprocessed as follows. The images are resized to $256 \times 256 \times 3$ to pass into the models in order to maintain uniformity across samples. Further, normalization in the range [0,1] is performed in order to enable efficient computation and neutralize the noise created by outliers. The four models: DenseNet169, ResNet101 and our model are trained with similar hyperparameters, with the only difference being the number of epochs, in order to prevent overfitting on the training split. 'Adam' is used as an optimizer with an adapting learning rate, initialized at 0.001. The batch size for training all the models was chosen to be 32. Since ternary classification is to be performed, categorical cross-entropy is chosen as the loss function.

To observe the effect of augmentation, each of the models is trained on two variations of the dataset. The differences between them are as follows:

- Augmented Set: The data is augmented in several ways. Horizontal flipping of images is enabled, with image inconsistencies filled with the numerical closest pixel value. The rotation range is kept as 30 degrees, while both the height and width shift range are set at 0.2. The zoom range and shear range are both equal to 0.15.
- Non-Augmented Set: No data augmentation is performed, other than automatic filling of image inconsistencies.

4.1 Evaluation Metrics

The testing accuracy and F1-score are chosen as the evaluation metrics. Along with these, a confusion matrix for each model is shown, which provides further detail about the classification of images into the three categories.

4.2 Results

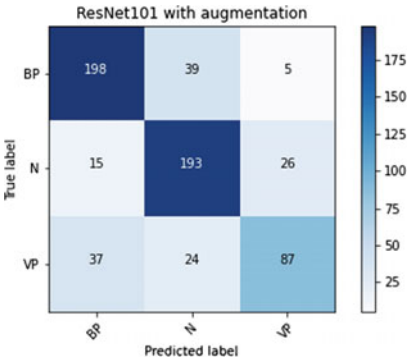
All the models are trained with a softmax layer at the end, with a ternary classification between bacterial pneumonia, normal and viral pneumonia. The accuracy and the F1-score for the test set are shown in Table 1, specifying the respective variations of the dataset, along with the number of epochs for which each neural network was trained. For Table 1, as well as all the classification matrices, bacterial pneumonia, normal and viral pneumonia are abbreviated as BP, N and VP, respectively, (Figs. 2, 3 and 4).

4.3 Observation

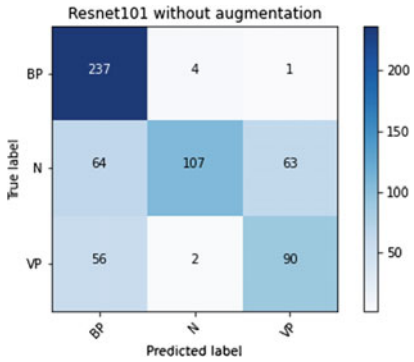
From the results in Table 1, it can be inferred that augmentation of data has a considerable effect on the evaluation metrics. In different models, overfitting was observed at different numbers of epochs, for both the augmented and non-augmented sets. To account for this difference, the number of epochs in the table for a particular model corresponds to the best performance in terms of evaluation metrics. For a given model, the observed accuracies and F1-scores are always higher when trained on the augmented set, when compared to the non-augmented set. Hence, data augmentation plays a significant role in order to prevent overfitting on the training data, which enables it to exhibit a better performance for testing data as well, which would be an ideal result considering a real-world scenario. The DenseNet169 classified the

Table 1 Comparison of accuracy on different architectures

Model information			Test accuracy	F1-score		
Model	Data split	Epochs		BP	N	VP
ResNet101	Augmented set	50	0.7660	0.80	0.79	0.65
	Non-augmented set	40	0.6955	0.79	0.62	0.60
DenseNet169	Augmented set	20	0.8717	0.90	0.88	0.81
	Non-augmented set	40	0.7417	0.79	0.72	0.55
Proposed model	Augmented set	50	0.8413	0.87	0.87	0.75
	Non-augmented set	50	0.6458	0.76	0.58	0.51



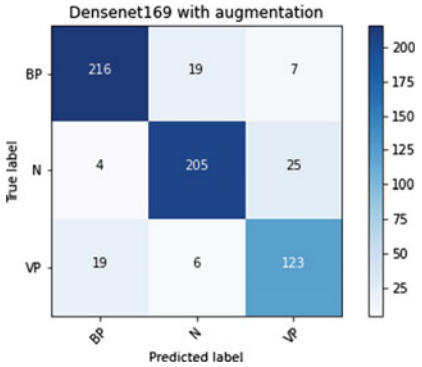
(a) ResNet101 results on the Augmented Set



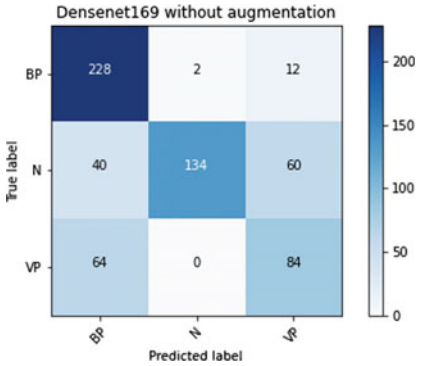
(b) ResNet101 results on the non-Augmented Set

Fig. 2 ResNet101 results on the data in augmented as well as non-augmented form

chest X-Ray images with accuracies of 87.17% , while our model displayed an accuracy of 84.13%. Our model outperformed the ResNet101, which had an accuracy of 76.6%. Therefore, it is observed that our model displays comparable results to various widely utilized neural network architectures, for the purpose of classifying chest X-Rays into three categories. It shows a simplistic architecture with a lesser number of layers, and no pretrained weights used for training purposes.

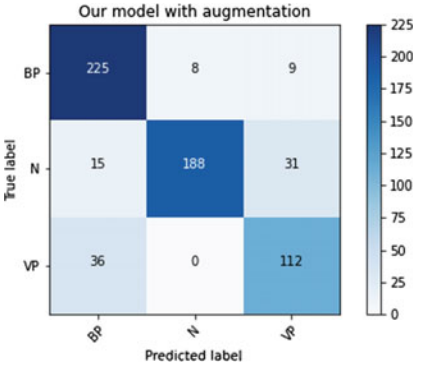


(a) DenseNet169 results on the Augmented Set

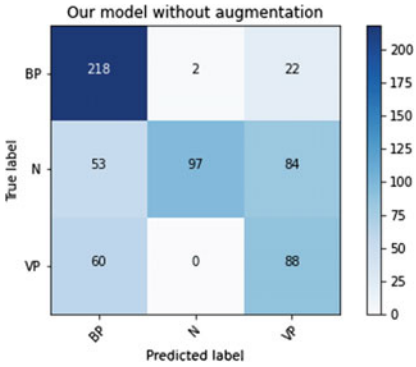


(b) DenseNet169 results on the non-Augmented Set

Fig. 3 DenseNet169 results on the data in augmented as well as non-augmented form



(a) Proposed Model results on the Augmented Set



(b) Proposed Model results on the non-Augmented Set

Fig. 4 Proposed model results on the data in augmented as well as non-augmented form

5 Conclusion

This paper has provided an exhaustive comparison of results obtained from various transfer learning models and our proposed architecture. All the models were trained under the same set of parameters, and out of all, Densenet169 with data augmentation provided better results even though it was trained for the least number of epochs as compared to others. Our proposed architecture, without pretrained weights and with a limited number of layers, outperformed resnet101 providing a decent accuracy on the test data, with augmentation performed on it. As observed in Table 1, it is evident that data augmentation had a significant effect on the performance, as it increased the dataset size and exposed the model to outliers, which created a generalized dataset

and subsequently gave higher accuracies for the testing data. The confusion matrices for each model also provided more details about false classifications, as well as class-wise correct predictions on the dataset. Rather than exclusively classifying chest X-Rays as either pneumonia positive or negative, our work classifies positive pneumonia cases into bacterial or viral pneumonia as well. For future work, apart from X-Rays, a dataset consisting of computed tomography (CT) scan images, with a larger quantity of training, testing and validation samples can be used for better results. The main aim remains to contribute to the field of medicine and to advance it with the help of AI and technology. In light of the ongoing pandemic, neural network architectures can also be created to classify cases into normal, pneumonia or COVID-19 positive samples.

References

1. Sharma, A., Raju, D., Ranjan, S.: Detection of pneumonia clouds in chest X-ray using image processing approach. In: 2017 Nirma University International Conference on Engineering (NUiCONE), Ahmedabad, pp. 1–4 (2017). https://doi.org/10.1109/NUICONE.2017.8325607.
2. Lung.org. n.d. Learn About Pneumonia (online). Available at: https://www.lung.org/lung-health-diseases/lung-disease-lookup/pneumonia/learn-about-pneumonia. Accessed 28 July 2020
3. Rajpurkar, P., Irvin, J., Zhu, K., Yang, B., Mehta, H., Duan, T., Ding, D., Bagul, A., Langlotz, C., Shpanskaya, K., Lungren, M., Ng, A.: CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning (2017)
4. Aledhari, M., Joji, S., Hefeida, M., Saeed, F.: Optimized CNN-based diagnosis system to detect the pneumonia from chest radiographs. In: 2019 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), San Diego, CA, USA, pp. 2405–2412 (2019). https://doi.org/10.1109/BIBM47256.2019.8983114.
5. Sharma, H., Jain, J., Bansal, P., Gupta, S.: Feature Extraction and Classification of Chest X-Ray Images Using CNN to Detect Pneumonia, pp. 227–231 (2020). https://doi.org/10.1109/Confluence47617.2020.9057809
6. Li, X., Chen, F., Hao, H., Li, M.: A pneumonia detection method based on improved convolutional neural network. In: IEEE 4th Information Technology, Networking, Electronic and Automation Control Conference (ITNEC). Chongqing, China, pp. 488–493 (2020). https://doi.org/10.1109/ITNEC48623.2020.9084734
7. Labhane, G., Pansare, R., Maheshwari, S., Tiwari, R., Shukla, A.: Detection of pediatric pneumonia from chest X-ray images using CNN and transfer learning. In: 2020 3rd International Conference on Emerging Technologies in Computer Engineering: Machine Learning and Internet of Things (ICETCE), Jaipur, India, pp. 85–92 (2020). https://doi.org/10.1109/ICETCE48199.2020.9091755.
8. Kermany, D., Zhang, K., Goldbaum, M.: Labeled Optical Coherence Tomography (OCT) and Chest X-Ray Images for Classification (2018). Mendeley Data, v2 https://doi.org/10.17632/rscbjbr9sj https://nihcc.app.box.com/v/ChestXray-NIHCC
9. Mikołajczyk, A., Grochowski, M.: Data augmentation for improving deep learning in image classification problem (2018), pp. 117–122. https://doi.org/10.1109/IIPHDW.2018.8388338
10. Ayan, E., Ünver, H.M.: Data augmentation importance for classification of skin lesions via deep learning. In: 2018 Electric Electronics, Computer Science, Biomedical Engineerings' Meeting (EBBT), pp. 1–4, IEEE (2018)
11. Russakovsky, O., et al.: Imagenet large scale visual recognition challenge. Int. J. Comput. Vision **115**(3), 211–252 (2015)

12. Yosinski, J. ,Clune, J., Bengio, Y., Lipson, H.: How transferable are features in deep neural networks? In: Advances in Neural Information Processing Systems, pp. 3320–3328 (2014)
13. Krizhevsky, A., Sutskever, I., Hinton, G.: ImageNet classification with deep convolutional neural networks. Neural Information Process. Syst. **25** (2012). https://doi.org/10.1145/3065386

Escalating Social Justice Through Effective Implementation of E-governance

K. Rajeshwar and Sonal Mobar Roy

Abstract The Constitution of India envisages for the right to social justice to all the citizens of the country, irrespective of caste, creed, sex, color, race and gender. Despite seventy years of Independence, the people struggle for making their voices heard, especially those coming from the most vulnerable sections of the society. The picture is more grim for rural sector in the country. Herein, a focused intervention is required for filling in the wide gap that exists between the elite and the poor, and therefore e-governance for grievance redressal comes into purview. E-governance provides effective online services to the public and promotes transparency and accountability in procedures, accelerates the response time and facilitates anywhere–anytime services to them using ICT. The authors make an attempt to explain how E-governance is required for exchange of information between citizens, corporates or government departments, helps in increasing revenue, re-structuring of administrative processes and providing quality services. The paper depicts how "Lokvani" proved effective in reducing the number of visits made by public to government offices for availing various services.

Keywords ICT · E-governance · Social justice

1 Introduction

The whole wave of computer applications has brought in an easier mode of access, communication, enhanced services, efficiency in storage and processing of data accurately for utilization to its users. It has also led to make people use and have better control of day-to-day services and revenue generation through some really efficient service portals. The government of India took note of the strengths of the ICT and

K. Rajeshwar (✉) · S. M. Roy
NIRD&PR, Hyderabad, India
e-mail: rajeshwar.nird@gov.in

S. M. Roy
e-mail: sonalmobarroy.nird@gov.in

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_18

strongly proposed 'Digital India' wherein a lot of emphasis was laid on digitalizing data.

Over the period of years it is observed that ICT has brought tremendous benefit to all stakeholders that includes individual, business groups, corporates and governments. It also being used for tabulation and processing of data for analytical leads to Decision Support System (DSS). The people are now more aware about computers and the advantages they have. They are now able to customize their needs in accordance with usage of technology, especially computers. This has actually accomplished to revamp the whole transactional system of the government procedure.

Consequently, quick and exact processing of data leads to better decision making, which in turn makes access to resources and its utilization more affordable which results in overall good governance. This delivers better access to information and other Private Public Partnership firms for well-organized service delivery to citizens in an effective way. By influx of technology, the people demand for transparency, accountability and a speedy decision to the filed in complaint. People have realized the importance of justice and the use of ICT that helps in achieving good governance and seeking justice through an easier platform. All the stakeholders, be it the people, the institutions, the society, etc. have recognized that E-governance is essential and with the use of ICT, wider participation which is inclusive in nature and deeper involvement of the masses is ensured.

1.1 Requirements of E-governance

The first prerogative is to have the infrastructure in place. Computerization of workstation is the basic mandate for an efficient E-governance system. It is noticed that a crucial necessity for achieving it. It is to be stressed here that a computer is the most important requirement for E-governance. In this regard a proper online work flow procedure and well connected systems are crucial. Proper trainings to be imparted to government employees for using systems and working on software applications. A user friendly interface eases work process, accessible service to be provided without any complexities. This helped in replacing the manual system with online. Consequently, the transparency of work being done and thus leads to march towards the e-governance.

1.2 Implementation of E-governance in India

Government of India has developed an action plan to implement and roll out the government functionalities through electronically by means of E-governance. Under this action plan, all the services are categorized into four models viz., Government to Government, Government to Citizens, Government to Business and Government to Employees for providing services. The service delivery mechanism is enhanced

Table 1 Population (aged 14 years and above) Able to operate a computer (%)

Gender	Age group (in years)							
	Rural				Urban			
	14–29	30–45	46–60	>60	14–29	30–45	46–60	>60
Male	22.70	6.40	2.20	0.60	53.60	30.90	20.90	10.60
Female	13.50	1.90	0.50	0.10	43.80	17.30	8.70	2.80
Persons	18.30	4.10	1.40	0.30	48.90	24.30	14.80	6.80

Source NSS KI (Report No. 575(71/25.2) (2015)

from national, state level to local level through citizen-centric and business-centric governance to empower the society and to increase the rural economy which aims for sustainable development. Under the E-governance plan, lot of activities such as Information Technology Act provides the legal framework for all the e-transactions, National Taskforce came into existence to look after the IT and software development, e-office application was developed and rolled out in all the ministries for avoiding delay in government functionality and for all the ministries, websites were developed with updated information to be disseminated for wider circulation and for the benefit of citizens.

The authors would like to stress here that E-governance has been successfully floated and executed in many parts of the country. However, it has a long way to go as more inclusive approach is required. People need to be made aware and trained. They need to be oriented for making a shift from paper mode to digital or electronic mode. A strategic roadmap needs to be designed for achieving the same (Table 1).

1.3 Success of E-governance in India

Most of the administrative work in India is rigged with authoritarian principles which have characterized it as slow and inefficient. It is again reiterated that a model needs to be introduced for doing away with the traditional mode of using paper and embarking on using digital mode. Improving of service delivery mechanism through providing a single window for operations thus leading to enhanced record keeping in a transparent and accurate manner is the need of the hour. There is a less number of procedural complexity and the processes have been considerably improved and become more transparent. A common man has become more aware about his rights and benefits. The newly introduced E-governance model has yielded good results. To further elaborate this point, the authors present the case of "Lokvani" here. They make an attempt to see how effective implementation of E-governance has helped people in grievance redress, thus meting out justice to them.

2 Objectives and Methodology

The authors have discussed the importance of social justice through E-governance application of Lokvani website which is implemented in Sitapur District of Uttar Pradesh State. The legal complaints of the filled in applications by the citizens of the Sitapur District are uploaded in the website for resolving. The secondary data collected is taken from the official of the Lokvani website during the period from 2014 to 15 to 2018-Jan, 2019. The data depicts how the complaints are resolved, disposal of complaints and how the Lokvani website is extensively being used by the public of Sitapur District, thus making it a success.

3 A Case Study of Lokvani

Just like equal distribution of wealth and resources is a must, similarly legal rights should also be equally distributed. There are areas of social justice in which immediate action is required for effective implementation of law [6]. There could be issues related to violation of rights, abuse-physical as well as mental, etc. People who fall victim to the above issues need a platform to raise their voices. They also expect government's support from the other for the redressal of their grievances. Herein, the Lokvani platform comes to their rescue. This initiative was conceptualized and started in 2004, by the district administration of Sitapur, Uttar Pradesh. As per Census 2011, in Sitapur, 11.84% live in urban regions of the District and the rest in rural. Average literacy rate in Sitapur District is 72.17% of which males and females are 76.44% and 67.41% respectively.

With assistance from National Informatics Centre (NIC), a society called Lokvani was constituted to implement the project autonomously. Lokvani centres were launched in 15 locations in the district. Existing cyber cafes and kiosks were identified and granted licenses. This would make the model quite self-sustainable. Login ids and passwords were given to the franchise owners and the website was hosted on NIC web server. The portal provides various services pertaining to government schemes, beneficiary list, access to digitized land records and details of funds released to respective nodal agencies. By charging a pre-fixed fee for financial sustainability the Lokvani centres were permitted to offer various services like issuance of certificates to the benefit of citizens.

3.1 Services Offered by Lokvani

Lokvani system promotes right to information by generating citizens' awareness towards their rights. It offers various services pertaining to Government Departments

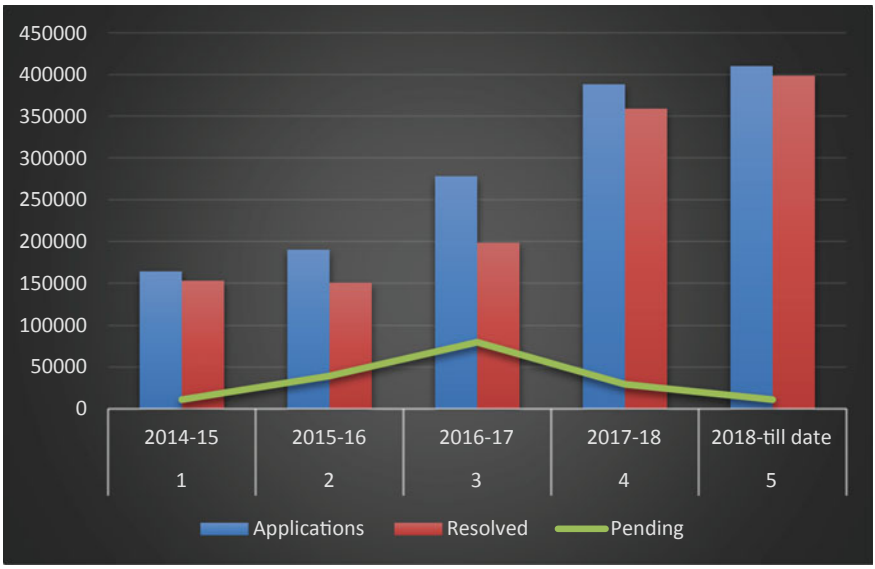


Fig. 1 Implementation & Status of Lokvani in Sitapur District

including District Administration, Land & Revenue, Development Department and Public Grievances.

The data in Fig. 1 portrays the execution status of Lokvani project at Sitapur district of Uttar Pradesh from 2014 till date. The data is sourced from the Lokvani website. Among them, 152,816 complaints were resolved. However, 1105 complaints are showing pending status. It is also evident that the number of people using this platform has increased over the period of years for grievance redressal. Though the number of complaints yet to be resolved are high, they have also increased in proportion over the years.

3.2 Status of Complaints

For the last 5 years, legal complaints registered by the citizens of Sitapur District is 2,75,668 in which 2,28,760 complaints are resolved. Figure 2 shows the status of year wise complaints registered through the Lokvani platform in Sitapur district. The authors have examined the data for the five years span, from 2014 to 2019. The total no. of complaints were 2,75,668 that were registered online and out of which 2,28,760 were resolved. In the table below, the authors have categorized the complaints as no. of complaints registered, no. of applications moved to the next level in 15 days' time period, complaints having difficulty in resolving, complaints noted as defaulters and complaints not having any status due to incorrect filling. The

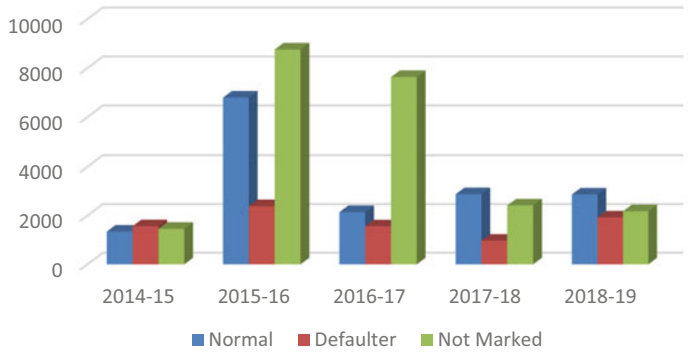


Fig. 2 Status of year wise complaints registered

data in Fig. 2 reveals that the last five years 46,908 total applications are uploaded in the Lokvani website.

Out of which 16,036 are depicted as normal that means the complaints are moved to the next level or resolved within 15 days' time, 8415 are represented as defaulter which means that the complaints are not moved to the next level within 15 days' time. 22,457 applications are shown as not marked status, which shows that the applications are not filled properly. It can be interpreted here that awareness campaign is required for the people if more participation of public is sought in order to submit the application correctly in the Lokvani portal. The relationship between the disposal complaints and the total complaints done is a positive high correlation which is 0.94 i.e., disposal and complaints resolving are going on in parallel and stands significant 1% level (Fig. 3).

For the last 5 years, overall 17.90% applications are pending in Lokvani website which is a very small percentage. This shows that the citizens of the Sitapur District

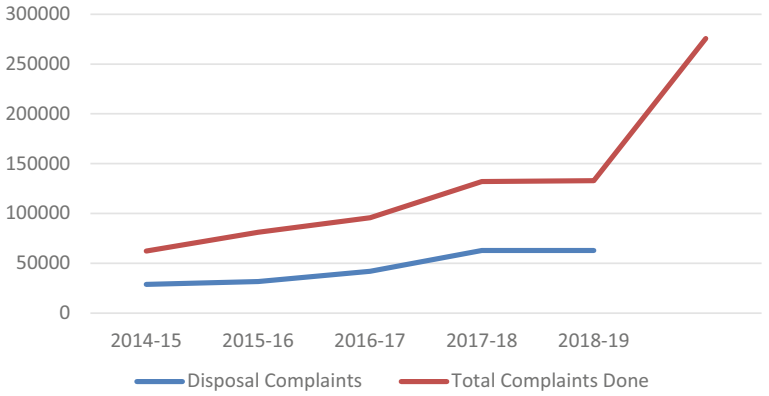


Fig. 3 Status of disposal and complaints done

are extensively using the Lokvani portal for registering the complaints and believe that the complaints will resolve at the earliest.

4 Lokvani: Grievance and Petitions Services

Citizens who are registered users on Lokvani platform can login and track their petition form a nearly kiosk, and also see the status using their mobile phones through IVRS and SMS mode. Color coding helps in knowing the status: white in the beginning and yellow if the petition remains unsolved for 4 days before the due date. It turns red when the due date expires and finally turns green when it is dealt with and disposed.

4.1 Process Flow of Lokvani

The Lokvani program is hosted on internet using SQL Server at the backend and ASP has been used for frontend of application. Proper authorization enables a user to a login and file a complaint. For Sitapur, the URL is https://darpg.gov.in/webcast/lokvani-uttar-pradesh and is supported by National Informatics Centre (NIC). Being in Hindi, the program is easily used by local people. Certain kiosks are identified and they draw support from NIC Delhi webserver and have a rate list fixed for their services.

To avail the services the person could not be educated or having the technical knowledge to register grievance in Lokvani. The operator of Lokvani center registers the complaint on behalf of the citizen and a copy of the same is acknowledged along with the complaint number. The database of the Lokvani stores the information and tracks the movement of the complaint. Whatever the complaints registered in Lokvani are monitored and sorted out by the District Magistrate office of Sitapur district. Further, the complaints are forwarded to the sections concerned. The name of the official who is designated to attend the complaint is uploaded on the server marking with a deadline. The petitioner can get these details within the next 2–3 days. Each officer is responsible for monitoring the status of complaints. It may be noted that the officers have become more attentive for following the defaulter meetings held every week. The introduction of color scheme for marking officer-wise complaint in the chart has proven effective for tracking the progress of the cases. Kiosks earning income through complaints register in Lokvani. A total of Rs. 25 expenditure incurs towards the register of complaint, printout of receipt, knowing the current status of the complaint and for final report. Lokvani will works efficiently between citizen and government to sort out the grievances and avoids the delay in providing the solutions.

4.2 Sustainability and Replicability of the Project

Lokvani project is a self-sustainable model in collaboration with Public Private Partnership (PPP) for long term financial viability. Any of the State can be replicated this model in order to avoid the delay of time in sorting out the grievances and increase the transparency and accountability of the government functionality towards good governance. The Lokvani society who take care of the management, operations, expansion and technical upgradation of the project can be set up in any of the district in the country. As a result, it will becomes a revenue generation for that district and employability will also increases as per the requirement. As per the secondary source of the data collected the Lokvani set up has been implemented in 59 district of Uttar Pradesh. The replications are carried out to other districts as per the orders of the state government is in progress.

5 Conclusion

As per the Lokvani, the services are catered to three stakeholders' viz., citizens, administration and kiosks operators or the small scale enterprises. An innovative step was taken wherein the existing cyber cafes were identified and turned into franchise centres. This led to making the model quire self-sustainable. It was also ensured that any kind of operational support from government would be avoided. Own source funds were created and that let to payment of the kiosk operators. As the system is pretty transparent, bottleneck were overcome in the system thus increasing the non-performances too to pull up their socks. The authors would like to opine that the Lokvani project has brought in a paradigm shift leading to an effective transparent governance system. A common man who is rigged with poverty, struggles with the challenges of illiteracy and economic constraints can also get access to justice. As expressed through the example of Sitapur, the model has emerged as a successful one and overcoming all detracts. An acknowledgeable takeaway from this project is raising awareness among the masses through persistent and focused campaigning for its effectiveness in harboring change.

References

1. Pathak, R.: Enabling efficient administration at the district level: a study of Lokvani project in Uttar Pradesh. Retrieved from https://sitapur.nic.in/lokvani/allabout/Lokvani_tiss.pdf 29 July 2012 (2008).
2. Essays, U.K.: E Governance Research Paper (2018). Retrieved from https://www.ukessays.com/essays/information-technology/a-research-paper-on-e-governance-information-technology-essay.php?vref=1
3. GO 876/78-2-2005-8 I.T./2005 at https://infotech.up.nic.in/hindi/anudesh/itgo.htm

4. Kumar, P., Kumar, D., Kumar, N.: E-governance in India: definitions, challenges and solutions. Int. J. Comput. Appl. **101**(16) (2014) (0975-8887)
5. Nagaraja, K.: E-governance in India: issues and challenges. IOSR J. Econ. Fin. (IOSR-JEF) **7**(5), 50–54 (2016). e-ISSN: 2321-5933, p-ISSN: 2321-5925, Ver. IV
6. Stronks, K., Toebes, B., Hendriks, A., Ikram, U., Venkatapuram, S.: World Health Organisation: Social Justice and Human Rights as a Framework for Addressing Social Determinants of Health, Final report of the Task group on Equity, Equality and Human Rights Review of social determinants of health and the health divide in the WHO European Region, WHO Regional Office for Europe, Denmark (2016)
7. Bhattacharya, S., Goswami (Ghosh), J.: Study of E-Governance: The Attractive Way to Reach the Citizens. JCA Special Issue on "2nd National Conference-Computing, Communication and Sensor Network" CCSN (2011)

REST API and Real-Time Notification of SIsKA-NG Mobile for the Academic Progress Information System

Gede Indrawan, I Gede Aris Gunadi, and I Made Subrata Sandhiyasa

Abstract The objective of this research is to develop mobile information system (MIS) for the academic progress information system. This MIS was called SIsKA-NG mobile where its client application was used to monitor the academic progress related to thesis research of the postgraduate student at the authors' study program. It involved the latest development in network and system, especially on the utilization of web service by using representational state transfer application programming interface (REST API) that was integrated with the existing SIsKA-NG web. Moreover, a real-time notification was developed by using a third-party service Firebase Cloud Messaging (FCM) for certain business processes, like account registration, examination registration, and examination schedule. All of those aspects constructed the back-end development while a certain cross-platform framework was used to construct the front-end development. Overall, this developed unique system was considered as the contribution of this research. Through the testing that involved System Usability Scale (SUS) evaluation, the average SUS score of 82.25, 79.25, and 77.5 was obtained, each from student, lecturer, and admin respondents. All of those scores are in grade A, A-, and B, respectively. They are higher than a score of 68.00 which means SIsKA-NG mobile was considered as a good usability product.

Keywords Academic progress · Real-time notification · REST API · Usability

1 Introduction

SIsKA-NG web [1] as an academic progress information system at the authors' study program is the improvement—in terms of many recent aspects accommodation—from the previous SIsKA [2, 3]. SIsKA itself is an Indonesian abbreviation of

G. Indrawan (✉) · I. G. A. Gunadi · I. M. S. Sandhiyasa
Universitas Pendidikan Ganesha, Singaraja Bali 81116, Indonesia
e-mail: gindrawan@undiksha.ac.id

I. G. A. Gunadi
e-mail: igedearisgunadi@undiksha.ac.id

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_19

"Sistem Informasi Kemajuan Akademik." So far, the existing SIsKA-NG web has already been able to manage the academic progress of postgraduate students related to their thesis research. However, by the web application nature (user has to access it to get information), it has major drawback on real-time delivery of information related to certain business processes inside and finally delaying response to that information. Postgraduate school management sees this situation will degrade quality of thesis management service by the study program.

This research contributed to the solution of that problem through the development of a mobile information system (MIS) that was called SIsKA-NG mobile, from the existing SIsKA-NG web. This solution was conducted since smartphone has already become a part of daily life, includes academic life [4, 5]. It involved the utilization of web service by using representational state transfer application programming interface (REST API) that was integrated with the existing SIsKA-NG web. Standard web service was used since its data and process distribution service between several applications must run well on different operating systems [6]. REST, also known as RESTful, was used by that web service because of its simplicity in the implementation [7–9]. Its communication uses HTTP and can be developed by using various programming languages [10]. Data response generated by REST is in JSON format which is very lightweight on data exchange [11].

Real-time delivery of information (notification) to SIsKA-NG mobile clients was developed for certain business processes by using third-party service Firebase Cloud Messaging (FCM). That FCM service communicates SIsKA-NG web server with its clients by using HTTP [12], and it does not burden both of them [13]. In case SIsKA-NG web server has certain business processes update, it will send the related data first to the FCM that will distribute that data to the clients [14]. This makes the users automatically receive the notification without opening SIsKA-NG mobile application on their smartphone (in the opposite nature to the web application).

This paper was organized into several sections; i.e., Sect. 1 described the problem background related to real-time delivery of information of SIsKA-NG. This section also described some related works; Sect. 2 contains research stages for SIsKA-NG mobile development by providing its design and related aspects; Sect. 3 covers SIsKA-NG mobile implementation and includes its System Usability Scale (SUS) evaluation; and finally, Sect. 4 consists of some important concluded points.

2 Research Method

The waterfall model was used for the development of SIsKA-NG mobile for the academic progress information system, which involves REST API and real-time notification. Although it was classified as a fairly old model, however, the waterfall model is still widely used and even tends to become a standard for other model developments [15].

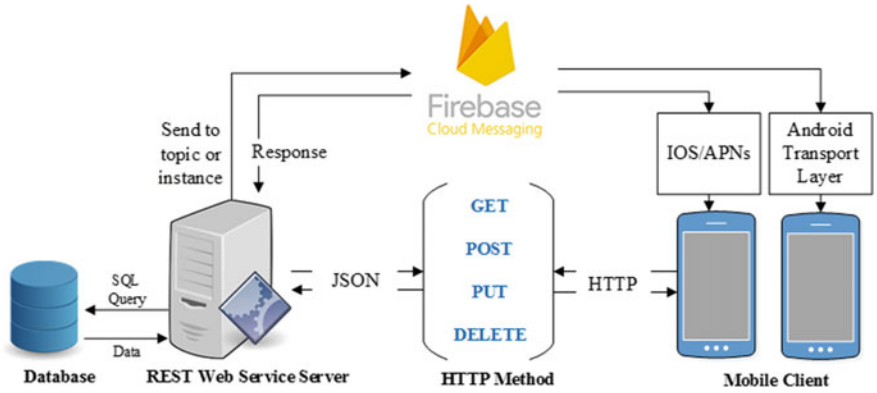


Fig. 1 The system architecture of SIsKA-NG mobile

2.1 Requirement Analysis and Definition

At this stage, the heuristic evaluation method [16] was used where the evaluators were involved to analyze user requirements since not all of SIsKA-NG web menu (business process) would be implemented by SIsKA-NG mobile.

2.2 System and Software Design

The system design provides SIsKA-NG mobile system architecture that was integrated with the existing SIsKA-NG web, while software design provides SIsKA-NG mobile user interaction. Figure 1 shows REST API architecture to identify the HTTP method sent by the client, which can receive, transmit, update, or delete data (GET, POST, PUT, or DELETE). When the client sends an HTTP request, the server will send a three-digit response [17] to that client with a header and possibly payload in the HTTP body.

Figure 2 shows the use case diagram of SIsKA-NG mobile, which is part of Unified Modeling Language (UML) used for software design [18, 19]. Three roles (actor) of user, i.e. student, lecturer, and administrator (the coordinator and secretary of the study program; and postgraduate school IT manager), each has different access right to 15 functions for interaction on SIsKA-NG mobile.

2.3 Implementation and Unit Testing

At this stage, the cross-platform development was chosen as the best solution than the native approach, in terms of time and cost, with the principle of "develop once and

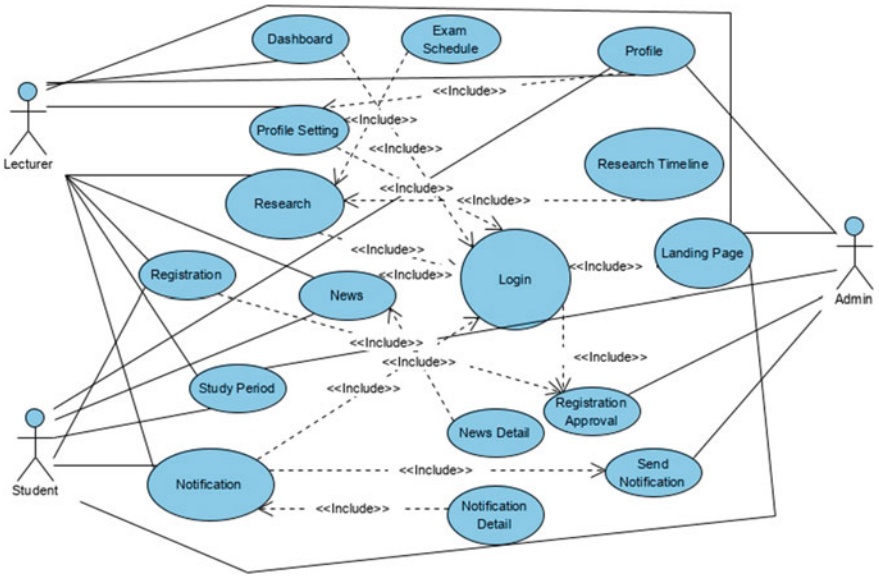


Fig. 2 The use case diagram of SIsKA-NG mobile

run anywhere" [20]. SIsKA-NG mobile client was implemented and tested by using programming language Dart based on the Flutter framework [21, 22] and supported by the PHP-based REST API web service as the back-end. Database MySQL was used by the existing SIsKA-NG web (see Fig. 1).

2.4 Integration and System Testing

At this stage, based on the requirements, all modules in the tree diagram of SIsKA-NG mobile (Fig. 3) were integrated and tested. The System Usability Scale (SUS) evaluation was used as part of the testing to find out direct user response since it has been widely used to determine the level of usability of applications on phones and tablets using the iOS and android platforms [23–25].

Testing involved the SUS evaluation. There are ten items of the SUS questionnaire [26] that were given to the respondents from the authors' postgraduate study program (covers all three roles described previously). The score ranges from 1 to 5 (from level "Strongly Disagree" to "Strongly Agree") were applied to those 10 items as follows: (1) "I think that I would like to use this system frequently"; (2) "I found the system unnecessarily complex"; (3) "I thought the system was easy to use"; (4) "I think that I would need the support of a technical person to be able to use this system"; (5) "I found the various functions in this system were well integrated"; (6) "I thought there was too much inconsistency in this system"; (7) "I would imagine that most

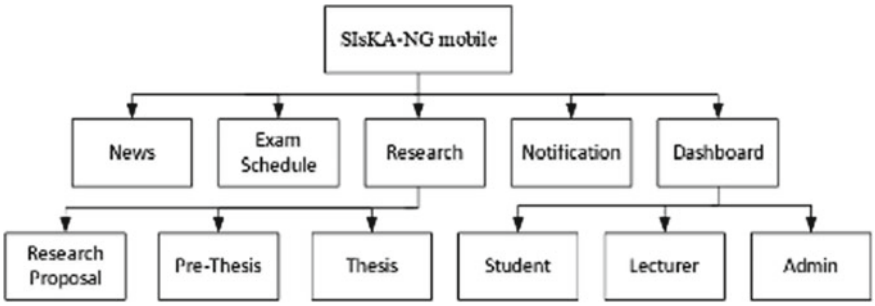


Fig. 3 Modules in the tree diagram of SIsKA-NG mobile

people would learn to use this system very quickly"; (8) "I found the system very cumbersome to use"; (9) "I felt very confident using the system"; and (10) "I needed to learn a lot of things before I could get going with this system."

3 Result and Analysis

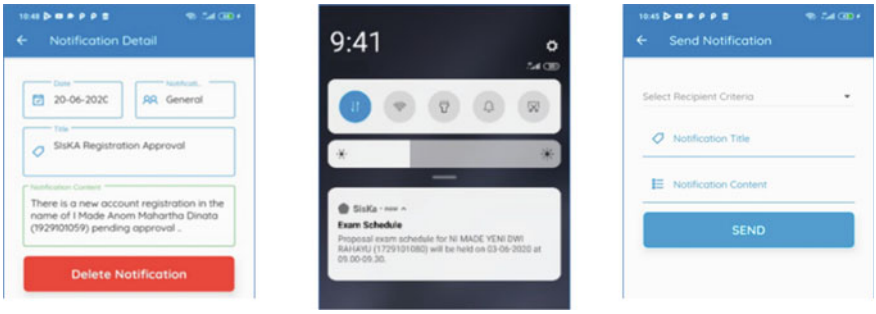
Several aspects related to SIsKA-NG mobile implementation and its SUS evaluation were analyzed in this section.

3.1 System Implementation

Figure 4a shows the initial page when SIsKA-NG mobile client is opened by the user. There is a login button to display the login page and a sign-up button to register

a) b) c)

Fig. 4 Pages of SIsKA-NG mobile: **a** sign up; **b** registration list; **c** login



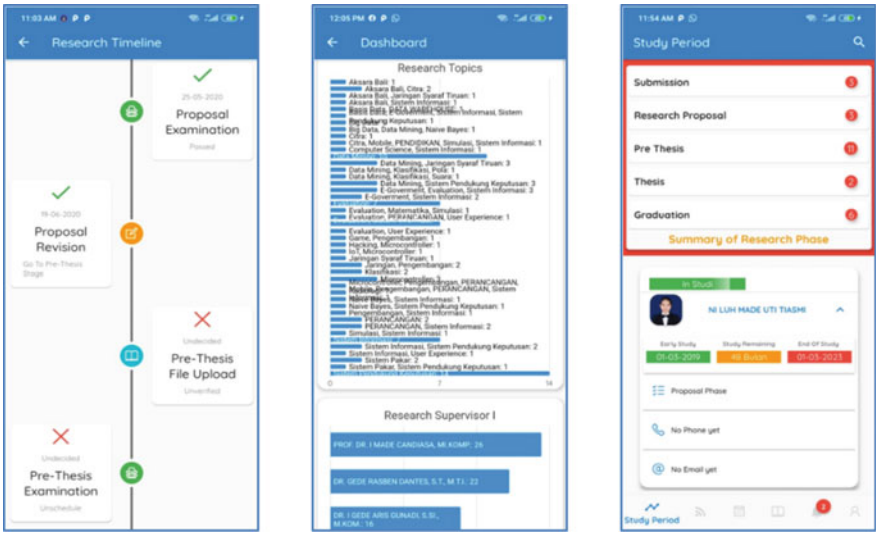
a) b) c)

Fig. 5 The real-time notification of SIsKA-NG mobile: **a** from the registration process; **b** from the examination schedule process; **c** manual notification user interface

where that registration can be approved or denied by the administrator (Fig. 4b). Figure 4c is the login page used by registered users to log into the application.

The most important automatic real-time notification feature was provided by SIsKA-NG mobile as a solution for the real-time delivery of information related to certain business processes, like account registration (Fig. 5a) and examination schedule (Fig. 5b). There is also notification management on the notification detail page that allows the user to delete the older notifications with the aim that the user can focus on the newer ones. For the testing purpose (testing FCM connectivity due to library or plugin code upgrade), manual real-time notification was provided and can be sent only by the administrator (Fig. 5c). An example for the automatic real-time notification of the examination scheduling business process is that the administrator can add the examination schedule on the SIsKA-NG web and then all of the related examinees (students) will receive the automatic real-time notification related to their examination schedule when the schedule is successfully created. In another example, if the students' study period ends in the end of the next semester, they will receive an automatic notification (a reminder) every week forward.

Figure 6a shows the research timeline page belonging to the student user where it will be opened first after login. It shows the stages that students must go through related to their thesis research, i.e., proposal submission, proposal examination, proposal revision, pre-thesis examination, pre-thesis revision, thesis examination, thesis revision, and finally the graduation. Figure 6b shows the dashboard page of the administrator that displays in the graph about certain information related to the student thesis research, i.e., the number of students, the number of researches per class, research stages per class, research topics, and the number of supervised students per lecturer. Figure 6c shows the dashboard page of the lecturers related to their supervised students. This page contains the number of students in the proposal, pre-thesis, and thesis stage, alongside with research timeline of each student. Hopefully, by this information, the lecturers can monitor easier the academic progress of their students related to their thesis research.



a) b) c)

Fig. 6 The pages of SIsKA-NG mobile: **a** student's research timeline; **b** administrator's dashboard; **c** lecturer's supervised students

3.2 System Testing

The testing involved the SUS questionnaire that was filled by the respondents. By using simple random sampling [27], ten respondents were selected from the authors' postgraduate study program with a validity of 75–80% [28], each from the student and the lecturer role. All three respondents from the admin role (study program coordinator, study program secretary, and postgraduate school IT admin) were also selected for the sake of evaluation comprehensiveness among all roles, even though did not reach the number of 6–14 respondents for the SUS [29].

The average SUS score of 82.25, 79.25, and 77.5 was obtained, each from the student, lecturer, and admin respondents. All of those scores are in grade A, A-, and B, respectively [24]. They are higher than a score of 68.00 [30] which means SIsKA-NG mobile was considered as a good usability product. Also, there is a pattern in those SUS scores where the lowest score came from the admin respondents. It has already been predicted since more business processes were involved that their complexities can reduce the satisfaction level.

4 Conclusion

SIsKA-NG mobile by using REST API has been successfully implemented without disrupting the operation of the existing SIsKA-NG web. Moreover, its integration with the existing SIsKA-NG web has added the most important real-time notification feature as a solution for the real-time delivery of information related to certain business processes, like account registration, examination registration, and examination schedule. It supports the quality improvement of thesis management service at the authors' postgraduate study program. Future work on security improvement would be on the HTTPS implementation which involves certain network configurations and certain application code adjustments.

Acknowledgements This work was supported by the Indonesian Ministry of Research and Technology through the Research Grant PTM No. 107/UN48.16/LT/2020.

References

1. Indrawan, G., Gunawan, I.M.A.O., Sariyasa, S.: The usability evaluation of academic progress information system (SIsKA-NG). Adv. Sci., Technol. Eng. Syst. J. **5**(2), 460–468 (2020)
2. Indrawan, G. et al.: SIsKA: Mobile based academic progress information system. In: 2nd International Conference on Innovative Research Across Disciplines, pp. 126–130 (2017).
3. Paramitha, A.A.I.I., Dantes, G.R., Indrawan, G.: the evaluation of web based academic progress information system using heuristic evaluation and user experience questionnaire (UEQ). In: 3rd International Conference on Informatics and Computing, pp. 1–6, (2018).
4. Sarwar, M., Soomro, T.R.: Impact of smart phones on society. Eur. J. Sci. Res. (2013)
5. Qian, K. et al.: Real-world relevant learning with android smartphones. In: 12th IEEE International Conference on Advanced Learning Technologies (2012).
6. Misbah, A., Ettalbi, A.: Towards a standard WSDL implementation of multiview web services. In: International Conference on Multimedia Computing and Systems (2017).
7. Belqasmi, F. et al.: SOAP-based versus RESTful web services: a case study for multimedia conferencing. IEEE Internet Computing (2012)
8. Janardanan, A. et al.: Android application for car wash services. In: International Conference on Emerging Trends and Innovations (2018).
9. Agocs, A., Goff, J.M.L.: A web service based on RESTful API and JSON Schema/JSON Meta Schema to construct knowledge graphs. In: International Conference on Computer, Information and Telecommunication Systems (2018)
10. Ma, S.P. et al.: Configurable RESTful service mashup: a process-data-widget approach. Appl. Math. Inf. Sci. (2015).
11. Vinothsaravanan, R., Palanisamy, C.: Extracting information from JSON database as simple as extracting in SQL using JSONiq. In: International Conference on Advances in Computing and Communication Engineering (2019).
12. Sanmorino, A., Fajri, R.M.: The design of notification system on android smartphone for academic announcement. Int. J. Interact. Mob. Technol. (2018)
13. Li, P. et al.: Implementation of cloud messaging system based on GCM service. In: International Conference on Computational and Information Sciences, (2013).
14. Moroney, L.: The Definitive Guide to Firebase. Apress, California (2017)

15. Mitchell, S.M., Seaman, C.B.: A comparison of software cost, duration, and quality for waterfall versus iterative and incremental development: a systematic review. In: 3rd International Symposium on Empirical Software Engineering and Measurement (2009).
16. Almarashdeh, I., Alsmadi, M.: Heuristic evaluation of mobile government portal services: an experts' review. In: 11th International Conference for Internet Technology and Secured Transactions (2017).
17. Arcuri, A.: RESTful API automated test case generation with Evomaster. ACM Trans. Softw. Eng. Methodol. (2019).
18. Petre, M.: UML in practice. In: International Conference on Software Engineering (2013).
19. Medvidovic, N. et al.: Modeling software architectures in the unified modeling language. ACM Trans. Softw. Eng. Methodol. (2002)
20. Latif, M. et al.: Cross platform approach for mobile application development: a survey. In: International Conference on Information Technology for Organizations Development (2016).
21. Mainkar, P., Giordano, S.: Google Flutter Mobile Development Quick Start Guide: Get Up and Running with iOS and Android Mobile App Development. Packt Publishing, Birmingham (2019)
22. Napoli, M.L.: Beginning Flutter: A Hands On Guide To App Development. Wiley (2019).
23. Sasmito, G.W., Zulfiqar, L.O.M., Nishom, M.: Usability testing based on system usability scale and net promoter score. In: 2nd International Seminar on Research of Information Technology and Intelligent Systems (2019).
24. Lewis, J.R.: The system usability scale: past, present, and future. Int. J. Hum.–Comput. Interact. (2018).
25. Pradini, R.S., Kriswibowo, R., Ramdani, F.: Usability evaluation on the SIPR website uses the system usability scale and net promoter score. In: 4th International Conference on Sustainable Information Engineering and Technology (2019).
26. Brooke, J.: SUS: a retrospective. J. Usability Stud. (2013)
27. Taherdoost, H.: Sampling methods in research methodology; how to choose a sampling technique for research. SSRN Electron. J. (2016)
28. Tullis, T.S., Stetson, J.N.: A comparison of questionnaires for assessing website usability. In: Usability Professional Association Conference (2004).
29. Katsanos, C., Tselios, N., Xenos, M.: Perceived usability evaluation of learning management systems: a first step towards standardization of the system usability scale in Greek. In: 16th Panhellenic Conference on Informatics, pp. 302–307 (2012).
30. Sharfina, Z., Santoso, H.B.: An Indonesian adaptation of the system usability scale (SUS). In: International Conference on Advanced Computer Science and Information Systems (2017).

Load Flow Analysis of 6 MW Biomass Power Plant

Dipesh B. Pardeshi, Anupama Deshpande, and Sachin V. Chaudhari

Abstract Financial specialists have begun moving toward inexhaustible wellsprings of vitality so as to fulfill the developing need for electrical force. The gathering of solar, wind vitality, and utilization of biomass for age of power diminishes our conditions on quick exhausting non-renewable energy sources. Biomass which is not just a reasonable vitality source, however, can likewise give natural manure. Biomass-based power age and dissemination framework is gradually picking up significance in creating nations like India because of the high accessibility of rural waste. The framework considered for this examination is a 6 MW biomass power plant on which the heap stream investigation was performed by utilizing the ETAP programming and applying Newton–Raphson technique for emphasis. The rundown of burden stream examination is assisted with deciding the estimations of voltages and force at various transports of the framework and recognizing their stacking condition.

Keywords Biomass power plant · Load flow analysis · Newton–Raphson method · ETAP

1 Introduction

Biomass-based force age is viewed as a significant aspect of the age and dispersion framework being arranged in India during the coming decade. This will not just increment the piece of the pie of sustainable power sources yet will likewise assist with decreasing CO_2 emanation because of the utilization of customary coal-based warm force creating stations [9]. It is assessed that vitality request in India will twofold by 2040 because of rising populace and modern interest. Because of restricted supplies of petroleum products, this vitality need can be satisfied by using the sustainable wellsprings of vitality (solar, wind, biomass, and so on). Among these

D. B. Pardeshi (✉) · A. Deshpande
Department of Electrical Engineering, Shri J.J.T. University, Jhanjhunu, Rajasthan, India

D. B. Pardeshi · S. V. Chaudhari
Sanjivani College of Engineering, Kopargaon, SPPU, Pune, Maharashtra, India

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_20

sources, biomass is reasonable to supplant petroleum products. For nations like India, which creates a lot of farming buildup, almost 200 Mt. reasonable biomass which is not just a vitality source, however, can likewise give natural manure, paper material, medications, and synthetics [4].

The utilization of horticultural waste to produce electrical force can make high possibilities for nations like India to satisfy their vitality needs by using the sustainable sources. By and by, these assets are not totally used, because of various difficulties. The structure of a biomass power plant is represented by the attributes of the fuel and is influenced by the expense and execution of the gear. Inappropriate structure may prompt unscheduled force stream which can make blockage in transmission organization and expands misfortunes in the transmission lines. To stay away from the weight of new venture for supplies and transmission lines, power stream or burden stream examines are significant. A proficient electric network is significant in the present current power market for dependable gracefully of intensity. In the vast majority of the force frameworks, the pace of disappointment increases because of sudden interest for power and unscheduled force stream which makes clog in transmission network along these lines expanding misfortunes in the transmission lines. For each 1% decrease in the receptive force, the electrical force framework encounters a 2% ruin in the responsive help, which further lessens voltage. As the transmission misfortunes begin rising, debilitating of receptive help to the framework makes the framework make progress toward responsive help, because of which the electrical influence framework encounters voltage breakdown and the framework arrives at a flimsy working point. To defeat these issues and to empower the purchasers to increase most extreme advantage of administrations, "Adaptable AC Transmission Systems" (FACTS) and "Disseminated age" innovation are utilized in the structure of new conveyance frameworks [15, 14].

To have better usage of existing gear and to maintain a strategic distance from the weight of new venture for supplies and transmission lines, it is important to control the responsive force so as to improve the nature of flexibly in an AC power framework [1]. It is anything but difficult to ascertain the receptive force devoured by the heap, yet it is hard to portray the responsive force produced or expended inside the organization, which is of significant concern. The motivation behind dispatching receptive force in a framework is to limit the genuine force transmission misfortunes [16]. The consistent-state complex voltages are created at all transports, through which dynamic and responsive influence streams in each transmission line and transformer of the organization can be determined utilizing the influence stream (load stream) examination [6, 8] With earlier data with respect to influence streams in the lines, it is conceivable to control the current frameworks and accomplish productive activity. Burden stream examination can likewise assist with making arrangements for new frameworks [5]. The point of this exploration is to play out the heap stream examination of a 6 MW biomass power plant utilized in this investigation and research the estimations of voltages at each transport of the organization notwithstanding the genuine and responsive force.

2 Modeling Approach

In a 3 Ø framework, the progression of dynamic and receptive force from the generators to the heaps happens through various organization transports and transmission lines which is named as the heap stream or force move through an organization. Burden stream study is a consistent state examination pointed toward deciding the estimations of voltages, current, and force (dynamic and receptive) moving through the framework under a given burden condition. Another significance of performing load stream or force stream study is on account of upkeep which assists with establishing that if a transmission line is controlled off for support or encounters a disappointment will the rest of the transmission lines in the framework have the option to deal with the additional heap without the requirement for expanding their evaluated qualities. Prior to the innovation of rapid PCs, load stream considerations were performed with the assistance of AC network analyzers. Notwithstanding, utilization of advanced frameworks has brought about more noteworthy adaptability and snappier activity [2].

Demonstrating of intensity plant frameworks encourages dynamic for venture designers, financers, and strategy creators just as gives vigorous examination capacities to vitality scientists. SAM (framework advisor model) is one such kind of intensity age displaying framework. SAM can be utilized to demonstrate framework-associated power frameworks utilizing biomass-, wind-, or sunlight-based vitality. In light of client determined sources of info, the SAM approach computes the framework's physical setup, its area, cost of establishment, and force age. It likewise causes estimations for private and business purchasers. Though SAM can be utilized to display biomass strong energized power age plant of any size and any blend, client needs to take care about the feedstock input and part cost which must be scaled by the framework being demonstrated [7].

Electrical transient analyzer program (ETAP) then again is a serious load flow module. Client can make and approve framework model effortlessly and acquire exact and solid outcomes. Implicit highlights like programmed gadget assessment, rundown alerts/admonitions, result analyzer, and keen illustrations make it the most productive load flow program accessible today. ETAP figures transport voltages, branch power elements, flows, and force streams all through the electrical framework. It considers swing, voltage directed, and unregulated force sources with various force frameworks and generator associations. It is fit for performing investigation on both spiral and circle frameworks. ETAP permits client to choose from a few unique techniques so as to accomplish the best figuring effectiveness and precision (ETAP specialized manual).

A few strategies have been created in business application programming bundles dependent on Guass–Seidel strategy, Guass disposal strategy, fast decoupled strategy, and Newton–Raphson (NR) strategy for examination of burden stream in a framework; these strategies include the fathoming of a lot of nonlinear mathematical conditions utilizing iterative methods. The dependability of the NR strategy for load flow

arrangement is relatively better than other burden stream techniques, as dissimilar to different strategies it can tackle cases that lead to difference [10, 12].

Aside from the utilization of programming's and cycle methods to display a force framework, a few calculations have likewise been attempted to play out the force stream investigation of intensity frameworks. A methodology joining efficient parallel genetic algorithm (EPGA) for ideal dynamic force arranging (OAPP) and straightforward fluffy rationale master rules for receptive force control (RPC) can be found in the writings to test and show the possibility of enormous force frameworks [9]. Joining strategies with creating advancements assists with increasing snappier comprehension of activities with the framework and play out its convenient investigation.

3 Load Flow Analysis of 6 MW Biomass Power Plant

Figure 1 shows the single line outline of a 6 MW biomass power plant situated in Maharashtra, India. The current framework comprises of 37 transports joining 35 burdens and three transformers. ETAP programming was utilized, and Newton–Raphson strategy was applied to play out the heap stream investigation. A sum

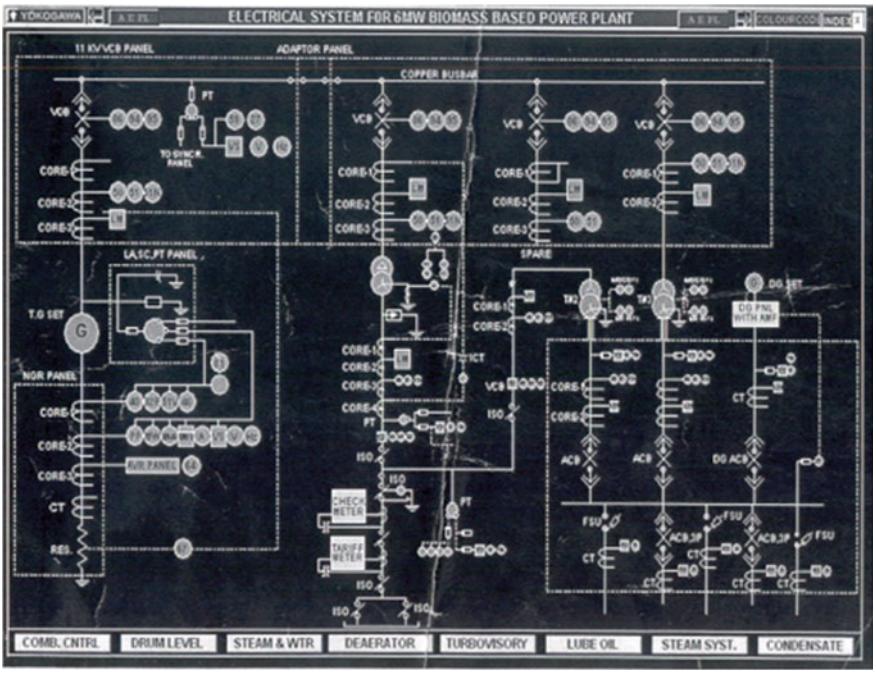


Fig. 1 Single line diagram of 6 MW biomass power plant

of 99 emphases were performed, and the aftereffects of the test can be found in the bordering sheets where the product plays out the computations for voltages, dynamic, and receptive force present at transports and transformers of the organization. The stacked states of various components of the organization are likewise created in the ETAP report. The created report, for example, burden stream examination, assists with dissecting the stacking state of the components in a framework. As can be found in the ETAP report, an outline of over-burden and under stacked components of the framework can support the vitality and force engineers/analysts to take choice about how to all the more likely use framework assets to their greatest limit if there should be an occurrence of under stacked components and how to secure gadgets for over-burden conditions.

4 Conclusion

Age of electrical force through horticultural waste has a potential in creating nations like India to fulfill the need of intensity necessity. The framework viable is a 6 MW biomass power plant situated in Maharashtra, India. Burden stream test was completed on the framework to examine the progression of dynamic and receptive force through different transports of the framework. ETAP programming was utilized to perform emphases on the framework components by utilizing Newton–Raphson technique. Basically over-burden and under stacked components of the framework were recognized through the rundown report produced by the ETAP programming. Age of intensity by the utilization of biomass will not just assistance to use horticultural waste yet in addition help to lessen our conditions on petroleum products for age of power.

Project: Biomass	ETAP 16.0.0C	Page: 1
Location: 6MW BPP		Date: 04-10-2020
Contract:		SN: 4359168
Engineer: D B Pardeshi	Study Case: LF	Revision: Base
Filename: PSII		Config.: Normal

Electrical Transient Analyzer Program

Load Flow Analysis

Loading Category (1): Design
Generation Category (1): Design
Load Diversity Factor: None

	Swing	V-Control	Load	Total
Number of Buses:	1	1	35	37

	XFMR2	XFMR3	Reactor	Line/Cable	Impedance	Tie PD	Total
Number of Branches:	3	0	0	34	0	0	37

Method of Solution:	Adaptive Newton-Raphson Method
Maximum No. of Iteration:	99
Precision of Solution:	0.0001000
System Frequency:	50.00 Hz
Unit System:	Metric
Project Filename:	PSII
Output Filename:	C:\ETAP 1600\PSII\Untitled.lfr

LOAD FLOW REPORT

Bus		Voltage		Generation		Load		Load Flow					XFMR
ID	kV	Ang	deg	MW	Mvar	MW	Mvar	ID	MW	Mvar	Amp	%PF	%Tap
ACTIVE POWER	0.433	94.648	-2.8	0	0	0	0	MAIN LV BUS	0.000	0.000	0.0	0.0	
BOILER AUX SPARE	0.433	94.648	-2.8	0	0	0	0	MAIN LV BUS	0.000	0.000	0.0	0.0	
* BUS-0002	11.000	100.000	0.0	0.738	0.566	0	0	VCB PANEL	0.738	0.566	48.8	79.4	
BUS-0003	11.000	99.991	0.0	0	0	0	0	VCB PANEL	-0.725	-0.568	48.3	78.7	
								BUS-0004	0.725	0.568	48.3	78.7	
BUS-0004	33.000	97.416	-1.4	0	0	0	0	BUS-0005	0.720	0.536	16.1	80.2	
								BUS-0003	-0.720	-0.536	16.1	80.2	
BUS-0005	33.000	97.414	-1.4	0	0	0	0	BUS-0004	-0.720	-0.536	16.1	80.2	
								RING MAIN UNIT	0.000	0.000	0.0	0.0	
								BUS-0013	0.720	0.536	16.1	80.2	
BUS-0006	11.000	99.996	0.0	0	0	0	0	VCB PANEL	-0.013	0.002	0.7	-98.7	
								BUS-0079	0.013	-0.002	0.7	-98.7	
BUS-0007	33.000	97.414	-1.4	0	0	0	0	RING MAIN UNIT	0.000	0.000	0.0	0.0	
								BUS-0009	0.000	0.000	0.0	0.0	
BUS-0008	33.000	97.414	-1.4	0	0	0	0	RING MAIN UNIT	0.000	0.000	0.0	0.0	
								BUS-0010	0.000	0.000	0.0	0.0	
BUS-0009	33.000	97.414	-1.4	0	0	0	0	BUS-0007	0.000	0.000	0.0	0.0	
BUS-0010	33.000	97.414	-1.4	0	0	0	0	BUS-0008	0.000	0.000	0.0	0.0	
BUS-0013	33.000	97.411	-1.4	0	0	0	0	BUS-0014	0.720	0.536	16.1	80.2	
								BUS-0005	-0.720	-0.536	16.1	80.2	
BUS-0014	33.000	97.410	-1.4	0	0	0	0	BUS-0013	-0.720	-0.536	16.1	80.2	
								BUS-0015	0.720	0.536	16.1	80.2	
BUS-0015	0.433	94.872	-2.9	0	0	0	0	MAIN LV BUS	0.714	0.504	1226.2	81.7	
								BUS-0014	-0.714	-0.504	1226.2	81.7	
BUS-0018	0.433	92.696	-2.1	0	0	0.060	0.045	TG AUX SPARE	-0.060	-0.045	103.1	80.0	
BUS-0019	0.433	93.379	-2.4	0	0	0.016	0.012	TG AUX SPARE	-0.016	-0.012	28.6	80.0	
BUS-0020	0.433	93.208	-2.3	0	0	0.022	0.016	TG AUX SPARE	-0.022	-0.016	38.7	80.0	
BUS-0021	0.433	88.697	-1.5	0	0	0.120	0.090	COOLING TOWER	-0.120	-0.090	226.0	80.0	
BUS-0022	0.433	88.697	-1.5	0	0	0.060	0.045	COOLING TOWER	-0.060	-0.045	113.0	80.0	
BUS-0023	0.433	88.697	-1.5	0	0	0.060	0.045	COOLING TOWER	-0.060	-0.045	113.0	80.0	
BUS-0024	0.433	89.146	-1.7	0	0	0.024	0.018	COOLING TOWER	-0.024	-0.018	45.0	80.0	
BUS-0025	0.433	88.825	-1.5	0	0	0.034	0.026	COOLING TOWER	-0.034	-0.026	63.9	80.0	
BUS-0026	0.433	93.139	-2.4	0	0	0.192	0.144	FUEL AND ASH HANDLING SUN	-0.192	-0.144	344.4	80.0	
BUS-0027	0.433	93.139	-2.4	0	0	0.096	0.072	FUEL AND ASH HANDLING SUN	-0.096	-0.072	172.2	80.0	
BUS-0028	0.433	93.000	-2.3	0	0	0.060	0.045	FUEL AND ASH HANDLING SUN	-0.060	-0.045	107.8	80.0	

Alert Summary Report

	%Alert Settings	
	Critical	Marginal
Loading		
Bus	100.0	95.0
Cable	100.0	95.0
Reactor	100.0	95.0
Line	100.0	95.0
Transformer	100.0	95.0
Panel	100.0	95.0
Protective Device	100.0	95.0
Generator	100.0	95.0
Inverter/Charger	100.0	95.0
Bus Voltage		
OverVoltage	105.0	102.0
UnderVoltage	95.0	98.0
Generator Excitation		
OverExcited (Q Max.)	100.0	95.0
UnderExcited (Q Min.)	100.0	

Critical Report

Device ID	Type	Condition	Rating/Limit	Unit	Operating	% Operating	Phase Type
ACDB,DCDB	Bus	Under Voltage	0.433	kV	0.410	94.6	3-Phase
BOILER AUX SPARE	Bus	Under Voltage	0.433	kV	0.41	94.6	3-Phase
BUS-0015	Bus	Under Voltage	0.433	kV	0.41	94.9	3-Phase
BUS-0018	Bus	Under Voltage	0.433	kV	0.40	92.7	3-Phase
BUS-0019	Bus	Under Voltage	0.433	kV	0.40	93.4	3-Phase
BUS-0020	Bus	Under Voltage	0.433	kV	0.40	93.2	3-Phase
BUS-0021	Bus	Under Voltage	0.433	kV	0.38	88.7	3-Phase
BUS-0022	Bus	Under Voltage	0.433	kV	0.38	88.7	3-Phase
BUS-0023	Bus	Under Voltage	0.433	kV	0.38	88.7	3-Phase
BUS-0024	Bus	Under Voltage	0.433	kV	0.39	89.1	3-Phase
BUS-0025	Bus	Under Voltage	0.433	kV	0.38	88.8	3-Phase
BUS-0026	Bus	Under Voltage	0.433	kV	0.40	93.1	3-Phase
BUS-0027	Bus	Under Voltage	0.433	kV	0.40	93.1	3-Phase
BUS-0028	Bus	Under Voltage	0.433	kV	0.40	93.0	3-Phase
BUS-0029	Bus	Under Voltage	0.433	kV	0.40	92.9	3-Phase

Critical Report

Device ID	Type	Condition	Rating/Limit	Unit	Operating	% Operating	Phase Type
BUS-0030	Bus	Under Voltage	0.433	kV	0.406	93.7	3-Phase
BUS-0031	Bus	Under Voltage	0.433	kV	0.40	93.2	3-Phase
BUS-0036	Bus	Under Voltage	0.433	kV	0.41	94.7	3-Phase
BUS-0039	Bus	Under Voltage	0.433	kV	0.41	94.9	3-Phase
COOLING TOWER	Bus	Under Voltage	0.433	kV	0.39	89.9	3-Phase
FUEL AND ASH HANDLING SPA	Bus	Under Voltage	0.433	kV	0.41	94.2	3-Phase
GEN-0002	Generator	Over Excited	0.182	Mvar	0.18	100.0	3-Phase
MAIN LV BUS	Bus	Under Voltage	0.433	kV	0.41	94.6	3-Phase
T3	Transformer	Overload	0.006	MVA	0.01	216.6	3-Phase
TG AUX SPARE	Bus	Under Voltage	0.433	kV	0.41	93.9	3-Phase

Marginal Report

Device ID	Type	Condition	Rating/Limit	Unit	Operating	% Operating	Phase Type
BUS-0004	Bus	Under Voltage	33.000	kV	32.147	97.4	3-Phase
BUS-0005	Bus	Under Voltage	33.000	kV	32.15	97.4	3-Phase
BUS-0007	Bus	Under Voltage	33.000	kV	32.15	97.4	3-Phase
BUS-0008	Bus	Under Voltage	33.000	kV	32.15	97.4	3-Phase
BUS-0009	Bus	Under Voltage	33.000	kV	32.15	97.4	3-Phase
BUS-0010	Bus	Under Voltage	33.000	kV	32.15	97.4	3-Phase
BUS-0013	Bus	Under Voltage	33.000	kV	32.15	97.4	3-Phase
BUS-0014	Bus	Under Voltage	33.000	kV	32.15	97.4	3-Phase
RING MAIN UNIT	Bus	Under Voltage	33.000	kV	32.15	97.4	3-Phase

SUMMARY OF TOTAL GENERATION , LOADING & DEMAND

	MW	Mvar	MVA	% PF
Source (Swing Buses):	0.738	0.566	0.930	79.37 Lagging
Source (Non-Swing Buses):	0.212	0.182	0.279	75.91 Lagging
Total Demand:	0.950	0.747	1.209	78.59 Lagging
Total Motor Load:	0.896	0.672	1.121	80.00 Lagging
Total Static Load:	0.000	0.000	0.000	
Total Constant I Load:	0.000	0.000	0.000	
Total Generic Load:	0.000	0.000	0.000	
Apparent Losses:	0.053	0.075		
System Mismatch:	0.000	0.000		

Number of Iterations: 3

References

1. Abdulkareem, A., Awosope, C.O.A., Orovwode, H.E., Adelakun, A.A.: Power flow analysis of Abule-Egba 33 kV distribution grid system with real network simulations. IOSR J. Electr. Electron. Eng. **6**(2), 67–80 (2014)
2. Dubey, A.: Load flow analysis of power systems. Int. J. Sci. Eng. Res. **7**(5), 79–84 (2016)
3. ETAP Technical Manual.
4. Gojiya, A., Deb, D.: On feasibility of biomass power plant with agricultural waste processing. In: IEEE (2017)
5. Hemanth Kumar, M.H., Vijayshree, G., Prakash, R., Shivsharannappa, G.C.: Load flow analysis of distribution generation system using Ieee 30 bus system. IJEEE **4** (2016)
6. Ikule, F.T., Ame-Oko, A., Idoko, E.: Load flow analysis using Newton Raphson method—a case study of South-West Nigeria 330 KV network. IJARIIE **5** (2019).
7. Jorgenson, J., Gilman, P., Dobos, A.: Technical Manual for the SAM Biomass Power Generation Model. National Renewable Energy Laboratory, Colorado (2011)
8. Kusekwa, M.A.: Load flow solution of the Tanzanian power network using Newton-Raphson method and MATLAB software. Int. J. Energy Power Eng. **3**(6), 277–286 (2014).
9. Mahdad, B., Srairi, K., Bouktir, T.: Optimal power flow for large-scale power system with shunt FACTS using efficient parallel GA. Int. J. Electr. Power Energy Syst. **32**(5), 507–517 (2008)
10. Mander, D.K., Virdi, G.S.: Result analysis on load flow by using Newton Raphson method. Int. J. Adv. Res. Electr. Electron. Instru. Eng. **6**(7), 5835–5844 (2017)
11. Pardeshi, D., Deshpande, A.: Possible modeling of small biomass power plant (June 26, 2020) available at SSRN: https://ssrn.com/abstract=3683050
12. Salgar, S.A., Mallareddy, C.H.: Load flow analysis for a 220 kV line-case study. Int. J. Eng. Res. Technol. **2**(5), 1–12 (2015)
13. Suhas Shirbavikar, S., Ashok, M., Babu Narayanan, M.: Dynamic analysis of renewable energy systems and their impact on smart grid. In: 16th National Power Systems Conference (2010)

14. Suresh, V., Sreejith, S.: Power flow analysis incorporating renewable energy sources and FACTS devices. Int. J. Renew. Energy Res. **7**(1) (2017)
15. Venkateswara Rao, B., Nagesh Kumar, G.V.: Optimal power flow by BAT search algorithm for generation reallocation with unified power flow controller. Int. J. Electr. Power Energy Syst. **68**, 81–88 (2015)
16. Youcef, A., Mohamed, B., Amine, L.A., Elmaouhab, A.: Optimal VAR control for real power loss minimization using differential evolution algorithm. Electr. Power Energy Syst. **66**, 262–271 (2015)

Enhancement of Performance in the Disk-Shaped Dual Band Microstrip Patch Antenna for the Advanced IoT Based 5G Applications

John Colaco, Anand Lokapure, and R. B. Lohani

Abstract In this paper, the authors have enhanced the performance in the various characteristics of disk-shaped microstrip patch antenna very much useful for the latest fifth-generation (5G) wireless communication and its advanced Internet of Things-based applications. The whole purpose of this research work is for designing and analyzing the disk-shaped dual-band patch microstrip antenna at the resonance frequency of 28.5 GHz useful for advanced IoT-based 5G applications. In this research work, authors have used Rogers RT/Duroid 5880 substrate material of dielectric constant $\varepsilon_r = 2.2$, the dielectric loss tangent of 0.0011, and thickness of the substrate of 0.72 mm. After simulating the design, the authors have obtained a good return loss of −21.5 dB at 28.5 GHz and −11.20 dB at 38.2 GHz, good bandwidth of 1.93 GHz, voltage standing wave ratio (VSWR) in the range $1 \leq \text{VSWR} \leq 2$, and high gain of 10 dBi.

Keywords IoT · Antenna · Microstrip · Circular patch · 5G · Return loss · Antenna radiation · Bandwidth · Dielectric substrate

1 Introduction

With the rise in advanced IoT-based applications and technologies, the need for high data rates has become a highly essential requirement in the field of wireless communications. This need is provided by 5G which is the advanced unifying Internet connectivity for future wireless fifth-generation system. It uses multiple input multiple output antenna for the device-to-device communications and beamwidth multiple

J. Colaco (✉) · R. B. Lohani
Department of Electronics & Telecommunication Engineering, Goa Engineering College, Ponda, Goa, India

R. B. Lohani
e-mail: rblohani@gec.ac.in

A. Lokapure
CSIR, National Institute of Oceanography, Panaji, Goa, India

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_21

access technologies [1]. Also, with the rise in demand for e-learning such as online lectures, online conferences, various broadcast services, and smart homes in this currently ongoing COVID-19 lockdown situation, the need for high-speed Internet has become an important factor. With the use of the now pathetic 4G communication, multiple problems are being faced by many professionals, students, and smartphone users concerning for streaming and downloading ultra-high definition videos. Therefore, seeing the problems of now 4G, we authors have developed a microstrip patch antenna at millimeter-wave bands to meet the need for 5G applications. This millimeter-wave band is the recent advanced technology in the field of wireless communication. However, it has a few challenges like high signal losses [2]. Using FEKO software, the authors have designed at a resonance frequency of 28.5 GHz a circular microstrip patch antenna for high efficiency, better bandwidth, and high gain. For highly advanced 5G applications like media and virtual experience, massive connectivity machine-type communication services types have been identified [3, 4]. Antennas have been found the most important devices in the field of advanced wireless communications [5]. The microstrip patch antenna is a metallic strip or patch placed on a dielectric substrate that can be designed for different mathematical shapes like triangular, elliptical, square, etc. The dielectric constants in a range of $2.2 \leq \boldsymbol{\varepsilon_r} \leq 12$ are available in the substrates used in microstrip patch antenna [6]. However, a thicker substrate originates at discontinuities by lowering permittivity [5]. In this proposed design, the advantage of using RT/Duroid 5880 substrate is that it gives low dielectric loss good for high-frequency and broadband applications [7]. The overall purpose of this research work is to design a microstrip patch antenna with microstrip transmission line feed to provide high bandwidth with good gain and antenna efficiency applicable for various advanced IoT-based 5G applications and analyze various electrical characteristics at 28.5 GHz frequency using FEKO.

The detailed survey has been conducted on 5G communications and compared with 2G, 3G, and 4G [8]. The survey is conducted on the Internet of Things technologies applicable to 5G networks [9]. The proposed square microstrip patch antenna is designed for the single band at 28 GHz applicable in 5G communications which are giving a high gain of 10.3 dB and high efficiency of 97% and high a bandwidth of 4.7 GHz [10]. The proposed circular patch antenna with coaxial feed and transmission line feed is designed and compared at 28 GHz giving a bandwidth of 0.7 GHz and 0.6 GHz, respectively [11]. Dual-band frequency compact microstrip patch antenna is presented for 5G communication [12].

2 Theory and Design

2.1 Theory

Figure 1 indicates the general view of the microstrip disk (circular) patch antenna design having a dielectric substrate plane with a disk-shaped patch, ground plane,

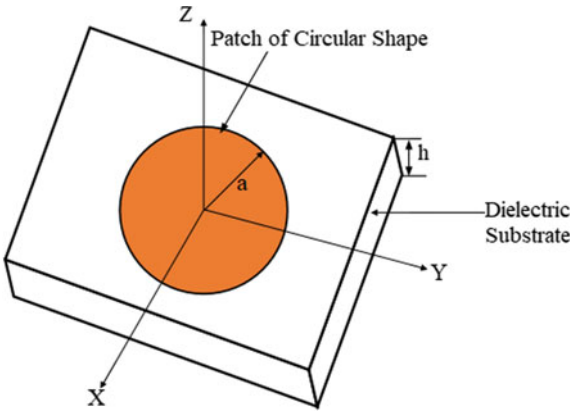


Fig. 1 General view of microstrip circular patch antenna

and microstrip feed line [6]. The patch radius is denoted by a. The high conducting materials, such as copper or gold, are used for the patch. This disk (circular) patch geometry of the microstrip patch antenna is the most preferred because of circular in shape as well as lesser in size as compared to the size of the other shapes like rectangular patch[13]. The radius of the circular patch denoted by a is determined by using the equations given as [6]

$$a = \frac{F}{\sqrt{\left\{1 + \frac{2}{\pi \varepsilon_r F}\left[\ln\left(\frac{\pi F}{2h}\right) + 1.7726\right]\right\}}} \tag{1}$$

where

$$F = \frac{8.791 \times 10^9}{f_r \sqrt{\varepsilon_r}} \tag{2}$$

In the above expressions, f_r is the resonant frequency measured in GHz, ε_r is the dielectric constant or relative permittivity of the substrate material, F is the logarithmic function, and the thickness of the dielectric substrate is denoted by h [6].

2.2 Design

In this design, we have used RT/Duroid 5880 substrate material with dielectric constant $\varepsilon_r = 2.2$, substrate height = 0.6 mm, and dielectric loss tangent = 0.0011. The resonant frequency of 28.5 GHz is used considering the 5G frequency spectrum range 26–40 GHz. Using the above-said design values, authors have calculated from Eqs. (1) and (2) the radius of the patch to approximately 2.0 mm. For obtaining dual band, the patch is cut into two parallel rectangular slots. Figure 2 is showing the

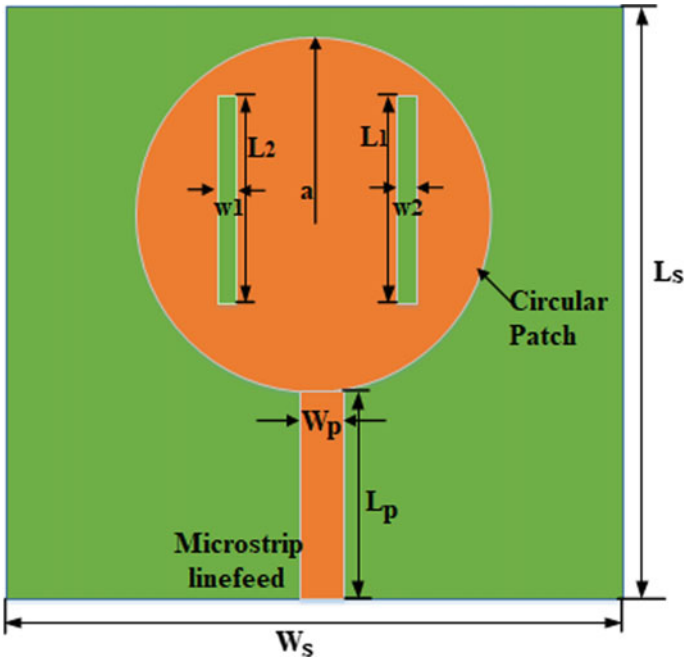


Fig. 2 Geometrical top view construction of proposed circular patch microstrip antenna

geometry of the two-dimensional top view of microstrip circular patch antenna. The other parameters used for the design are given in Table 1.

The values of design parameters calculated are given in Table 1.

Table 1 Antenna parameters used for design

Parameter symbols	Details of parameters	Value (mm)
W_s	Substrate width	10
L_s	Substrate length	10
W_f	Microstrip feed line width	0.72
L_f	Microstrip feed line length	1.9
W_1	Slot 1 width	1.9
L_1	Slot 1 length	0.18
W_2	Slot 2 width	1.9
L_2	Slot 2 length	0.18
a	Circular patch radius	1.94

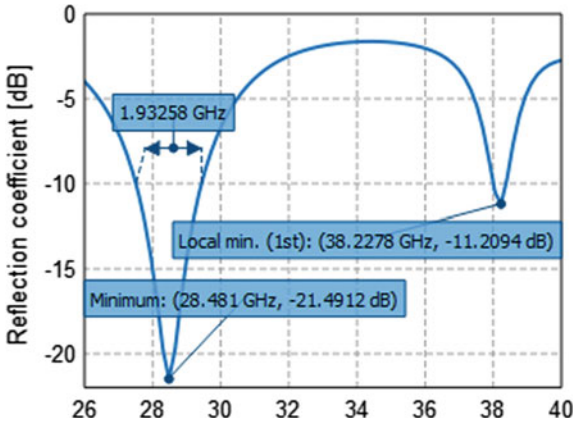


Fig. 3 Graph of *S*-parameter (return loss)

3 Simulation and Result Analysis

3.1 Return Loss (S-parameter)

The return loss, in other words the reflection coefficient for the microstrip antenna for the advanced IoT-based 5G applications, should not be smaller than −10 dB as it should have a return loss of higher than −15 dB for enhanced performance of an antenna [10]. In Fig. 3, the proposed dual-band antenna has a reflection coefficient of −21.5 dB and improved bandwidth of 1.93 GHz at resonance frequency 28.5 GHz which will help advanced IoT-based 5G applications. Also, the proposed antenna has better return loss and better bandwidth than [14].

3.2 Voltage Standing Wave Ratio

For the best requirements in the advanced IoT-based 5G applications, the value standing wave ratio should lie in the range $1 \leq \text{VSWR} \leq 2$ [11]. As seen in Fig. 4, the value of the standing wave ratio for the proposed antenna has 1.2 which is satisfying the requirement.

3.3 3D Radiation Pattern

Figure 5 shows a 3D pattern of radiation specifying the energy radiation of an antenna with a good high gain of 10 dBi. This high gain represents the good strength of the signal needed for a long range of advanced wireless communication like 5G at very high data rates [4]. The proposed antenna has a comparatively higher gain than [15].

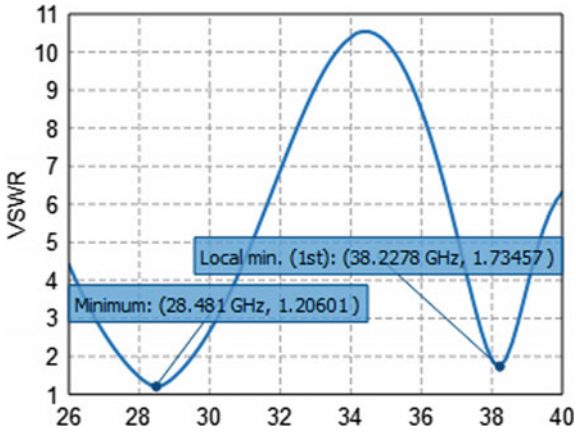


Fig. 4 Graph of standing wave (voltage) ratio

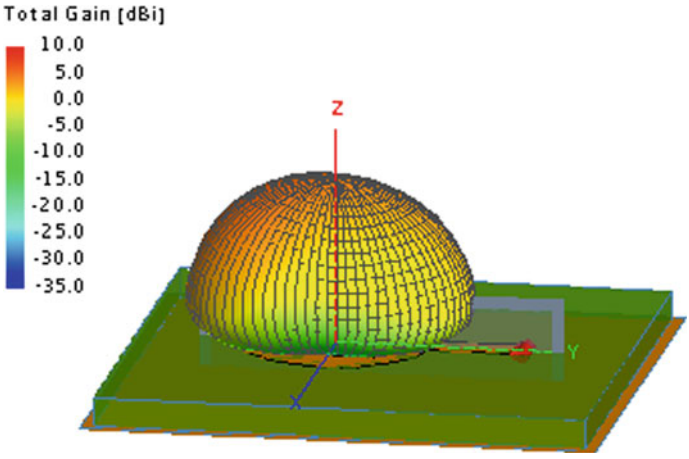


Fig. 5 Graph of antenna gain with 3D radiation pattern

3.4 Half-Power Beamwidth with 2D Radiation Pattern

Beamwidth is the separation of the angle indicating the width of the radiated beam in degrees. Figure 6 shows the graph of 2D radiation pattern, and angle of separation for half-power beamwidth for the proposed antenna is 106.479° which is good for IoT-based 5G applications and is more than [4].

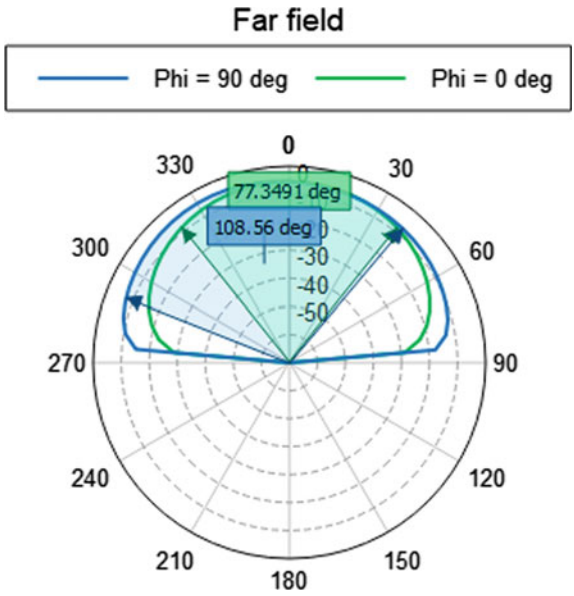


Fig. 6 Graph showing half-power beamwidth for phi = 0° and phi = 90° with 2D radiation pattern

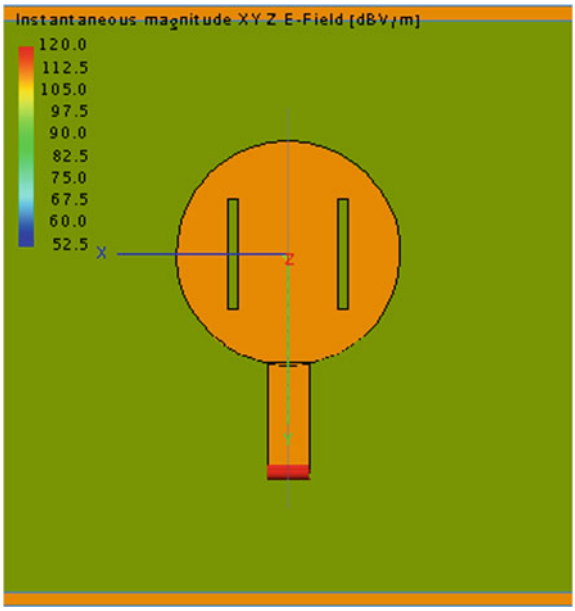


Fig. 7 Graph of the near field

3.5 Electric Field Intensity on Antenna

From Fig. 7, it is observed that the electrical field becomes zero at the center of the patch, maximum at one side, and minimum at the other side. The maximum and minimum electric field changes their side instantaneously according to the phase of applied signal, and fields extend to the outer periphery giving rise to fringing electric fields and cause radiation on the patch useful for IoT-based 5G applications [16]. This is because the distribution of electric field intensity is in its fundamental mode (cavity theory) as transverse magnetic field distribution (TM10) mode, where the variation of electric field in the *Z*-direction is considered negligible along the *Z*-axis (perpendicular direction), in the direction of *Y* is also negligible along *Y*-axis (direction of impedance width) and in *X*-direction is one minimum to the maximum along *X*-axis (direction of resonance length) [16].

4 Conclusion and Future Research Work

Using FEKO software, the proposed microstrip circular patch antenna has been designed, and the same is implemented. From our simulation and results analysis, it is to conclude that the designed antenna has the good gain of 10 dBi for long-range of wireless communication, enhanced bandwidth of 1.93 GHz for IoT enabled advanced 5G applications, return loss of –21.5 dB, required voltage standing wave ratio of 1.2, high beam-width for good signal strength and good electric field intensity on the patch for high radiation. However, the limitations of this proposed antenna are lesser bandwidth and return loss. The future work is to obtain better bandwidth and gain by configuring antenna with metamaterials structures as these structures are being artificially engineering to produce unusual physical electromagnetic characteristics that are so far not found in nature such as storing and reradiating energy. Also, design using arrays and other shaped patches can be implemented and compared for performance in terms of various characteristics.

References

1. https://www.connectzone.in/what-is-5g.php.
2. Zhang, J., Ge, X., Li, Q., Guizani, M., Zhang, Y.: 5G millimeterwave antenna array: design and challenges. IEEE Wirel. Commun. (2017)
3. SK Telecom, S.: Telecom's 5G Architecture Design and Implementation Guideline (2015)
4. Colaco, J., Lohani, R.: Design and implementation of microstrip circular patch antenna for 5g applications. In: 2nd International Conference on Electrical Engineering. IEEE, Turkey (2020)
5. Thaher, R.H.: Single and multiband UWB circular patch antenna for wireless communication applications. Am. J. Electromagn. Appl. **3**(3), 16–23 (2015), https://doi.org/10.11648/j.ajea.20150303
6. Balanis, C.A.: Antenna Theory: Analysis and Design (2005).

7. https://rogerscorp.com/advanced-connectivity-solutions/rt-duroid-laminates/rtduroid-5880-laminates.
8. Gupta, A., Kumar, R.: A survey of 5G network: architecture and emerging technologies, recent advances in software defined networking for 5G networks. IEEE Access, pp. 1206–1232 (2015).
9. Akapakwu, G.A., Silva, B.J., Hancke, G.P., Abumahfauz, A.M.: A survey on 5G networks for Internet of things: communications technologies and challenges. IEEE Access **6** (2017)
10. Faisal, M., Gafur, A., Rashid, S., Z., Shawon, M.O.F., et al.: Return loss and gain improvement for 5g wireless communication based on single band microstrip square patch antenna. In: 1st tinternational Conference on Advances in Science, Engineering and Robotics Technology. IEEE, Dhaka, Bangladesh (2019).
11. Kumar, S., Kumar, A.: Design of circular patch antennas for 5G applications. In: Second International Conference on Innovations in Electronics, Signal Processing and Communications. IEEE, Shillog, India (2019).
12. Zhang, Y., Zheng, H., Gao, B., Tang, C., Liu, R., Wang, M.: A compact dual-band antenna for 5G application. In: Cross Strait Quad-Regional Radio Science and Wireless Technology Conference (CSQRWC). IEEE, Taiyuan, China (2019).
13. Goyal, R.K., Modani, U.S.: A compact microstrip patch antenna at 28 GHz for 5G wireless applications. In: 3rd International Conference and Workshops on Recent Advances and Innovations in Engineering. IEEE, Jaipur, India (2018)
14. Noor, M.U.S., Khattak, M.I.: Design and analysis of dual-band microstrip patch antenna array for 5G cellular communication networks with improved radiation characteristics. In: 2nd International Conference on Latest Trends in Electrical Engineering and Computing Technologies. IEEE, Karachi, Pakistan (2019).
15. Gaid, A.S.A., Qaid, O.A.S., Ameer, M.A.A., Qaid, F.F.M., Ahmed, B.S.A.: Small and bandwidth efficient multi-band microstrip patch antennas for future 5G communications. In: Saeed F., Mohammed F., Gazem N. (eds) Emerging Trends in Intelligent Computing and Informatics. (IRICT). Advances in Intelligent Systems and Computing, vol. 1073. Springer, Cham. Malayasia (2019).
16. Kumar, V.M., Sujith, Enhancement of bandwidth and gain of rectangular microstrip patch antenna. Thesis NIT Rourkela

A Survey on Predicting Player's Performance and Team Recommendation in Game of Cricket Using Machine Learning

Milon Biswas, Tajim Md. Niamat Ullah Akhund, Md. Kawsher Mahbub, Sikder Md. Saiful Islam, Sadia Sorna, and M. Shamim Kaiser

Abstract Machine learning in sports analytics is a hot field in computer science. Using machine learning algorithms, we can predict the outcome of a game or performance of teams or individual players and building new strategies for upcoming competitions. Cricket is one of the foremost popular games in the world. Choosing the right player is one of the most challenging work for all kinds of sport and no exception in cricket. In the field of machine learning, several algorithms are used for prediction and classifications. Machine learning algorithms like linear regression, support vector machine, random forest, and naive Bayes with linear and polynomial kernel showed good results to predict the runs scored by a batsman and runs given by a bowler. In this work, we explored the techniques that have been applied to solve the challenges in cricket.

Keywords Machine learning · Team recommendation · Cricket · Survey · Players · Performance

1 Introduction

Cricket is very popular game in the world, and Cricket World Cup 2019 was one of the most watched sporting events in the world. People of every cricket playing country are very concerned about the national cricket squad before any match or any series. Various factors inherit play while selecting a team. In the field of machine learning,

M. Biswas (✉) · Md. K. Mahbub · S. Md. Saiful Islam · S. Sorna
Department of Computer Science and Engineering, Bangladesh University of Business and Technology, Dhaka, Bangladesh
e-mail: mskaiser@juniv.edu

T. Md. Niamat Ullah Akhund
Department of Computer Science and Engineering, Daffodil International University, Dhaka, Bangladesh

M. Shamim Kaiser
Institute of Information Technology, Jahangirnagar University, Dhaka, Bangladesh

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_22

several algorithms like linear regression, random forest, support vector machine, and naive Bayes are used for prediction and classifications. Feature selection may be a vital part of a machine learning project. There are several features for a player selection like experience, matches played against an opponent, highest score, strike rate, total number of 100's, total number of the '50s, total number of runs scored by the player, average runs scored, number of innings played, ball faced per matches, habitat or away, specific format specialist, players ratings, number of 4's per match, number of 6's per match, number of runs scored by the player against the opponent, and fielding performance. During this research, we will attempt to analyze every feature to form a player selection.

Though Ha-Du-Du is the national sport of Bangladesh but cricket is the most popular game in Bangladesh. Most of the citizens of this country are crazy about cricket, and they keep their eyes to the player's selection committee. So, whether he/she is a broadcaster, player or die-hard fan, one would want to know, who is going to play the next games in advance. That is why, people want to do a prediction about cricket. In our country, cricket team selection processes are dependent on humans. Moreover, human judgment cannot be relied upon because sometimes, they made mistakes. Players selection for a team in cricket is vital to its success. If a team wants to win a game, they must choose their best players for the game. Sometimes, selection committee struggles with the selection procedure, and they face lots of confusion before the match about players selection. For example, which player to pick or drop, which batsman or bowler should play in which position, and so on. So, Bangladesh cricket council (BCB) needs an automated system for team combination using machine learning. Furthermore, the Bangladesh cricket team does not need to fully dependent on human judgment any more for the team recommendation. This system will help the Bangladesh cricket council (BCB) to know all player's fitness, strength, weakness, measures player performance, and rank the players also enable them to decide for player's selection.

2 Literature Review

Authors of [19] proposed a two-phase framework for player selection and team formation. Md. Muhaimenur Rahman et al. [15] they analysis of Bangladesh One Day International (ODI) cricket data using machine learning and showed the importance of some features, and they have found the predicted outcome. They collected seventeen features and analyzed those features with several machine learning algorithms like naive Bayes, KNN, random forest, SVM, decision tree and predicted the outcome. They got their best result using the latest version of the decision tree algorithm that is C5.0 and successfully got the accuracy of 63.63% before starting the game, 72.72% for first innings and 81.81% for second innings. Passi et al. [14] applied some machine learning algorithms. They set some parameters used some equations to generate players rating to evaluate players performance. They analyzed the player's characteristics and stats and predicted the batsman's and bowlers' per-

formance separately in ODI matches using supervised machine learning. They got 90.74% accuracy for run and 92.25% for predicting wickets taken by a bowler in random forest classifier, and it is the most accurate classifier also achieved an accuracy of 51.45% for run predicting and 68.78% for wicket predicting. To predicate a player's performance, Minhazul Abedin et al. [1] analyzed supervised classification models as random forest, K-nearest neighbors (KNN), support vector machine (SVM), and decision tree. Random forest predicts the winning percentage with the highest accuracy of 92.61% when the amount of training dataset is 90%, and KNN predicts the least accuracy of 73.73% when the amount of training dataset is 90%. Kalpdrum Passi et al. [13] analyzed different characteristics, abilities, and statistics using four machine learning algorithms where random forest got the highest accuracy. To predicate team selection, Chetan Kapadiya et al. [12] analyzed different machine classifiers like naive Bayes, decision tree, random forest, SVM, and weighted random forest. Weighted random forest predicts the highest accuracy of the percentage that is 93.73%. Naive Bayes predicts 58.12% of accuracy, decision tree predicts 86.50% of accuracy, random forest predicts 92.25% of accuracy, and SVM predicts 68.78% of accuracy. S. R. Iyer et al. [9] used neural networks to prediction of an athlete's performance in cricket team selection, and they got maximum 86% accuracy in batman's selection and maximum 80% accuracy in blower selection with neural networks. Manuka Maduranga et al. [8] analyzed different types of approaches to predict a match outcome in the cricket domain. They using historical cricket data and collective knowledge approach. Here, collective knowledge approach predicts the highest accuracy of 87%. Sandesh Bananki et al. [10] analyzed the SVM model with linear and nonlinear poly and RBF kernels. Here, SVM with RFB kernel predicts the accuracy of 75%, precision of 83.5% and recall rate of 62.5%. Linear kernel predicts the accuracy of 70.83%, and poly kernel predicts the accuracy of 68.75%. Sankalp Rane et al. [18] used logistic regression, support vector, random forest, decision tree where random forest gave the best result. Madan Gopal et al. [11] analyzed different machine classifiers like SVM, random forests, logistic regression, decision trees, and KNN. In today's day-to-day life, machine learning is used everywhere like medical [4], robotics [3], remote sensing [6], management of disabled people [2], EVM [17], Security [16], farming [7], agriculture, and IoT [5]. That is why if we can use it in game, it may result good.

3 Dataset Understanding

The models for the selection of an optimal team contained some features. The parameters feature which affects a player's performance given following. The main factors are batting, bowling, and all-rounder attributes. In batting attributes, features like who batted first, batting position, no. of innings played, runs, strike rate, centuries, the fifties, zeros, batting average, and so on. In bowling attributes, the considered features are who balled first, bowling average, bowling strike rate, four/five wicket haul, and no. of innings. A part of the dataset is shown in Fig. 1. Again, in all-rounder

	Player	Span	Mat	Runs	HS	Bat Av	100	Wkts	BBI	Bowl Av	5	Ct	St	Ave Diff
2	SR Tendulkar (INDIA)	1989-2013	664	34357	248*	48.52	100	201	5/32	46.53	2	256	0	1.98
3	DPMD Jayawardene (Asia/SL)	1997-2015	652	25957	374	39.15	54	14	2/32	62.92	0	440	0	-23.77
4	KC Sangakkara (Asia/ICC/SL)	2000-2015	594	28016	319	46.77	63	0	-	-	0	609	139	-
5	ST Jayasuriya (Asia/SL)	1989-2011	586	21032	340	34.14	42	440	6/29	35.66	6	205	0	-1.52
6	RT Ponting (AUS/ICC)	1995-2012	560	27483	257	45.95	71	8	1/0	47.50	0	364	0	-1.54
7	MS Dhoni (Asia/INDIA)	2004-2019	538	17266	224	44.96	16	1	1/14	98.00	0	634	195	-53.03
8	Shahid Afridi (Asia/ICC/PAK)	1996-2018	524	11196	156	23.92	11	541	7/12	32.78	10	167	0	-8.86
9	JH Kallis (Afr/ICC/SA)	1995-2014	519	25534	224	49.1	62	577	6/54	32.14	7	338	0	16.95
10	R Dravid (Asia/ICC/INDIA)	1996-2012	509	24208	270	45.41	48	5	2/43	41.80	0	406	14	3.61
11	Inzamam-ul-Haq (Asia/ICC/PAK)	1991-2007	499	20580	329	43.32	35	3	1/0	24.00	0	194	0	19.32
12	TM Dilshan (SL)	1999-2016	497	17671	193	38.16	39	154	4/4	44.16	0	242	3	-6.00
13	M Muralitharan (Asia/ICC/SL)	1992-2011	495	1936	67	9.26	0	1347	9/51	22.86	77	203	0	-13.60
14	SR Waugh (AUS)	1985-2004	493	18496	200	41.65	35	287	5/28	35.56	3	223	0	6.09
15	MV Boucher (Afr/ICC/SA)	1997-2012	467	10469	147*	29	6	1	1/6	6.00	0	953	46	23.00
16	CH Gayle (ICC/WI)	1999-2019	462	19321	333	38.79	42	257	5/34	36.66	3	235	0	2.13
17	Wasim Akram (PAK)	1984-2003	460	6615	257*	18.73	3	916	7/119	23.57	31	132	0	-4.83

Fig. 1 A part of Dataset

attributes, features like batting hand, bowling hand, the strength of the opposition, matches played, wickets, runs scored, strike rate, and average. There are some others factors like ground, venue, opponent, day/night effect, economy, pitch, and weather are being considered.

4 Methodology

In every machine, learning-based work data is the most valuable part, and we can say that it's work is the heart of the whole process. So, collecting applicable data from a trust able source is the most important duty. They collected data from espncrickinfo and crickbuzz which are the most familiar and reliable source of the cricket statistical data. After collecting the all data extract and store necessary data from the collected historical data. Select a model and train the machine test the result, and finally, they got a result. In this work, we have analyzed the accuracy level of prediction of various machine learning algorithm in cricket. We have considered KNN, SVM, naive Bayes, decision tree, random forest, and weighted random forest machine learning algorithms in this work. The flow of work is mentioned in Fig. 2.

5 Results Analysis

For prediction, in all of the existing systems, the most used algorithms are naive Bayes, random forest, SVM, decision tree, and weighted random forest. Most of the researcher used random forest algorithm, naive Bayes. But there are very few papers where the weighted random forest has been used and got better accuracy than random forest. Consider those data they applied different kinds of algorithms like SVM, random forest, KNN, naive Bayesian, decision tree, and so on. Here is some

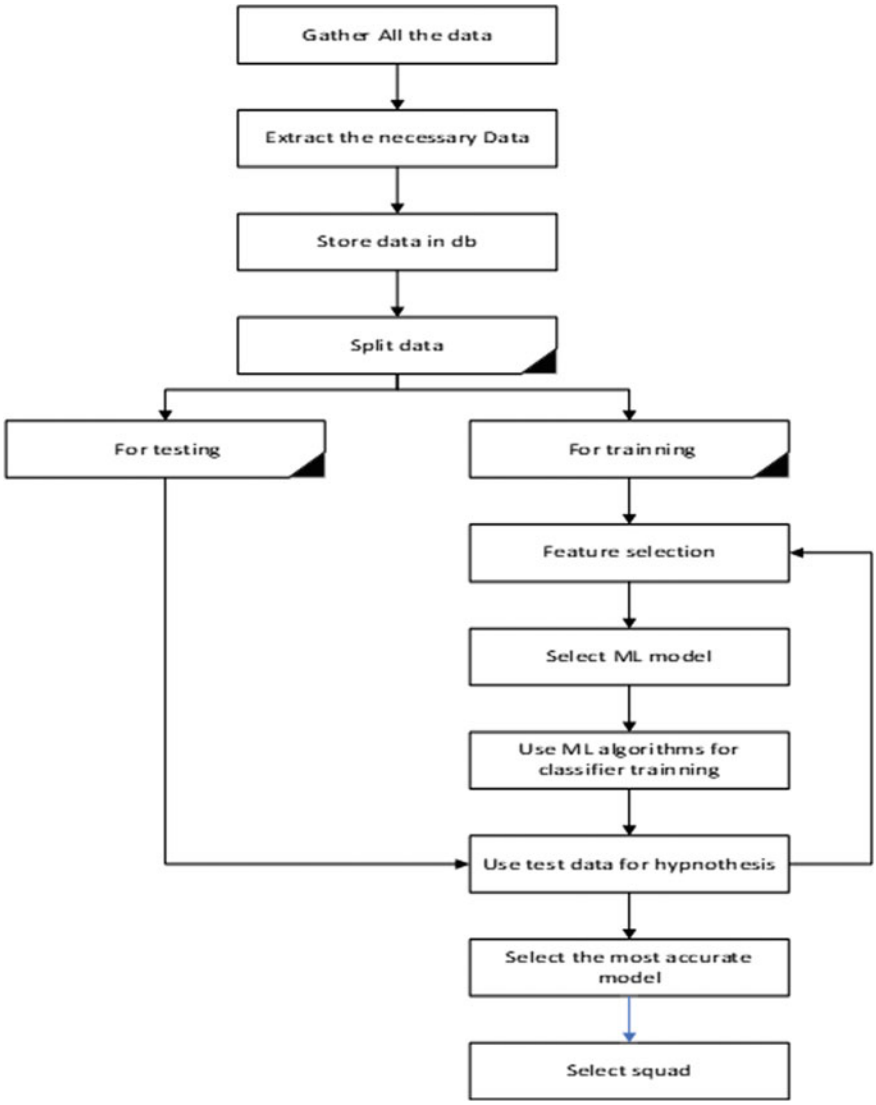


Fig. 2 Flow of work

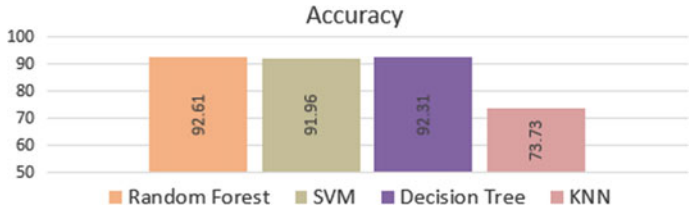


Fig. 3 Accuracy between RF, SVM, DT, and KNN machine learning algorithms

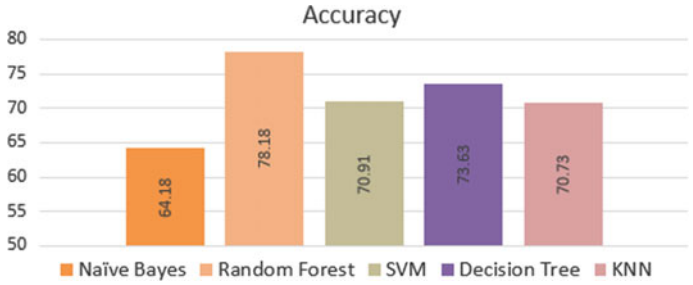


Fig. 4 Accuracy between naive Bayes, RF, SVM, decision tree, and KNN

Fig. 5 Accuracy between naive Bayes, RF, SVM, decision tree, and weighted RF machine learning algorithms

output that analysis using those previous data. From Fig. 3, Muhaimenur Rahman et al. [15] used different types of algorithm, and we see that random forest provides highest accuracy, and naïve Bayes provides lowest accuracy.

From Fig. 4, Minhazul Abedin et al. [1] analyzed, supervized classification models, and got the highest accuracy using random forest which is 92.61% and Naïve Bayes provided the lowest accuracy that is 64.18%.

From Fig. 5, Chetan Kapadiya et al. [6], they applied feature extraction to extract the most useful features and got an accuracy of 92.25% and the least accuracy of 58.12%.

Here, the above-mentioned model or algorithm considered batting, balling and all rounder and compared to the other algorithm weighted random forest algorithm provided the highest accuracy of the result that is 93.73% and SVM provided the least accuracy which is not suitable for this type of prediction.

6 Conclusion

The selection of players for a sports team is a complex task. Find a balance cricket team for is hard to choose, but performance of player can show way for environment condition, pitch condition scored by players, easy to choose in team to play. All of researcher considered various factors and tried to find the player's strengths and weaknesses against the opponent team. All of the work helped to know all player's fitness, strength, weakness, measures player performance, and rank the players also enable them to decide for player's selection. Based on the strength and weaknesses, they recommend a team combination. As a result, a cricket team does not need to fully dependent on human judgment any more for the team recommendation.

References

1. Abedin, M.M., Urmi, S.R., Mozumder, M.T.I., Samiur, M., Rahman, A.F.: Forecasting the outcome of the next odi cricket matches to be played
2. Akhund, T.M.N.U.: Designing and implementation of a low-cost wireless gesture controlled robot for disable people
3. Akhund, T.M.N.U.: Study and implementation of multi-purpose iot nurse-bot
4. Akhund, T.M.N.U., Mahi, M.J.N., Tanvir, A.H., Mahmud, M., Kaiser, M.S.: Adeptness: Alzheimer's disease patient management system using pervasive sensors-early prototype and preliminary results. In: International Conference on Brain Informatics. Springer, Heidelberg, pp. 413–422 (2018)
5. Akhund, T.M.N.U., Rahman, M.H.: Bat banisher: An approach to create a high frequency ultrasound system to protect agricultural field from bats
6. Akhund, T.M.N.U., Sagar, I.A., Sarker, M.M.: Remote temperature sensing line following robot with bluetooth data sending capability
7. Hasan, M.A., Akhund, T.M.N.U.: An approach to create iot based automated smart farming system for paddy cultivation
8. Hatharasinghe, M.M., Poravi, G.: Data mining and machine learning in cricket match outcome prediction: missing links. In: 2019 IEEE 5th International Conference for Convergence in Technology (I2CT). IEEE, pp. 1–4 (2019)
9. Iyer, S.R., Sharda, R.: Prediction of athletes performance using neural networks: an application in cricket team selection. Expert Syst. Appl. **36**(3), 5510–5522 (2009)
10. Jayanth, S.B., Anthony, A., Abhilasha, G., Shaik, N., Srinivasa, G.: A team recommendation system and outcome prediction for the game of cricket. J. Sports Anal. **4**(4), 263–273 (2018)
11. Jhanwar, M.G., Pudi, V.: Predicting the outcome of odi cricket matches: a team composition based approach. In: MLSA@ PKDD/ECML (2016)
12. Kapadiya, C., Shah, A., Adhvaryu, K., Barot, P.: Intelligent cricket team selection by predicting individual players' performance using efficient machine learning technique

13. Passi, K., Pandey, N.: Predicting players'performance in one day international cricket matches using machine learning. In: Computer Science and Information Technology (CS & IT) (2017)
14. Passi, K., Pandey, N.: Increased prediction accuracy in the game of cricket using machine learning. arXiv preprint arXiv:1804.04226 (2018)
15. Rahman, M.M., Shamim, M.O.F., Ismail, S.: An analysis of Bangladesh one day international cricket data: a machine learning approach. In: 2018 International Conference on Innovations in Science, Engineering and Technology (ICISET). IEEE, pp. 190–194 (2018)
16. Sarker, M.M., Akhund, T.M.N.U.: The roadmap to the electronic voting system development: a literature review. Int. J. Adv. Eng. Manage. Sci. **2**(5), 239465
17. Sarker, M.M., Shah, M.A.I., Akhund, T.M.N.U., Uddin, M.S.: An approach of automated electronic voting management system for Bangladesh using biometric fingerprint. Int. J. Adv. Eng. Res. Sci. **3**(11), 236907 (2016)
18. Shetty, M., Rane, S., Pandita, C., Salvi, S.: Machine learning-based selection of optimal sports team based on the players performance. In: 2020 5th International Conference on Communication and Electronics Systems (ICCES). IEEE, pp. 1267–1272 (2020)
19. Tavana, M., Azizi, F., Azizi, F., Behzadian, M.: A fuzzy inference system with application to player selection and team formation in multi-player sports. Sport Manage. Rev. **16**(1), 97–110 (2013)

Skip Salp Swam Algorithm for Feature Selection

N. B. Arunekumar and K. Suresh Joseph

Abstract Feature selection is a prominent component of the machine learning pipeline where the features that may deteriorate the performance of the learning algorithm are curtailed retaining the valuable features alone. The selection algorithm's predominant role of sampling the least subset of features that aid the algorithm to achieve high accuracy becomes hectic as the number of features increases. This challenge of sampling the subset of features from the given set is accomplished using the swarm intelligence algorithms. The skip salp swarms algorithm inspired by the skip connections of neural networks is proposed to overcome the deficiency in the exploration and the exploitation of the SSA. The proposed algorithm is tested over 22 datasets from the standard repositories which possess varying dimensions and properties. The proposed algorithm is tested against 6 predominant meta-heuristic algorithms and it is observed that the proposed algorithm outperforms all the other algorithms over 86% of the datasets.

Keywords Wrapper feature selection · K-Nearest Neighbor · Machine learning · Salp swarm algorithm · Computational intelligence · Skip connections

1 Introduction

The dimensions or features of data to be analyzed for classification or pattern recognition are enormously huge. The feature selection is an essential pre-processing step to eliminate the redundant or irrelevant features that may have adverse effects on the classification. A small increase in number of features will adversely increase the complexity of the feature selection algorithm. With N number of features, the total probable subsets are combinations of features. Being a NP- complete problem feature selection cannot be accomplished using simple heuristic methods like brute force. The feature selection algorithms employ various modalities to retain the best

N. B. Arunekumar (✉) · K. S. Joseph
Pondicherry University, Puducherry, India

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_23

features alone. The datasets experimented within this work range from small dimensions with minimum dimensions of 19 to large datasets that range to dimensions of 11,340. All the datasets used are downloaded from the standard machine learning repository of UCI [1] and ASU for feature selection. The vital contributions of this paper include.

- Introducing an encircle parameter to increase exploration and exploitation behavior
- Using skip connections to propagate information on food to all the members
- Using both neighbors positions for updating the follower's position in the chain.
- Applying the proposed SSSA for the feature selection problems.

2 Related Works

Feature selection algorithms can be clustered under three different groups in general which are filter wrapper and embedded. The filter algorithm picks the features with respect to the relationship among the data itself. The CBF [2] measures the correlation among the features and ranks according to it. The other algorithms are mutual information, ReliefF, F-score, and Laplacian Score [3]. Second variant is the wrapper algorithm which uses the feedback from the learning algorithm and could retain high accuracy due to the effect of utilization of feature as a single entity and mining co-existence of features. No free lunch theorem [4] also states that no single algorithm would be a universal solution to solve all the problems. For the conversion from continuous to binary space in wrapper algorithms methods like naïve binary threshold method with respect to the particle position, transfer function method constituting S-shaped and the V-shaped transfer functions maps can be used. A total of 8 transfer functions with 4 on each type were utilized on PSO [5]. The other recent binary algorithms are Binary butterfly algorithm [6], binary ant lion algorithm [7], binary grey wolf algorithm. The transfer function maps the values to the same probability of flipping at all the stages overcoming it time variant transfer functions were introduced over dragonfly [8] algorithm. The bioinspired algorithms are a subset of meta-heuristic algorithms that mimic the patterns found in nature. Predominantly used evolution-based algorithms are genetic algorithm, ant colony optimization, swarm-based algorithms are the particle swarm optimization [9], bat algorithm [10],cuckoo search [11], grey wolf optimizer [12], grasshopper optimization algorithm [13], Whale optimization algorithm [14], which perform astoundingly good in nature because of the cooperative behavior they depict. The bioinspired algorithms have been promising than the gradient-based algorithms which are slow and often has the possibility of getting trapped with the local optimum on searching a non-convex search area. In the original paper, its clearly depicted that the SSA [15] has outperformed several other meta–heuristics algorithms in both large and small-scale problems.

3 Preliminaries

The salp swarm algorithm [15] depicts the chaining and the motion in sea salps. The salps are translucent creatures that feed around the organic matter in the sea and float around in chains. This behavior where the chain members or followers attach themselves as a long chain is portrayed in the SSA. In the first iteration the salps are initialized randomly. Each salp position is given with the index i and dimension j. The head of the salp chain which is the first salp particle repositions itself in accordance to the food position in each dimension with respect to Eq. 1 where is the food position in the dimension j, and are random numbers between range [0, 1], and are the upper and the lower bounds respectively. The parameter is the control parameter that switches between exploration and the exploitation. The binary variants proposed based on the salp swarm algorithm are binary salp swam algorithm [16] which utilizes transfer functions and crossover scheme and improved salp swarm algorithm [17] which introduces an inertial parameter that induces the exploratory behavior. The bSSA investigated the utilization of all the 8 transfer functions and the crossover. The crossover operator [18] is one among the prime method utilized in the genetic algorithms which helps the offspring to adsorb the traits from the parents. bSSA used a single point crossover which was analogues to the follower reposition. The bSSA algorithm is used for the comparison against the proposed algorithm. The knn algorithm is non-parametric and works on the basis of classifying a sample with respect to the class of its neighbors with count of k using Euclidean space as metric

$$s_j^i = \begin{cases} F_j + z * \left(\left(ub_j - lb_j\right) * r_1 + lb_j\right) r_2 \geq 0.5 \\ F_j - z * \left(\left(ub_j - lb_j\right) * r_1 + lb_j\right) r_2 < 0.5 \end{cases} \tag{1}$$

$$s_j^i = \frac{1}{2}\left(s_j^i + s_j^{i-1}\right) \quad i \geq 2 \tag{2}$$

4 Proposed Methodology

In the original salp swarm algorithm, the whole chain follows a single head and repositions itself restricting propagation of information about the food's position till the end of the chain and thus detaining the exploitation around the original food source. The other major setback is the lack of exploration in the chain's followers. By just getting half a way from its original position the salps may follow the course of its predecessor and premature-clustering due to the blinded chaining effect is seen.

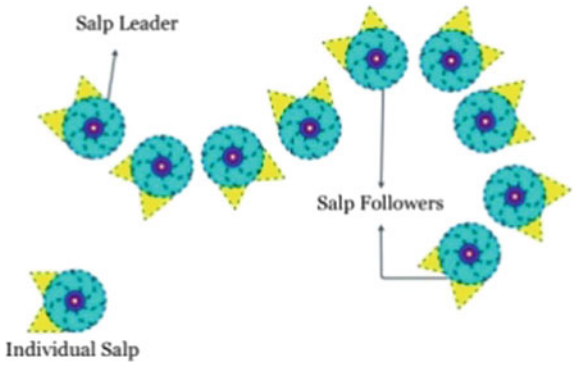


Fig. 1 Salp chain

4.1 Encircling Based Head Position Update

See Fig. 1.

The lack of exploration which is evident in the salps is averted using the encircling parameter introduced in GWO added to both the head position update and follower position updating equations. The modified encircling parameter A is given by Eq. 3 with parameter z and a random variable between 0 and 1. The new modified head position update equation with the encircling parameters for all *i*th salps with is given in Eq. 4

$$A = \frac{2 * z * r_3 - z}{2} + 1 \tag{3}$$

$$s_j^i = \begin{cases} A * F_j + z * \left(\left(ub_j - lb_j \right) * r_1 + lb_j \right) & r_2 \geq 0.5 \\ A * F_j - z * \left(\left(ub_j - lb_j \right) * r_1 + lb_j \right) & r_2 < 0.5 \end{cases} \tag{4}$$

4.2 Skip Connections

To avert the fault in information dissemination about food position till the end of chain the Skip connections which are inspired from the neural networks are being introduced. These skip connections drive back gradients to the inner layers during the back propagation in the neural networks. Analogous to it, Instead of using a single connection between the food position to chain members through head, skip connections from the food position to inner salps is also enabled. In this work, all the odd-numbered salps are updated using the new modified head position Eq. 4 and the even number are updated with the modified follower position with the embedded encircling parameter and dual neighbor guidance update equation as in Eq. 5. which calculates the centroid between predecessor and the successor. The distance between

```
Random initialize all salps s_j^i within the bound
While (current iteration < = Maximum iterations)
  Evaluate each salp's fitness
  Food  F = salp with best fitness
  Compute  z
      for each i  in salp swarm s_i
       if  (i %2 ==1 )

                    Utilize Eqn.4 to update the head position

       end if
    end for
    for each i  in salp swarm s_i
       if  (i  == last salp )
          Utilize Eqn.2 to update its position.
       else if ( i %2 ==0 && i !=  last salp )
        Calculate centroid ,velocity and update follower positions using Eqn.5
       end if
     end for

     Reposition salps which are away from the bound
end
   Return F
```

Fig. 2 SSSA algorithm pseudo code

the centroid and the salp is calculated as the velocity based on which the new position is derived. The overall algorithm of the proposed skip salp swarm algorithm is given in Fig. 2

$$\left.\begin{aligned} \text{cent}_j^i &= \tfrac{1}{2}\left(A * s_j^{i-1} + A * s_j^{i+1}\right) \\ \text{vel}_j^i &= \text{abs}\left(\text{cent}_j^i + s_j^i\right) \\ s_j^i &= \text{centroid}_j^i - \text{vel}_j^i \end{aligned}\right\} \quad i\%2 == 0 \tag{5}$$

5 Results and Discussion

The feature selection problem holds two objectives that are extremely contradictive in nature where the accuracy must be high and the number of feature chosen must be less as in Eq. 6. The () are the hyper parameters of weightage for the error rate and the no. of feature selected. The parameter denotes the number of features selected by the current solution and denotes the total number of features. The weights that are inherited from the literature are and. The feature vector mask holds only the binary values which is

threshold with respect to the value 0.5 The datasets with different dimensionality will be useful to thoroughly examine the capability of algorithm. In the KNN algorithm k is given as 5 and 10 swarm agents were considered. The performance of the algorithms is accessed using the metrics such as the mean accuracy, mean fitness, and the mean of the number of features selected. The mean of each metric is calculated from 30 independent runs. The total dataset is split into 80% training and 20% testing. On each independent run, a new set of testing and training data are sampled and the same set is subjected to each algorithm to ensure fair comparison among the algorithms. The algorithms which are considered for comparison are given the same hyper parameters as specified in literature.

$$\text{Fitness} = \varphi * \delta(E) + \alpha * \frac{|k|}{|K|} \tag{6}$$

From Table 1 and normalized metrics comparison in Fig. 3 it is evident that out of the total 22 datasets, features chosen by SSSA have been fruitful over 19 i.e. 86% percent of datasets over existing algorithms.

The SSSA has achieved least number of features in 10 of the 22 datasets. Considering individually as the least it has achieved over 63% on ALO, 86% over PSO, 91% over GOA, and 100% over all the other algorithms. Even though some of the algorithms have least number of features they failed to achieve the accuracy attained by the SSSA. Evidently, those algorithms would be suffering from premature convergence or will possess the inability to avoid local optima. The next aspect to be taken care of is the stability of the algorithm. This can be measured with the least standard deviation of the accuracy over the 30 runs. From the Fig. 4 it's clearly visible that the SSSA has the least deviation over the other algorithms in most of the datasets and thus could be termed more stable than the existing algorithms which are compared.

6 Conclusion

In this work, a skip connection based SSA algorithm Skip Salp Swam Algorithm SSSA has been proposed which is inspired from the skip connection used in deep learning models to backpropagate the error in to the initial layers were the gradients get lost. Along with the skip connections the other enhancements of introducing encircling parameter and chaining using both the neighbor instead of the predecessor alone was also proposed. The proposed algorithm has been tested with 22 different datasets collected from standard repositories. The observations are, the algorithm unlike the ALO which gets stuck in local minima by choosing least number of features but with low accuracy as in Fig. 3 SSSA avoids local minima and attains the global minima required by the fitness function. The future enhancements of this work is to experiment the algorithm for various engineering problems and using transfer functions that can still enhance the algorithm for feature selection problems.

Table 1 Comparison of Average Accuracy, Fitness and No.of dimensions over 30 runs

Dataset		Wine	Clean1	Vechicle	Heart	Ionosphere	Arcene	Sonar	Breastew	Soybean	Libras	Spambase	Waveform
Total	**Dim**	**13**	**167**	**18**	**13**	**34**	**10,000**	**60**	**29**	**60**	**90**	**57**	**40**
KNN	Acc	0.697	0.869	0.523	0.661	0.841	0.834	0.787	0.783	0.953	0.764	0.801	0.809
bGWO	Dim	7	111	9	7	18	5623	35	20	12	54	44	32
	Fit	0.045	0.062	0.281	0.166	0.094	0.115	0.110	0.158	0.004	0.172	0.079	0.169
	Acc	0.960	0.944	0.721	0.838	0.910	0.889	0.894	0.847	1.000	0.832	0.929	0.838
PSO	Dim	4	**70**	6	5	13	4653	22	13	8	36	**28**	23
	Fit	0.037	**0.043**	0.288	0.161	0.075	0.107	0.078	0.151	0.002	0.154	0.078	0.165
	Acc	0.966	**0.960**	0.712	0.841	0.928	0.897	0.925	0.852	1.000	0.848	0.926	0.839
ALO	Dim	5	80	5	5	**6**	**1208**	19	15	6	33	32	34
	Fit	0.041	0.072	0.298	0.167	0.077	0.071	0.096	0.169	0.002	0.171	0.084	0.176
	Acc	0.963	0.932	0.702	0.835	0.924	0.929	0.906	0.835	1.000	0.832	0.921	0.831
GOA	Dim	6	82	8	5	15	4984	28	14	14	44	29	**21**
	Fit	0.107	0.112	0.355	0.223	0.125	0.136	0.159	0.202	0.011	0.207	0.106	0.203
	Acc	0.896	0.892	0.646	0.778	0.878	0.868	0.844	0.801	0.993	0.796	0.898	0.800
SSSA	Dim	**4**	74	**5**	**5**	6	1569	**19**	**13**	**2**	**30**	31	26
	Fit	0.024	0.052	**0.237**	**0.125**	**0.058**	**0.068**	**0.076**	**0.147**	**0.001**	**0.146**	**0.073**	**0.165**
	Acc	0.979	0.952	**0.763**	**0.877**	**0.943**	**0.933**	**0.927**	**0.856**	**1.000**	**0.857**	**0.932**	**0.840**
bSSA	Dim	5	92	8	5	15	5497	30	15	11	47	33	25
	Fit	**0.022**	0.071	0.253	0.135	0.091	0.113	0.100	0.158	0.003	0.173	0.081	0.173
SSA	Dim	4	78	7	5	13	4837	25	14	7	38	30	23
	Fit	0.024	0.058	0.252	0.133	0.082	0.106	0.081	0.148	0.002	0.155	0.076	0.168
	Acc	0.979	0.946	0.749	0.870	0.921	0.898	0.922	0.856	1.000	0.848	0.928	0.836
Dataset		**Arrhythmia**	**Semeion**	**Hepatitis**	**CNAE**	**Hill valley**	**Lung-cancer**	**Lymphoma**	**TOX _171**	**CLL_SUB_111**	**COIL 20**	Avg Rank	Rank

(continued)

Table 1 (continued)

Dataset		Wine	Clean1	Vechicle	Heart	Ionosphere	Arcene	Sonar	Breastew	Soybean	Libras	Spambase	Waveform
Total	**Dim**	**279**	**265**	**19**	**856**	**100**	**56**	**4026**	**5748**	**11,340**	**1024**		
KNN	Acc	0.612	0.970	0.744	0.873	0.560	0.567	0.907	0.689	0.519	0.971	7.9	8
bGWO	Dim	183	170	8	664	67	29	2072	3731	7060	655	7	7
	Fit	0.304	0.020	0.147	**0.085**	0.371	0.133	0.067	0.154	0.325	0.018	4.8	6
	Acc	0.700	0.986	0.856	**0.922**	0.632	0.871	0.938	0.851	0.678	0.0.989	4.3	6
PSO	Dim	115	115	6	**396**	46	18	1757	2692	5368	440	3	2
	Fit	0.286	0.018	0.120	0.103	0.362	0.098	0.072	0.129	0.283	0.013	3.1	3
	Acc	0.715	0.987	0.882	0.901	0.639	0.905	0.931	0.874	0.719	0.92	3.0	2
ALO	Dim	124	124	**4**	755	**28**	13	**223**	**1574**	**1058**	**187**	3	3
	Fit	0.314	0.022	0.117	0.102	0.377	0.111	0.046	0.132	0.220	0.009	4.2	4
	Acc	0.687	0.983	0.884	0.906	0.622	0.891	0.954	0.870	0.779	0.993	4.3	5
GOA	Dim	139	134	9	428	50	26	1976	2866	5652	507	5	5
	Fit	0.351	0.031	0.172	0.215	0.412	0.246	0.078	0.228	0.390	0.024	7.0	7
	Acc	0.650	0.974	0.831	0.788	0.589	0.756	0.926	0.774	0.612	0.981	7.0	7
SSSA	Dim	**103**	**109**	4	647	32	**10**	3106	174	2326	234	2	**1**
	Fit	**0.280**	**0.017**	**0.089**	0.090	**0.359**	**0.068**	**0.044**	**0.099**	**0.216**	**0.008**	**1.1**	**1**
	Acc	**0.721**	**0.988**	**0.913**	0.917	**0.640**	**0.933**	**0.957**	**0.903**	**0.784**	**0.994**	**1.1**	**1**
bSSA	Dim	154	147	8	722	54	26	2075	3225	6244	562	6	6
	Fit	0.314	0.023	0.098	0.118	0.388	0.099	0.065	0.152	0.315	0.018	4.5	5
	Acc	0.689	0.983	0.905	0.890	0.613	0.905	0.940	0.852	0.687	0.987	4.2	4
SSA	Dim	128	124	6	416	45	20	1876	2799	5539	468	3	4
	Fit	0.291	0.019	0.092	0.121	0.375	0.070	0.066	0.127	0.293	0.014	3.0	2
	Acc	0.711	0.986	0.911	0.882	0.626	0.933	0.938	0.876	0.709	0.991	3.0	2

Bold significance on experimentation over 30 runs and 22 datasets the mean of proposed SSSA in terms of accuracy, fitness and dimensions selected is better than other algorithms

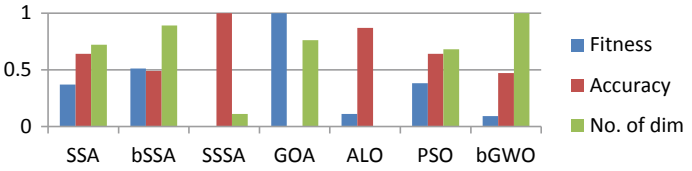


Fig. 3 Comparison of Normalised metrics

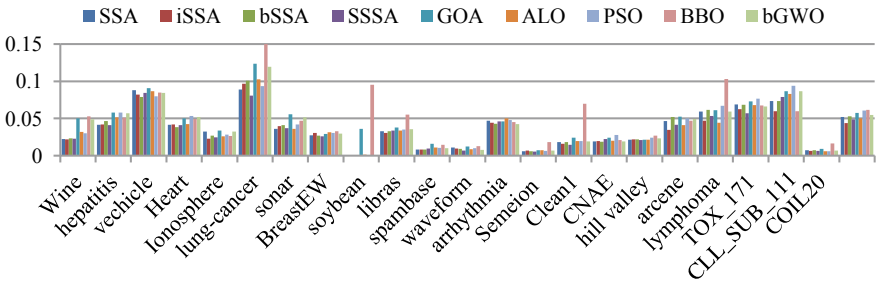


Fig. 4 Standard deviation over accuracy

References

1. Dua, D., Karra Taniskidou, E.: {UCI} machine learning repository (2017)
2. Hall, M.A., Hall, M.A.: Correlation-based feature selection for discrete and numeric class machine learning, pp. 359–366 (2000)
3. He, X., Cai, D., Niyogi, P.: Laplacian score for feature selection. Adv. Neural Inf. Process. Syst., pp. 507–514 (2005)
4. Wolpert, D.H., Macready, W.G.: No free lunch theorems for optimization. IEEE Trans. Evol. Comput. **1**(1), 67–82 (1997)
5. Mirjalili, S., Lewis, A.: S-shaped versus V-shaped transfer functions for binary particle swarm optimization. Swarm Evol. Comput. **9**, 1–14 (2013)
6. Arora, S., Anand, P.: Binary butterfly optimization approaches for feature selection. Expert Syst. Appl. **116**, 147–160 (2019)
7. Emary, E., Zawbaa, H.M., Hassanien, A.E.: Binary ant lion approaches for feature selection. Neurocomputing **213**, 54–65 (2016)
8. Mafarja, M. et al.: Binary dragonfly optimization for feature selection using time-varying transfer functions. Knowl.-Based Syst., pp. 1–20 (2018)
9. Kennedy, J., Eberhart, R.: Particle swarm optimization. In: Proceedings of ICNN'95—International Conference on Neural Networks, vol. 4, pp. 1942–1948
10. Arunekumar, N.B., Kumar, A., Joseph, K.S.: Hybrid bat inspired algorithm for multiprocessor real-time scheduling preparation. In: 2016 Iternational Conference on Communication and Signal Processing, pp. 2194–2198 (2016)
11. Rajabioun, R.: Cuckoo optimization algorithm **11**, 5508–5518 (2011)
12. Mirjalili, S., Mirjalili, S.M., Lewis, A.: Grey Wolf optimizer. Adv. Eng. Softw. **69**, 46–61 (2014)
13. Saremi, S., Mirjalili, S., Lewis, A.: Grasshopper optimisation algorithm: theory and application. Adv. Eng. Softw. **105**, 30–47 (2017)
14. Mirjalili, S., Lewis, A.: The Whale optimization algorithm. Adv. Eng. Softw. **95**, 51–67 (2016)
15. Mirjalili, S., Gandomi, A.H., Mirjalili, S.Z., Saremi, S., Faris, H., Mirjalili, S.M.: Salp swarm algorithm: a bio-inspired optimizer for engineering design problems. Adv. Eng. Softw. **114**, 163–191 (2017)

16. Faris, H., et al.: An efficient binary Salp Swarm Algorithm with crossover scheme for feature selection problems. Knowl.-Based Syst. **154**(January), 43–67 (2018)
17. Hegazy, A.E., Makhlouf, M.A., El-Tawel, G.S.: Improved salp swarm algorithm for feature selection. J. King Saud Univ.—Comput. Inf. Sci. (2018)
18. Kaya, M.: The effects of two new crossover operators on genetic algorithm performance. Appl. Soft Comput. **11**(1), 881–890 (2011)

A Novel Approach to Recommend Skincare Products Using Text Analysis of Product Reviews

Urjita Bedekar and Gresha Bhatia

Abstract The objective of this research is to develop a system that can categorize products according to their specific features with the help of product-based reviews. This system can be used to recommend suitable products according to their category. This system also avoids recommendation of those products which have harmful side effects as scrutinized with the help of sentimental analysis discerning negative reviews. Every user who searches for skincare products typically targets his requirement, and a recommendation system should be able to cater to such users. In this paper, we have proposed a model that integrates K-means clustering and topic modeling (LDA) to identify the specific topics included in the review text of a particular product. To further classify products concerning the topic, we have also used coherence scores and similarity scores to evaluate the model based on latent Dirichlet allocation (LDA). The coherence score would be a measure of semantic similarity between the words of topics which will be useful for assessing the quality of topics as predicted by the algorithm. The results turn out promising as products can be filtered by sentiment analysis and categorized according to their topics. The system can be used to not only approve suitable products but can also discard those products which can have harmful side effects, thereby eliminating the chance of dangerous side effects to a considerable extent.

Keywords Topic modeling · Text classification · Natural language processing · Recommendation system

U. Bedekar (✉) · G. Bhatia
Vivekanand Education Society's Institute of Technology, Mumbai, India
e-mail: 2016.urjita.bedekar@ves.ac.in

G. Bhatia
e-mail: gresha.bhatia@ves.ac.in.com

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_24

1 Introduction

Online product reviews are a potential source of information for retailers as well as consumers. They are key sources of real-time feedback from customers. Based on the results of text categorization using clustering as well as topic modeling, various skin concerns are identified by extracting high-frequency words from review text as in [1]. By framing the product categorization task using natural language processing, models can generate uninhibited categories which are not limited to the supervised labels from the training dataset, which can also be applied to names of products as in [2], which has similar architecture as used in [3]. In the proposed system, not only the product category but also the role of the product is taken into consideration through sentiment analysis. Review text can be exploited to get the only those specific topics that represent the category of the product. This can be helpful to address skin concerns through the recommendation of the product which belongs to the same category. Review text has been converted into the bag of words before processing as in [3]. To filter global background words that do not point to a specific category, clustering is used to get all the words that are common in many groups. Clustering can help in getting groups that relate to similar topics. Topic modeling using linear Dirichlet allocation [4] can be further used to get topics that consist of similar high-frequency words. Clustering can have general words such as good or really in many clusters. Specific words can be fetched to represent the product for its category which can be evaluated using various topic coherence measures as in [1]. This system will serve many purposes, such as enabling consumers to quickly extract the key topics covered by the reviews without having to go through all of them, which will ultimately help users to reach their goal. This will increase the usability score of an application. This will help the sellers to get consumer feedback in the form of topics (extracted from the consumer reviews) which will be useful to increase the sale of desirable products. This system will also help sellers to rank the products depending on their usability on a web page which will save time searching for the product while also eliminating unsuitable products for users.

2 Need of the System

In recent years, web services have become increasingly popular, and user reviews and rating of the product play a crucial role in the sale of the product. When it comes to skincare products, one must be extra cautious about the products as they can have a wide range of effects on various skin types. Skincare product recommendations not only need to be tested for suitable products but also those which would be perfect for the user's requirements. It is observed that when users search for skin lotion on a Web site, they also read reviews paying extra attention to those which are familiar to them. Here, familiarity includes physical aspects such as skin type or skin concerns that match with the user. People having sensitive skin often experience redness after

using new products, and so such people pay extra attention to those reviews which have positive remarks about the sensitivity of the product. For example, if the user has dry skin and wants to purchase a product for the same, the user will look for reviews that have explicitly stated that the product considerably improved the dry skin of the reviewer. However, since it is difficult to gauge exact information about the effect of a particular product on a particular skin type, an intelligent system capable of finding key insights from these reviews will serve for both the consumers and the sellers.

3 Literature Survey

Many product recommendation systems have been developed so far. There are various ways to categorize products. Umaashankar et al. [5] have developed taxonomy classification for clothing product categorization, similar to a hierarchical technique used in [6, 7] and [8]. Yeh et al. [9] have used association reasoning neural networks to discover customer's behavior; however, it lacks the universal approach which can be applied to all the products. Wirojwatanakul and Wangperawong [10] have worked on multi-label classification; however, mislabelled products can lead to inaccuracy in their system. Zhang et al. [11] explains various deep learning techniques which can aid recommendation, but this can encounter the same drawback as [10]. Alashkar et al. [12] have spectacularly used a knowledge base that models the relationship between the facial style attributes and makeup style attributes for the recommendation of cosmetics. Nakajima et al. [13] have proposed their system based on compatibility between users and cosmetic ingredients present in various skincare products. Also, the system assumes that each element of a product contributes to some effect such as 'whitening' or 'reducing acne.' However, the combination of ingredients may create unforeseen consequences. The method proposed by Seo et al. [14] can be used as a guide for recommendation of cosmetics according to consumers' skin concerns. They have also used consumer sentiments as customized cosmetic recommend indexes primarily focusing on marketing and advertising of the products which are preferred the most by users. The system proposed by Moe and Aung [15] uses cross-domain recommendation to link problems domain and cosmetics domain. They have also presented a valuable approach for building ontologies based on skin problems and cosmetics. The system proposed by Matsunami [16] has highlighted the importance of reviews in purchasing decisions mainly because of the potential for irritation with unsuitable products. The automatic review rating feature has also been implemented for the recommendation of cosmetics. Matsunami et al. [17] have identified user groups who have similarity for particular cosmetic review. Our system has considered the behavior of cosmetic products for various skin types as we have used topic modeling in identifying topics of reviews using a dataset as in [18] and [19] which included review text as well as the rating for products. Xie and Xing [20] have proposed a multi-grain clustering topic model which has achieved a great performance by integrating clustering and topic modeling.

4 Methodology

The flowchart of the proposed system has been shown in Fig. 1. Review text is preprocessed to generate words using the bag of words (BOW) model which is then followed by clustering and topic modeling. Since LDA evaluates the model using bag of words, the order of the phrase is not taken into consideration. Ordering is taken into consideration while performing sentiment analysis. Polarity score is used to identify consumer's sentiment, which is a crucial factor for discarding products having harmful side effects.

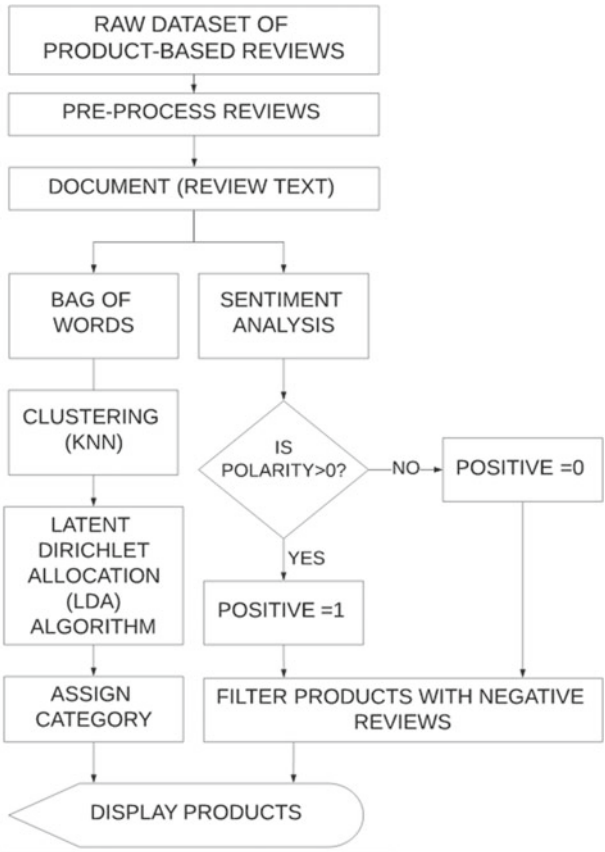


Fig. 1 Flowchart

4.1 Prepare Text for Analysis

Review text for skincare products has been extracted from a database of Amazon product reviews for items in the 'beauty products' category. After dropping tuples that appeared duplicative, 198,502 tuples were included in the training model. The text was then tokenized. Tokens that have fewer than three characters are removed. Words are lemmatized and stemmed by reducing to their root form. All stop words are removed. Bigram and trigram models are created. These models approximate the probability that one word would be followed by another. Thus, our model can have 'nail polish' as a particular meaning phrase in the 'manicure' topic. Successive bigrams form long phrases. Before topic modeling, we convert the tokenized and lemmatized text to a bag of words—is a form of the dictionary where the key is the word and value is the frequency of that particular word in the entire corpus. In classic document clustering approaches, documents are usually represented with a bag of words (BOW) model which is purely based on raw terms. Clustering gives rise to clusters of various topics, which are not interrelated. However, they can contain the words which are shared among clusters such as use, really and good. Based on the results obtained after clustering, text can be represented based on global topics as found in many clusters, and specific topics are found in a particular cluster. Further LDA can be used to group the topics of a particular cluster such as skincare which consists of specific words used to represent particular skin concerns or hair concerns. Also, common hair concerns such as frizzy or dry hair can be found out, and products that are effective on the same can be shortlisted using the same topic. For this, we use the latent Dirichlet allocation (LDA) algorithm to identify the skin concerns based on high-scoring words. For example, 'oily' and 'dry' are among high-scoring words in the 'skincare' cluster. The products can be categorized with the help of these high-scoring words, and their effect can be evaluated using sentiment analysis.

Raw data (Review Text):

Very oily and creamy. Not at all what I expected… ordered this to try to highlight and contour and it just looked awful!!! Plus, took FOREVER to arrive.

After preprocessing the data:

['oily,' 'expect,' 'order,' 'try,' 'highlight,' 'contour,' 'look,' 'awful,' 'take,' 'forever,' 'arrive'].

4.2 Clustering Using KNN

Document clustering is extensively used in document classification and summarization. Specifically, document clustering enables us to extract local topics specific to each document cluster(dry) and global topics (nice, great, buy) shared across clusters. Clustering can help us to identify the groups of high-frequency words in a document collection and subsequently enable us to identify words specific to each group in local topics (shampoo, conditioner) and words shared by all groups in global topics

Table 1 Clusters formed from the dataset

Cluster 1	Cluster 2	Cluster 3	Cluster 4
Use	Color	Hair	Skin
Product	Nail	Use	Use
Love	Brush	product	Face
Good	Polish	Conditioner	Product
Work	Love	Shampoo	Feel
Great	Great	Dry	Dry
Buy	Use	Leave	Make
Smell	Coat	Make	Cream
Really	Look	Feel	Really
Look	Good	Work	Lotion

(use, make), thus exploiting the grouping structure of documents. These fine-grained topics can facilitate a lot of utilities. For instance, we can use the group-specific local topics to classify products further according to their subgroups. Global topics can be used to filter common words and get specific words that can give an insight into the category of product. The data was divided into four clusters, where the first consists of global words, and the remaining have a mixture of global as well as specific words. Topics of skincare products could be recognized by applying LDA over cluster 4 in Table 1.

4.3 Topic Modeling Using LDA

Latent Dirichlet allocation (LDA) is a probabilistic model that considers that various topics constitute a document (review text in this case). The basic idea is that documents are represented as random mixtures over latent topics, where a distribution over words characterizes each topic. LDA does not consider sequences of words, but rather it uses a bag of words to represent a document. The documents are represented as random mixtures over latent topics, where a distribution over words characterizes each topic. β is a Dirichlet distribution matrix that represents the probability of a particular topic containing a specific word. α represents the topic distribution per document. A lower value of α indicates an unbalanced distribution. A higher value of α indicates a balanced distribution. φ denotes word distribution for any topic. Our goal is to find an α which represents our document as a probabilistic model of topics. α is calculated for all the documents (reviews) using φ and β. In the case of skincare products, we were interested in mapping all words related to a topic together. Figure 2 shows some of the words contained in the topics for 'usability' and 'smell' of products, respectively. Many other words related to skin type like oily and greasy have been grouped into a single topic 'oily' which can be used for filtering according to skin type. Words like itching and swelling were paid extra attention to discarding unsuitable products. Standard topic models lack the mechanism to model the grouping behavior among documents as they may include global high-frequency words in many topics, thereby local words and global words are mixed and cannot

```
(8,
 '0.050*"brush" + 0.018*"use" + 0.014*"hold" + 0.012*"set" + 0.012*"small" + '
 '0.011*"easy" + 0.011*"get" + 0.010*"come" + 0.009*"well" + 0.009*"make"'),
(9,
 '0.109*"smell" + 0.083*"scent" + 0.040*"nice" + 0.027*"fragrance" + '
 '0.027*"strong" + 0.020*"good" + 0.017*"pleasant" + 0.014*"light" + '
 '0.014*"spray" + 0.013*"fresh"'),
```

Fig. 2 Topics after performing LDA

be distinguished. Topics in Fig. 2 can be labeled as 'usage' and 'perfume.' Perfume can form a different category of products for which sentiment analysis can be used to recommend products with a good fragrance.

4.4 Sentiment Analysis

The sentiment is assigned for each review with a value from −1.0 (negative) to 1.0 (positive) with 0.0 being neutral. The subjectivity is a value from 0.0 (objective) to 1.0 (subjective). Subjective reviews are removed while unbiased reviews are used to identify the sentiment of the reviewer. Those products with positive reviews can be further filtered and ranked for their polarity score. To create the recommendation system, we wanted to leverage the skincare concern categories and the review sentiment scores together. Specifically, we wanted to find the products that had reviews that spoke positively regarding the different concerns. We matched review sentences with positive sentiment scores to the four categories based on the presence of the topic modeling words. For example, a sentence that said 'I loved this product for my sensitive skin.' would be flagged as a positive sentence for the 'sensitive' category for that product. This product will be recommended to those users who have sensitive skin.

5 Results

We have used the dataset which contains product reviews and metadata from Amazon, including more than 100,000 reviews spanning May 1996–July 2014 for analysis. Based on the results of text categorization using clustering as well as topic modeling, various skin concerns are identified by extracting specific high-frequency words from review text. Various skin concerns such as acne or pimples can be similarly addressed using that particular category of products with positive reviews.

5.1 Coherence Score

The coherence score is for checking the quality of the topics as learned by the trained model. The results indicate that coherence score for the LDA model has also increased from 0.49 to 0.52 after including the model of K-means clustering. With the number of topics, coherency score increases to an extent till it reaches a point where it starts to flatten, with this it makes a better sense to select the model which gives the highest score before drop. In this case, $K = 7$ would give better results as shown in Fig. 3. Score is typically $0 < x < 1$. Topic coherence can be defined as how interpretable a topic is based on the degree of relevance between the words within the topic itself.

$$\text{Coherence score(extrinsic)} = \log[\text{P(word1, word2)}/\text{P(word1).P(word2)}]$$

$$\text{P(word1, word2)} - \text{Probability of both words word1, word2 being present in the random document}$$

$$\text{P(word1)} - \text{Probability that word1 is present in the document}$$

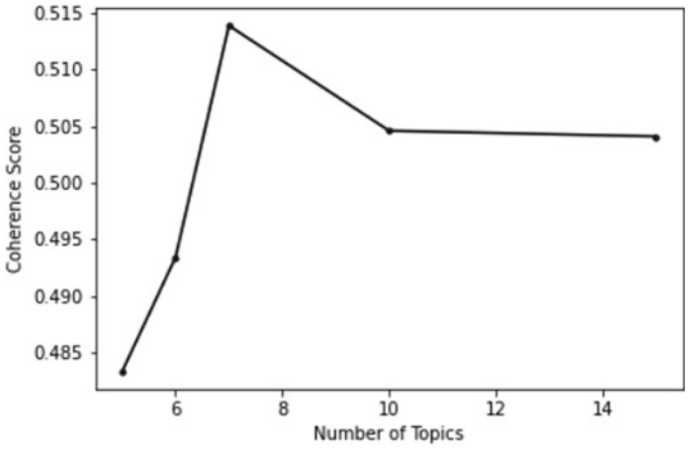


Fig. 3 Coherence score versus number of topics (K)

5.2 Topic Distribution

Topic distribution is the proportion of words in a document that belongs to the topic.

$$\text{P(topic T|document D)} = \text{the proportion of words present in unseen document D that are currently assigned to topic T.}$$

The unseen document consisting of the sentence—'I absolutely loved this product for dry skin'—was compared before and after using *K*-means clustering for document similarity with the topics as part of testing the model as shown in Table 2.

Table 2 Comparison of only topic modeling and clustering along with topic modeling

	K-means + LDA	LDA
Maximum topic distribution	0.76	0.51

Figure 4 shows all the products with positive reviews for the dry skin category. Further searching more about ASIN to verify if the products were actually suitable for dry skin, ASIN-B000052YKI redirected to 'Purpose Dual Treatment Moisture Lotion with SPF 15 4 fl oz (120 ml).' We can clearly observe that the products which turned out to have value are actually the products for dry skin with an overall good rating.

```
print(new_df_drySkin.head())
```

```
          asin                                reviewText_processed
477   B000052YKI  i have been using purpose products ever since ...
766   B000052YMR  this is a great cream for me to use in the win...
4613  B0000ZHH5G  vanicream works great on my skin  it keeps my ...
6588  B0001EKWPI  i've tried many soaps for sensitive skin and f...
6577  B0001EKWPI  doctor recommended dove for sensitive skin for...
```

Fig. 4 Products filtered for dry skin

6 Conclusion

The proposed system can categorize products very effectively and can further filter the products with overall negative reviews. This system works by forming groups of words that are closely related and then form topics of words consisting of only specific words that can represent a category of product. The products having overall positive reviews are finally recommended. This system can be connected to search engines to recommend suitable products. With this system, it is not only possible to filter products which would not be suitable for a particular category, but it is also possible to filter based on skin concerns. The reviews with the topics are identified and are further mapped onto their respective product. The main advantage of this system is its cross-disciplinary nature which makes it applicable for any type of product. This system can aid the filtration of products and can also help to rank the products, which will increase the usability of any system. This system can be used in various online shopping applications to gauge effective products for various categories. By analyzing the text of reviews, it is possible to unravel highly reliable and effective components of the product present in reviews.

References

1. Röder, M., Both, A., Hinneburg, A.: Exploring the space of topic coherence measures. In: WSDM 2015—Proceedings of the 8th ACM International Conference on Web Search and Data Mining, pp. 399–408 (2015). https://doi.org/10.1145/2684822.2685324.
2. Li, M.Y., Kok, S., Kok, S.: Unconstrained product categorization with sequence-to-sequence models. In: eCOM@SIGIR. CEUR Workshop Proceedings, vol. 2319.
3. Armand, J., Grave, E., Bojanowski, P., Mikolov, T.: Bag of tricks for efficient text classification. arxiv.org/1607.01759 (2017)
4. Blei, D.M., Ng, A.Y., Jordan, M.I.: Latent Dirichlet allocation. J. Mach. Learn. Res. **3**, 993–1022 (2003)
5. Umaashankar, V., GirishShanmugam, S., Prakash, A.: Atlas: a dataset and benchmark for e-commerce clothing product categorization. arxiv.org/1908.08984 (2019)
6. Dumais, S., Chen, H.: Hierarchical classification of web content. In: Proceedings of the 23rd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval. pp. 256–263. ACM (2000).
7. Cevahir, A., Murakami, K.: Large-scale multi-class and hierarchical product categorization for an E-commerce giant. In: Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers, pp. 525–535.
8. Shen, D., Ruvini, J.D., Sarwar, B.: Large-scale item categorization for ecommerce. In: Proceedings of the 21st ACM International Conference on Information and Knowledge Management, pp. 595–604. ACM (2012).
9. Yeh, I.-C., Lien, C., Ting, T.-M., Wang, Y.-Y., Tu, C.-M.: Cosmetics purchasing behavior—an analysis using association reasoning neural networks. Expert Syst. Appl. **37**, 7219–7226 (2010).
10. Wirojwatanakul, P., Wangperawong, A.: Multi-label product categorization using multi-modal fusion models. arxiv.org/1907.00420 (2019).
11. Zhang, S., Yao, L., Sun, A., Tay, Y.: Deep learning based recommender system: a survey and new perspectives. ACM Comput. Surv. **1**(1), Article 1. Publication date: July 2018.

12. Alashkar, T., Jiang, S., Fu, Y.: Rule-based facial makeup recommendation system. In: 2017 IEEE 12th International Conference on Automatic Face & Gesture Recognition, pp. 325–330.
13. Nakajima, Y., Honma, H., Aoshima, H., Akiba, T., Masuyama, S.: Recommender system based on user evaluations and cosmetic ingredients, pp. 22–27 (2019). https://doi.org/10.1109/INCIT.2019.8912051.
14. Seo, B., Kim, K., Yoo, I., Park J., Park, D.: Development of diagnostic method and algorithm for skin type based on consumer language and sentiment. In: 2018 3rd Technology Innovation Management and Engineering Science International Conference (TIMES-iCON), Bangkok, Thailand, 2018, pp. 1–3. https://doi.org/10.1109/TIMES-iCON.2018.8621838
15. Moe, H.H., Aung, W.T.: Building ontologies for cross-domain recommendation on facial skin problem and related cosmetics. Int. J. Inf. Technol. Comput. Sci. (IJITCS) **6**(6) (2014)
16. Matsunami, Y., Ueda, M., Nakajima, S.: Explaining item ratings in cosmetic product reviews. In: Proceedings of the International MultiConference of Engineers and Computer Scientists 2016, vol. I, IMECS 2016.
17. Matsunami, Y., Okuda, A., Ueda, M., Nakajima, S.: User similarity calculating method for cosmetic review recommender system. In: Proceedings of the International MultiConference of Engineers and Computer Scientists 2017, vol. I, IMECS 2017.
18. McAuley, J., Targett, C., Shi, J., van den Hengel, A.: Image-based recommendations on styles and substitutes. SIGIR (2015).
19. He, R., McAuley, J.: Ups and downs: modeling the visual evolution of fashion trends with one-class collaborative filtering, WWW (2016)
20. Xie, P., Xing, E.: Integrating document clustering and topic modeling. arxiv.org/1309.6874 (2013)

Hand Motion and Face Recognition System for Visually Disabled People

Suvarna Nandyal and Maleka Anjum

Abstract Keen security framework has gotten essential in current everyday life. The proposed security framework has been developed to prevent breakdowns in deeply secure regions, such as home conditions with less use of force and a more dependable independent safety device for both intruder recognition and entryway security. The past techniques have seen the utilization of accelerometers, which are fixed on the hand as they are viewed as the best device for completing such practices. In our work, the visitor is identified by a standard web camera to capture the image. If the face recognizes the people, they will be greeted by name and the door will be unlocked. Doors will not be unlocked if they are not identified. For Door Access, the system will perform detection and recognition. Home Automation Model utilizing the online site is intended to give simplicity of control of home apparatuses for individuals, especially the old or those genuinely unfit to perform everyday exercises effectively. Electronic computerization discusses straightforwardly with raspberry pi to control home appliances. The proposed technique, permits clients to deftly and conveniently control various family apparatuses with basic signals.

Keywords Gray scaling · Filtering · Gesture recognition · Convolutional neural network · Image Classification · Home appliances

1 Introduction

The point of the home mechanization framework is to assemble a gadget that can screen home apparatuses utilizing any of the two techniques doled out: Signal based and electronic. Impaired or older individuals who can't walk need an easy method for getting to things around them that must be done reliably and productively. This structure consolidates mechanization and innovation. Conventional home robotization frameworks are not suitable for old networks or impaired individuals. Home mechanization frameworks are utilized to control home machines by means of

S. Nandyal (✉) · M. Anjum
Department of Computer Science & Engg, PDA College of Engineering, Kalaburgi, India

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_25

controller (cell phone). Electronic computerization and signal based robotization give an advantage to individuals who are truly incapable to deal with everyday undertakings adequately. Many motion following innovations that generally include hand gloves connected to sensors or uncommon capacity gloves.

The strategy for facial acknowledgment and discovery is completed by the primary part investigation (PCA) approach and, rather than utilizing sensors, the interloper location is practiced by performing picture handling on caught video information outlines and by estimating the contrast between the recently caught outline and the running casing regarding pixels in the caught outlines. This is an independent assurance framework that has been created utilizing the Raspberry Pi electronic advancement board and run on a battery power gracefully, remote web access utilizing a USB modem. Auto Police e-Complaint enrollment has been practiced by sending security infringement cautioning sends to the closest police headquarters email address. The proposed framework is all the more impressive, more solid and expends far less information and force than the other current frameworks.

2 Related Work

This paper [1] recommends a technique where characterizations are utilized for the distinguishing proof of signs and discoveries that are completed naturally. The ongoing picture is in a RGB design that should be meant a dim scale as a delivering apparatus. This changed picture is then meant a multi-goal picture utilizing a Gabor change. This changed picture or controlled picture is then used to recover the qualities of a changed Gabor. In this article, MATLAB Rb is utilized to display the proposed signal language arrangement framework with 2 GB of RAM and the Intel Pentium Core 2 team processor. The dataset is taken from an open source permit.

In paper [2], wiping out the tedious procedure of ordering a solitary signal, orders a few motions in their strategy. This innovation has a wide assortment of employments, for example, gesture-based communication acknowledgment, touchless vehicle helps frameworks and gaming gadgets. They have built up a profound organization based Long Short-Term Memory (LSTM) developed by Encoder-Decoder engineering that unequivocally characterizes the arrangement of developments in one go. Their innovation is successful to the point that even with insignificant information, they can produce dependable discoveries taking things down a notch. It is a multilayer, fluffy, neural-network based classifier framework. The all-out guideline of accessible methodologies should be contemplated and contrasted with the conclusion of the ideal arrangement. A gadget with ideal execution, minimal effort, an ideal blend of techniques and results against various settings ought to be picked. ANFIS is the favored type of their plan.

In paper [3], An entryway lock access framework, comprising fundamentally of three subsystems, was proposed: face ID, facial acknowledgment and programmed entryway access control. Face discovery is completed utilizing the Principal Component Analysis (PCA). Attributable to the guideline of the microcontroller, the

entryway would open naturally for the distinguished client. In the opposite side, the caution rings for the obscure person. The drawback of this technique is that the information pictures are shot ceaselessly from a web camera before the 'stop video' button is squeezed [1]. Someone is relied upon at the area to check the photographs or status of the gadget of an unapproved individual and to make further move. The (PC) is appended to the microcontroller, the entire gadget would not run if the Machine crashes or the non-work. In paper [4–6] proposed gesture controlled robot using image processing and obtained 85 % accuracy.

3 Proposed System

Figure 1 shows the complete architecture of the proposed work.

The most basic part of the work is the right plan of the equipment. The electronic parts of the paper are effectively perceived and include the association of different equipment gadgets to the Rasperry Pi advancement board through GPIO pins and different other usable ports without the need to weld the wires for the association. The

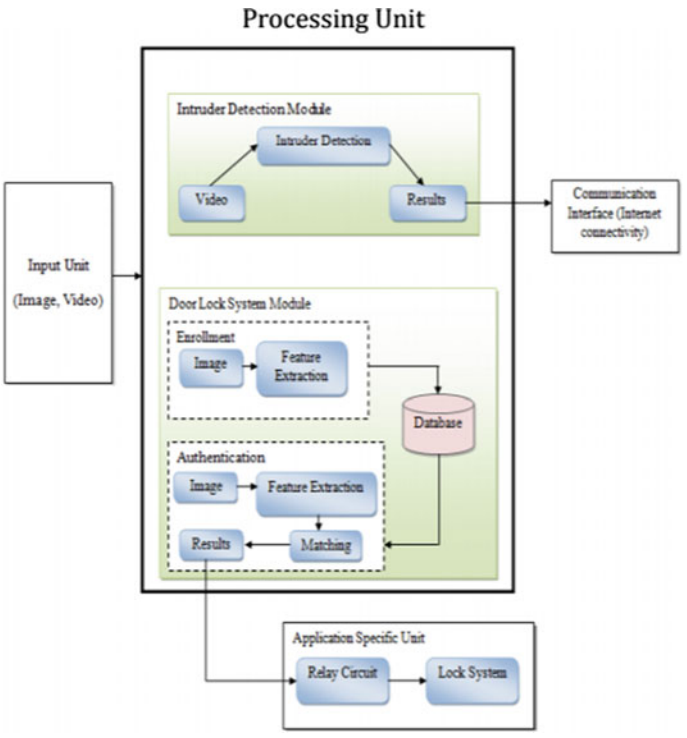


Fig. 1 Proposed system architecture

accompanying wiring graph Fig. 3 outlines how to associate and fuse the fundamental equipment parts of the proposed gadget, for example, Electronic Lock, Push Button Switch, Relay, Raspberry Pi Screen, Raspberry Pi Monitor, Web Camera, Wi-Fi Receiver Transmitter, Power Bank Source with Internet Data Card Dongle.

The system detects the gestures provided by the user as input and controls the home appliance. The main goal is to provide portability and allow people who are blind, deaf and stupid to control different appliances with ease and comfort [7, 8]. Also, because of the increase in the number of industrial and home appliances that need to be controlled, control methods are needed. Using an android application [9], the gesture input from the user is captured and sent to the raspberry pi (which acts as a microcontroller) and the raspberry pi then operates on the appliances' respective functionality. For image classification, the system uses a Convolutional Neural Network algorithm. CNN is used for issues such as recognition of patterns and images.

4 Implementation

Figure 2 shows that the modules are associated with the raspberry pi with a middle hand-off board. The hand-off module is utilized so as to shield Pi from being worn out, since Pi endures up to 5 V.

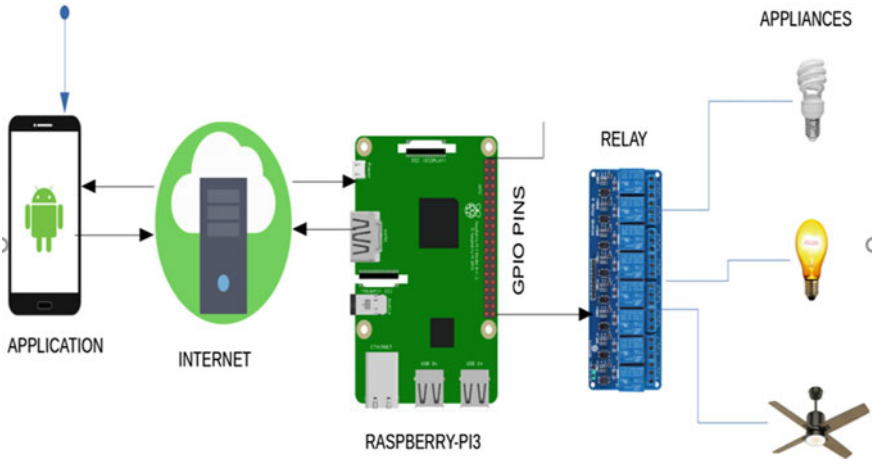


Fig. 2 Hardware connectivity of appliance

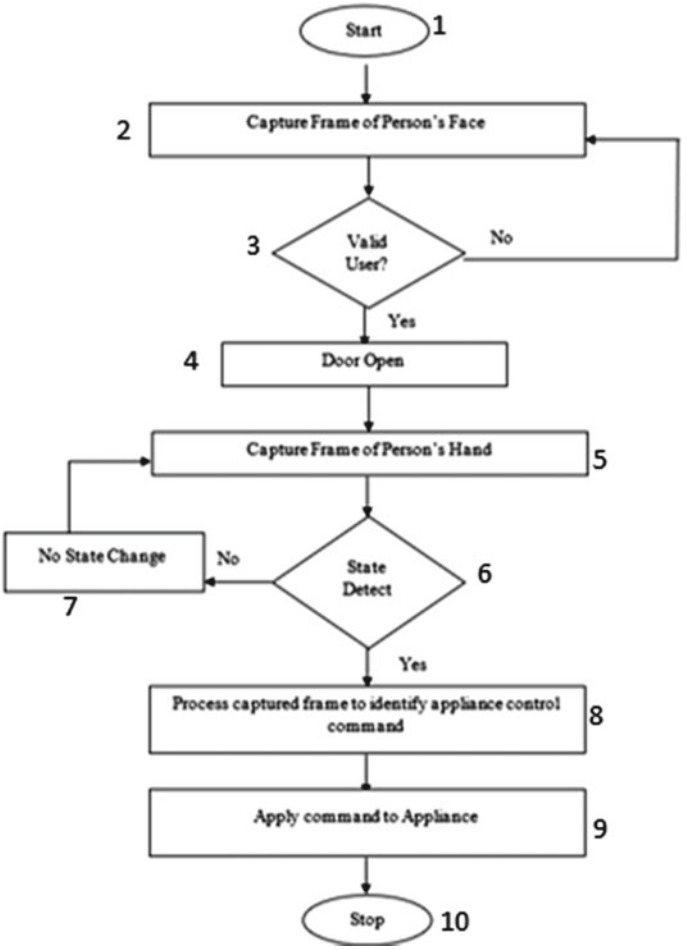


Fig. 3 Flow diagram of the system architecture

4.1 Work Flow of the System

A regular level of home security developed to provide our home IOT-based face affirmation with security can be executed. To get the image to recognize the visitor, a standard web camera. It's a technique which recognizes the visitor. If the visitor perceives the face [10], it will invite them by name and open the name of the gateway. The gateway will open if they are not recognized. Identification and affirmation for door access will be carried out by the framework (Fig. 3).

Fig. 4 Actions assigned to the gestures

4.2 Appliance Control Based on the Classified Image

For desired actions to perform, the user can send gestures via the android app. The devices can be controlled by using the on/off buttons on the web application in an additional way, as shown in Fig. 4.

5 Results and Discussion

In this work, we focused on the equipment needed for our structure for Windows 2010 and web architecture. To run our gadget that integrates at least 4 GB of Ram, Raspberry pi 3 model B, 32 Gb SD card, jumper wires, hand-off board, lamp, and fan, we need some equipment. We need network interfaces such as Ethernet, Wi-Fi and gadget interfaces, such as android phones.

In Fig. 5, USER1 request for access to enter, access is granted as USER1 is authorized person to enter home, whereas in Fig. 6 USER4 access is denied as he is not an authorized person to enter.

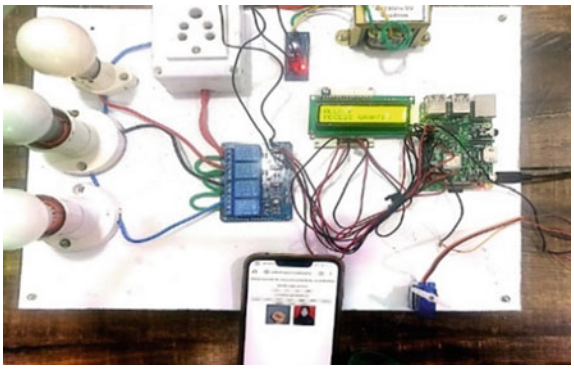


Fig. 5 Showing USER1 access granted

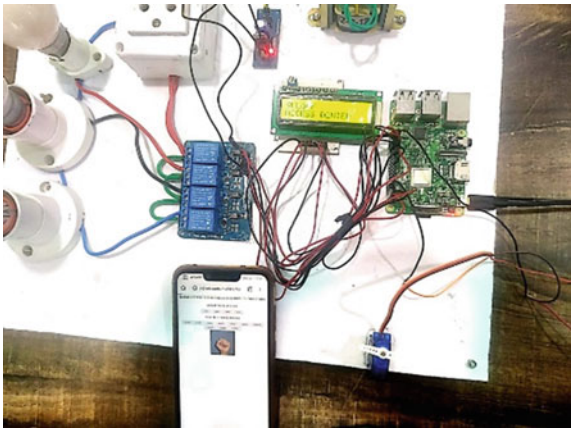


Fig. 6 Showing USER4 access denied

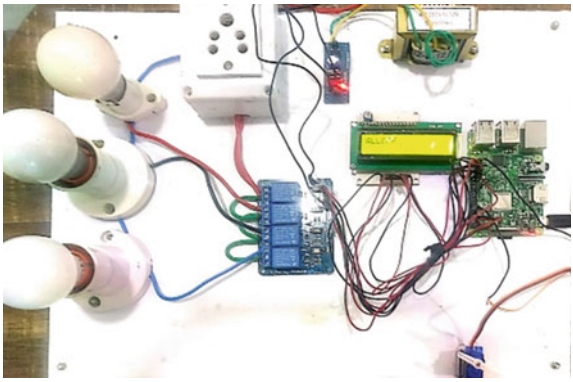


Fig. 7 Showing ALL OFF

In Fig. 7, the authorized user selects ALL OFF option and all appliances will be off, whereas in Fig. 8 user selects ALL On option.

6 Conclusions

This work investigates an Android gadget for the ID of signs by hand. The gadget has been effectively evolved and tried on ios. This strategy is useful for an acrid and numbskull holding a webcam-associated Android PC to screen machines. Also, it is furnished with facial acknowledgment and the gatecrashers discovery highlight

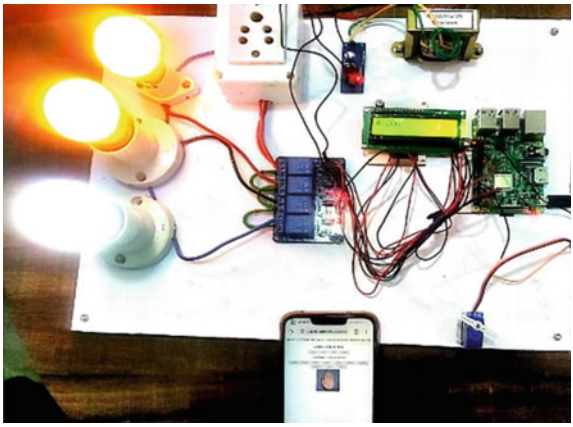


Fig. 8 Showing ALL ON

in this arranged entryway access framework. This innovation has been effectively checked by testing the entrance control of the home entryway lock based on the facial acknowledgment measure by testing in script facial pictures. For home computerization for the outwardly impaired, remote gear is utilized. Truly influenced individuals in this strategy utilize home PCs adequately so they can utilize the hardware securely. This instrument is helpful to use on the gadgets; the controller framework would supersede this framework by squeezing the switch, for which the home PC is likewise proper.

References

1. Neethu PS, Suguna R, Sathish D (2017) Real time hand gesture recognition system. Int J Emerg Technol Eng Res (IJETER) 5(3)
2. Wankhade KA, Zade GN (2014) Sign language recognition for deaf and dumb people using ANFIS. Int J Sci Eng Technol Res (IJSETR) 3(5)
3. Kakde MU, Rawate AM (2016) Hand gesture recognition system for deaf and dumb people using PCA. Int J Eng Sci Comput 6(7)
4. Kaura HK, Honrao V, Patil S, Shetty P (2013) Gesture controlled robot using image processing. Int J Adv Res Artif Intell (IJARAI) 2(5)
5. Khan RZ, Ibraheem NA (2012) Hand gesture recognition. Int J Artif Intell Appl (IJAIA), 3(4)
6. Sun J-H, Ji T-T, Zhang S-B, Yang J-K, Ji G-R (2018) Research on the hand gesture recognition based on deep learning. In: 12th international symposium on antennas, propagation and EM theory (ISAPE) Year: 2018 | Conference Paper | IEEE
7. Rishad EK, Vyshakh CB, Shahul SU (2017) Gesture controlled speaking assistance for dump and deaf. In: National conference on discrete mathematics & computing (NCDMC)
8. Billinghurst M, Piumsomboon T, Bai H (2014) Hands in space: gesture interaction with augmented-reality interfaces. IEEE Comput Gr Appl 34(1)
9. Raheja JL, Singhal A, Sadab, Chaudhary A (2014) Android based portable hand sign recognition system. Int J Sci Eng Technol Res (IJSETR) 3(5)

10. Tripathi S, Sharma V (2011) Face detection using combined skin color detector and template matching method. Int J Comput Appl 26(7)
11. Hatwar PD, Wahile NA, Padiya IM (2017) Hand gesture recognition using ANFIS. Int J Emerg Technol Eng Res (IJETER) 5(3)
12. Ghotkar AS, Khatal R, Khupase S, Asati S, Hadap M (2012) Hand gesture recognition for Indian Sign Language. In: International conference on computer communication and informatics 2012 IEEE

Price Prediction with the Sectorial Effect on Stock Market Due to COVID-19

Chetan Gondaliya, Hiral Patel, and Ajay Patel

Abstract In this paper, the examination explores the effect of the COVID-19 pandemic period brought about by the COVID-19 to the financial exchange of India. The stock market is one of the investing factor for people to get maximum benefit. By investing in stock market is risky in nature. Lots of analytical aspects and insights required for investing in stock. The examination inspects the degree of the influence of the COVID-19 on the Indian securities exchange and whether the market response would be the equivalent in pre-and post COVID-19 period. Market Model Event study philosophy is utilized. This paper also focus on sectorial effect on stock market. In this context majorly IT, Pharmaceutical and Infrastructure based stock companies are explored. Stock Exchange (BSE) are chosen aimlessly with the end goal of the examination and data gathering. The effect analyzing time duration slot taken for study is from January 2020 to 15 August 2020. In this period the market had affected with lots of unusual things like lockdown. Some of sector may improve the financial aspects and some of goes down. The outcomes demonstrate the market responded emphatically with expressively average +ve abnormal behavior in nature. The model utilize the ROI for this mentioned period and also forecast the future price of selected stock. According to proposed work, the sectorial effect, ROI margin and price prediction will be performed.

Keywords COVID-19 · Linear regression · Marginal effect · ROI · Sectorial effect · XGBoost

C. Gondaliya (✉) · A. Patel
AMPICS, Ganpat University, Mehsana, Gujarat, India
e-mail: cpg01@ganpatuniversity.ac.in

A. Patel
e-mail: ajaykumar.patel@ganpatuniversity.ac.in

H. Patel
DCS, Ganpat University, Mehsana, Gujarat, India
e-mail: hrp02@ganpatuniversity.ac.in

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_26

1 Introduction

COVID-19 is an unpredictable situation, aroused in major countries of the globe. This pandemic has not only affected human being but it has also influence economy worldwide. Our Government is trying to balance and to boost the economy of the country by proposing or declaring the various kinds of special packages from the emergent funds.

Indian stock market is one of the significant pillars of Indian economy. It works on global movement and sees upwards and downward situations on the daily basis. Corona virus has highly affected our share market hence the effects are observed in our Economical activities. First corona virus case was reported on 30th January in India and suddenly Government has ordered the lockdown on 24 March, 2020. Is it late to ordered lockdown? It may be possible to slow the spreading of corona virus in the population [1].

India is a second highest population country in the world where health facility is not as much good, poor cleaning area. COVID-19 outbreaks are affected to many sector of Indian Government. Due to lockdown, all the business activity is halt down, global economy was crashed, major descent in oil prices, unemployment rate is increased in almost countries of the world. There is no medicine, vaccine for corona virus, lockdown is only one way to pause the corona virus spread in the country, which have adverse effect on GDP growth, economy of Indian market and share market. It will lead to a negative sentiment towards stock market, which have reported major drop in share market on march to April 2020.

Stock market prices is not only depending on historical prices of stocks, but also affect the sentiment of current market. Stock price will reflect to all publicly available information. The purpose of this study to observe the influence of the lockdown on the stock market.

2 Literature Survey

Corona virus outbreaks has caused major growing concern about the economic consequences. Equity market is depending on changes in expected returns not in changes in expected future growth rate [2].

Demonetization step was taken by Indian Government on November 8, 2016 for get rid of the black money from the country. As a result of this event, cash flow activity was suddenly declined. Author has used three methods for review the short-term impact of demonetization on equity market. It has review the different days of stock price after demonetization announcement [3].

The lockdown has major impact on supply chain of different industry. They have used two methods: (1) Fixed effect (2) Random effect (3) OLS (Ordinary least square) It has finding effect of weekly quarterly and as per given period of time for comparing the price of stock market.

In this paper, we are going to utilize python to play out certain information investigation of the Indian financial exchange and its responses to COVID-19. We will attempt to comprehend the effect of certain components like the expanding number of affirmed cases in the nation, google pattern examination of individuals scanning for COVID-19, relationship with DJIA (Dow Jones Industrial Average), and so forth. A compelling method of empathetic the fallouts of occurrences like the escalation of COVID-19 is to consider as factor value. These effect changes catch current longings. Viably, reserve markets stretch ongoing, high effect studies in concerns to future aspects [4].

The COVID-19 duration was proclaimed contagion by WHO, its shortage of lessons recognized with COVID-19 effect on financial exchange, as its new worldwide turn of events. A study has been completed to dissect the effect of COVID-19 on various economies, yet the number is still little [5–7].

Ramelli and Wagner [8] look at market responses to the COVID-19 on the universal exchange and money related arrangements of the firm and found a negative ramification for globally arranged US firms, particularly those with China presentation and the US; markets moved hotly at the point when the infection spread across Europe and the US [8].

Nguyen and Pham [9] analyze the relationship between search-based feeling and Vietnamese stock showcase returns. The creator inferred that the sentiment induced impact is chiefly determined by negativity. What's more, idealistic speculators appear to defer making their venture move until the market adjusts itself. The creator didn't found any measurable proof that hopefulness influences the market returns, however, it gives an establishment to their case that hopeful gathering might be potential speculators who search wit [10, 11].

The expectation to take part in the business sectors. Again data assumes a key job for more interest of the new speculators in the market that eventually influences the stock returns [9].

Lee and Brahmasrene [12] research the relationship between the macroeconomic factors and stock cost of Korea in the short run just as the since quite a while ago run and found the proof of since quite a while ago run harmony connection between chosen macroeconomic factors, for example, cash flexibly, modern creation file, swelling, trade rates and financing costs and Korean securities exchange, while, cash flexibly and momentary loan cost are not identified with stock costs in the short-run. Be that as it may, conversion scale is decidedly identified with stock costs in the short-run. The creator additionally found that there is no convincing proof that the outside stuns for example, local or worldwide budgetary emergency affect the stock value elements in the Korean securities exchange [12].

3 Methodology

An example of 6 organizations are chosen by area pharmaceutical, foundation and IT. The information identified with stock costs are taken from BSE. The duration chose for the examination was January to September 2020. The BSE 50 record shutting costs for the individual time frame are taken. The occasion date is characterized as the day first case of covid was reported and requested in Jan 2020. This methodology accept that the data was known to the market from that day. During the lockdown period, the stock cost of a sum of 1 month are taken half days preceding the occasion day, occasion day (t1).

In this exploration study, the proposed model occasion analysis system has been utilized to quantify the impact of the COVID-19 on the securities exchange and costs. The proposed model exploited by Fama [13] expect a direct connection between the profits of the security to the profits of the market portfolio. This model has been picked in light of the fact that it is incredible in distinguishing unusual returns when contrasted with other expound procedures like file model, and it is likewise liberated from reactions of capital resource evaluating model. The explanation behind utilizing this model is on the grounds that it has been demonstrated to yield substantial outcomes [14]. Market model philosophy is a moderately incredible philosophy material under a wide assortment of circumstances (Earthy colored and Warner 1985).

This paper focus on the forecasting and sectorial effect on stock. The proposed model work as per following Fig. 1.

The proposed model applied for modelling the dramatically effect of COVID-19 on stock. As we all know the people also suffering with money for daily rifting life line so for investing on stock is also crucial. This is generalize model but also consider the factors affected due to covid-19 for world while. For an example on such duration the transportation is keep quit still by opening lockdown so such consideration also have been keep in mind to develop model. This proposed model is consider the sectorial impact because every sector has different rise and fall during this pandemic.

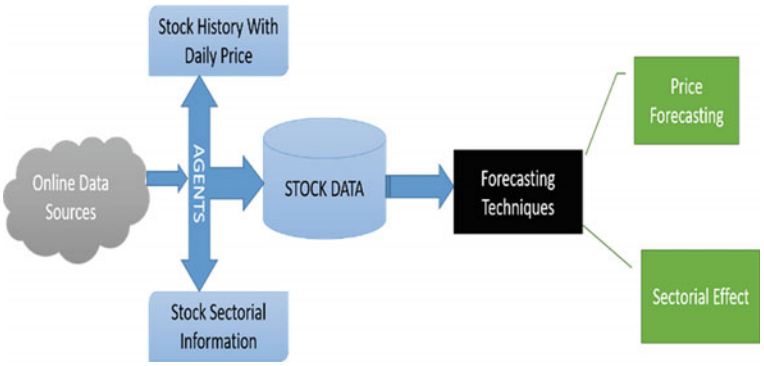


Fig. 1 Proposed model

The model usually deal with such an accidental situations. The generalize model represented and the outcome is based on time frame given predictive values.

The model applied with forecasting techniques and compare the results. The model also intuits the sectorial effect on stock market. In this model two prediction techniques are applied. First is Linear Regression and other is XGBoost (Extreme Gradient Boosting).

1. Linear Regression: Linear Regression is a one of the easiest and prevalent machine learning algorithm of supervise learning. It is mainly used for statistical analysis and predicting the contiguous depended variable on any numerical data pattern. In simple linear regression, single independent variable is used for estimation of single dependent variable where in linear regression, you can use more than one independent variable to predict one dependent variable. In Linear regression, line of paramount fit is adjacent to data points that have strong correlations. It is observed that there is little difference between observed data and best fit line where there is a large difference between observed data and best fit line.
2. XGBoost: Angle boosting is a procedure to change over feeble students to solid students, in an iterative style. The name XGBoost alludes to the building objective to push the restriction of computational assets for supported tree calculations. Since the time its presentation in 2014, XGBoost has demonstrated to be an extremely amazing AI strategy and is normally the do by calculation in many ML rivalries. We will prepare this model on the train data, tune its hyperparameters utilizing the approval set, lastly apply the XGBoost model on the test set and report the outcomes. Evident highlights to utilize are the balanced shutting costs of the keep going for n no of days, just as the volume of the keep going N days.

4 Model Implementation and Result Discussion

The model is implemented through Python advance libraries. The linear regression is applied to predict the price (Figs. 2, 3 and 4).

After representing all the graph, the sectorial effect also represented. Here from one sector one graph is represented with both model LR & XGBoost Regressor. Same the XGBoost repressor work well.

5 Conclusion

By applying the LR model and XGBoost Repressor methods to the selected sector wise stock the result is mentioned in Table 1. By seeing the price graph the IT and Pharma sector had a positive impact of COVID-19. Also by performing the experimental study on sectorial stock the price movement of pharma sector based

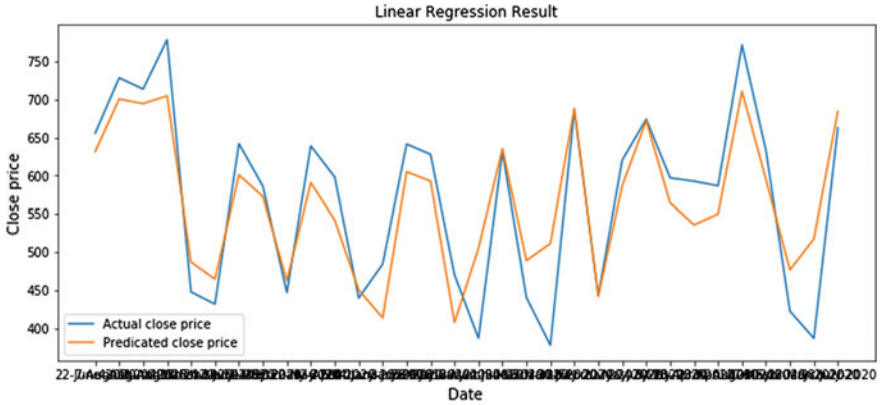


Fig. 2 Cipla LR model results

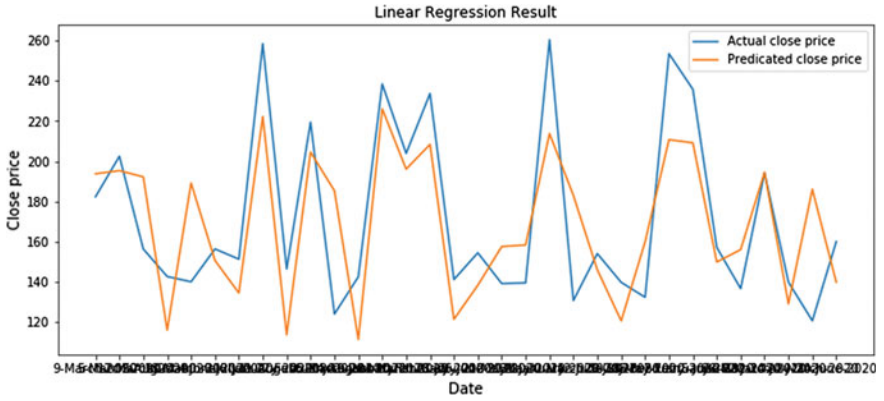


Fig. 3 DLF LR model results

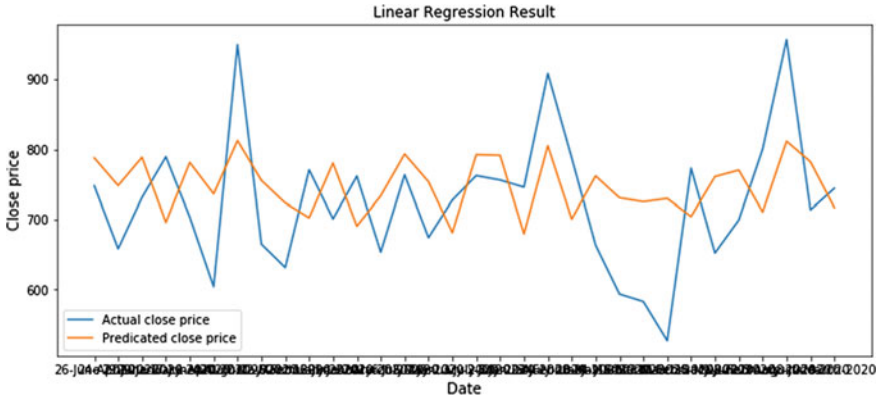


Fig. 4 INFOSYS LR model results

Table 1 Result tunning

Stock name	Applied techniques	Error rate	R2 score
Cipla	LR	MAE: 34.037093535876494 MSE: 1884.891276668207 RMSE: 43.41533457971051	R2: 0.7908314474363111
Cipla	XGBR	MAE: 8.844686126708982 MSE: 167.211629489276 RMSE:12.931033581631285	R2: 0.9867333571473805
Sun Pharma	LR	MAE: 20.506294061387358 MSE: 672.1126015881239 RMSE: 25.92513455294155	R2: 0.6046625174699142
Sun Pharma	XGBR	MAE: 8.623699569702154 MSE: 137.310789934639 RMSE:11.717968677831452	R2: 0.9576681916972009
Infosys	LR	MAE: 85.10610809280237 MSE: 8960.801711298416 RMSE: 94.66151124558712	R2: 0.8156091732315329
Infosys	XGBR	MAE: 19.007185745239255 MSE: 766.5569371288778 RMSE:27.686764656219367	R2: 0.9454587578038609
TCS	LR	MAE: 165.94244781872882 MSE: 35323.19896689452 RMSE:187.94466996138655	R2: 0.773715902310364
TCS	XGBR	MAE: 39.25653457641606 MSE: 2765.7529281720717 RMSE: 52.59042620260908	R2: 0.8934830703822343
DLF	LR	MAE: 23.153914146637398 MSE: 691.8409337374945 RMSE: 26.30286930617066	R2: 0.4942341745881188
DLF	XBGR	MAE: 4.4371897697448714 MSE: 42.02437339765689 RMSE: 6.482620874126211	R2: 0.979719627545672
Godrej	LR	MAE: 110.74135455477992 MSE: 17,249.8388775759 RMSE:131.33864198161902	R2: 0.15991570811882272
Godrej	XBGR	MAE: 21.91278648376465 MSE: 1137.703189674835 RMSE: 33.72985605772481	R2: 0.9559880775063512

stock had really most positive impact. Other sector IT is also influenced positively. The Infrastructure based companies get reverse effect. By comparing the R2 score of predictive modelling techniques XGBoost Repressor gives more accurate results. By adding sentiments, the proposed model can be more accurate.

References

1. Alam, M.N., Alam, M.S., Chavali, K.: Stock market response during COVID-19 lockdown period in India: an event study. The J. Asian Financ., Econ., Bus. **7**(7), 131–137 (2020)
2. Gormsen, N.J., Koijen, R.S.: Coronavirus: Impact on stock prices and growth expectations. University of Chicago, Becker Friedman Institute for Economics Working Paper, (2020-22) (2020)
3. Raza, S., Munir, S.: Did demonetization affects indian stock market? A short-term analysis. A short-term analysis (January 19, 2020).
4. Ozili, P., Arun, T.: Spillover of COVID-19: impact on the global economy. SSRN Electron J (2020). https://doi.org/10.2139/ssrn.3562570
5. Ducharme, J.: World Health Organization Declares COVID-19 a 'Pandemic.' Here's What That Means. Time Online Newspaper, Retrieved April 20, 2020, from https://time.com/5791661/who-coronavirus-pandemic-declaration/ March 11 (2020)
6. DH Web Desk: Coronavirus India update: State-wise total number of confimed cases, deaths on April 17. Deccanherald Online Newspaper, Retrieved April 20, 2020 (2020), https://www.deccanherald.com/national/coronavirus-india-updatestate-wise-total-number-of-confirmed-cases-deaths-onapril-17-826340.html
7. Baker, S.R., Bloom, N., Davis, S., Kost, K., Sammon, M., Viratyosin, T.: The Unprecedented Stock-Market Reaction to COVID-19, KelloggInsight. Kellogg School of Management, Northwestern University, Chicago, IL (2020)
8. Ramelli S, Wagner AF (2020) Feverish Stock Price Reactions to COVID-19, Swiss Finance Institute Research Paper No. 20-12. https://doi.org/10.2139/ssrn.3550274
9. Nguyen, D.D., Pham, M.C.: Search-based sentiment and stock market reactions: an empirical evidence in Vietnam. J. Asian Financ., Econ. Bus. **5**(4), 45–56 (2018). https://doi.org/10.13106/jafeb.2018.vol5.no4.45
10. Olsen, A.A., McLaughlin, J.E., Harpe, S.E.:. Using multiple linear regression in pharmacy education scholarship. Curr. Pharm. Teach. Learn. (2020).
11. Iqbal, Z. et al.: Int. J. Eng. Res. Appl. www.ijera.com 3(6), 855–867 (2013). ISSN: 2248-9622,
12. Lee, J.W., Brahmasrene, T.: An exploration of dynamical relationships between macroeconomic variables and stock prices in Korea. J. Asian Financ., Econ. Bus. **5**(3), 7–17 (2018). https://doi.org/10.13106/jafeb.2018.vol5.no3.7
13. Fama, E.F.: Effient capital markets: a review of theory and empirical work. J. Financ. **2**, 383–417 (1970)
14. Salamudin, et al.: Intangible assets valuation in the Malaysian capital market. J. Intell. Capital **11**(3), 391–405 (2010). Emerald Group Publishing Limited. https://www.emeraldinsight.com/1469-1930.htm

Lane Detection for Visual Assistance of Self-driving Vehicles for Structured and Unstructured Roads

Vishal B. Pattanashetty, Venkatesh Mane, Sanjana S. Hurkadli, Nalini C. Iyer, and Shweta Kore

Abstract Safety is a very critical aspect of any module designed. In automotive domain, ensuring safety of drivers and co-passengers is very vital, along with this ensuring that no harm is caused by the malfunctioning behavior of components and long-life of automotive parts is required. Driver assistance is a combination of software and hardware system application to help the driver in assisting while driving, thus designed to provide safety and prevent harm. Driver assistance includes many applications in assisting parking, lane detection, dent detection, pot-hole detection, etc. We have implemented a combined solution for lane detection, vehicle detection, and steering assistance providing information of lane curvature and direction. The system is intended to automate, enhance safety, and ease out the driving process. The methodology implemented can be extended to other applications. The driving test for driver's license with specific parameters can be used to test the driver's performance automatically. With obstacles detected, it becomes easy to avoid collision by providing a warning system.

Keywords Object detection · Lane detection · Self-driving vehicles · Camera calibration

V. B. Pattanashetty (✉) · V. Mane · S. S. Hurkadli
KLE Technological University, Hubballi, India
e-mail: vishalbps@kletech.ac.in

V. Mane
e-mail: mane@kletech.ac.in

N. C. Iyer · S. Kore
JSS Dharwad, Dharwad, India
e-mail: nalini.c@kletech.ac.in

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_27

1 Introduction

According to the census carried out by the World Health Organization globally, there were around 1.25 million traffic fatalities in the year 2013 and keeps increasing exponentially year after year. So millions of lives could be saved around the world every decade with fully autonomous vehicles. Research shows about 94% of the fatality rates caused due to human error are reduced due to autonomous vehicles. Road deaths due to the accident wipe out a city per year that is over 1.3 lacks per year. Sixty percentage deaths are of the young people in age of 18–35. Sadly, India loses 3% of its GDP due to road accidents.

People are becoming aware of the need for autonomous vehicles due to many reasons, for example, human error often involves a lack of vigilance. They are often distracted whether reason being texting, picking up a call, having a mind elsewhere apart from driving, or paying attention to the back of the seat. The accidents led many engineers to come up with solutions for the future also with the fact that the efficient driving of self-driving vehicles leads to a decrease in carbon emissions. Being impaired or having harsh driving conditions also plays a major role in human error which leads to accidents. Hence, the solution to avoid these accidents is through integrating assist systems in automobiles which results in self-driving vehicles. A simple definition of a self-driving car is that it is capable of visualizing, sensing, and navigating itself through different sensors integrated into the automobile and not seeking any human input. Self-driving cars obtain data through various sensors like cameras, RADAR, etc., and process the following through many computer vision techniques.

This paper is organized into six sections, second section deals with literature survey where it will try to understand various techniques and tools used for lane detection, then in next section, we look for design methodology adopted, then experiments, then to results, and finally will come to see the future implementations.

2 Literature Survey

In the paper 'Self-driving vehicles,' a self-driving car is also referred to as a robot-car, driverless-car that is capable of taking inputs from its surroundings through different sensors integrated onto it, process the data and guide the driver or feed it to other systems.

In 2017, semantic instance segmentation approach focused on tackling the problem with discriminative loss function, so that the pixels of the same instance lie closer and are easy to cluster. VPGNet tackles cases of lane detection and marking in condition of low light, rain, and night. With this network, vanishing point is estimated with a single pass for the classification of lane. In 2019, with cascaded CNNs, the application of real-time lane detection is focused with two cascaded neural networks

with first one for extracting boundary instances, followed up a classification for estimating the classes of multiple lanes. ERFNet is focused as row-wise classification task of lane detection. It deals with more over lane-marking detection. In [5], the unstructured road is taken into consideration by dividing the regions using mean-shift segmentation and Bayes rule to estimate the points for the boundaries of the lane. In [6], here also, the unstructured road is detected by using vanishing point estimation. Other lane detection being still based the basic idea revolving around training networks to detect lane, segmentation, lane marking or classification, and clustering.

With over a million deaths a year worldwide caused mainly due to human error, traveling by car is considered to be one of the deadliest modes of transportation. Therefore, self-driving cars eliminate the possibility of human error that causes accidents as well as other fatalities and injuries.

There are no fully automated vehicles legally operating in the world, but there are several self-automobiles with different variations implemented in trucks, cars, and other automobiles that have different assist systems and are usually self-driving prototypes. Though still in its primary stage, self-driving technology is implemented in almost every automobile for ease and could radically change the transportation system, business, and economy of a country. According to automaker and technology company, level 4 self-driving cars could be in the markets in the next upcoming years.

3 Design Background

Human error is a major factor in increasing road accidents especially when distracted due to common reasons like picking a call or paying attention to your surroundings instead of the road. Also, allowing the visually impaired to have the convenience of traveling without aid for others. Addressing the issue requires monitoring of the surroundings through different sensors to obtain input data which is processed through different machine learning algorithms and computer vision techniques. This intern results in an assist system that is integrated into the automotive vehicles that guide the individual accordingly. Hence, visual assistance can be provided by creating a visual assist system that includes path guidance through lane detection and warnings through object detection, thus enabling the driver to help in the above-mentioned problems thus saving lives.

4 Experimental Setup

4.1 Evolved Computer Vision Approach

Camera Calibration Image Distortion Removal The first step is to calibrate the camera. This is done due to the distortions arising due to the physical properties of the lens. Therefore, we need to find the calibration matrix for the required lens along with the distortion coefficient for the camera that was used to take a video of the road. The distortions are caused due to the convex nature of the lens that curves the light rays entering the pinhole of the camera which distorts the video. Therefore, lines appear distorted that are straight in the real world. Fortunately, we can correct these distortions by measuring them. We use multiple images of a reference object which is not distorted. Here, we take the multiple images of a chessboard on flat surface taken by a camera. The OpenCV has a suitable method called find checkerboard corners that identifies black and white squares and intersection. We choose the chessboard since it provides high contrast images. After reverse-engineering the points, it finds the distortions traced to the reference image (Fig. 1).

Our Approach for ROI Estimation This defines a way to estimate the proper ROI and points for the perspective transform. A sliding kernel of size ($z \times z$) is used on the bottom of the image and slided horizontally to estimate the start points, and this gives the bottom ROI. The result of thresholding the image gives the estimation of dividing the image horizontally to estimate the ROI of top points. With the bottom ROI points as reference and diving the image into half, and sliding the window on top, but this time from the middle, gives the top ROI points. These ROI can be used for both Hough transform and perspective wrap accordingly. The kernel size that we used here is 5×5 (Fig. 2).

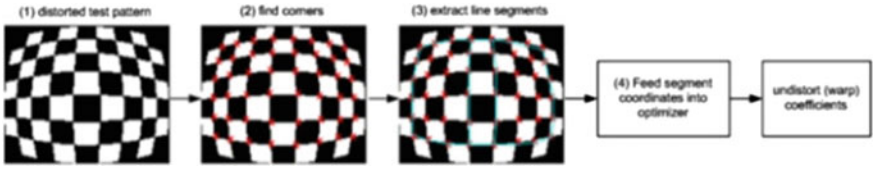


Fig. 1 Camera calibration using checkerboard

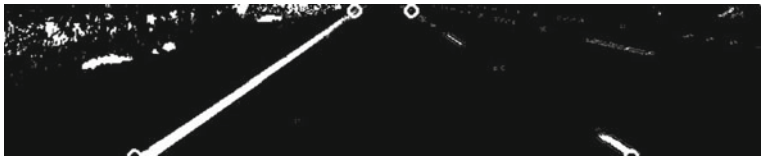


Fig. 2 ROI estimation

Fig. 3 Perspective transform from vehicle to sky view

Perspective Transformation After the correction of the video, we have the undistorted video of the road in the view of the vehicle. We can transform the video or warp the video from vehicle to sky view, where the location coordinates are necessary. To obtain the coordinates, only one frame from the video is necessary through the vehicle's perspective and the locations of desired perspective points. We can refer these points as source and the destination points (Fig. 3).

Gradient and Color Threshold Here, we are not entailing altogether the data from all the pixels of a road picture image to find the curvature of the road. The color of the road lane lines is yellow and white. So we can veneer the yellow and whites shades of the picture image. There are some cases where the white and yellow lines are not lanes or the lanes have different appear distinct from white and yellow, then we use what we call as a Sobel operator. A Sobel operator trials the level of variation in rate and its path amid two locations/points in an image. This is also denoted to as gradient or derivative. The color is represented in the HSL format. The hue rate is the superficial depiction of red, green, and blue. The overload value depicts whether the color is bright or dull. Lightness depicts how closer to the color the white is. The yellow lanes can be separated with a combination of lightness and saturation, while the white ways are singled out obligating high lightness rate irrespective of saturation and hue.

Curve Fitting Since lanes are usually curved in nature. We can represent these lanes using a second-degree polynomial function $x = y^2 + By + C$. We find the coefficients for each lane using the built-in function polyfit. The function needs to be feed with points which in return give the coefficients of a polynomial for a specified curve to best fit the given points. To decide which pixels are part of the lane, a basic algorithm can be implemented. Histograms of all columns of the lower half of frames of a video are taken and can get two similar graphs with two peeks. The peaks indicate the x position of the base of the lane. Therefore, it can be used as a starting point. To obtain the points, the sliding window technique can be used which uses one window on top of the other which in return follows the lane in the frame. The pixel inside 'window' is marked as "pixel of interest." In order to obtain an indication of the base-point of the next window, we average the values of pixels and give these points to the polyfit function. This results in obtaining the coefficients of second-degree polynomial which can be used to detect the curvature of lanes. To approximation

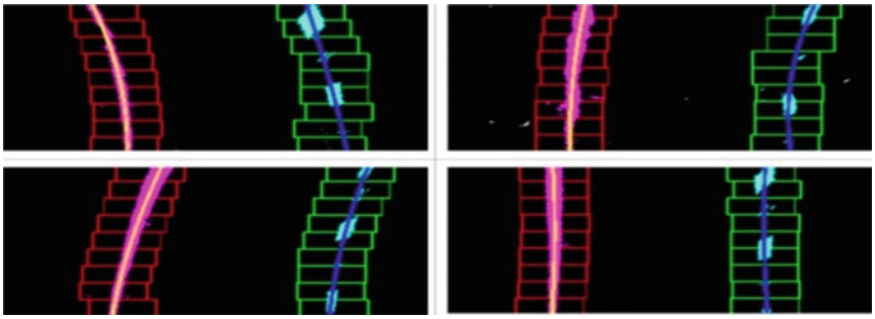


Fig. 4 Using sliding window technique

Fig. 5 Lane projection and object detection viewed on frame

the vehicle location, we can compute the lane width in pixels at the bottommost sides of the image (nearby to the camera). Approximately this remained 1000 pixels extensive. That infers that individually pixel resembles to 3.7 m in actual measure. Then, we can calculate the quantity of pixels that the image focus is offset from the lane center and exclaim the offset in meters (Fig. 4).

Lane Projection and Object Detection to View Vehicles The object detection is obtained through TensorFlow libraries which consist of a pretrained neural network. With the information now gathered, we can display the radius of the curvature of lanes and guide the driver in direction or give them proximity warning according to object detection (Fig. 5).

4.2 Lane Detection Using CNN Approach

In this method, we address the problem of lane detection by teaching the machine to learn the pattern of lane. This approach holds good for both structured and unstructured roads. It is based on the idea of pixel-wise segmentation. But here, we are

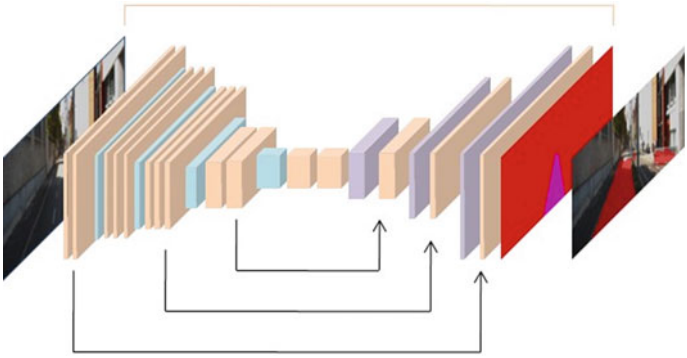


Fig. 6 Lane detection model

Fig. 7 **a** Sample of input image, **b** sample of ground truth image

interested only in the road class for lane detection. The model is trained to learn the pattern of lane and output the label of pixel-wise classification, the parameters of the training are tuned with the output of addition of classification label and input and validated with the expected output. The dataset used is Kitti road dataset. Validating on the addition output gives the advantage of predicting efficient pixel information and avoiding loss of information. Softmax function is used for classification label and mean square error for output error estimation (Fig. 6).

In the model, the pink layer is being the convolution and blue max pooling, so that the local information is ignored at first. The green being the up sampling to output the label of the dimensions of input. The skip connections in black are needed to restore the local information (Fig. 7).

5 Result

5.1 Result of Experimental Setup A

The project provides a reliable solution to:

Fig. 8 Result showing lane detection, object detection, and curvature information

Finding the lane of the road, determining the curvature of the road, detecting the vehicles on the road, warning the driver of nearby vehicles. Here, the radii of curvature of the road are estimated from the formula taking the tangent of the assumed circle drawn on the curves of the road. The curvature here is considered as average of both the radii (Fig. 8).

$$\text{Curvature} = \frac{(\text{Left ROC} + \text{Right ROC})}{2}$$

Result of Experimental Setup B (Fig. 9)

6 Conclusion

The project's original idea and goals have been achieved to the desirable level with advanced lane, object, and curvature detection, thus helping a big role in self-driving vehicles. Because we are able to finding the lane of the road. Determine the curvature of the road, detecting the vehicles on the road and warning the driver of nearby vehicles. The project has also extended its application to detecting the nearby vehicles that pass by the source vehicle, thus alerting the driver of the surrounding. The project holds a major contribution in automating the vehicle industry with computer vision approaches for noncritical tasks of vehicle at a much efficient costs compared to the more expensive sensors.

Fig. 9 Result showing **a** input image, **d** output images

References

1. Pavaskar, S., Budihal, S.: Real-time vehicle-type categorization and character extraction from the license plates. In: Mallick, P., Balas, V., Bhoi, A., Zobaa, A. (eds.) Cognitive Informatics and Soft Computing. Advances in Intelligent Systems and Computing, vol. 768. Springer, Singapore (2019). https://doi.org/10.1007/978-981-13-0617-4_54
2. Suneeta, V.B., et.al.: Facial expression recognition using supervised learning. In: Smys, S., Tavares, J., Balas, V., Iliyasu, A. (eds.) Computational Vision and Bio-inspired Computing, ICCVBIC 2019. Advances in Intelligent Systems and Computing, vol. 1108. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-37218-7_32
3. Bhagyashree, K., Iyer, N.C., Arlimatti, T., Shet, R.M., Preeti, P.: Object detection techniques using deep learning: a survey for real-time applications. Test Eng. Manag. **82**(Jan/Feb 2020), 5485–5493 (2020)
4. Dhanakshirur, R.R., Pillai, P., Tabib, R.A., Patil, U., Mudenagudi, U.: A framework for lane prediction on unstructured roads. In: Thampi, S., Marques, O., Krishnan, S., Li, K.C., Ciuonzo, D., Kolekar, M. (eds.) Advances in Signal Processing and Intelligent Recognition Systems, SIRS 2018. Communications in Computer and Information Science, vol. 968. Springer, Singapore (2019). https://doi.org/10.1007/978-981-13-5758-9_16
5. Iyer, N.C., Pillai, P., Bhagyashree, K., Mane, V., Shet, R.M., Nissimagoudar, P.C., Krishna, G., Nakul, V.R.: Millimeter-wave AWR1642 RADAR for obstacle detection: autonomous vehicles. In: Innovations in Electronics and Communication Engineering. Lecture Notes in Networks and Systems, vol. 107. Springer, Singapore (2020)

Review of System Development Life Cycle (SDLC) Models for Effective Application Delivery

Oluwaseyi Ezekiel Olorunshola and Francisca Nonyelum Ogwueleka

Abstract There are different system development models, tools, and applications that have been designed and developed before the use of computers for processing of information, and there is an increase in demand of software with cheaper cost, having more functionality, faster delivery, and of high quality than how it was previously. Therefore, there is need to know the methods and their applications that would conform to the organizational requirements for successful system deployment. Each method has constructive criticisms with various advantages and disadvantages to the system that are of importance for deploying software in a manner such that it helps in good decision making on a chosen method for delivery within deadline and proper quality. This paper explained four commonly used system development methods namely; waterfall, iterative, agile, and rapid application development (RAD). It thereafter categorically assesses these methods on a comprehensive set of features and made an alternative analysis of the applicability of the methods based on system development life cycle tools which are requirement, design, implementation, and testing. This paper describes the system development life cycle (SDLC) tools and applications for successful system deployment and the constructive comparison that should serve as a tool in model selection for system development.

Keywords System development life cycle · Development models · System development · Constructive comparison

1 Introduction

SDLC is a sequential process which explains the main stages and activities of development cycles. It is used by system analysts, developers, and designers to plan and implement application and delivery of the systems or products on time and within an inexpensive budget. With vast development methodologies that we have today, it is often a difficult task to select a fitting plan of action. Sometimes, it is inescapable

O. E. Olorunshola (✉) · F. N. Ogwueleka
Computer Science Department, Nigerian Defence Academy, Kaduna, Nigeria

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_28

to compare multiple methodologies in order to go for the best methodology to fit in a single project.

SDLC covers the identification of any need to work out a business requirement, developing a business case for the proposed solution and also the implementation of the system solution. The aim of this paper is to describe the SDLC methodologies, tools, and applications for successful system deployment.

2 Literature Review

According to Land et al. [7], the project life cycle comprises of all the activities of the project, while SDLC focuses on realizing the product requirements. Unhelkar [12] said that designers, analysts, programmers, end users, and management information system (MIS) managers are confronted with a wide array of methods and are poorly informed on their utility and desirability. Therefore, a comprehensive approach is important for an understanding of the attributes and appropriateness of these methods. The initial concepts of SDLC were originated in the 1960s to develop large-scale functional business systems in an age of large-scale business conglomerates. In the earliest days of computer programming, the only models that were used to develop complex things were in manufacturing and construction industries [1]. Thus, it made a lot of sense that the organized approaches used in those industries should be applied in developing computer systems as well.

SDLC is a very traditional development process which goes in a sequential manner from start to finish. Some overlapping procedures are inevitable such as testing and refining (which might not be available in some methods). Nevertheless, crossing between major phases is not common.

Over the last fifty years, computer systems have been taking important roles in corporations. From sending mails with postmen to sending email via the Internet, from filling paper applications to electronic applications, from auditing financial logbooks to the spreadsheets stored in the enterprise systems, every aspect is closely related with information technology [2].

3 System Development Life Cycle (SDLC)

Decades back, software development concepts came up and new perceptions and designs have emerged on customer-oriented applications/software and solutions to problems. Each approach has its advantages and disadvantages, strengths, and weaknesses. It is realistic that one single solution can no longer fit in man organizations due to different backgrounds, structures responsibilities, desires, and goals. Different organizations have various policies and procedures that work for them and as such will require different approaches and solutions. Yet, shared aims on each system development stage can be found. There should not be many variations on how the works

are described, organized, and managed with different organization backgrounds and requirements. Therefore, the modern SDLCs are adequately flexible to be used across different types of businesses, products, and services.

Not limited to the listed models, there are various models used in the SDLC process some common models are waterfall, iterative, agile, and rapid application development (RAD) models.

The various SDLC models are designed for different types of projects. Each model follows a series of unique steps that best fits its project type to ensure the success of system development [3]. Waterfall, iterative, agile, and rapid application development (RAD) are identified as the most popular models being used in the industry, and they will be introduced one by one in details as follows:

3.1 Waterfall Model and Its Phases

Waterfall model is the first, best known, and mostly used methodology. It is a step by step life cycle model that is simple to use and understand. Each phase must be completed before another phase starts. Figure 1 shows the different phases of a waterfall model.

Requirement. This phase mainly focuses on communicating with business users to gather and analyze requirements. Project managers try their best to understand and analyze the business, and capture all the details of the user's needs [4].

Design. With the business case documentation in hand prepared in requirement phase, business analysts evaluate and start on the logical design by making use of the information and requirements that are collected by the project managers. Based on the high-level design which has fulfilled all the user requirements, system analysts transform the high-level design to the physical design which put hardware and software technology into consideration. The system architecture is defined at design phase as well.

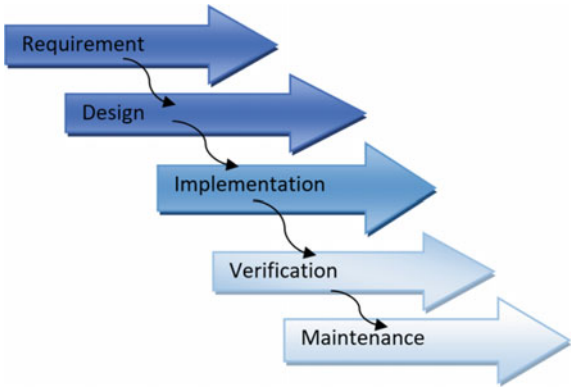


Fig. 1 Phases of a waterfall model

Application of waterfall model: Waterfall model should be selected as the SDLC methodology to use when the project has limited amount of unconfirmed requirements, fully documented system, and static or standard technologies, and management can provide enough resources and experts to pick up the role at each phase. Waterfall is not suitable for complex projects.

Tiwari and Rathore [10] mentioned some advantages of waterfall methodology as:

- Easier to manage as there is a clear schedule for each stage that gives clear milestones.
- Provide extensive documentations.

Tiwari and Rathore [10] stated some disadvantages of waterfall methodology to be:

- Does not have scope change or requirement change.
- Does not preview the product until the deployment phase.

3.2 Iterative Model and Its Phases

Iterative model works on the simplified requirements which are found in system or application requirements. The "product" is iteratively enhanced and evolved to the final product for deployment. It is called a "build" for each iteration. At each build, design changes and new functionalities are added to the product. Figure 2 shows the phases of an iterative model.

Requirement. Same as waterfall model.

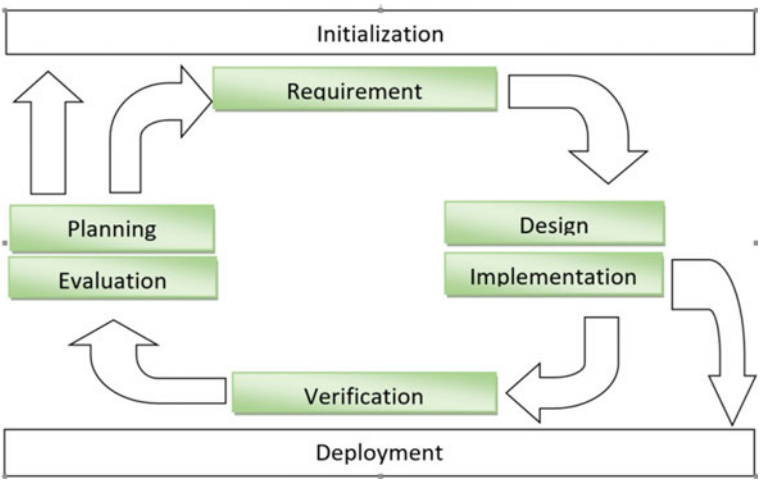


Fig. 2 Phases of the iterative model

Design. Same as waterfall model. However, there is the design which holistically records how the system is going to be implemented. In addition, subset of designs can be modified after every round of build. Therefore, the subset of designs is not finalized until reaching the deployment phase [5].

Applications of Iterative Model. Appropriate situations for using iterative model as the SDLC methodology are: when major requirements are defined but the minor details might evolve as time goes on, new technologies are being used and there is a learning curve for the solution developer to learn, etc.

Tousif et al. [11] mentioned some advantages of iterative methodology as:

- Easier to start on a complex project.
- Preview of the project periodically is possible.

Tousif et al. [11] stated some disadvantages of iterative methodology to be:

- There is need for more management work to ensure that each build meets the standard overlapping implementation could be chaotic.
- There is need for more involvement of business users.

3.3 Agile Model and Its Phases

Agile model extends the advantages of iterative model and aims at user satisfaction and product adaptability by rapid delivering of product. From requirement phase to deployment phase, agile model breaks the product into smaller builds. Thus, each build contains some new features, and for the very last build, it contains all the required features of the software or system [6]. Figure 3 shows the phases of an agile model.

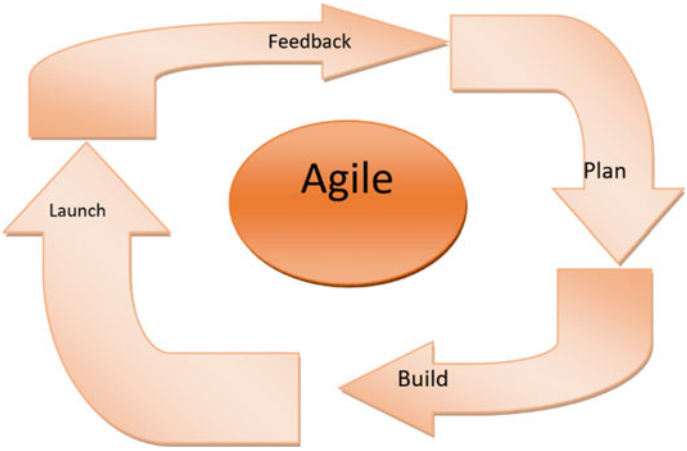


Fig. 3 Phases of the agile model

Requirement. As requirements cannot be gathered completely at the beginning, close relation with business users is necessary to gather feedbacks after every release.

Design. System requirement specification and design specification documentation are expected to be short and simply listing out what is covered in the current build. Agile model does not pay too much attention to the documentation as waterfall and iterative models. Although the same set of documentations is expected to be ready at each phase, the information that can be found in each documentation is very limited [7].

Applications of Agile Model. Appropriate situations for using agile model as the SDLC methodology are when there is no detailed information provided from business users, when the project is features driven, when product requirements change dynamically, and when there is close collaboration within the team and close relationship with business users.

Dennis et al. [4] mentioned some advantages of agile methodology as:

- Realistic approach—what you need the most, implement first; what is less important, implement last.
- Less documentations and rules to follow: more flexibility.

Dennis et al. [4] stated some disadvantages of agile methodology to be:

- Very high risks for maintenance and extendibility.
- Not suitable for complex and core projects.

3.4 Rapid Application Development (RAD) Model and Its Phases

Rapid application development (RAD) focuses on gathering user requirements through workshops, tests on the pre-released prototypes conducted by users, and then reusing the prototypes to further develop the product. RAD model is basically assembling the working parts together to generate the product in a limited amount of time for business users to quickly provide feedback regarding the requirements [8]. Figure 4 shows the phases of a RAD model.

Applications of RAD Model. RAD model should be selected as the SDLC methodology to use when business users tend to change their requirements from time to time and delivery can be done part by part.

Kosta et al. [6] mentioned some advantages of the RAD methodology as:

- Can tolerate frequent requirement change.
- Measurable progress.

Kosta et al. [6] stated some disadvantages of the RAD methodology to be:

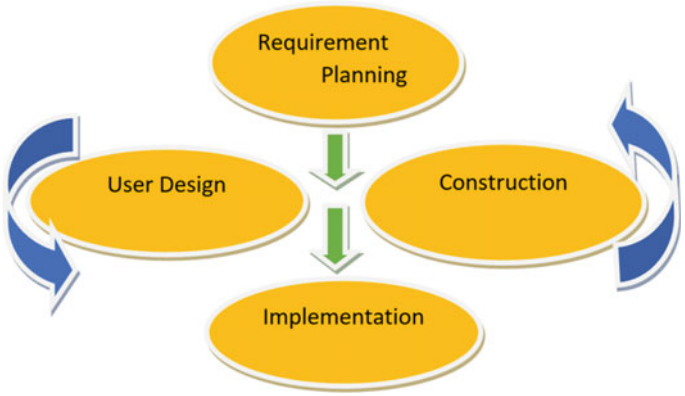


Fig. 4 Phases of the RAD model

- Could have performance or technical issues with automation.
- Costly automation tools.

4 Deduction

The waterfall model is best suitable when the requirements are well known and the processes are well defined. The team may have done a similar project in the past, so they are familiar with the processes [9]. The iterative model is suitable for large projects when the duration of the project is long and there are speculated changes that may take place during the project's lifecycle [10, 11]. The agile model is very suitable when the requirements changes and there is need for a feedback loop between the customers and management. It not suitable for complex and large projects and for a large team [11, 12]. RAD is best suitable when the team's technical skills are not advanced and when there is need for user involvement and modularization of the project [13]. Table 1 shows the comparison of the waterfall, iterative, agile, and RAD models in terms of the requirements, cost, duration, user involvement, probability of success, team expertise, flexibility to change, and risk involvement.

5 Conclusion

This paper has identified four major SDLC models and their applications as well as some of their numerous advantages and disadvantages. For those in most common use, a comprehensive evaluation was made to enhance successful system deployment. Different organizations, however, have the responsibility of determining which of these methods would be used to meet their needs based on their policies, structures,

Table 1 Comparison of the waterfall, iterative, agile, and RAD models

S/N	Features/Models	Waterfall	Iterative	Agile	RAD
1.	Requirements	Easily understood	Easily understood	Easily understood	Easily understood
2.	Cost	Low	Low	Very high	Low
3.	Duration	Long	Very long	Short but varies	Short
4.	User involvement	Beginning of the project	At each phase	High throughout the project	Beginning of the project
5.	Success probability	High	High	Very high	Good
6.	Team expertise	High	High	Very high	Medium
7.	Flexibility to change	Very low (rigid)	High	Rigid	High
8.	Risk involvement	High	Can be managed	Reduced	Very low

and goals. Selecting a particular method could be a difficult task for a company; however, it is recommended that two or more methods can be used on a single project. This could give the added advantages of both methods. However, the cost of implementing such a hybrid of methods would be higher.

References

1. Balaji, S., Murugaiyan, M.S.: Waterfall vs. V-model vs. agile: a comparative study on SDLC. Int. J. Inf. Technol. Bus. Manag. **2**(1), 26–30 (2012)
2. Berger, H., Beynon-Davies, P.: The utility of rapid application development in large-scale, complex projects. Inf. Syst. J. **19**(6), 549–570 (2009)
3. Berra, Y.: Software Development Life Cycle (SDLC). DePaul Univ. 2 12, 2012. https://condor.depaul.edu/jpetlick/extra/394/Session2.ppt (2012)
4. Dennis, A., Wixom, B.H., Roth, R.M.: Systems Analysis and Design. Wiley & Sons (2018)
5. Highsmith, J., Cockburn, A.: Agile software development: the business of innovation. Computer (Long. Beach. Calif) **34**(9), 120–127 (2001)
6. Kosta, E., Pitkänen, O., Niemelä, M., Kaasinen, E.: Mobile-centric ambient intelligence in health-and homecare—anticipating ethical and legal challenges. Sci. Eng. Ethics **16**(2), 303–323 (2010)
7. Land, S.K., Smith, D.B., Walz, J.W.: Practical Support for Lean Six Sigma Software Process Definition: Using IEEE Software Engineering Standards, vol. 70. Wiley (2012)
8. Petersen, K., Wohlin, C., Baca, D.: The waterfall model in large-scale development. In: International Conference on Product-Focused Software Process Improvement, pp. 386–400 (2009)
9. Tiruneh, T., Mishra, M.K.: Analysis and performance evaluation of requirement elicitation techniques. Int. J. Comput. Sci. Eng. **6**(4), 118–123 (2018)
10. Tiwari, S., Rathore, S.S.: A methodology for the selection of requirement elicitation techniques **3**(2), 212–217 (2017). arXiv Prepr. arXiv1709.08481
11. Tousif, R., Khan, M.N.A., Riaz, N.: Analysis of requirement engineering processes, tools/techniques and methodologies. Int. J. Inf. Technol. Comput. Sci. **5**(3), 40–48 (2013)

12. Unhelkar, B.: The Art of Agile Practice: A Composite Approach for Projects and Organizations. Auerbach Publications (2016)
13. Yousuf, M., Asger, M.: Comparison of various requirements elicitation techniques. Int. J. Comput. Appl. **116**(4) (2015)

Assessing the Impact of Information and Communication Technology (ICT) on Work Efficiency of Graduates

Oluwaseyi Ezekiel Olorunshola and Francisca Nonyelum Ogwueleka

Abstract Information and communication technology circumscribes all forms of technology involved in information dissemination. ICT has been used in different sectors which include oil and gas industry, banking sectors, educational sectors, real estate, marketing, and so on. This work seeks to assess how much ICT bears on the work efficiency of graduates. This research was carried out through a survey. One hundred and one respondents responded to the questionnaire. From the result, it was seen that all of the respondents were in one accord that ICT has improved their work efficiency.

Keywords Information · Communication · Technology · ICT · Work efficiency · Graduate

1 Introduction

Most graduates today are familiar with information and communication technology (ICT). Everyone uses it in his or her daily activities one way or the other. However, the fact that everyone uses ICT does not essentially correlate with its efficient usage.

ICT has become an indispensable part of our society today. Its value in people's lives has increased in high geometric rates, and the trend is expected to continue. "Information Technology is the new oil." ICT literacy has become a functional, fundamental prerequisite for people's work as well as their social and personal interactions. Enhancing work efficiency of graduates could be achieved using ICT as it could be used to enhance teaching and learning process, improve communications, enhance collaborations, and increase productivity [1].

ICT is basically the effort of our society to teach all citizens: current and emerging, beneficial knowledge and skills in communications devices and computing, the software that make them work, applications that run on them, and systems that are built with them [2].

O. E. Olorunshola (✉) · F. N. Ogwueleka
Computer Science Department, Nigerian Defence Academy, Kaduna, Nigeria

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_29

ICT is a very broad field and as such certain topics or technologies. For this research, "ICT can be believed to mean all networking components, applications, devices, and systems that are combined to let organizations and people to interact within the digital world" [3].

Based on the overview of literature on ICTs, the collective phenomena of ICT connectivity and individual work productivity were a research gap worth exploring.

2 Review of Related Studies

ICT has affected and enhanced work efficiency in many ways. The agriculture, library, and higher education sectors are major sectors that have been impacted. Palvalin, Lönnqvist, and Vuolle in their research work on the analysis of the impacts of ICT on productivity of knowledge work used a case study in the medium European tele-operator that provides ICT services for the consumer and enterprise markets and found that ICT can be used to eradicate non-value-adding tasks or make them more proficient. They also concluded from their research that ICT-based productivity benefits are highly dependent on how well the ICT service is utilized [4].

Makura examined students' perceptions of ICT usage by students and staff at a South African University. They discovered among other things, that students observed ICT, especially the computer as having positive impact on their academic performance and other curricular issues [5].

Adesote and Fatoki [6] examined the influence of ICT in teaching and learning of History in the Senior Secondary School in the twenty-first century. They concluded that the correct use of ICT can impact and modify traditional methods of learning and teaching of history, hence ensuring education of high quality [6].

De Wet et al. investigated the impact of ICT on the work of employees and their individual lives. It was discovered that although ICT is generally perceived as having positive impact, employees should make a deliberate decision in the management of time spent on ICT to reduce the negative impact on their work and individual lives. It was also discovered that the prevailing positive occurrences of ICT outweigh the negative ones, and findings also insinuate that as the quantity of communication increased, the standard of conversations declined [3].

Buttar focused on discussing about the opportunities and challenges posed by fusing of ICT in various aspects of higher education. They concluded that the use of ICT for academic purposes not only upgrades classroom teaching and learning process, but also provides the convenience of e-learning. They also pointed out that it is essential for the teachers or trainers to be made to adopt technology in their teaching styles to provide pedagogical and educational gains to the learners [7].

Lwoga, Sife, Busagala, and Chilimo analyzed and proposed a way ahead on creating ICT awareness, expertise, and literacy through public universities in Tanzania. It was also recommended that universities should evaluate the systems used to improve their community's ICT awareness, expertise, and literacy in order to bridge the digital divide gap [8].

Sulisworo majored the research work to pursue the ICT contribution to knowledge management that enhance higher education institute competitiveness [9].

Aristovnik focused on reviewing some previous fact-finding exercises appraising ICT efficiency and its impact on educational outcomes as well as different conceptual and specialized issues related to performance measurement in some European Union (EU) and Organization for Economic Co-operation and Development (OECD) countries. The study also showed that majority of the countries under examination hold great potential for a heightened efficiency in ICT and for improvement in their educational outcomes [10].

3 Research Methodology

This work is a survey research focused on collecting primary data with questionnaires. The sample population consists of both military and civilian students of Air Force Institute of Technology (AFIT), Kaduna. The random sampling technique was used in selecting students from various departments of AFIT.

A structured questionnaire titled "An Assessment of the Impact of ICT on work efficiency in Nigerian Tertiary Institution" was used. It consists of three sections A, B, and C. Section A sought data on the population's bio data, Section B requested data on factors that determine the assessment of communication efficiency through modern technology in the administration of school, while section C contains questions used to determine the frequency of ICT usage. The respondents were required to tick each item in the appropriate column that mostly represents their opinion based on their degree of agreement or disagreement with the statement.

The questionnaire was administered using Google form, and paper copies of the questionnaire were also distributed. A hundred and three (103) samples of the questionnaire were distributed to the randomly selected students, and all of them were used.

3.1 Data Analysis Techniques

The statistical package for the social sciences (SPSS) was used in analyzing the data collected from the questionnaires. The research questions were assumed using percentages after construction of contingency table for the items with regard to the research questions. The Likert scale with strongly agree (5), agree (4), neutral (3), disagree (2), and strongly disagree (1). The Likert scale for frequency was also used to determine the frequency of usage of ICT. It has the value of never (1), seldom (2), sometimes (3), frequently (4), and always (5). The mean of the variables were gotten, and a decision mean score value was used to determine if the variables are to be accepted (if the mean score is greater than the decision mean score of 3) and

rejected if the mean score is lesser than the decision mean score of 3. The mean was used to answer the research questions.

Formula 1 was used to obtain the mean score:

$$X = \frac{fx}{f} \tag{1}$$

where

X mean score
f frequency of response
x the numeric value.

4 Data Analysis and Result

This section discusses the analysis of the data gotten.

4.1 Analysis of the Socio-demography of the Population Collected

Table 1 shows the socio-demography survey. The result of this survey based on gender, level of education, degree of computer literacy, and frequency of computer usage is shown in the table.

4.2 ICT Efficiency

Figure 1 shows the data visualization of the result from the analysis.

From Fig. 1, it is observed that strongly agree in the variable "ICT increases the accessibility to knowledge" has the highest frequency of responses. Most respondents voted for agree and strongly agree for the variables, while only few (<15%) voted for neutral for each of the variables.

Table 2 shows the decision for each of the questions asked. A decision score of mean 3 was used to determine if a question is to be rejected or accepted.

As seen in Table 2, all the variables are accepted proving that ICT has improved the efficiency of graduates. All the variables have a mean score greater than 3 with 3.87 being the lowest mean with the variable "The use of media increased your efficiency at work" and 4.703 being the highest mean score with the variable "ICT increases the accessibility of knowledge?".

Table 1 Socio-demographic survey

	N (%)
Gender	
Male	90 (88.2%)
Female	12 (11.8%)
Educational qualification	
Secondary school	7 (7.1%)
HND	2 (2.0%)
Bachelor	77 (77.8%)
Masters	13 (2.0%)
Computer literate	
Yes	101 (96.2%)
No	1 (1.0%)
Computer training	
None	5 (4.9%)
Formal	79 (76.7%)
Non-formal	19 (18.4%)
*Access to computer**	
Home	93 (90.3%)
Internet café	20 (19.4%)
Faculty library	8 (7.8%)
Faculty e-learning center	12 (11.7%)
Access to internet	
Yes	101 (98.1%)
No	2 (1.9%)
Frequency of internet usage	
Always-daily	80 (79.2%)
Occasionally-once a week	5 (5.0%)
Often-several times a week	16 (15.8%)
Frequency of email usage	
Never	1 (1.0%)
Always-daily	51 (49.5%)
Occasionally-a few time a year	39 (37.9%)
Often-weekly	1 (1.0%)
Sometimes-monthly	11 (10.7%)
Personally or professionally use of computer	
Yes	96 (93.2%)
No	7 (6.8%)

4.3 Frequency of Usage of ICT at Work

Figure 2 shows the data visualization of the frequency of usage of ICT. From Fig. 2, it is observed that the respondents' majority always utilize ICT applications for professional and personal work, and they always make research using the Internet.

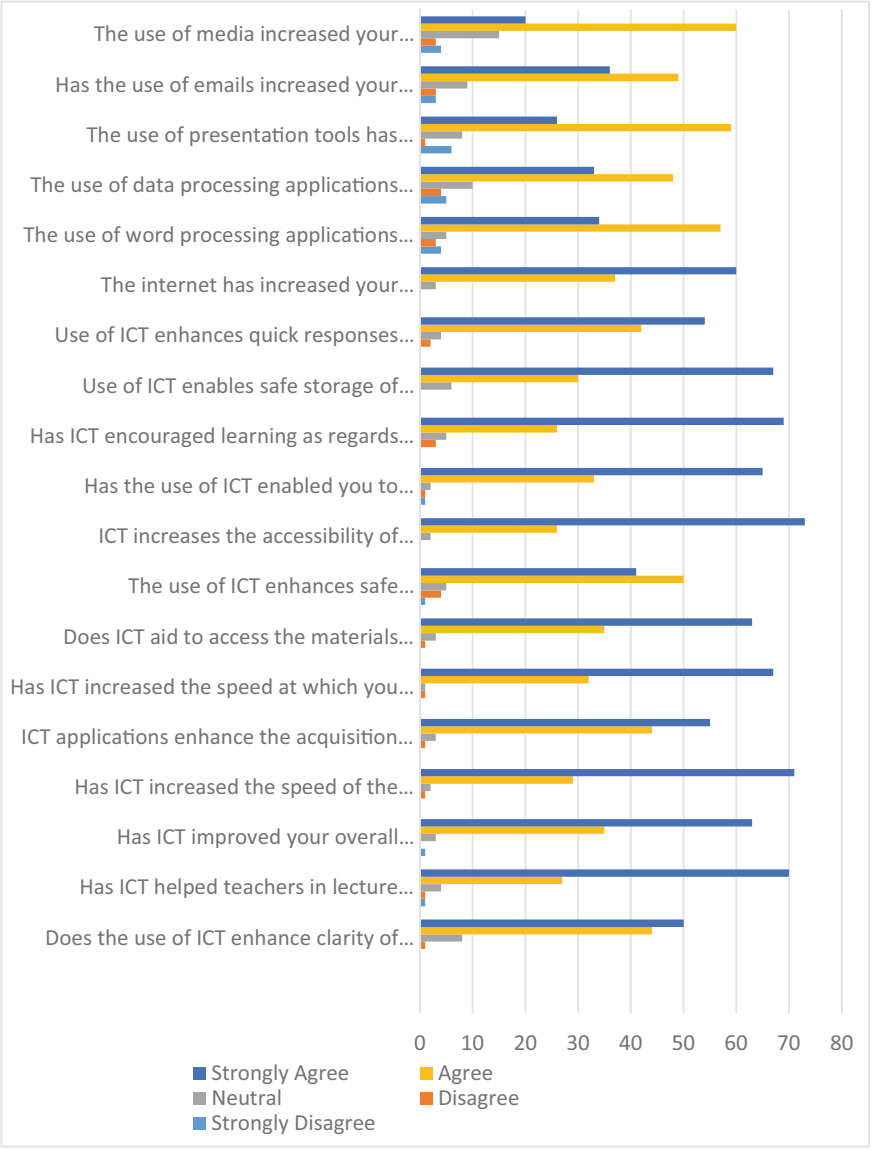


Fig. 1 A data visualization of ICT in enhancing work efficiency

Most respondents voted for frequently and always for the variables, while only few (10–35%) voted for neutral for each of the variables.

Table 2 Decision of respondents

Variables	*N*	Total	Mean	Decision
Does the use of ICT enhance clarity of expression to your superior or subordinate at work?	103	452	4.388	Accepted
Has ICT helped teachers in lecture delivery?	103	473	4.592	Accepted
Has ICT improved your overall productivity (or efficiency) at work?	102	465	4.559	Accepted
Has ICT increased the speed of the dissemination of information to both your superiors and subordinate?	103	479	4.650	Accepted
ICT applications enhance the acquisition of basic skills needed at work	103	462	4.485	Accepted
Has ICT increased the speed at which you perform your assigned task?	101	468	4.634	Accepted
Does ICT aid to access the materials needed to perform your role?	102	466	4.569	Accepted
The use of ICT enhances safe transmission of confidential information to the right destination	101	429	4.248	Accepted
ICT increases the accessibility of knowledge	101	475	4.703	Accepted
Has the use of ICT enabled you to network with other people to improve your productivity?	102	465	4.559	Accepted
Has ICT encouraged learning as regards your role?	103	470	4.563	Accepted
Use of ICT enables safe storage of important documents against natural disaster and theft	103	473	4.592	Accepted
Use of ICT enhances quick responses from your superior and subordinate	102	454	4.451	Accepted
The Internet has increased your efficiency at work	100	457	4.57	Accepted
The use of word processing applications have increased your efficiency at work	103	423	4.107	Accepted
The use of data processing applications has increased your efficiency at work	100	400	4.000	Accepted
The use of presentation tools has increased your efficiency at work	100	398	3.980	Accepted
Has the use of emails increased your efficiency at work?	100	412	4.120	Accepted
The use of media increased your efficiency at work	102	395	3.87	Accepted

Where N = number of respondents for the variable, Total = total number of the answers (N * value of the variables), and Mean = mean of the variable (Total/N)

5 Conclusion

The use of ICT in the twenty-first century has brought about efficiency among graduates. This study assessed the impact of ICT on the work efficiency of graduates in Nigeria using Air Force Institute of Technology (AFIT), Kaduna as case study. This study clearly shows that ICT applications and components like presentation

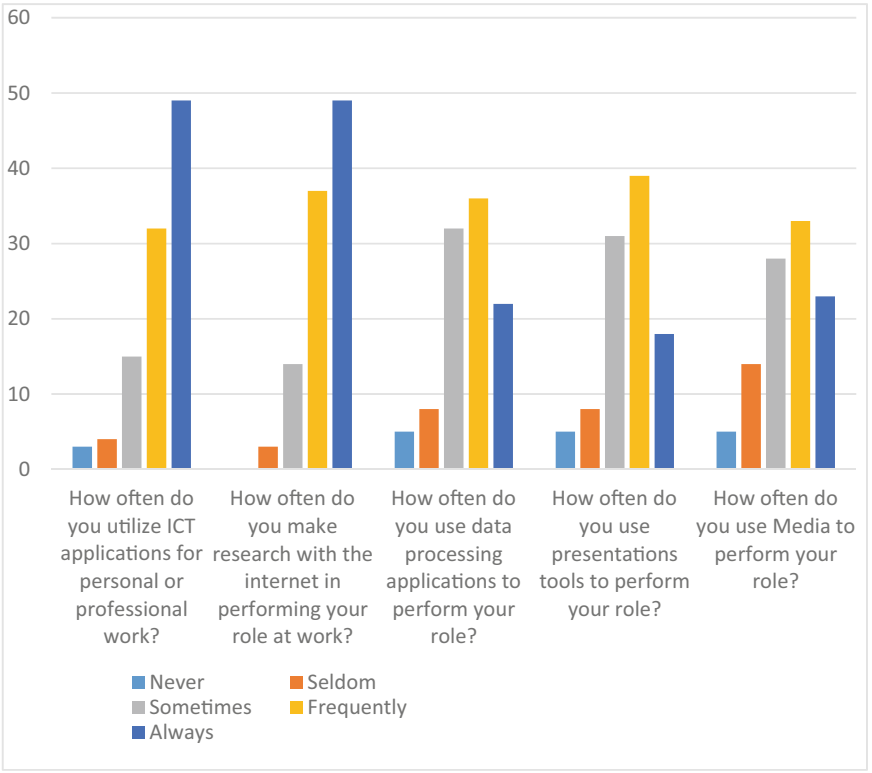


Fig. 2 Frequency of ICT usage. Where N = number of respondents for the variable, Total = total number of the answers (N * value of the variables), and Mean = mean of the variable (Total/N)

tools, word editor, and data analysis and so on have a high frequency usage, and it was generally agreed among respondents that ICT has enhanced the work efficiency among graduates. The mean of each variable is more than the decision mean score.

It is recommended for further study that this survey should cover more tertiary institution. Primary and secondary schools should also be captured in this survey to assess ICT efficiency in the educational sector. Considering the frequency of the usage of ICT and how it has enhanced work efficiency, it is recommended that university administration should provide more ICT applications (e.g., Internet, data processing applications, word processing applications, computer components, and so on) to her staff and students.

References

1. Danner, R.B., Pessu, C.O.A.: A survey of ICT competencies among students in teacher preparation programmes at the University of Benin, Benin city, Nigeria. J. Inf. Technol. Educ. Res. **12**(1), 33–49. https://doi.org/10.28945/1762
2. Shukla, S., Lakhmani, A., Agarwal, A.K.: A review on integrating ICT based education system in rural areas in India. In: 2016 International Conference System Modeling & Advancement in Research Trends (SMART), pp. 256–259 (2016). https://doi.org/10.1109/SYSMART.2016.7894531
3. De Wet, W., Koekemoer, E., Nel, J.A.: Exploring the impact of information and communication technology on employees' work and personal lives. SA J. Ind. Psychol. **42**(1), 1–11 (2016). https://doi.org/10.4102/sajip.v42i1.1330
4. Palvalin, M., Lönnqvist, A., Vuolle, M.: Analysing the impacts of ICT on knowledge work productivity. J. Knowl. Manag. **17**(4), 545–557 (2013). https://doi.org/10.1108/JKM-03-2013-0113
5. Makura, A.H.: Students' perceptions of the use of ICT in a higher education teaching and learning context: the case of a South African University. Mediterr. J. Soc. Sci. **5**(11 SPEC. ISSUE), 43–47 (2014). https://doi.org/10.5901/mjss.2014.v5n11p43
6. Adesote, S., Fatoki, O.: The role of ICT in the teaching and learning of history in the 21st century. Educ. Res. Rev. **8**(21), 2155–2159 (2013). https://doi.org/10.5897/ERR2013.1617
7. Buttar, S.S.: ICT in higher education. Int. J. Soc. Sci. **2**(1), 1686–1696 (2016). https://doi.org/10.22445/avsp.v4i1.4397
8. Lwoga, E.T., Sife, A., Busagala, L., Chilimo, W.: The Role of Universities in Creating ICT Awareness, Literacy and Expertise: Experiences from Tanzanian Public University, p. 34. Makerere University, Uganda (2013)
9. Sulisworo, D.: Enabling ICT and knowledge management to enhance competitiveness of higher education institutions. Int. J. Educ. **4**(1), 112–121 (2012). https://doi.org/10.5296/ije.v4i1.1207
10. Aristovnik, A.: The impact of ICT on educational performance and its efficiency in selected EU and OECD countries: a non-parametric analysis. Turk. Online J. Educ. Technol. **11**(3), 144–152 (2012)

Soybean Crop Non-beneficial Insect Identification Using Mask RCNN

Vivek Tiwari, Shailendra Gupta, Priyadarshini Roy, Chinky Karda, Shalini Agrawal, Manorama Singh Rathore, Swati Jain, and Anuj Pal

Abstract Since the beginning of globalization, agriculture has been the backbone of the Indian economy. However, agriculture is a labor-intensive job and is prone to natural disasters like drought, scanty rainfall, and pest attacks. The major threat to crops is insects and pests. They cause huge damage to crops, thereby causing losses in crop yield. Therefore, it is important to protect crops from insects for which it is important to identify insects properly and use targeted insecticides and control measures. Hence, this paper proposes a deep learning solution for the identification and localization of non-beneficial insects of soybean using the mask RCNN framework. The presented model creates bounding boxes and segmentation frame for each insect present in the picture. It is built on feature pyramid network (FPN) and uses Resnet101 as the backbone. The model has been evaluated against accuracy, precision, recall, $F1$-measure, AUC, and mean average precision.

V. Tiwari (✉) · P. Roy · C. Karda · S. Agrawal · M. S. Rathore
DSPM International Institute of Information Technology (IIIT), Naya Raipur, India
e-mail: vivek@iiitnr.edu.in

P. Roy
e-mail: priyadarshini18100@iiitnr.edu.in

C. Karda
e-mail: chinky18100@iiitnr.edu.in

S. Agrawal
e-mail: shalini18101@iiitnr.edu.in

M. S. Rathore
e-mail: manorama18101@iiitnr.edu.in

S. Gupta
Globus Engineering College, Bhopal, India

S. Jain
Government J. Yoganandam Chhattisgarh College, Raipur, India
e-mail: sjcscghedrp@gmail.com

A. Pal
Sagar Institute of Science, Technology & Research (SISTec-R), Bhopal, India
e-mail: junaamity@gmail.com

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_30

Keywords Deep learning · Mask RCNN · Localization · Segmentation masks · Insect detection · Smart agriculture · AI in smart farming

1 Introduction

Agriculture is the lynchpin of the Indian economy and contributes 18% to the GDP [1]. It provides livelihood to the farmers, and a population of 7.2 billion feeds on this agricultural harvest [2]. The major challenge faced by farmers is protecting crops from pests and insects [3]. These concerns motivated us to build a real-time application for insect identification and localization. The mask RCNN [4] is a deep learning model that does classification, localization, and pixel-level segmentation of insects. Mask RCNN is an improvement over the faster RCNN [5], giving an output of pixel-level mask along with object detection. In mask RCNN, the CNN feature map generated by faster RCNN, that is, the region proposals are produced by a fully convoluted network known as region proposal network, which is placed on top of the feature map. These bounding boxes are being forwarded to fast RCNN [5] to get the classification and tightening of the bounding box done. It added a branch to faster RCNN, which produces a binary mask, so that the classification and bounding box regression can happen simultaneously. But the regions obtained by RoI pooling were a bit misaligned as compared to the original input image, so RoI align was used [6]. In RoI align, the stride is not quantized, thereby giving the more accurate result, and the exact value of input features is evaluated using bilinear interpolation by pooling them at four regularly sampled locations applying either max or average over the four values. Also, in mask RCNN [7], feature extraction is made by the ResNet-FPN backbone (in our project ResNet101), which gives results in increased accuracy and speed. In this study, 256 top-down layers have been employed to create the feature pyramid. The working of mask RCNN is depicted in Fig. 1.

2 Short Literature Survey

Insect identification has recently gained popularity in the research community due to the significant advances in deep learning. Firstly, Venugoban and Ramanan [8] prepared a very small dataset containing 200 images with ten classes. A proposal of a larger dataset then of 1440 images equally divided into 24 classes of insects that harm the crops was made by Xie et al. [9]. In 2018, a new dataset consisting of 563 samples of ten different classes of insects was prepared by Deng et al. [10] to perform the classification of their dataset. They trained a support vector machine (SVM) on manually selected features. Before the advent of deep learning, using support vector machines for classification was a very common approach and still performs at par on small datasets as evident in the work of Rani and Amsini [11]. However, in the last few years, convolutional neural networks (CNNs) have surpassed

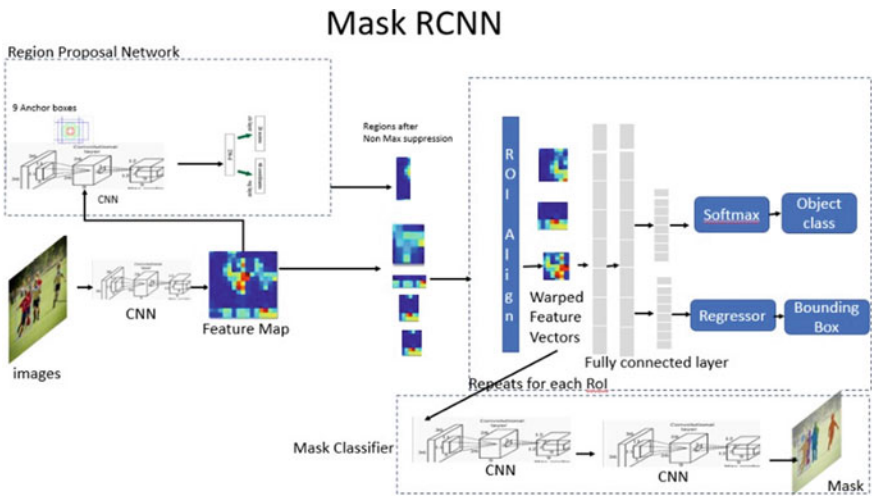


Fig. 1 Workflow of mask RCNN deep learning model

these old techniques in areas like image classification, image segmentation, and object detection, particularly on large datasets. Dawei et al. [12] used CNNs to classify pests. They used the pretrained model AlexNet [5] to classify a portion of the dataset used by Deng et al. and achieved high accuracy by using this transfer learning approach. Lately, Wu et al. [13] provided a very meticulous dataset called IP102 dataset, containing around 75,000 insect images with 102 classes. Hence, Ren et al. [14] and Wu et al. [13] used this dataset to CNNs and came up with their own architecture for classification.

3 Methodology

The proposed work classifies, localizes, and masks the insect input image by training the convolutional neural network. The classifier predicts the insect class, localizer identifies the insect location in the image, and mask highlights it. There were considered four soybean non-beneficial pests, namely Eocanthecona bug, tobacco caterpillar, red hairy caterpillar, and larva Spodoptera [15]. Matterports mask RCNN [4] has been implemented, which is written in Python and uses TensorFlow and Keras. The model has been trained on a system having 32 GB RAM, Geforce GTX 1080 Ti GPU, CUDA 10.1, driver Version 430, Python 3, TensorFlow-GPU 1.5, and Keras 2.2.4.

1. *Dataset Information*
   Initially, there were 3048 number of total images for four classes of insects developed by IIIT, Naya Raipur team, and Indira Gandhi Krishi Vishwavidyalaya (IGKV) Raipur [16, 17]. The composition of the four bugs is:

   1. Eocanthecona bug—1198,
   2. Tobacco caterpillar—990,
   3. Red hairy caterpillar—1001,
   4. Larva Spodoptera—635.

   This number is insufficient for training and further performance. Data augmentation is done to increase the dataset size [18]. The following was the final composition of all four classes of insects:

   1. Eocanthecona bug—1796,
   2. Tobacco caterpillar—1802,
   3. Red hairy caterpillar—1799,
   4. Larva Spodoptera—1777.

2. *Data Preprocessing*

   - **Annotation**: Data annotation is the task of labeling the data of various forms like text, image, audio, videos, etc. [19]. After a model has processed enough annotated data, it can start to recognize the same patterns when presented with new, unannotated data.
   - **Augmentation**: Augmentation is a technique of expanding the dataset by applying simple or complex transformations like flipping, scaling, color saturation, etc., to the existing data [20]. For generalizing a deep learning model, it is necessary to provide variations in data as well as ensure balanced sets of data in each class.
   - **Train and Test split**: The data is split into train and test set [21]. The training composition has to be maintained in such a way so as to avoid any bias toward any particular insect in the result: Split ratio training: test :: 80% : 20%.

3. Training the Model

   - **Transfer learning**: A research problem in machine learning uses the output and information obtained from solving a problem and then utilizing it on a different but related problem [22]. COCO dataset [23] on which the mask RCNN was originally trained, and COCO weights were obtained from it, which are being used to start the training of our model.
   - **Import the required libraries**: Libraries, i.e., the os, json, skimage.draw, datetime, numpy, cv2, and matplotlib.pyplot, are imported into *final.py* script, which is the training code. Also, define the path for pretrained COCO weights, and the logs directory was the trained weights which will be saved.

- **Set the configuration parameters**: According to our dataset, the input parameters values for the configuration of the model, such as the batch size, GPU count, validation steps, and learning rate have to be set. The parameters set during training specifically (*config.py* file) are:
  - GPU COUNT = 1
  - IMAGES PER GPU = 8 *# Batch size*
  - NUM CLASSES = 1 + 4 *# background + 4 insects*
  - IMAGE MIN DIM = 128 *# Set the limits of the small side*
  - IMAGE MAX DIM = 128 *# determines the image shape*
  - RPN ANCHOR SCALES = (8, 16, 32, 64, 128) *#anchor side in pixels*
  - TRAIN ROIS PER IMAGE = 32 *# training ROIs per image*
  - STEPS PER EPOCH = 100 *# Use a small epoch since the data is simple*
  - VALIDATION STEPS = 5 *# use small validation steps (epoch is small)*
- **Loading dataset**
  1. Create train and val (validation) dataset objects
  2. Add classes: Here, A, B, C, D are the class names and 1, 2, 3, 4 their corresponding class ids
  3. Parse json file containing annotations and extract file name, file path, width, height, bounding box coordinates, and class id of the object for each image.
  4. Add image to the dataset
  5. Prepare the dataset for use using dataset.prepare().
- **Set the training parameters**: The number of epochs was set at 49. The layers can be either 'heads' or 'all.' 'heads' are used when the dataset is in small scale or when the model can work well with less training. It trains the region proposal network, classifier, and mask heads of the network. There were used 'heads' for training dataset up to the 25th epoch then switched to 'all' as the need to work with a huge dataset. So the proposed model requires extensive training to trains all layers.
- **Training optimizer**: Optimizers update the weight parameters to minimize the loss function. The optimizer is being used SGD with momentum. SGD with momentum leads to faster converging by accelerating gradient vectors in the right direction [24].

4 Result and Discussion

As the training is over, it creates *mask RCNN insects 0049.h5* as the output weight. The accuracy and performance of the obtained output weight on the validation set containing 2125 images are tested using the confusion matrix, ROC curve, and mAP of the model. Here are the following results.

- **Confusion Matrix**: It is a table or matrix which is used to visualize the performance of a classifier or the classification algorithm. The multiclass classification matrix is depicted in Table 1 for the proposed model.
  Classification Report: Shows the model's accuracy and precision, recall, and $F1$-score on the validation set for all classes, which is calculated using the confusion matrix and represented in Fig. 2.
  Accuracy is the number of correctly classified insects into its respective class divided by the total number of insects in the dataset to be classified [25]. *The precision* of a class measures out of all the instances predicted to be of that class how many actually belong to that class [21]. *Recall* of a class measures out of all the instances in the dataset that actually belong to that class how many have been predicted to belong to that class. *F1-score* is the harmonic mean of precision and recall.

Table 1 Multi-class classification matrix

	BG	A	B	C	D
BG	0	0	0	0	0
A	13	539	1	2	0
B	60	1	426	0	0
C	55	19	3	483	1
D	14	1	1	6	500

	precision	recall	f1-score	support
0	0.00	0.00	0.00	0
1	0.96	0.97	0.97	555
2	0.99	0.87	0.93	487
3	0.98	0.86	0.92	561
4	1.00	0.96	0.98	522
accuracy			0.92	2125
macro avg	0.79	0.73	0.76	2125
weighted avg	0.98	0.92	0.95	2125

Fig. 2 Classification report of the model

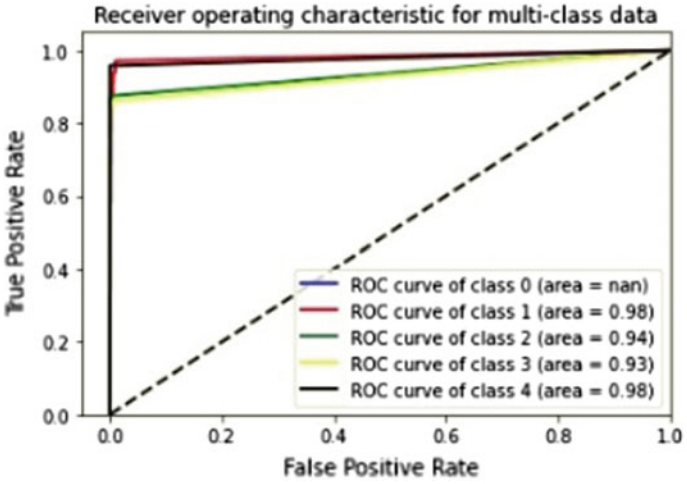


Fig. 3 ROC curve as a measure

$$\text{Accuracy} = (\text{TP} + \text{TN})/(\text{TP} + \text{TN} + \text{FN} + \text{FP}), \text{Precision} = \text{TP}/(\text{TP} + \text{FP})$$
$$\text{Recall} = \text{TP}/(\text{TP} + \text{FN}), F1\,\text{Score} = (\text{Precision} * \text{Recall})/(\text{Precision} + \text{Recall})$$

Macro-average is the mean of a performance measure of all classes. *The weighted average* is the weighted mean of a performance measure of all classes.

- **Receiver operating characteristic (ROC) curve and area under the curve (AUC)**: ROC is a curve used for visualizing the diagnostic ability of a binary classifier system by varying its discrimination threshold. When area under curve is one, it indicates that the model has a good measure of reparability, and AUC near 0 indicates the worst measure [21]. The ROC curve for the presented work is represented in Fig. 3. Furthermore, the model output on real images is shown in Figs. 4, 5, 6, 7, 8, and 9 with projected bounding box with probability.

5 Conclusion

Traditional detection and conventional eye observation methods are inefficient for large crops. We proposed a conceptually flexible and general framework, mask RCNN, for crop insect identification, localization, and masking. It integrates object detection along with instance segmentation, i.e., classifies each pixel of the detected object into predefined categories. The key point about mask RCNN is to tell what is in an image as well as where it is by generating a mask. It efficiently detects an insect and generates a segmentation of a mask for it. This reduces human intervention and accelerates the process of insect detection. At present, proposed model is trained

with four insects; it can be further extended to classify more insects, thus helping the farmers to identify the insect easily and correctly and use effective pesticides (by investigating type and population of insects). This will enable them to protect their crops more successfully. Further, it can be made easily accessible to them by integrating this model into a mobile-based application.

Output images with a bounding box

Fig. 4 Classification probability (Class A): 0.997

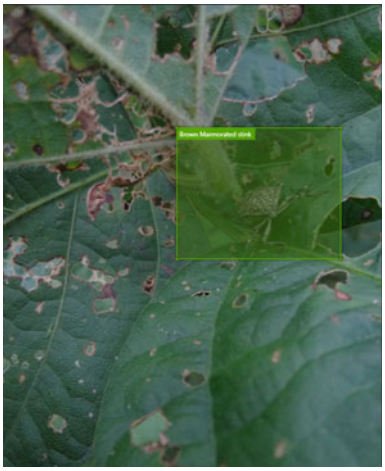


Fig. 5 Classification probability (Class A): 1.000

Fig. 6 Classification probability (Class C): 0.991

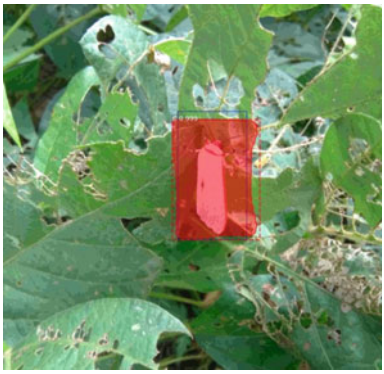


Fig. 7 Classification probability (Class C): 0.999

Fig. 8 Classification probability (Class D): 0.998

Fig. 9 Classification probability (Class D): 0.993

Acknowledgements This project was funded (via sanction letter no. RGPV/TEQIP-III/2019/04/624/65) under Collaborative Research Scheme (CRS), TEQIP-III, RGPV, Bhopal, which is supported by National Project Implementation Unit (NPIU), a unit of Ministry of Human Resource Development, Government of India.

References

1. Kamilaris, A., Prenafeta-Boldú, F.X.: Deep learning in agriculture: a survey. Comput. Electron. Agric. **147**, 70–90 (2018)
2. Agnihotri, N.P.: Pesticide consumption in agriculture in India-an update. Pestic. Res. J. **12**(1), 150–155 (2000)
3. Chen, Y., Adena, G.B., Mafra-Neto, A., Keogh, E.: Flying insect classification with inexpensive sensors. J. Insect Behav. **27**(5), 657–677 (2014)
4. Gamage, H.V.L.C., Wijesinghe, W.O.K.I.S., Perera, I.: Instance-based segmentation for boundary detection of neuropathic ulcers through Mask-RCNN. In: International Conference on Artificial Neural Networks, pp. 511–522. Springer, Cham (2019)
5. Zimmermann, R.S., Siems, J.N.: Faster training of Mask R-CNN by focusing on instance boundaries. Comput. Vis. Image Underst. **188**, 102795 (2019)
6. Choudhary, M., Tiwari, V., Venkanna, U.: Iris presentation attack detection based on best-K feature selection from YOLO inspired RoI. Neural Comput. Appl. (2020). https://doi.org/10.1007/s00521-020-05342-3
7. Cheng, T., Wang, X., Huang, L., Liu, W.: Boundary-preserving mask R-CNN (2020). arXiv preprint arXiv:2007.08921
8. Venugoban, K., Ramanan, A.: Image classification of paddy field insect pests using gradient-based features. Int. J. Mach. Learn. Comput. **4**(1), 1 (2014)
9. Xie, C., et al.: Automatic classification for field crop insects via multiple-task sparse representation and multiplekernel learning. Comput. Electron. Agric. **119**, 123–132 (2015)
10. Deng, L., et al.: Research on insect pest image detection and recognition based on bio-inspired methods. Biosyst. Eng. **169**, 139–148 (2018)

11. Rani, R.U., Amsini, P.: Pest identification in leaf images using SVM classifier. Int. J. Comput. Intell. Inf. **6**(1), 248–260 (2016)
12. Dawei, W., et al.: Recognition pest by image-based transfer learning. J. Sci. Food Agric. **99**(10), 4524–4531 (2019)
13. Wu, X., et al.: IP102: A large-scale benchmark dataset for insect pest recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 8787–8796 (2019)
14. Ren, F., Liu, W., Wu, G.: Feature reuse residual networks for insect pest recognition. IEEE Access **7**, 122758–122768 (2019)
15. Musser, F.R., Catchot, A.L., Davis, J.A., Herbert, D.A., Lorenz, G.M., Reed, T., Reisig, D.D., Stewart, S.D.: 2012 soybean insect losses in the southern US. Midsouth Entomol. **6**, 12–24 (2013)
16. Tiwari, V., Saxena, R.R., Ojha, M.: InsectBase: soybean crop insect raw image dataset_V1 with bounding boxes for classification and localization. figshare. Dataset (2020). https://doi.org/10.6084/m9.figshare.13077221
17. Tiwari, V., Saxena, R.R., Ojha, M.: InsectBase: soybean crop insect processed (cropped) image dataset_V1 for classification. figshare. Dataset (2020). https://doi.org/10.6084/m9.figshare.13078883
18. Tiwari, V., Saxena, R.R., Ojha, M.: InsectBase: soybean crop insect augmented cropped image dataset for classification. figshare. Dataset (2020). https://doi.org/10.6084/m9.figshare.13084208.v1
19. Kashyap, R., Tiwari, V.: Energy-based active contour method for image segmentation. Int. J. Electron. Healthc. **9**(2–3), 210–225 (2017)
20. Choudhary, M., Tiwari, V., Venkanna, U.: An approach for iris contact lens detection and classification using ensemble of customized DenseNet and SVM. Future Gener. Comput. Syst. **101**, 1259–1270 (2019)
21. Kunal, S., Saha, A., Varma, A., Tiwari, V.: Textual dissection of live Twitter reviews using naive Bayes. Procedia Comput. Sci. **132**, 307–313 (2018)
22. Pan, S.J., Yang, Q.: A survey on transfer learning. IEEE Trans. Knowl. Data Eng. **22**(10), 1345–1359 (2009)
23. Lin, T.-Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P., Lawrence Zitnick, C.: Microsoft coco: common objects in context. In: European Conference on Computer Vision, pp. 740–755. Springer, Cham (2014)
24. Gulli, A., Pal, S.: Deep Learning with Keras. Packt Publishing Ltd. (2017)
25. Choudhary, M., Tiwari, V., Venkanna, U.: Iris liveness detection using fusion of domain-specific multiple BSIF and DenseNet features. IEEE Trans. Cybern. (2020). https://doi.org/10.1109/TCYB.2020.3005089

Induction Heating Using Three-Phase Variable Frequency Converter

Happy L. Shah, Khushali A. Joshi, Nirali A. Solanki, Shreya A. Vyas, and Nilesh K. Jaiswal

Abstract There are numerous methods for melting of metal, out of which, induction heating is the most convenient one. It is a process in which an object (metal) is heated by the principle of electromagnetic induction. In the conventional method of heating that is by using single-phase power supply, there is a zero instantaneous power twice a cycle which results in low efficiency and discontinuous method of heating; therefore, by using three-phase power supply, there is a smooth power delivery and faster operation of heating, as three-phase system has a continuous power throughout the complete alternating cycle, and one of the major advantages of using three-phase supply system is that both melting of metal and stirring action take place in a single unit. It is a method which is not widely used in industries, they generally use single-phase supply for melting, and for stirring of melted metal, different method is employed. By using three-phase supply at a frequency of 1 kHz and maintaining the *V*/*F* ratio constant, speed control of induction motor at a smoother rate is also possible. Thus, by controlling speed, the metal can be heated at required temperature, and the method also results in efficient heating.

Keywords Three-phase work coil · Power electronics converter · Microcontroller device

1 Introduction

In the process of induction heating, the effect of the current which is induced by electromagnetic action in the charge is employed. The developed heat depends on the power taken by the charge, and also, the power which is drawn is directly equivalent to the square of the voltage and inversely depends upon resistance. Now, to melt the charge, sufficient amount of heat is required, and it is only possible by producing higher flux and frequency. The frequency and magnetic flux keep the resistance to a lower value and maintain voltage to a much higher rate. Therefore, it is easily possible

H. L. Shah (✉) · K. A. Joshi · N. A. Solanki · S. A. Vyas · N. K. Jaiswal
SVIT Vasad, Vasad, India
e-mail: Shahhappy35@gmail.com

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_31

to treat magnetic materials compared to the nonmagnetic materials just because of their higher permeability [1]. Basically, induction furnace is of two types depending upon construction of magnetic core. They are core type furnace and core less furnace [2].

Coreless Furnace: The conventional method which is used in core less furnace consists of primary coil, the refractory lining, the frame which includes supports and a tilting mechanism. The primary winding of a coil is made up of hollow conductors, it is connected to single-phase AC supply, and charge is put into the crucible which is generally used as a core. The material of the crucible is of thermal refractory grading, and as there is no magnetic core present in the furnace, high frequency current is used. Now, due to high current in primary winding, there are high copper losses, which increases temperature of coil. Hence, coil is made up of hollow conductors, and water is circulated to keep the coil cool [2].

The conventional method discussed above has single-phase AC supply as a power source and takes more time to melt the metal, and it also has a zero instantaneous power twice a cycle which results in discontinuous form of heating. Now, when three-phase AC supply is given to the primary winding, rotating magnetic field is produced which creates eddy current in the charge. These eddy currents heat the charge to the melting point of metal and produce electromagnetic force which creates stirring action in the charge which is required for obtaining uniform quality of metal. Hence, by using three-phase supply, the time which is required to melt the metal is decreased which in turn increases the efficiency of heating, there is precise control of temperature by variable frequency, also the faster heat treatment avoids the chances of deformation of object, and the stirring action which is required for uniform heating is obtained in a single unit. Figure 1 shows the block diagram of an induction heating process.

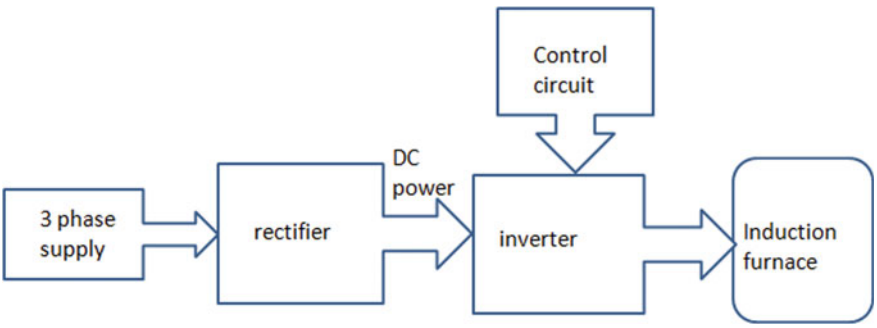


Fig. 1 Affiliated block diagram

2 Design, Analysis and Implementation of Projected Hybrid Topology

Our motive is to implement an advanced technology in a heating process by which the efficiency and proper melting can be obtained. By using microcontroller device, the frequency can be easily varied, and we can get required operation of inverter and furnace.

2.1 Implementation of Control Circuit

1. 8051 Microcontroller: The 8051 microcontroller unit is provided with 5 V DC supply, it can be directly provided by 5 V battery or a 5 V adapter, and it is connected with VCC pin of microcontroller, the ground pin of MCU is connected with ground, and port pins of MCU are connected with the gate of inverter circuit [3]. The unit is shown in Fig. 2a. The port pins of port 0, that is P0.0, P0.1, P0.2, P0.3, P0.4 and P0.5, are connected with Ig1, Ig2, Ig3, Ig4, Ig5 and Ig6 of opto-isolator circuit Fig. 2b.
2. Isolation circuit: The gate pulses generated by microcontroller are then given to the isolation circuit for providing the isolation between inverter and microcontroller unit, annd individual buffer circuit and opto-isolator circuit is used for each gate drives as shown in Fig. 2b. Isolation is provided to protect the microcontroller unit from exposure to high voltages from the power circuit [4]. Microcontroller programming is used to generate six different gate pulses at a frequency of 1 kHz, and it is shown in Fig. 2d.

As shown in the Fig. 2a, the port pins of 8051 microcontroller unit are individually connected with the isolation circuit that is with F-1, F-2, F-3, F-4, F-5, F-6 through buffer and resistance. The output of isolation circuit is directly given to the three-phase inverter which operates on 180° mode and converts fix DC supply into variable frequency AC supply. Figure 2c shows the basic inverter circuit of three-phase bridge inverter. As shown, there are six power switches, and the switches are operated (opened and closed) periodically in proper series to produce the desired output waveforms. The rate of switching determines the output frequency of inverter.

2.2 Implementation of Power Circuit

The three-phase alternating current supply is given to the six pulse rectifier that converts variable alternating current into fix direct current with filtering out the ripple and provides smooth DC supply. The DC output of rectifier is then given as an a input to the three-phase inverter which converts the fixed DC supply into the variable alternating supply with the variable frequency [5], and the high frequency

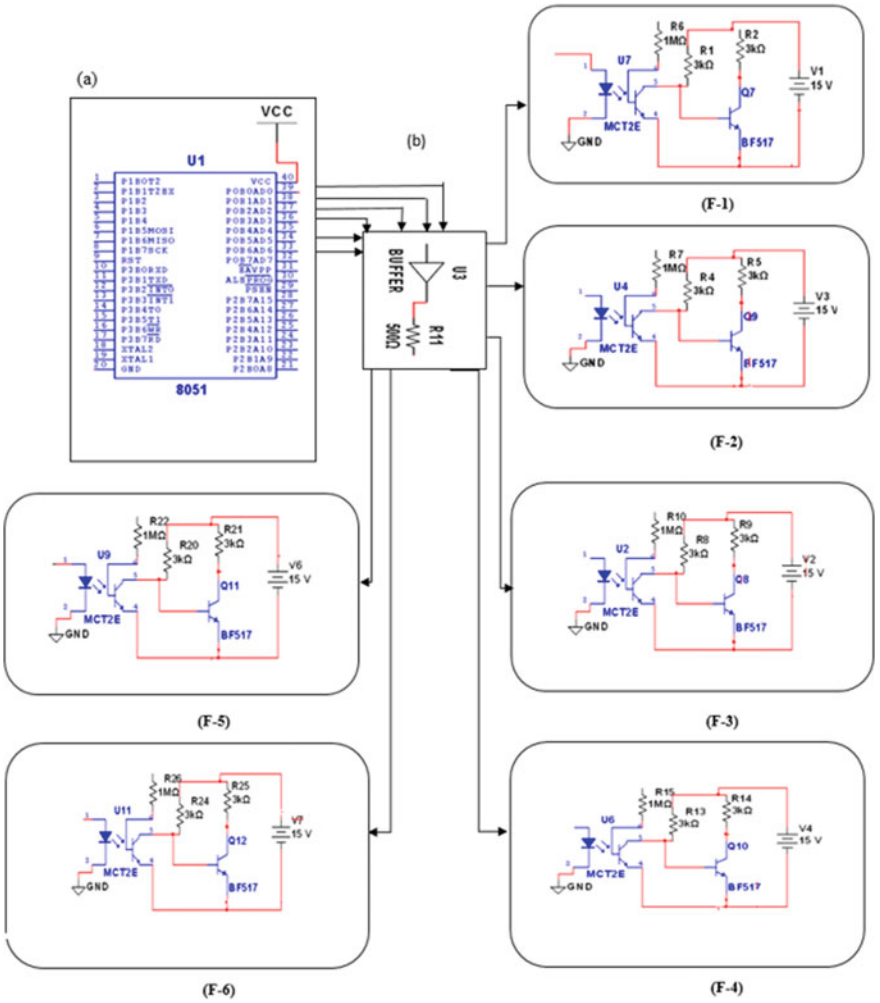
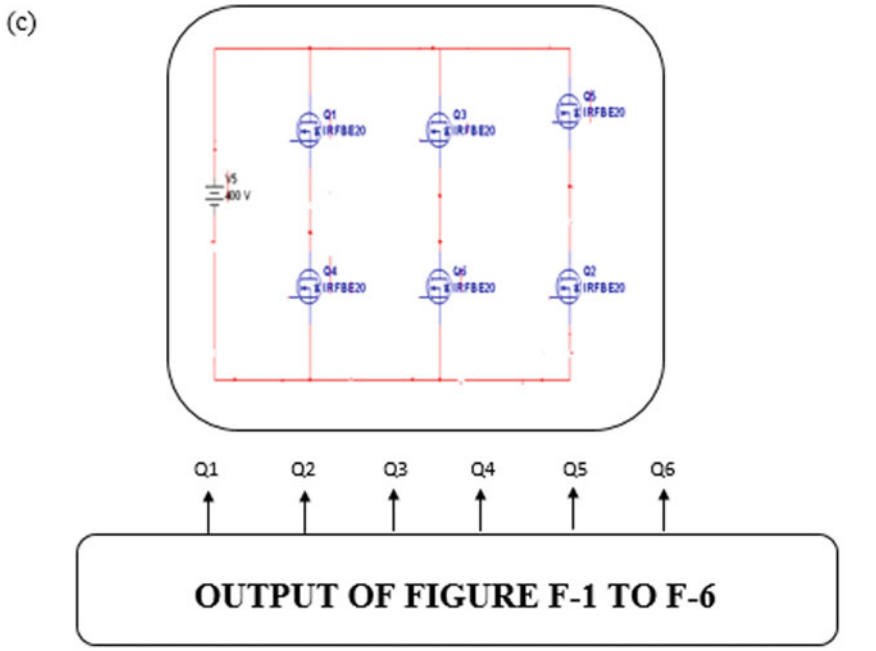


Fig. 2 **a** Microcontroller unit **b** isolation circuit **c** three-phase inverter circuit. **d** Simulation of gate pulse for three-phase inverter in Keil

gate pulses are generated by the microcontroller unit and are given to the input of gate circuit of inverter [3].

The output of inverter which is variable frequency AC supply acts as an input to three-phase induction motor. The three-phase variable frequency AC supply is then given to induction motor which generates the rotating magnetic field, and this field links with the workpiece placed inside the crucible and induces emf by phenomenon of Faradays law of electromagnetic induction [6]. Now, this emf induces induced current in the workpiece which melts the metal which is placed inside the crucible (here, rotor is removed from the induction motor, and graphite material crucible



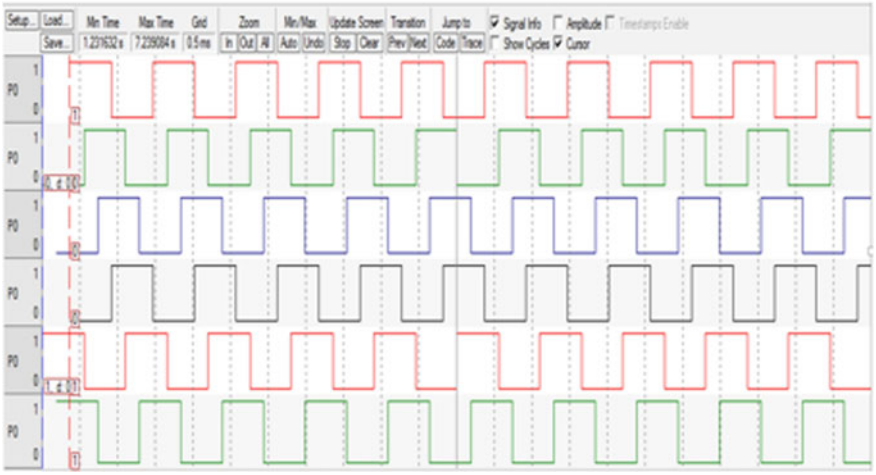


Fig. 2 (continued)

is used). The induced current also opposes the primary current of stator. So, now according to Lenz's law, the direction of induced current will be such that it opposes the induced emf (cause) which is induced as a result of relative motion present between rotor conductors and rotating magnetic field (rmf). Thus, to oppose the relative motion, the torque is produced in the motor, the force is experienced by the

crucible, and due to which, it starts rotating and tries to catch up the speed of rotating magnetic field [6]. The stirring of metal and the melting takes place in the single unit, and there is no need of providing any other heating technique for stirring action [7]. To avoid damage to three-phase work coil from heating, thermal insulation is provided between coil and crucible [8].

A distinguishing feature of alternating currents is that they are highly concentrated on the outside of a conductor. This is called the skin effect, and due to this phenomenon, the induced current which is produced in the material to be heated is large on the outward side of the conductor and diminishes towards the centre (in lesser proportion) [9]. Thus, on the outside of the material, most of the heat is generated. The skin effect is featured by its so-called penetration depth *D*. The penetration depth is defined as the thickness of the layer, in which 87% of the power is developed when measured from outside of the conductor [10].

$$\xi = \sqrt{(2\varrho/\omega\mu)}$$

The electrical frequency has significant effect on the process of induction heating. The desirable frequency range is obtained by the workpiece dimensions, material type, coil and object arrangement and the suitable depth of penetration [11]. As we increase the frequency, penetration of heat decreases. The frequency between 500 Hz and 10 kHz which is a normal range, is known as medium frequency, and the frequency between 100 kHz and 2 MHz is called high range [8]. For rapid heating of small objects, frequency in the scale of 100–450 kHz is required to produce high energy of heat for melting, or the same range of frequency can melt the skin of large parts. When deep penetration of heat is required, low frequency is essential which gives extended range of heating cycles, and the frequency range should be between 5 and 30 kHz [1].

During high frequency range, most transfer of energy takes place during the process of induction heating, usually when the depth of current penetration is low. It has been constituted that the recommended frequency must be such that the depth of penetration should not be greater than 1/8th times the diameter of metal to be melted. The furnace power rating relies upon the charged capacity, melting time of material and the material type to be melted. Metals having different melting temperature and resistance are melted at high frequency in induction furnace [9] (Fig. 3).

3 Design of Three-Phase Work Coil-Based Induction Furnace

3.1 Design of Induction Furnace to Melt a Workpiece

Now, assume a cylindrical workpiece to be melted having a similar shaped crucible.

Let Hm = height of the workpiece

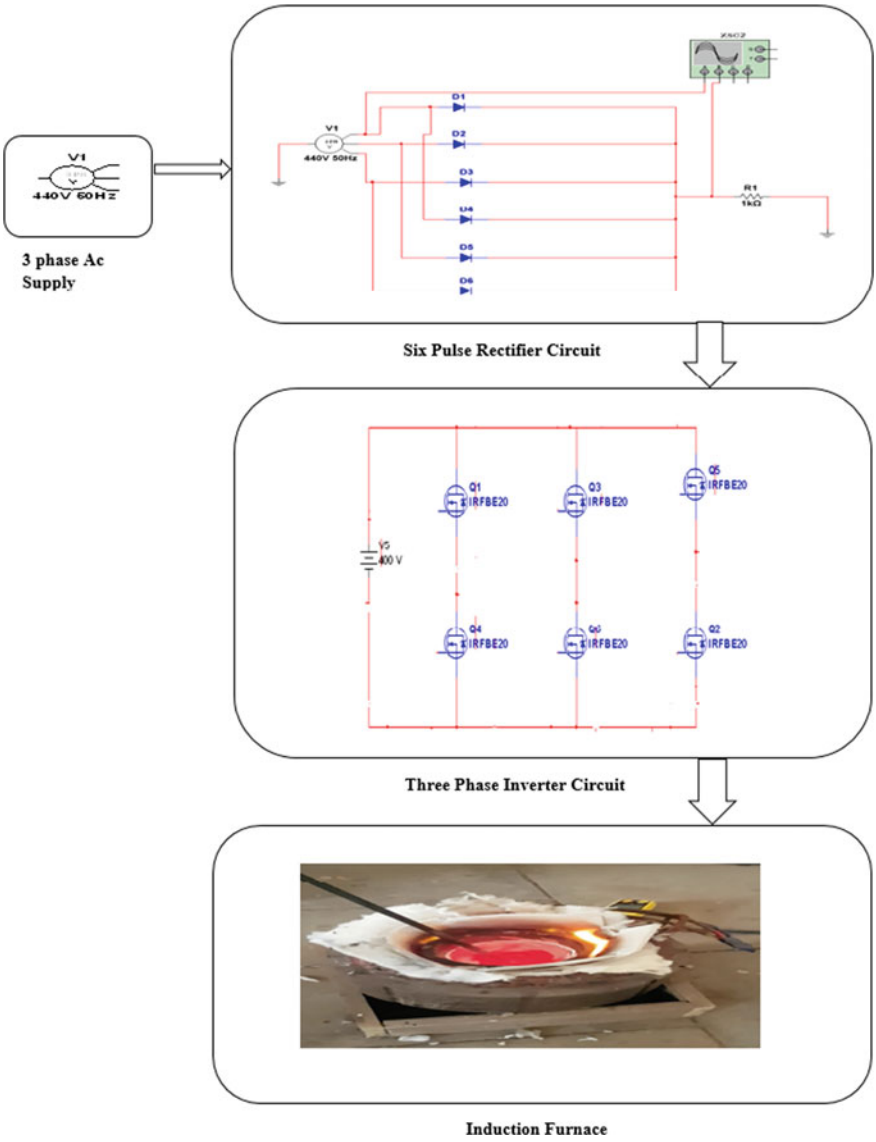


Fig. 3 Power circuit

Dm = diameter of the workpiece
The requirement of energy for melting the workpiece is given by

$$\mathbb{Q}_{\text{th}} = MC\Delta T + ML_{\text{f}}$$

where

$\mathbb{Q}_{th}$ the heat energy calculated in joules
M mass of the workpiece to be melted
T the change in temp in °C
L_{f} latent heat of fusion.

The requirement of power to generate required amount of heat energy in a specific time can be calculated by the relation

$$P = \mathbb{Q}_{th}/t$$

whereby

t the required time taken to attend maximum temperature.

- Power requirement similar to above can also be calculated by using the magnetic flux density (B_P) and the induced emf in the workpiece. By coordinating power equations mentioned above, electrical parameters of the furnace can be determined as

$$\varphi = B_P * A$$

Emf induced in the workpiece of metal due to changing magnetic field is calculated as

$$e = \mathrm{d}\phi/\mathrm{d}t$$

Induced emf's average value is

$$E = e/\sqrt{2}$$

The resistance of each strip

$$R = \rho l/a$$

The emf in the metal to be melted causes the eddy current to flow through the outer layer given by Ohm's law as

$$I = E/R$$

The eddy currents which are induced perpendicular to the magnetic lines passing through the layers of workpiece cause the power dissipation within the workpiece which can be formulated by

$$P = I^2 R$$

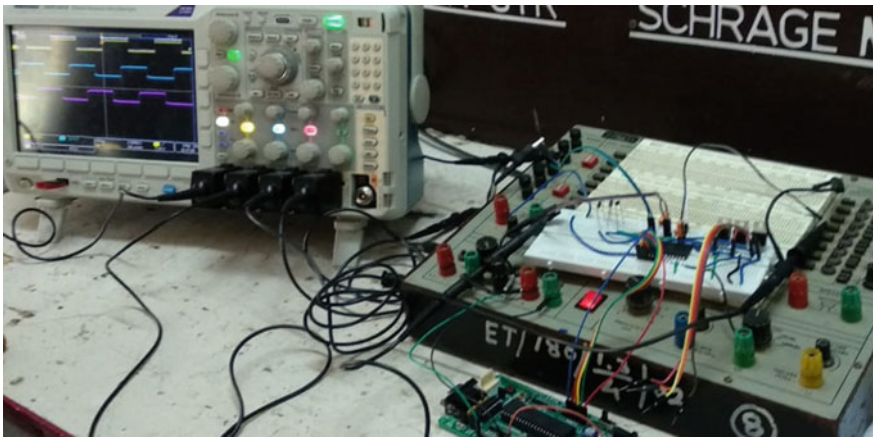


Fig. 4 Testing and output of control circuit

4 Testing of Power and Control Circuit

4.1 Mounting of Control Circuit Using Driver IC-IR2130

See Fig. 4.

4.2 Mounting of Power Circuit

See Fig. 5.

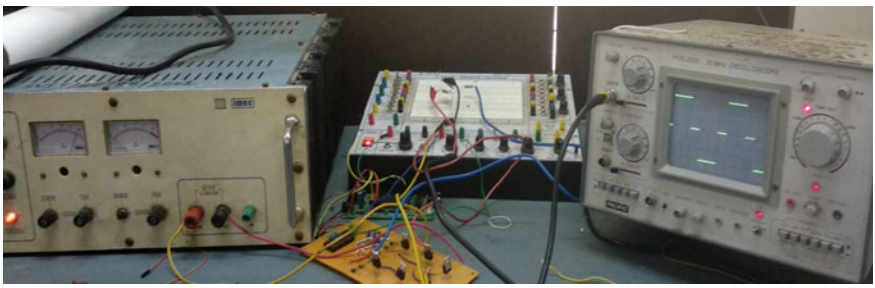


Fig. 5 Evaluation and output of power circuit

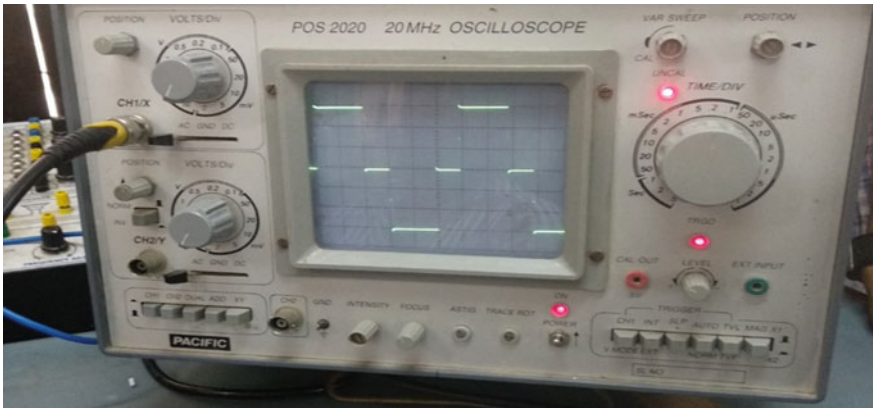


Fig. 6 Line voltage depiction

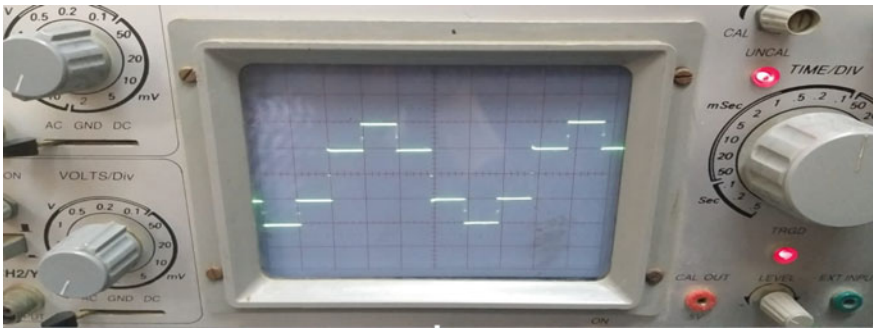


Fig. 7 Phase voltage depiction

4.3 Line Voltage of Inverter

See Fig. 6.

4.4 Phase Voltage of Inverter

See Fig. 7.

4.5 Augmented System

See Fig. 8.

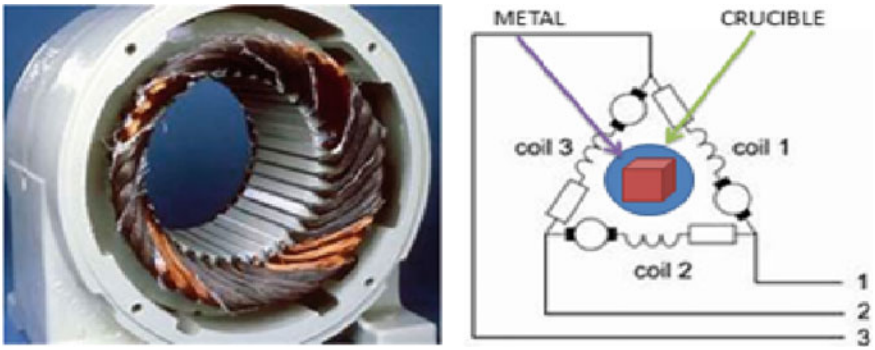


Fig. 8 Developed system

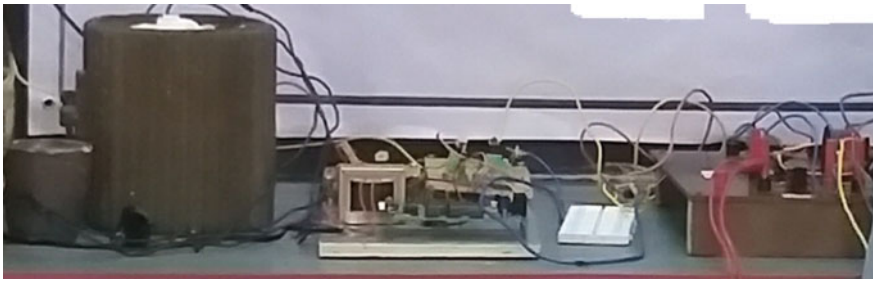


Fig. 9 Final circuit implementation

4.6 Full Circuit

See Fig. 9.

5 Conclusion

From all inclusive research and implementation, it has been analysed that from this method there are following outcomes.

1. Power consumption can be controlled.
2. No wastage of energy through heat loss.
3. Cost effective. Electricity consumption is many times lower than the other types of heating, so cost for electricity is lower.
4. Less time than other types of heating
5. Higher yield: The oxidation losses which are very significant in production economics are very less due to absence of combustion sources.

6. Flexibility: There is no need of molten metal to start medium frequency coreless induction melting furnace.
7. Natural Stirring: The natural stirring process results in homogeneous melt, especially when furnace is operated under medium frequency range.
8. Cleaner melting: There is no bi-product which is produced due to cleaner melting environment.
9. Compact installation: High melting rates are easily obtained from small furnaces.
10. Energy conservation: Energy efficiency is significantly better than combustion processes, and generally, induction melting ranges from 55 to 75%.
11. Scrap is reduced or eliminated by the use of precise controls.
12. Temperatures are accurately controlled to meet the specific requirements of the application.
13. This method is very much suitable for the production of high grade alloy steels, as there is no contamination of charge in addition to the composition of metals.
14. There is a possibility vacuum heating which is necessary for precious metal melting.
15. There is no presence of dirt, smoke and noise.

In addition to above, the melting of precious metals can be done using induction heating as it provides accurate control of temperature, and also, the deep hardening and surface hardening of metals is easily possible. It is used in vast application such as in process of annealing, brazing and to make the alloys from the melting of different metals.

References

1. Alshaikhli, A.K.M., Faris, F.H.: Design and construction of coreless furnace, Int. J. Sci. Eng. Res. **5**(1), 5518 (2014) ISSN 2229
2. Ajwalia, R.P., Upadhyay, P.A.: Utilization of electrical energy industrial application of induction heating. Pergamon press, London (1969)
3. Mazidi, M.A., Mazidi, J.G., McKinley, R.D.: The 8051 Micro controller: Pearson New International Edition; Pearson Education (2013)
4. Singh, M.D., Khanchandani, K.B.: Power electronics improvement in melting efficiency of induction heating. IEEE Conf. Publ. **149**, 28–33 (1977)
5. Tudbury, C.A.: Basics of induction heating industrial electrical furnaces and applications, 2nd (edn.) New York (1960)
6. Bimbhra, P.S.: Electrical machinery research of induction heating, Vimal R Nakum, Nobal Group of Institution
7. Green, D.: Coreless induction furnace and metal bath. IEEE J. London, Int School Sci. Biol. Pearson (1965)
8. Induction and Dielectric Heating. Electricity and Productivity Series, N6. British Electrical Development Association, pp. 8–9 (1962)
9. Rudnev, V., Loveless, D., Cook, R., Micah Black Handbook of Induction Heating
10. Sawhney, A.K.: Electrical Machine Design
11. Bala, K.C.: Design Analysis of Electric Induction Furnace for Melting Aluminium Scrap

Modeling and Simulation of an Automotive RADAR

Preeti S. Pillai, Raghavendra Shet, Nalini C. Iyer, and Sahana Punagin

Abstract An autonomous car, well known as self-driving or driverless car, is a vehicle that is skilled in sensing its environment with various sensors and moves with little or no human input. It makes use of Radar, Lidar and camera for its function. The use of Automotive RADAR is harmless to human health as its power is less than 10mW. Validating ADAS in real-life traffic conditions is prolonged process which leads to safety critical conditions and high costs. To speed up its development, virtual driving simulation tools are being used. The RADAR System working at 77GHz is simulated in the MATLAB simulation platform and presented in this paper. It also includes waveform generation, antenna characterization and processing algorithms for range and speed determination. For the development of algorithm, realistic sensor models and simulation tools are required to generate input data. Using this algorithm, given the driving scenario as input data for the ego vehicle, RADAR scans in its field of view. The radar signal is computed by incorporating radar systems specifics, like antenna pattern, transmitted power, etc. Basic radar signal processing methods then extract generic object information to be used for the following processing unit. The developed RADAR simulator is used in real-time driving simulators for the growth of driving assistance systems and in turn used in the radar signal processing techniques. This paper presents the development of RADAR sensor model using MATLAB-based driving simulator.

Keywords RADAR · Automotive radar · FMCW · ADAS

1 Introduction

Many vehicles these days have advanced driver assistance systems inbuilt in them. ADAS is an electronic system that assists the driver while driving. ADAS features include automated lighting, electronic stability control, anti-lock braking, adaptive cruise control, collision avoidance, traffic warnings, lane departure warning and blind

P. S. Pillai (✉) · R. Shet · N. C. Iyer · S. Punagin
K.L.E. Technological University, Hubli, India
e-mail: preeti_pillai@bvb.edu; preeti_pillai@kletech.ac.in

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_32

spot detection. Adaptive cruise control (ACC) is an existing cruise control system for vehicles that automatically adjusts the vehicle speed to maintain a safe distance from vehicles ahead. RADAR, Lidar or a camera setup are few on-board sensors from which the control action is decided [1, 2]. This system only provides some help to the driver but does not drive the car by itself. This paved way for autonomous vehicles [3]. It is expected that 3 of 4 cars on the road will be driverless by 2040.

Automotive radar is used to find the obstacles like vehicles and pedestrians, in the surrounding area of the car [4, 5]. It basically works on the principle of reflection of electromagnetic waves. The transmitter emits out radio waves that strike an obstacle and reflects back to the receiver. By scheming the direction in which the radio waves are sent and received, the distance of an object can be detected. This is achievable with the help of steerable antennas that automatically direct or receive signals simultaneously from different directions.

Automotive radars are mainly categorized into three types, i.e., long-, medium- and short-range radars. Long-range radars are used in measuring the distance and speed of the vehicles ahead. Medium-range radars are used for detecting objects in case of cross traffic alert systems [6]. Short-range radars are used for sensing in the district of the car to aid systems while parking. It is also used for blind spot detection. Figure 1 shows the different types of radar used widely. Radar sensors are used in many automotive applications and above all in driving assistance systems. To make sure such advanced safety systems work properly, it is essential to test them. Validating those systems in real traffic is time consuming and can lead to high costs [7, 8]. To speed up the development of advanced assistance systems, virtual driving simulation tools are used. Simple geometrical models without radar system specifics are often used to model the radar sensor in simulations tools [9, 10].

Usually, two different frequency bands, 24 GHz band and 77 GHz band, are largely used in automotive radars [11]. The 77 GHz band offers elevated performance. The

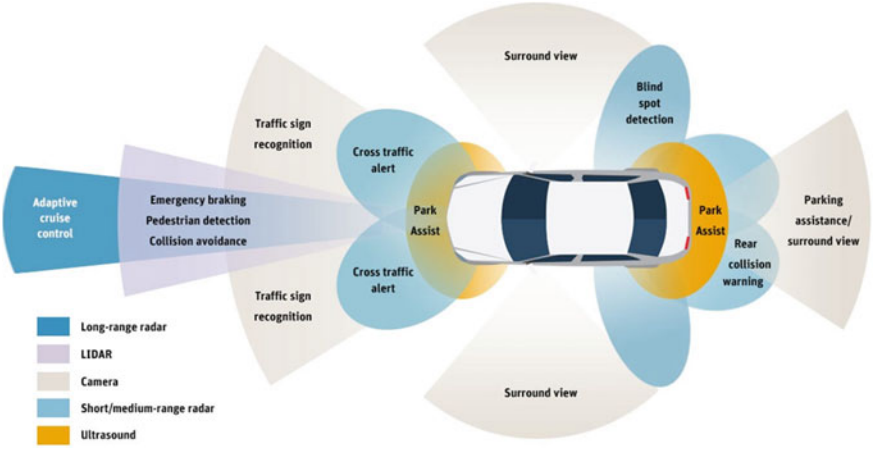


Fig. 1 Applications of different types of RADAR

24 GHz band radars are effortless to develop, but they are large in size, making it complex to integrate them in a vehicle. Therefore, radars which operate at 24 GHz require around three times larger antennas than radars operating at 77 GHz band, to attain the same performance [12]. The most important element for moving to higher frequencies is that, at high frequency, radars yield a better resolution at lower cost.

2 Methodology

The general block diagram of an automotive RADAR is shown in Fig. 2. It is used to estimate the range and velocity of the objects detected in the field of view. The algorithms to be tested need to be provided with information of the vehicle environment, like distance, velocity and azimuth angle of the detected vehicle. Therefore, a generic radar sensor model is developed. The traffic scenario input information of the radar sensor model is provided by the virtual driving simulation environment, and the output of the model is used for further simulation steps. This model is suitable for object detection in real time and can be parameterized. Following this, the model is separated into different modules [1]. The first module is the FMCW generator model. It transmits the chirp signal. This wave hits the target in an environment within the field of view of the sensor. This information is used and merged with the characteristics of an antenna in the algorithm. The combination of the geometric model with the antenna beam yields to a radar signal that depends on the position of the object, which leads to the received power of the object by using the radar equation. The received signal is called as an echo signal. This information is processed to estimate the number of objects in the field of view of the sensor.

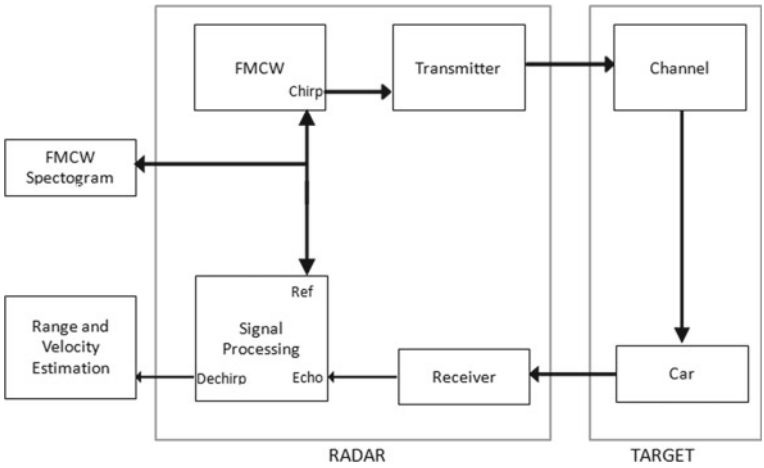


Fig. 2 Functional block diagram of the RADAR

2.1 Generation of FMCW

In pulsed radar systems, an echo pulse is received after a long time, and it is considered to be response of wrong sequence of pulse. Continuous wave (CW) radar systems, when compared to classical pulse waveforms, have an advantage of low measurement time and face less difficulty for its computation. There are two classes of continuous waveforms. Firstly, the linear frequency modulated (LFM) wave is used to resolve two tiny targets which are positioned at long range with a very little separation between them. The second type is well known as frequency shift keying (FSK) continuous waveform. This waveform cannot resolve two targets in the range but has a very high velocity resolution. A new waveform is design for automotive applications which are based on the combination of above-mentioned two CW transmit signals, and it is called as frequency modulated continuous waveform (FMCW) [13]. This intertwined technique of LFM and FSK continuous waveforms provides extremely small measurement time and simultaneously offers high range and velocity resolution. Figure 3 shows the FMCW radar. It consists of the synthesizer which generates the chirp. It is the heart of the FMCW radar. Chirp is basically a continuous sine wave whose frequency increases frequency, f_c ($f_c = 77$ GHz) bandwidth or beat frequency, linearly with respect to time. Chirp is characterized by the start f_b and the sweep time, δ_t [14].

$$\delta_t = \frac{(5.5)(2)(\text{Maximum Range})}{c} \tag{1}$$

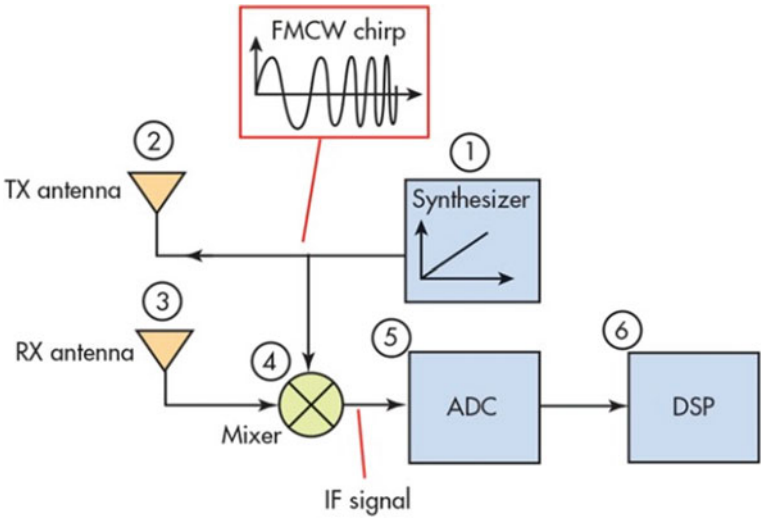


Fig. 3 Functional block diagram of the RADAR

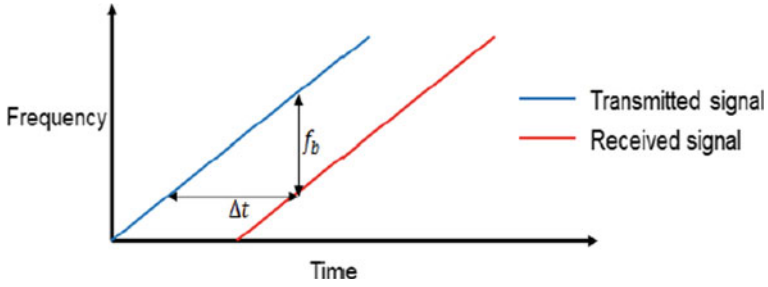


Fig. 4 Transmitted and received signal for the single object detected

Here, c is the speed of light. The range resolution is set to 1 m. The sweep bandwidth is calculated in one way that is calculated using range resolution as

$$f_{\mathrm{b}} = \frac{c/2}{1} \tag{2}$$

The slope S of the chirp defines the rate at which chirp ramps up. This is called as sweep slope.

$$\text{Sweepslope, } S = \frac{\text{Sweep bandwidth, } f_{\mathrm{b}}}{\text{Sweep time, } \delta_t} \tag{3}$$

The whole system is dependent on the parameters, bandwidth and slope.

The transmitted and received signal for the single object detected is shown in Fig. 4 [13]. The range, R, is calculated as

$$\text{Range, } R = \frac{c * \delta_t}{2} \tag{4}$$

To extract both range and speed of the target car, this model uses Root-MUSIC algorithm. It helps in computing frequencies and powers of sinusoids [15].

2.2 Radar and Target Parameters

Parameters of Radar and target are given in Tables 1 and 2, respectively. Experimental results verify the theory [16].

1. *Radar parameters*: A radar system makes use of radio frequency electromagnetic signal which are reflected from the target in order to determine information regarding that target. In any radar system, the signal transmitted and received will reveal many characteristics. Based on a simple model of a phased array system, radar parameters are discussed below [13].

Table 1 Radar parameters

Parameters	Value
Operating frequency	77 GHz
Maximum target range	200 m
Range resolution	1 m
Maximum target speed	230 km/h
Sweep time	7.33 ms
Sweep bandwidth	150 MHz
Maximum beat frequency	27.30 MHz
Sample rate	150 MHz

Table 2 Target parameters

Parameters	Value
Car distance	43 m
Car speed	96 m/h

2. *Target parameters*: Target is the detected car. In this experiment, the moving target with speed 96 km/h is considered. Other parameters are mentioned in the table.

3 Simulation Results

All the simulations of an automotive radar are done in the platform of MATLAB using different toolboxes, namely phased array system toolbox and automated driving toolbox [17]. Phases array system toolbox is helpful to develop the model of radar transmit and receive hardware. It also helps in signal processing, object detection and estimation. Automated driving toolbox models the vehicle motion using the driving scenario object. Multi-object tracker object is used to track the synthetic vehicle detections. The results of this work are a concept of generic radar model, which is individually adjustable by the characteristics of a radar system. The FMCW waveform used in the radar system is set up by defining the radar parameters [13]. Figure 5 shows the FMCW signal which is varying with frequency with respect to time axis. It also shows the spectrogram of the same FMCW signal. The spectrum analyzer shows the spectrum of each received sweep and also its de-chirped signal which is as shown in Fig. 6. It is to be noticed that channel 1 in Fig. 6 is the received signal and has a wideband which sweeps through the entire bandwidth and the channel 2 is the de-chirped signal and has a narrowband.

From the Range-Doppler response as shown in Fig. 7, it is observed that the car in front is a little more than 40 m away and appears quite static as the source and the target are both moving. This is likely to happen because the radial speed of the car relative to radar is only about 4 km/h, which is a mere 1.11 m/s. In this paper, the

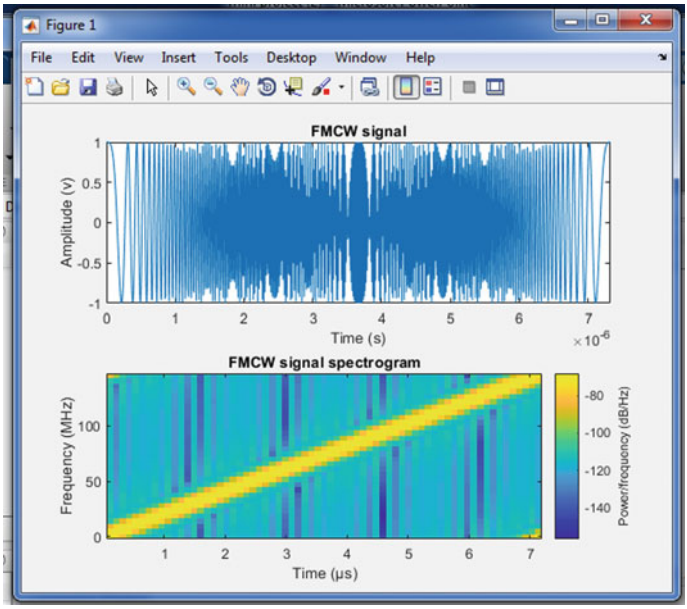


Fig. 5 FMCW signal and its spectrogram

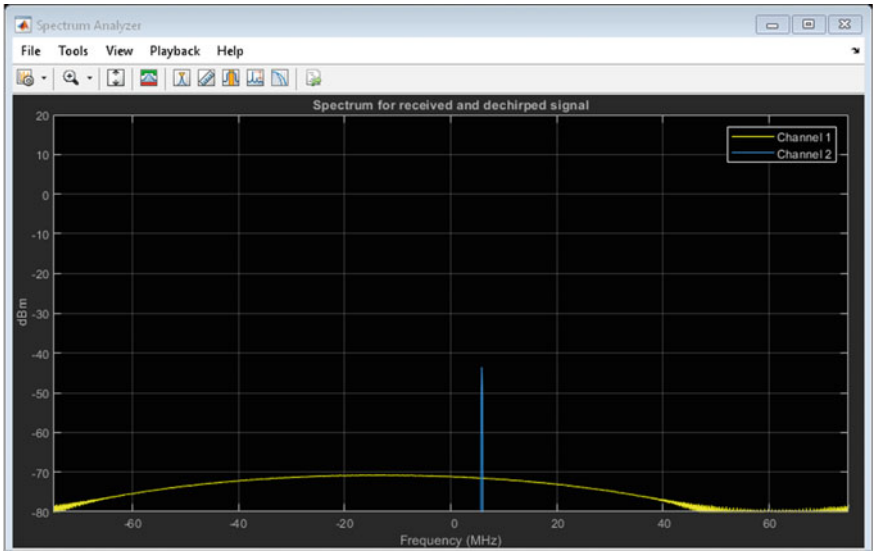


Fig. 6 Spectrum of received and de-chirped signal

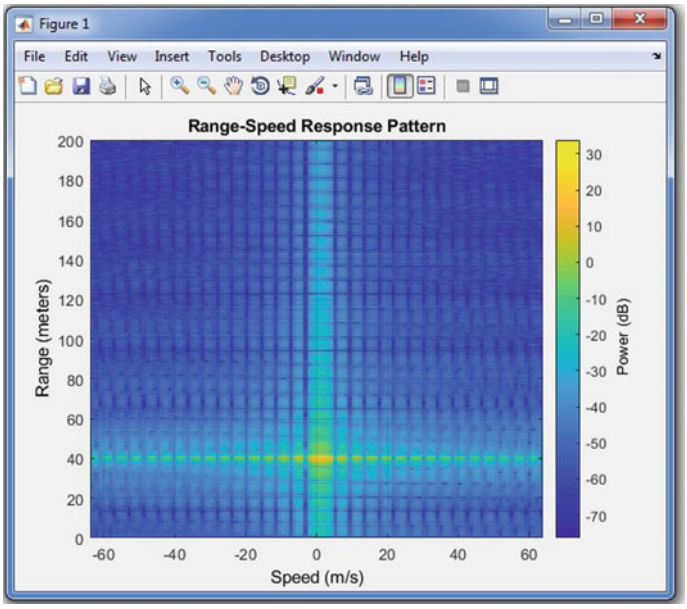


Fig. 7 Range-Doppler response pattern for the object estimated at a particular distance

Root-MUSIC algorithm is used to extract both the beat frequency and the Doppler shift [18]. Figure 8 shows the radar detection, and it tracks for the three target vehicles at 1.1 s of simulation time. The upper-left side of the plot shows the chase camera view of the driving scenario from the ego vehicles point of view shown in blue color. The reference speeds of the vehicles are given as follows: Ego vehicle is traveling at 80 km/h, and the yellow, orange and purple color cars are traveling at 100 km/h, 110 km/h and 130 km/h, respectively.

On the right side of Fig. 8, we have the bird's eye view of the scenario. The vehicles, direction and tracks are in the ego vehicle's coordinate reference frame. The SNR estimation is also shown just beside the detection. The estimated location of the vehicle by the tracker is marked with black squares along with the track's ID. The estimated velocity of each vehicle is marked in black line directed toward the vehicle's velocity. The length of the line is directly proportional to the estimated speed, longer the line higher is the vehicle's speed with respect to the ego vehicle. The purple car with ID2 shows the longest line, whereas the yellow car with ID1 has the shortest. This shows that the tracked speeds are in consistent with the modeled vehicle speeds discussed previously in Table 2.

The plot on the lower-left side of the figure shows the radar images generated by the signal processing. The upper plot shows the distribution of received radar echo in range and radial speed. Here, all three vehicles are observed. The lower plot shows the special distribution of received target echoes in range and angle. Again, all three targets are present, and their locations match with what is shown in the bird's eye

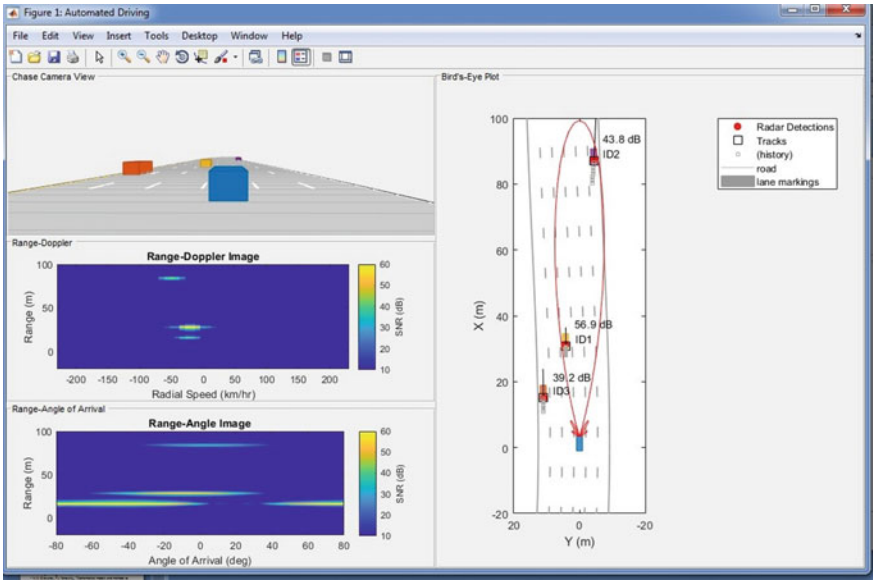


Fig. 8 Chase camera view and bird's eye plot

plot. Due to its closeness to the radar, the orange car can still be detected in spite of the large beam forming losses due to its position well outside of the beam's 3 dB beam width. These detections have generated a track (ID3) for the orange car. For validation, a part of real-time radar data is obtained using AWR1642 automotive radar. Radar data is obtained as a hex.

The validation of this complete model is done in real time. The radar data is obtained within the maximum range of 50 m. The object is detected at the range of 20 m as shown in Fig. 9.

4 Conclusion

This paper has roofed the modeling and simulation of FMCW radar system for automotive applications using MATLAB-based tools. It includes FMCW waveform generation, antenna characterization and also range and speed determination of the target. This helps to overcome collisions and also lend a hand to uphold fuel efficient flow of traffic. The validation of real-time radar data is also made. Advancements have pick up the pace with the participation of the top car developers including BMW, Ford, Audi and most recently Google and Tesla entering the ring.

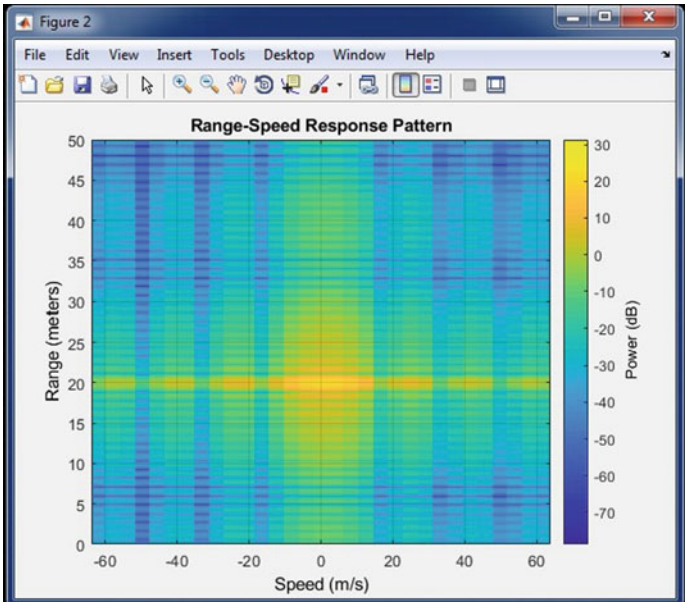


Fig. 9 Validating the target at 20 m with maximum detection range of 50 m

References

1. Carpenter, S.: Autonomous Vehicle Radar: Improving Radar Performance with Simulation. High Frequency Electronics, ANSYS
2. Pattanashetty, V.B., Iyer, N.C., Viswanath, H.L.: Intellectual conveyance structure for travellers. In: Saeed K., Chaki N., Pati B., Bakshi S., Mohapatra D. (eds.) Progress in Advanced Computing and Intelligent Engineering. Advances in Intelligent Systems and Computing, vol. 564. Springer, Singapore (2018). https://doi.org/10.1007/978-981-10-6875-1_37
3. Pattanashetty, V.B., Iyer, N.: Smart driving assistance using Zigbee. No. 2015-28-0105. SAE Technical Paper (2015)
4. Iyer, N.C., et al.: Millimeter-wave AWR1642 RADAR for obstacle detection: autonomous vehicles. In: Saini, H., Singh, R., Tariq Beg, M., Sahambi, J. (eds.) Innovations in Electronics and Communication Engineering. Lecture Notes in Networks and Systems, vol. 107. Springer, Singapore (2020). https://doi.org/10.1007/978-981-15-3172-9_10
5. Pattanashetty, V.B., Iyer, N.C., Viswanath, H.L., Kore, S.: Inclusive device to connect people with the surrounding World. In: Fong, S., Akashe, S., Mahalle, P. (eds.) Information and Communication Technology for Competitive Strategies. Lecture Notes in Networks and Systems, vol. 40. Springer, Singapore (2019). https://doi.org/10.1007/978-981-13-0586-3_49
6. Iyer, N.C., Gireesha, H.M., Shet, R.M., Nissimgoudar, P., Mane, V.: Autonomous driving platform: an initiative under institutional research project. Procedia Comput. Sci. **172**, 875–880 (2020). https://doi.org/10.1016/j.procs.2020.05.126
7. Grgic, M.: Generic radar model for automotive applications. Masters thesis, Graz University of Technology
8. Shraddha, B., Subhas, M., Nikita, P., Preeti, P., Bhagyashree, K., Shet, R., Jyoti, P., Iyer, N.: Model based learning of linear integrated circuit course: a practical approach. J. Eng. Educ. Transf. (2016). https://doi.org/10.16920/jeet/2016/v0i0/85549

9. Krnfelt, C., Pden, A., Bazzi, A., El Haj Shhad, G., Abbas, M.: 77 GHz ACC Radar Simulation Platform
10. Pattanashetty, V.B., Iyer, N.C., Dinkar, A., Gudi, S.: Self-coordinating bus route system to avoid bus bunching. In: Satapathy, S., Bhateja, V., Joshi, A. (eds.) Proceedings of the International Conference on Data Engineering and Communication Technology. Advances in Intelligent Systems and Computing, vol. 469. Springer, Singapore (2017). https://doi.org/10.1007/978-981-10-1678-3_50
11. Rohling, H., Meinecke, M.-M.: Waveform Design Principles for Automotive Radar Systems. Technical University of Hamburg-Harburg, Germany
12. Glazunov, N.M., Yanovsky, F.J.: Mathematical Models and Methods to Radar
13. Maier, F.M., Makkapati, V.P., Horn, M.: Environment Perception Simulation for Radar Stimulation in Automated Driving Function Testing
14. Moran, B.: Mathematics of Radar
15. Nikita, P., et al.: Active learning in electronic measurements and instrumentation course through hands-on. J. Eng. Educ. Transform. [S.l.] (2016). ISSN2394-1707. Available at: https://journaleet.org/index.php/jeet/article/view/85543. Date accessed: 29 Sept 2020. https://doi.org/10.16920/jeet/2016/v0i0/85543
16. AWR1642 Datasheet by Texas Instruments
17. Phased Array System Toolbox. A white paper by MathWorks
18. Bhagyashree, K., Iyer, N.C., Arlimatti, T., Shet, R.M., Preeti, P.: Object detection techniques using deep learning: a survey for real-time applications. TEST Eng. Manage. **82**(Jan/Feb 2020), 5485–5493 (2020)

Effect of Data Augmentation on the Accuracy of Convolutional Neural Networks

Pushkar Bhuse, Bhavesh Singh, and Purva Raut

Abstract Data augmentation is one of the most famous techniques to increase the efficiency of a deep learning model. It involves using conceptual techniques to generate new data samples from existing ones. In the context of image processing, there are a plethora of options for performing data augmentation, for example, rotating images, changing hue, altering greyscale, etc. The main reason why data augmentation is used to reduce overfitting and increase the accuracy of a model. Convolutional neural networks (CNNs), conversely, are the most famous deep learning models used in the domain of image processing, image classification in particular. However, not much is known about the impact of various data augmentation techniques on the performance on various CNNs. In this paper, we analyse the effect of data augmentation on two different types of CNNs of varying dimensions. The aim of this study is to observe the impact of data augmentation techniques on the performance of CNNs of varying complexities. In order to bolster our findings, we used three different datasets for the use case of image classification and tried to compare the results for both CNNs. We analysed the performance of both types of CNNs over varying learning rates and epochs in order to compare our findings across varying hyperparameters.

Keywords Deep learning · Convolutional Neural Network · Data augmentation · Image processing · Hyperparameters · Image classification

1 Introduction

With reference to image processing, data augmentation (DA) is one of the most ubiquitous techniques to reduce overfitting and improve model accuracy. It encompasses a suite of techniques that enhance the size and quality of training datasets

P. Bhuse (✉) · B. Singh · P. Raut
Department of Information Technology, Dwarkadas J. Sanghvi College of Engineering, Mumbai, India
e-mail: bhavesh.singh@djsce.edu.in

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_33

in order to make it more versatile for deep learning models that use it to make predictions. Shorten et al. in [1] discuss the various DA techniques and its impact on deep learning models specifically in the domain of computer vision. In their survey, they discussed various algorithms including geometric transformations, colour space augmentations, kernel filters, mixing images, random erasing, feature space augmentation, adversarial training, generative adversarial networks, neural style transfer and meta-learning.

Convolutional neural networks (CNN) are one of the most famous deep learning models adopted for tasks pertaining to image processing. One of the main applications of CNN falls in the subdomain of image classification. In [2], Neha S. et al. show the performance of various CNNs for the task on image classification on ImageNet, CIFAR10 and CIFAR100 image datasets and also discuss the performances of individual CNNs across the three datasets. Luke Taylor et al. discuss the various DA techniques that can be used to increase the performance of CNNs in [3]. In CNNs, DA is useful because it increases the amount of data which helps the model in learning efficiently and also because the model gets trained with various outliers, thereby preventing overfitting and thus making the model more accurate.

In this paper, we analyse the working of mainly two types of CNNs. The main difference between the two CNNs is the number of convolutions included in the network. The performance of these networks is compared both with and without DA, and its change in efficiency is recorded. It is observed that the effect of DA on both these networks is different. In order to further bolster the findings, the CNNs were trained on three different image datasets, namely CIFAR-10, SVHN and Fashion MNIST. The same process is carried out over varying learning rates and epochs in order to cover a large number of scenarios as well as to understand the effect of DA on the accuracy of CNNs.

The rest of the paper is ordered as follows, Sect. 2 focuses on the work done with regards to monitoring impact of DA techniques on deep learning models in general and CNNs in particular. In Sect. 3, a brief description of all three datasets used in this experiment is provided. Section 4 describes the components pivotal to the experiment, mainly DA techniques used and the two CNN models implemented in this study. Section 5 comprises the methodology followed in order to observe our findings and make inferences. Section 6 explains the inferences drawn from the conducted experiments and also clearly states the behaviour of each CNN model across various scenarios. Section 7 summarizes the findings shown in Sect. 6 and states the conclusion of the research and suggests its future scope.

2 Related Works

Large amount of research has gone into DA, its optimization techniques and its implications. Attempts were aimed at the effect of various DA techniques on deep learning neural networks. In [4], AlexNet was trained on the CIFAR and ImageNet datasets and observed the change in model performance after using various DA

techniques like flipping, cropping, shifting, etc. It helped in inferring DA techniques like cropping, flipping, WGAN, rotation generally performed better than the other methods and that how combinations of these methods help improve performance of the network. Wang et al. [5] used a comparison of multiple solutions to the problem of DA in image classification and observed these results on various datasets. Taylor et al. [3] experimented the impact of generic DA techniques on a simple CNN model trained on coarse-grained data. They inferred that geometric augmentation methods outperformed photometric methods when it is training on a coarse-grained dataset. In [6], it showed that a model trained using DA was more robust, adaptive and overall better performing than the methods like dropout and weight-decay.

In [7], multiple data augmentation techniques were presented to improve the classification of images by CNNs. It included a unique DA technique where new images of high perceptual quality combining contents of base images were formed. These techniques were tested on three medical case studies. Sebastian et al. [8] investigated the use of augmenting data with synthetically created samples. They used data-wrapping as the technique to generate additional training data. They found that while it is possible to perform generic augmentation in feature space, augmentation in data-space provides a greater benefit for improving the classifiers performance.

However, the effectiveness of DA subject to the complexity of a CNN is not discussed much. There is a need to understand whether CNN models on which the augmented data is being trained will actually benefit from the augmentation or no. This bolsters the motivation to conduct this research. In this paper, a clear comparison is made between the effectiveness of DA techniques on CNN models of different complexities.

3 Datasets

All three datasets used in this research are of the domain of image classification. The datasets used in this research were CIFAR10, SVHN and Fashion MNIST.

3.1 CIFAR-10

The CIFAR-10 is an image dataset which contains a labelled set of 60,000 32×32 colour images in ten classes, with 6000 images per class. There are 50,000 training images and 10,000 test images.

3.2 SVHN

SVHN is a real-world image dataset similar in flavour to MNIST (e.g. the images are of small cropped digits). It incorporates an order of magnitude more labelled data of over 600,000 digit images and comes from a real-world problem of recognizing digits and numbers in natural scene images.

3.3 Fashion MNIST

Fashion MNIST is a dataset of Zalando's article images that consists of a training set of 60,000 examples and a test set of 10,000 examples. Each example is a 28×28 grayscale image, associated with a label from ten classes. It shares the same image size and structure of training and testing splits.

4 Architecture

4.1 Data Augmentation

Limited data is a major problem in applying deep neural networks like convolution neural networks. These networks require a huge amount of data for training and to provide better performances, and getting this dataset can be costly. Data augmentation is one of the techniques where the data is just duplicated by applying some variations to it and hence enlarged. These include applying some filters, zooming, shearing images, etc., to the original image. But however, augmentation can be counterproductive if it produces images very dissimilar, thereby losing the key features of the original images, and hence, this technique must be carefully applied. Data augmentation can also be helpful in preventing overfitting the model and improving the overall accuracy of the model.

In this research, the following data augmentation techniques were applied:

- Random Rotation: In this technique, the original input images are rotated randomly in anti-clockwise and clockwise directions by a certain degree α. In this study, α was set to 15°.
- Random Affine: This technique allows you to specify random affine translations of the image (scaling, rotations, translations, and/or shearing or any combination.
- Zooming: This technique is used to focus on particular pixels in an image. Scaling factor of 0.2 was used in our system
- Random Horizontal Flips: If the technique is applied, it randomly flips the input images.

- Colour Jitter: This technique comprises various filters that are applied on the input images with a certain threshold. Various filters such as of brightness, hue, saturation and contrast with a threshold value of 0.2 were used in transforming the input images.

4.2 Convolutional Neural Network

A convolution neural network is a deep neural network which has multiple layers and takes in an input image, extracts the important features from it which can either be used for classifying the image or for some other purpose. In the convolution layer, various features maps are mapped on the input image, and convolution operation is applied. Feature maps are just filter matrices which are responsible for extracting different features from an input image. These filter maps are learned automatically by the network to extract important features from the image. The pooling layers section would reduce the number of parameters when the images are too large but retain important information. Pooling can be of different types such as average, max or sum pooling. After this step, the two-dimensional image vector is flattened into a single large vector which is passed through dense layers. Dense layers are a series of linear neural layers. A deep convolution neural is the one which consists of many layers of convolution, pooling together one after the other followed by dense layers.

In this study, we have used two different convolution models, to show the impact of data augmentation on the overall accuracy of the system.

4.3 Shallow Convolutional Model

This model uses three convolution layers consisting of 32 filters and filter-size 3 with relu activation for the first three layers. As shown in Fig. 1, the convolutional neural networks apply a filter to an input to create a feature map that summarizes the presence of detected features in the input. Pooling is not applied because the image size is already very small. All the inputs to these layers are batch normalized which makes them have a mean of zero and standard deviation of one. The output of the final convolution layer is applied to the linear layers. The final layer has a Softmax activation to give out a probability distribution over ten different classes.

4.4 Deep Convolutional Model

In Fig. 2. The CNN consists of three sequential models. With each sequential model consisting of two convolution layers with filter-size 3, one max pooling layer. The number of filters, i.e. feature maps for the convolution layers in each sequential model

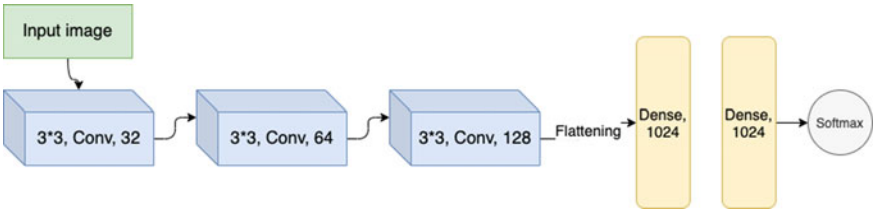


Fig. 1 Shallow Convolutional Model Architecture

doubles with starting from 32 with a padding of size 1 applied to the input image. Padding is mainly because the pixels in the corners and edges are not given much weight while processing, and hence by padding, important information at edges and corners are shifted further inwards. The relu activation function is applied and batch normalized and then passed through a max pooling layer with a filter size of 2. The output of the last sequential model is passed through linear neural layers which is distributed over ten classes with Softmax as an activation function.

4.5 Hyperparameters

In this research, mainly the variation of two hyperparameters was considered, namely learning rate and number of epochs for which the model was trained. In order to cover multiple permutations, a grid search was employed to observe the performance of the two CNNs for varying values of these two hyperparameters. As supported by [9], learning rate is a very important factor in improving the accuracy of a deep learning model. In this research, each CNN was mainly trained on three different LR values, 0.01, 0.001 and 0.0001. A similar grid search was employed to also observe a number of epochs that would be more suitable to demonstrate our results more clearly.

5 Methodology

This research was a highly iterative process. It involved trying out multiple combinations of CNN architectures, DA techniques and hyperparameters simultaneously. In order to carry forward the research, each dataset was preprocessed and used by both the CNNs to make predictions.

The first step included no DA. The same process was carried out on both CNNs after implementing the data augmentation techniques mentioned above. In every combination, all possible combinations of hyperparameters, i.e. LR and number of epochs, were taken into consideration. The variation of accuracy was observed for

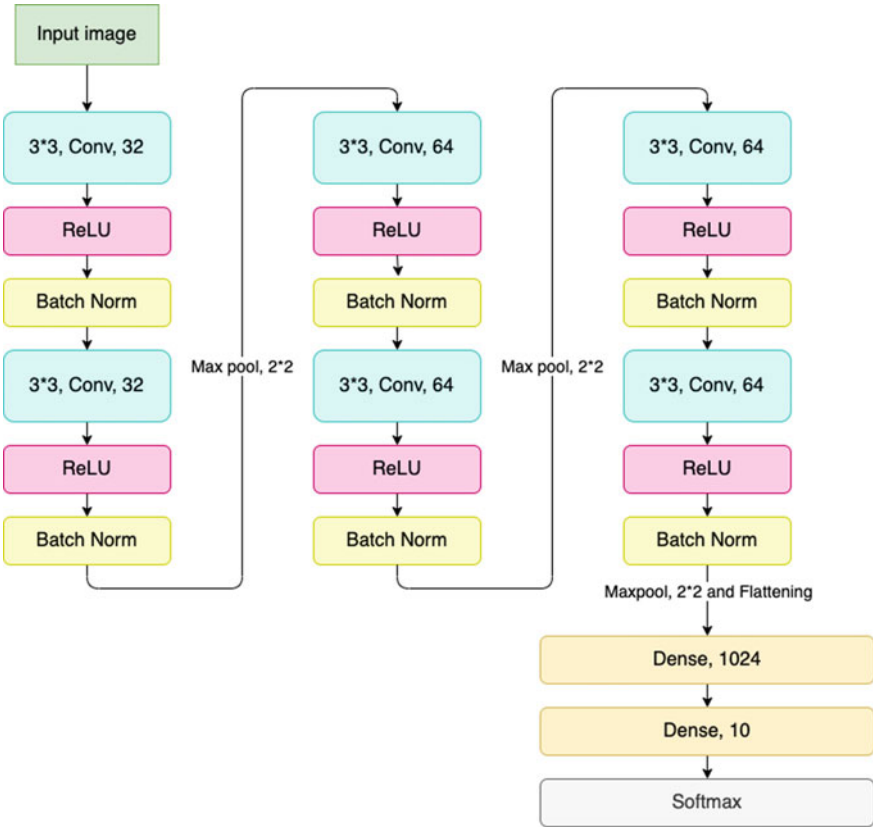


Fig. 2 Deep convolutional model architecture

each CNN with and without DA for a wide array of hyperparameter combinations, and its progression was recorded.

In the first stage of experimentation, DA was ignored, and this data was applied as it is on both the CNNs. In the second stage, however, the preprocessed data is augmented and then fed to the above-mentioned CNN models. In order to draw the most appropriate inferences, a comparison is made between the performances of each type of CNN model before and after the application of DA. If the performance of a CNN model (basic or complex) increases (in terms of prediction accuracy), data augmentation could have had a positive impact, whereas if accuracy degrades, it can be inferred that data augmentation did not help in improving the overall performance of the model.

6 Results

As described in Sect. 5, both CNNs were employed on each dataset to check for performance enhancement with or without data augmentation techniques explained in Sect. 4.1. Given below is the elaboration of the performance of both CNNs on individual datasets. In order to cover all cases, the variation of accuracy is shown with respect to learning rate and number of epochs as well. The accuracy shown in the graphical representation refers to the testing accuracy.

6.1 CIFAR-10

As shown in Fig. 3, on training the complex CNN on the CIFAR dataset even without data augmentation, the accuracy of the model remained comparatively high. By increasing the number of epochs, the effect on the model accuracy was negligible, whereas decreasing the learning rate improved accuracy. By the inculcation of data augmentation mechanisms, although slightly, the accuracy of the model did improve. The way learning rate and number of epochs affected the model remained the same.

In the case of the basic CNN model, it can be conserved that without the application of data augmentation techniques the model displayed a maximum testing accuracy of 0.86. Decreasing the learning rate increased the performance of the model, and in most cases, increasing the number of epochs caused the model to overfit. However, when data augmentation was employed, and the maximum accuracy of the CNN fell to 0.69. Since overfitting was curbed by DA, increasing the number of epochs now increased the model performance. The impact of learning rate remained the same. It could be concluded from these observations that data augmentation in a basic CNN reduced its performance while improving it in a complex CNN.

6.2 SVHN

As observed in Fig. 4, one quality consistent with the SVHN dataset across all implementations is that increasing the number of epochs in the training phase improves model performance. In case of the complex CNN, it could be observed that data augmentation helped in increasing its performance by a small level. Since the model performs considerably better without DA, the addition of data augmentation improves its performance slightly. However, addition of data augmentation helps in reducing overfitting at higher learning rates (0.001).

In the case of a basic CNN model, an anomalous behaviour similar to the one shown by the CIFAR-10 dataset is observed. Without data augmentation, the model is capable of achieving an accuracy of WRITE ACC. But the addition of data using DA, model performance at all learning rates deteriorates. Even the problem of overfitting

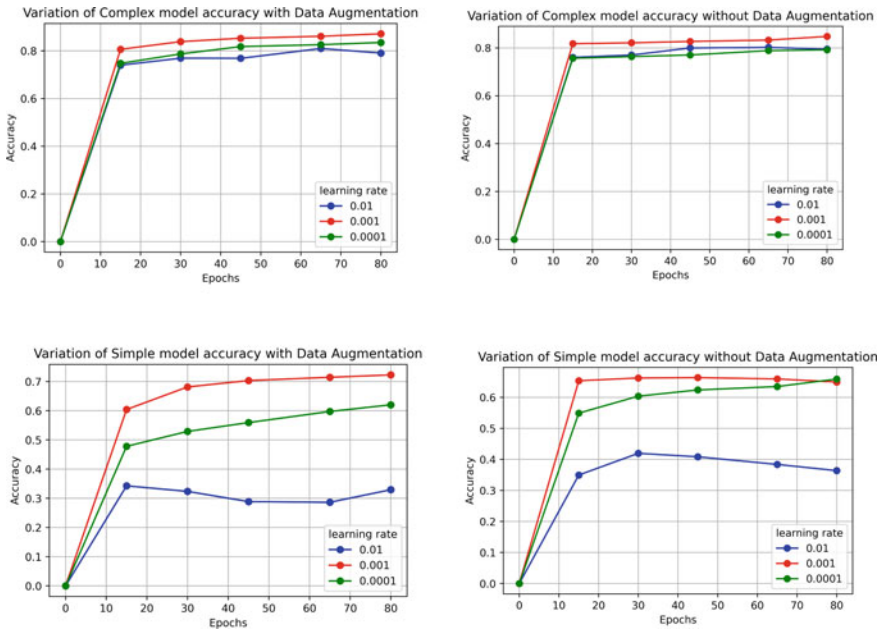


Fig. 3 Variation of testing accuracy of the complex and basic CNN with and without DA on CIFAR-10 dataset

which was solved by DA in the complex model does not help in the case of the basic CNN when the learning rate is 0.01. The variation of accuracy with increasing the number of epochs remains the same in both situations.

6.3 Fashion MNIST

As seen in Fig. 5, in both cases, the complex CNN performs quite similar with and without DA. It can be observed that in almost all cases, an increase in the number of epochs helps in improving model performance. Data augmentation has virtually no impact on the performance of the complex CNN. By varying the learning rate, no significant change in model accuracies was encountered. In both cases, model accuracies reached a maximum accuracy of mention accuracy.

The observations regarding the basic CNNs are aligned with those seen in the CIFAR-10 and SVHN dataset. It is seen that model accuracy deteriorates by the inclusion of augmented data. Learning rate variation has the same impact in both cases. Since DA helps in reducing overfitting, the model can train for a larger number of epochs without loss in accuracy.

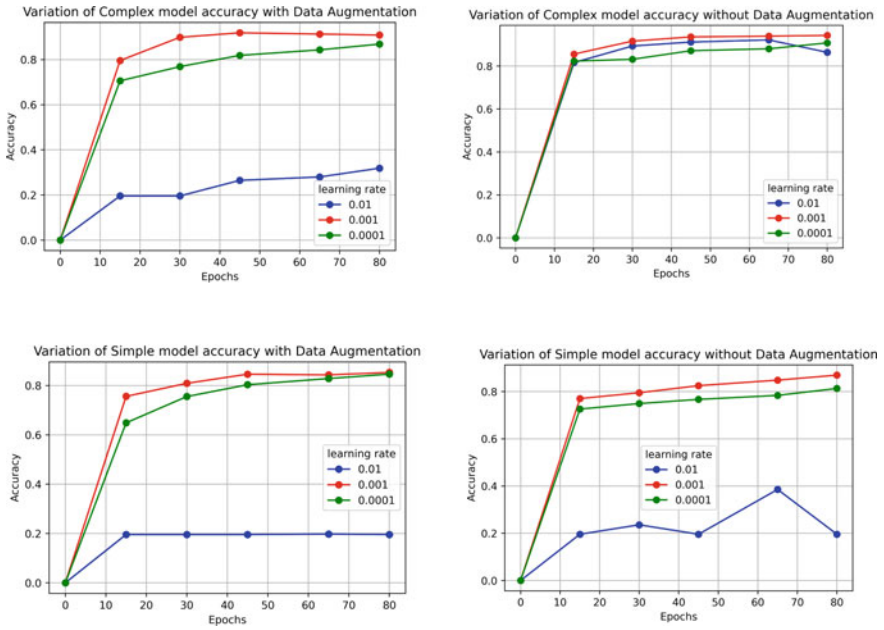


Fig. 4 Variation of testing accuracy of the complex and basic CNN with and without DA on SVHN dataset

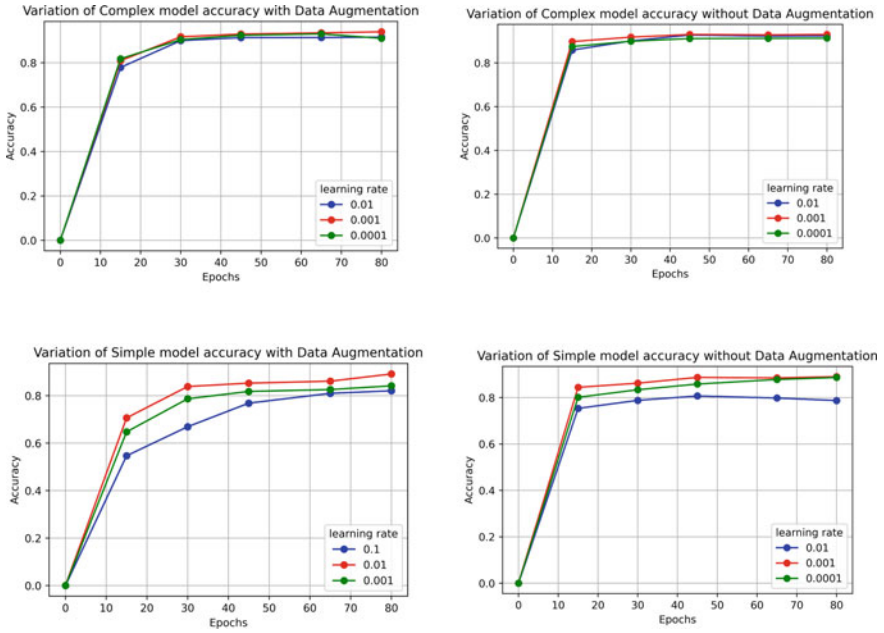


Fig. 5 Variation of testing accuracy of the complex and basic CNN with and without DA on Fashion MNIST dataset

7 Conclusion

The graphical representation of the effect of DA techniques, as observed on all three datasets, on both the complex and basic CNN model depicts some clear inferences. As expected, it can be inferred that DA helps in overcoming overfitting problems as well as improving model accuracy in the complex CNN model. It could also be inferred that in the cases where the complex model has considerably high performance and low overfitting, data augmentation causes no or in some cases a negative impact on the models performance. Having said that, DA also helps in cases when there is a need to train the model for a larger number of epochs.

However, the same cannot be said in the case of the basic CNN model. Anomalously, DA techniques had quite a detrimental impact on the performance of these models throughout the three datasets. During training, inclusion of DA delayed model convergence by a great extent and also did not seem to counter overfitting. Throughout the three datasets, the basic CNN model performed better without data augmentation as compared to with it. As shown in Sect. 6, for all datasets, the CNN model of lower complexities underperformed whenever DA was used. This anomalous behaviour caused the difference between the impacts of DA techniques on the performance of CNNs with different complexities.

The behaviour shown by the CNNs mentioned above and the anomalous behaviour of DA on their performance encourages further research about this behaviour. There can be general trends that can be followed by observing the effect of DA on different CNNs by gradually varying their complexity. This could help in mathematically understanding the relation between model complexity and DA. The same experiment can be carried out in a more focused manner by taking into consideration individual DA techniques and their impact on CNN performance. Lastly, the veracity of these observations should be tested on different datasets and use cases to make a more cogent inference.

References

1. Shorten, C., Khoshgoftaar, T.M.: A survey on image data augmentation for deep learning. J. Big Data **6**, 60 (2019). https://doi.org/10.1186/s40537-019-0197-0
2. Shijie, J., Ping, W., Peiyi, J., Siping, H.: Research on data augmentation for image classification based on convolution neural networks. Chinese Automation Congress (CAC). Jinan, pp. 4165–4170 (2017). https://doi.org/10.1109/CAC.2017.8243510
3. Perezand, L., Wang, J.: The effectiveness of data augmentation in image classification using deep learning. arXiv preprint arXiv:1712.04621 (2017)
4. Hernández-García, A., König, P.: Further advantages of data augmentation on convolutional neural networks. In: Kůková, V., Manolopoulos, Y., Hammer, B., Iliadis, L., Maglogiannis, I. (eds) Artificial Neural Networks and Machine Learning - ICANN 2018. ICANN 2018. Lecture Notes in Computer Science, vol. 11139. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-01418-6_10

5. Mikołajczyk, A., Grochowski, M.: Data augmentation for improving deep learning in image classification problem. International Interdisciplinary PhD Workshop (IIPhDW). Swinoujście, pp. 117–122 (2018). https://doi.org/10.1109/IIPHDW.2018.8388338
6. Wong, S.C., Gatt, A., Stamatescu, V., McDonnell, M.D.: Understanding data augmentation for classification: when to warp? CoRR, abs/1609.08764 (2016)
7. Truex, S., Liu, L., Gursoy, M.E., Yu, L., Wei, W.: Demystifying membership inference attacks in machine learning as a service. IEEE Trans. Services Comput. (2019)
8. Krizhevsky, A.: Learning Multiple Layers of Features from Tiny Images (2009)
9. Netzer, Y., Wang, T., Coates, A., Bissacco, A., Wu, B., Ng, A.Y.: Reading Digits in Natural Images with Unsupervised Feature Learning NIPS Workshop on Deep Learning and Unsupervised Feature Learning (2011)

Ontology for Preliminary Detection of COVID-19

Poly Sil Sen, Shabnam Banerjee, and Nandini Mukherjee

Abstract During the recent COVID-19 pandemic, healthcare workers are trying to manage this medical emergency at war footing. It is the need of the time to have a COVID-19 detection procedure that can be used to stothe kiosk-based operation is shownp the spread among the general people. This paper proposes an ontology-based technique to detect and preliminarily diagnose suspected COVID-19 patients without direct intervention of healthcare professionals. The proposed procedure works as a part of remote healthcare services and tries to respond timely to the situation. The ontology is used for handling general people, rather than patients, to detect COVID-19 primarily.

Keywords Ontology · COVID-19 · Remote health framework · Data model · Workflow · Sensor

1 Introduction

"Coronavirus disease (COVID-19) is an infectious disease caused by a newly discovered coronavirus" [12]. As known from the website of WHO COVID-19 outbreak situation, 33441919 confirmed cases of COVID-19 have been reported worldwide; 1003497 deaths have been reported; 235 countries, areas or territories with COVID-19 cases are reported as on September 30,2020 [13]. COVID-19 has emerged as a health emergency for the world and after 9 months of its initiation of spread, the situation requires to think about new methodologies to deal with. COVID-19 has been accepted as a pandemic. Efforts are being made to make the situation better for

P. Sil Sen (✉)
Techno Main Salt Lake, Kolkata, India

S. Banerjee
Tata Medical Centre, Kolkata, India

N. Mukherjee
Jadavpur University, Kolkata, India
e-mail: nmukherjee@cse.jdvu.ac.in

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_34

common people and people who are dealing with the situation from the front line, i.e., healthcare professionals.

Healthcare procedures include a normal method of investigation and detection of diseases suffered by the patients. A person suffering from some medical complaint visits a doctor. A treatment episode starts with this. In the current situation, this procedure is not sufficient, as early identification and detection of persons suffering from COVID-19 are needed. The next important reason for not following the standard procedure is that healthcare units are overflowing with patients, lacking in sufficient number of healthcare persons and proper infrastructure to manage the huge number of patients.

On the other hand, many healthcare applications are available as mobile applications or as IoT-enabled applications to support the well-being of general people. These applications are particularly useful for remote diagnosis of diseases and advising the patients remotely regarding treatment. Our previous work focuses on the development of a real-time application, based on sensor, mobile and cloud technologies that provide an integrated environment for the patients and remotely located medical professionals to interact in meaningful manner and to create a treatment episode for delivering basic healthcare services to the patients. The application has been deployed in a rural environment in West Bengal, India. The deployment of the application was made in kiosks set up in villages, where health assistants collect the patient's complaint-related data and transmit the data to the doctor over the Internet. The doctor, on the other side of the system, studies those data in real time and generates prescription or puts forward advice related to the treatment of the patient [7].

Recently, some mobile applications have come up in the market to perform the initial diagnosis of a COVID-19 patient. These applications may be connected to governmental agencies to detect and take action on a COVID-19 patient. However, to the best of our knowledge, none of these applications are based on any formal process of application development, they are focused on few specific tasks only and therefore are not extensible for inclusion of future scope.

A healthcare application requires to incorporate the domain knowledge in its functioning. Ontology can help to build a knowledge repository of a domain that can be used for automating certain tasks. Thus, for an IoT-enabled application development, where domain knowledge plays an important role, an ontology can be used in understanding, interpreting, using, validating the domain knowledge and capture, store and manage data accordingly. In our earlier work [9], we have defined an ontology for our remote health framework. However, in the current scenario, the ontology is to be enhanced to deal with the present COVID-19 pandemic situation.

This paper focuses on presenting a formal method to be followed for any IoT-enabled remote healthcare delivery system for initiating the treatment of a patient in non-contact mode with contact tracing and necessary screening. The paper follows the WHO norm and particularly focuses on the *history taking* and the *screening process*. In any diagnosis-related procedure, obtaining accurate history is a critical first step, and in most cases (at least 70% of time), a doctor is able to make the diagnosis based on the history alone [10]. It is observed that previously suggested procedures for history taking are unable to manage the pandemic situation. Therefore,

in this work, we propose to enhance the ontology presented in [9] to ensure that a person visiting a health kiosk running our application may be questioned on the basis of the domain knowledge, and in case the person is suspected to have COVID-19, he or she may be advised to go for further investigation.

In Sect. 2, works done in the field of ontology and COVID-19 are discussed. In Sect. 3, an overview of the kiosk-based IOT-enabled application is provided. In Sect. 4, a workflow of the COVID-19 detection procedure is shown and explained. Section 5 presents the proposed modified ontology. Finally, Sect. 6 concludes.

2 State of the Art

An ontology has been proposed for COVID-19 in [6] for extensive medical research on this disease and to study the explosion of COVID-19-related data meticulously for further understanding and research of the disease. The ontology follows principles of OBO Foundry and is aligned with Basic Formal Ontology and ISO/IEC standard 21838-2 top-level ontology. This ontology represents different coronavirus diseases that cover cause of the disease, transmission of the disease, distribution and control of the diseases, study of the virus, gene-related study in this context, interaction of the virus with host, prevention, diagnosis and treatment of the disease. It is a continuing work [4].

The COVID-19 Surveillance Ontology is an ontology at the application level for monitoring COVID-19 cases. Apart from this, it monitors different associated respiratory conditions from data of medical record systems of different make [2].

The ontology that covers COVID-19 case records of WHO provides "semantic references" to the different questions/answers of the related form. This ontology accommodates various clinical signs, symptoms and questionnaires related to clinical assessment and information of different other diseases. It covers WHO–19-Rapid-CRF assessment questions related to clinical or research perspective. It allows different types of questions and includes admission, discharge and death forms [3].

In order to enhance the research related to infectious diseases and address all related topics like prevention, diagnosis, treatment, cure, vaccine and other domains like public health, a study of all related standards and repetitively reused ontologies, like the Infectious Disease Ontology (IDO) which is developed following the guidelines of Open Biomedical Ontologies (OBO), is done in [1]. Many other ontologies like Vaccine Ontology, etc. are studied in this paper.

COVID-19 ontology and natural language processing is used in [5] to detect fake data related to COVID-19. Reasoning in description logics (DLs) is used to detect inconsistencies between reliable medical sources and not-so-reliable ones.

Case-based reasoning (CBR) is used to detect new cases of COVID-19 positives early by using existing COVID-19 case database [8]. It diagnoses cases of COVID-19 positives and negatives and predicts the pace with which COVID-19 spread will happen.

3 Overview of Kiosk-Based IoT-Enabled Application

Our kiosk-based application has been developed for the villages in India where resources like Internet connectivity and electricity are big constraints. The application has two main components. On one side, there are kiosks deployed in the villages and running the patient-side component of the remote healthcare delivery application. There are five modules in this section which include *sensor data acquisition,* symptom collection, *history taking,* knowledge-based query generation and *collection of patients' responses and finally* transmission to the cloud. The other component is deployed at the doctor's side, from where the doctor can log in, check patients' vitals and other responses and generate prescriptions.

In the present mode of kiosk operation, health assistants (rural girls) have been trained to use the information and communication technologies (ICT) and to collect necessary data from the patients using sensors and using Android-based mobile phones and tablets. Our Android app is installed on the mobile device. On the other side of the application, doctors can view the complaint of the patient, along with other details like vitals, history and even old prescriptions and relevant reports. Based on these data and through remote interactions over the Internet, doctors can put forward their advice for each patient and prescribe medicines. A high-level overview of the kiosk-based operation is shown in Fig. 1a.

4 Proposed Workflow

In the proposed workflow, two issues have been addressed—(1) for each patient, irrespective of whether the patient is symptomatic, or asymptomatic, his or her contact history and travel history are traced. (2) On the other hand, a health worker is remotely screened for any symptoms.

As mentioned earlier, *history taking is an important step. However, a medical professional usually collects the data related to the chief complaint, history of present illness, past medicalsurgical history, systemic review, family history, allergy history, etc.*

Screening is also another necessary step for healthcare applications. "Screening is defined as the presumptive identification of unrecognized disease in an apparently healthy, asymptomatic population by means of tests, examinations or other procedures that can be applied rapidly and easily to the target population [11]." In our remote healthcare application, this screening is to be carried out remotely using eHealth sensors fitted onto the patient's body.

Our kiosk-based application has been designed to collect history and sensor data based on normal protocols as described above. Similarly, no screening process has been included for the health assistants in the kiosk-based remote healthcare delivery application. Thus, the current application needs to be enhanced to include the

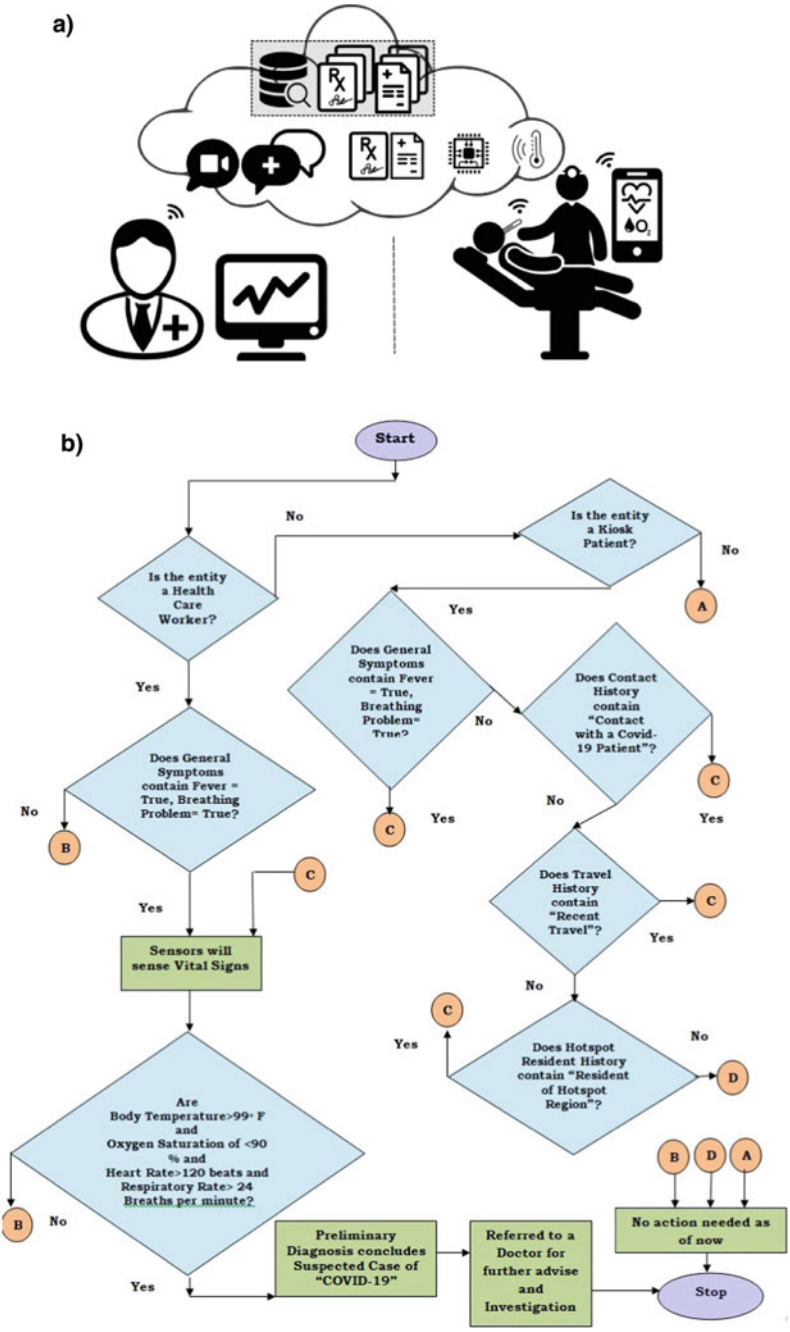


Fig. 1 **a** Kiosk operated rural healthcare delivery, **b** workflow of screening process for preliminary detection of COVID-19

COVID-19 protocols as defined by WHO. A workflow for identifying suspected COVID-19 patients is needed.

The proposed workflow as depicted in Fig. 1b shows that when a normal person or a health worker visits the kiosk, he or she undergoes the screening process to find whether he or she has fever or breathing problem. If it is no, for a normal person, then it is checked whether he or she has come in contact of with COVID-19 patient or has a recent history of travel or he or she is a resident of hotspot region or not. If these conditions are true for a person, then it is more likely that the person is infected with COVID-19. If any of the beforementioned conditions are true for normal person and if a health worker has fever or breathing problem, then screening is performed. In the screening process, the following conditions are checked:

1. Body Temperature >99 F and
2. Oxygen Saturation <90%
3. Heart Rate >120 beats
4. Respiratory Rate >24.

If the above conditions are found to be true, then the person under screening is said to be a suspect of having COVID-19. The patient is then sent for investigation and for further advice from a doctor.

5 Proposed Modified Ontology

As already mentioned, a formally defined ontology is required for a remote healthcare application. The earlier ontology proposed in [9] is inadequate to incorporate the requirements of COVID-19 screening. This is because, in case of COVID-19 screening, non-contact mode is necessary and contact tracing is important. Hence, in this section, the modification to the existing ontology for remote health framework using sensor and cloud is explained in detail. The proposed modified ontology will help in performing screening of normal people and healthcare workers with specific symptoms and normal people with certain conditions true .

In this section, entities or classes are indicated with capitalized words. Association of classes or entities is called object property. The convention used to name the object properties in this work and our previous work [9] is as follows:

1. “has” / “contains” relation is indicated by including the respective string within the name of the object property.
2. First and starting letters of words of entity names to indicate the associated entities are included in the name of the object property.
3. Sometimes “_” is used to separate different parts of the name belonging to different entities.

Data attributes of an entity are called data properties. Some of the entities, data properties and object properties of our previous work of ontology on remote health framework [9] are reused in this paper. *Thing* is the root entity of this ontology and any other ontology.

In this section, all the newly introduced concepts are shown in the figures with red rectangles or squares. The newly introduced concepts are discussed below.

If the workflow of the screening process just discussed above is studied carefully, the first new entity can be found. *Screening* is a type of *ClinicalProcess* which is a type of *Process* which is a type of *Method* as shown in Fig. 2a.

Three types of *Person* are discussed in the flow chart. *HealthCareWorker* is newly introduced. *KioskPatient* is the alias of a normal *NonHealthCareWorker* as shown in Fig. 2b. In this context, it is to be noted that in the current pandemic situation, it is very important to identify different types of health workers and other related front-line workers working in the field of health care. According to WHO [14], *Doctor*, *LaboratoryTechnician*, *Nurse* are of the type of entity *DirectHealthServiceProvider*. *PhysicianEmployed* belongs to the type of the entity *HealthServiceProviderInOtherSector*. For example, a physician employed in a place of mining or an organization belongs to this category. There are *HealthManagement andSupportWorkers*, like driver of an ambulance, clerks in a hospital and other support staff. *HealthServiceProviderInHealthSector* has two types *HealthManagement andSupportWorker* and *DirectHealthServiceProvider*. *HealthCareWorker* entity has two types *HealthServiceProviderInHealthsSector* and *HealthServiceProviderIn OtherSector*. All such persons and other types of persons are related to the entity *Person* as shown in Fig. 2b.

The next needed entity is *GeneralSymptoms*. Clinical symptoms are obtained directly from the patient. In our previous work [9], we have already defined *GeneralSymptoms* which belongs to *ClinicalSymptoms* and which is a type of *Clinical InformationEntity* as shown in Fig. 2d. In this work, two data properties or attributes, *fever* and *breathingProblem* of *GeneralSymptoms* entity, are added to decide whether a *HealthCareWorker* or a *NonHealthCareWoker* is supposed to be screened for COVID-19 or not.

According to the workflow, three types of history or medical history are studied. These are *ContactHistory*, *TravelHistory*, *HotSpotResidentHistory*. *PresentHistory* which is a type of *MedicalHistory* which is a type of *EventCollection* which is in turn type of *Collection*. *MedicalHistoryAll* is also another type of*EventCollection* that contains all types of *MedicalHistory* 3a.

The *Person* entity is associated with the *MedicalHistoryAll* with the *hasPMHAll* object property (association of entities). The *MedicalHistoryAll* alias *HistoryAll* alias *MhistoryAll* entity is discussed in our previous paper [9]. *MedicalHistoryAll* is associated with all types of History by *containsHAll* object property.

PresentHistory is associated with *ContactHistory*, *TravelHistory*, *HotSpot ResidentHistory* as shown in Fig. 3c by the object property *containsPrH_CH_TH_ HRH*. Each of the lastly mentioned entities has an attribute containing details of the specific history type, it belongs to. These attributes are used in the screening process for checking contact with COVID-19 patient, travel and staying at hotspot area.

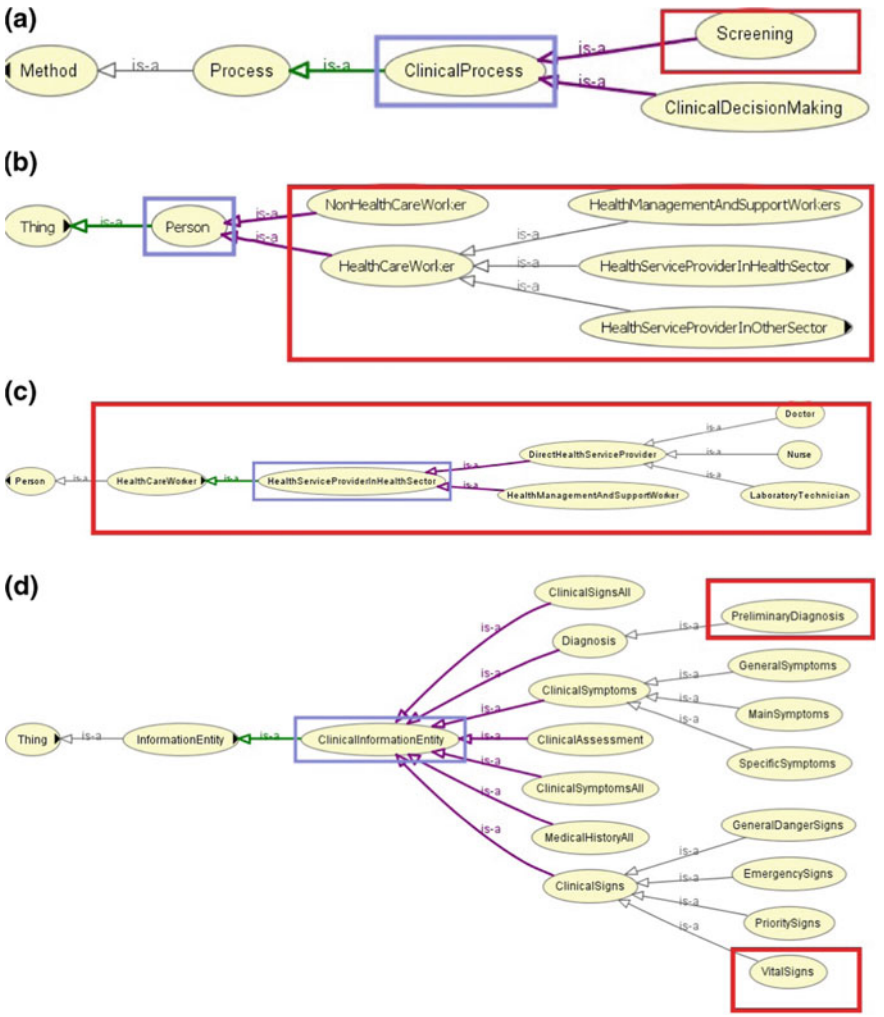


Fig. 2 **a** Screening, **b** Non healthcare workers, **c** Healthcare workers, **d** Clinical information entity, **e** Sensor types

In the workflow, *Sensor* senses few parameters which are *VitalSigns*. Clinical signs are observed or checked in a person. *ClinicalSignsAll* which is a *ClinicalInformation Entity* is related to all types of *ClinicalSigns* by the object property *hasSignsAll*. *VitalSigns* is a type of *ClinicalSigns* as shown in Fig. 2e. There are many data properties of *VitalSigns* like *heartRate*, *respiratoryRate*, *respiratoryEffort*, *bloodPressure*, *bodyTemperature*, *oxygenSaturation*. Among these data properties, *heartRate*, *respiratoryRate*, *bodyTemperature* and *oxygenSaturation* are used to decide whether a *Person* is preliminary detected with COVID-19.

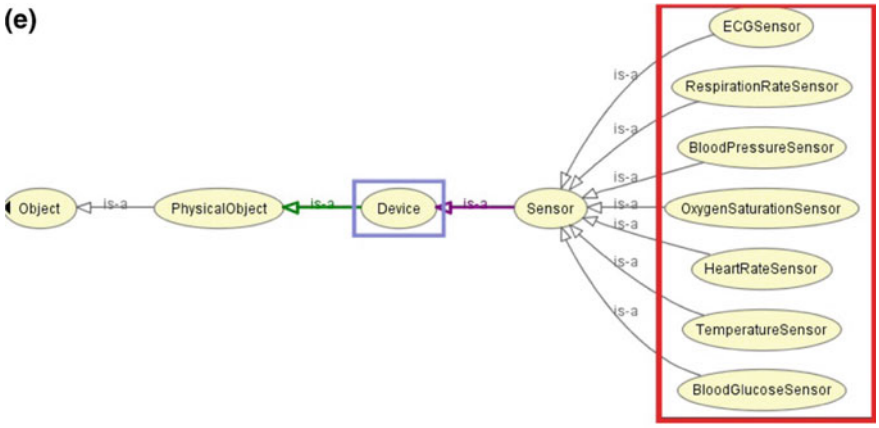


Fig. 2 (continued)

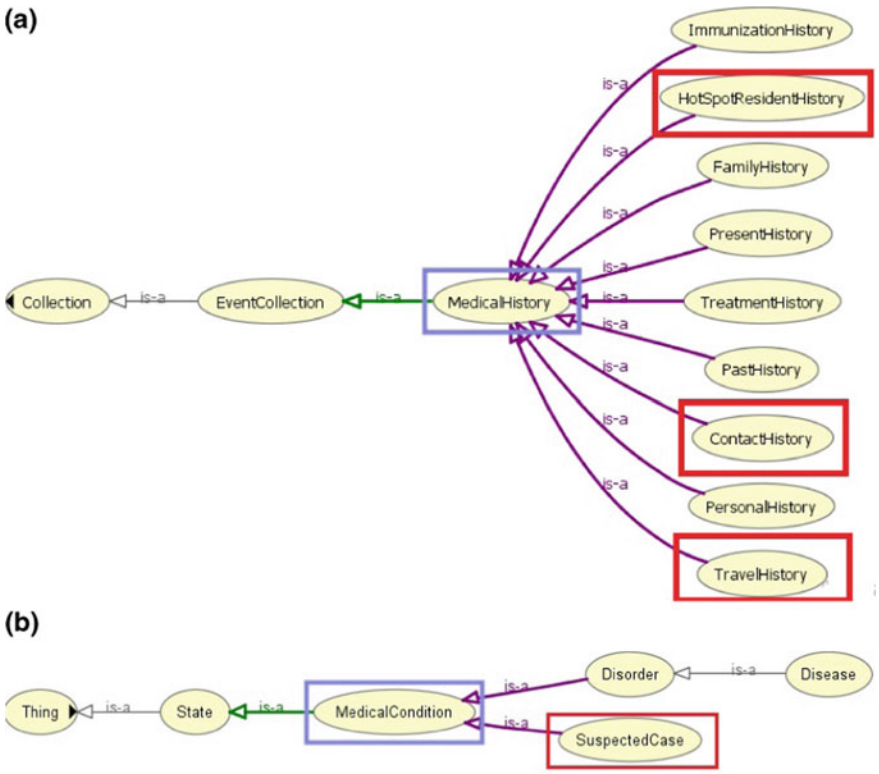


Fig. 3 **a** Medical history, **b** Medical condition, **c** Sensor types, **d** Needed object properties

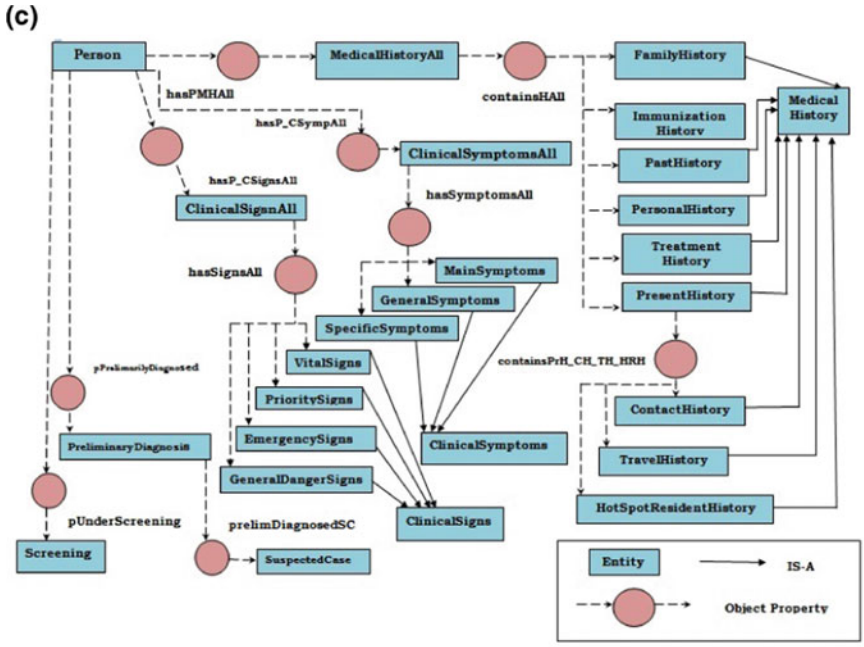


Fig. 3 (continued)

In the workflow, suspected case is found to be Covid-19 positive when certain medical constraints are found positive. A *SuspectedCase* is a *MedicalCondition* which is a *State* as shown in Fig. 3b. A *State* is a *Thing*. If all conditions of preliminary detection of COVID-19 are found to be true, then the *PreliminaryDiagnosis* (which has *prelimDiagnosisDetails* data property) is related to the entity *SuspectedCase* by *containsPD_SC* object property. *SuspectedCase* has the *sCDetails* data property value to be "COVID-19."

PreliminaryDiagnosis is a type of the entity *Diagnosis* which is a type of *ClinicalInformationEntity* as shown in Fig. 2d.

Sensor has many types including *BloodGlucoseSensor* ,*BloodPressureSensor*, *ECGSensor*, *HeartRateSensor*, *OxygenSaturationSensor*, *RespirationRateSensor*, *TemperatureSensor* [15] as shown in Fig. 2e. Among these sensors, *HeartRateSensor*, *OxygenSaturationSensor*, *RespirationRateSensor*, *TemperatureSensor* are used in the current process of preliminary detection of COVID-19.

Detail of the entities of our previous work [9] is provided in the following table 1.

Details of the newly introduced and needed object properties are listed in Table 2. The introduction of the new object properties is needed as before the detection procedure started from a patient, and now, the detection procedure is starting with a *Person* who can be a *HealthCareWorker* or a *NonHealthCareWorker*. Association of the multiple entities in an object property can be found in Figure 3c.

Table 1 Details of needed entities of previous ontology

Sl	Entity name	Description
1	*MedicalHistoryAll*	A record about a person's health
		It contains various types of illness and treatment details etc
		It contains different types of *MedicalHistory* with the help of the object property *containsHAll*
2	*GeneralSymptoms*	It is a type of *ClinicalSymptoms* which is patient's words about the problem
		ClinicalSymptoms is a type of *ClinicalInformationEntity*
		That is a type of *InformationEntity* Which is a type of *Thing*
		ClinicalSymptomsAll contains all types of *ClinicalSymptoms* via the object property *hasSymptomsAll*
3	*ClinicalSigns*	*ClinicalSigns* is a type of *ClinicalInformationEntity* that is a type of *InformationEntity*
		which is a type of *Thing*
		ClinicalSignsAll contains all types of *ClinicalSigns* via the object property *hasSignsAll*
		ClinicalSigns are signs observed by a Doctor at the time of a meet
4	*Sensor*	*Sensor* is a type of *Device*. *Device* is a type of *PhysicalObject*
		PhysicalObject is a type of *Object* which is in turn a *Thing*
5	*ClinicalProcess*	*ClinicalProcess* which is a type of *Process*
		It is a type of *Method* which is a type of *Thing*
6	*MedicalCondition*	*Disorder* is a type of *MedicalCondition* which is a type of *State*
		State is a *Thing*

Table 2 Details of object properties

Sl	Object property name	Represents the association of
1	*hasPMHAll*	some *Person* with *MedicalHistoryAll*
2	*hasP_CSignAll*	Some *Person* with *ClinicalSignsAll*
3	*hasP_CSympAll*	Some *Person* with *ClinicalSymptomsAll*
4	*pPreliminarilyDiagnosed*	Some *Person* with *PreliminaryDiagnosis*
5	*pUnderScreening*	Some *Person* with *Screening*
6	*containsHAll*	*MedicalHistoryAll* with different types of *MedicalHistory* not associated with *PresentHistory*
7	$containsPrH\_CH\_TH_HRH$	*PreentHistory* with *ContactHistory*, *TravelHistory*, *HotSpotResidentHistory*
8	hasSignsAll	*ClinicalSignsAll* with different types of *ClinicalSigns* (Defined in our previous work [9])
9	*hasSympAll*	*ClinicalSymptomsAll* with different types of *ClinicalSymptoms* (Defined in our previous work [9])
10	*prelimDiagnosedSC*	*PreliminaryDiagnosis* with *SuspectedCase*

All the object properties are shown in Fig. 3c. To reduce complication of the diagram, the inverse properties are not shown.

6 Conclusion

In this paper, a modified ontology is presented for screening of people with certain medical conditions, before meeting any doctor. Through the experience of the current work and going through the current situation, it is understood that a fixed approach of storing health data using the available healthcare applications is not correct. Approaches of data storage, knowledge representation and data standards must change with change in situations and requirements. Health data standards can have a role, if a person comes under the supervision or care of a care person or a care service provider. In the present situation, it will be a life-saving approach to preliminarily detect a COVID-19 infected person to be COVID-19 positive before laboratory investigation or visit/ encounter with healthcare person. In this way, he or she can go for further advice, investigation and care to care providers. This can only happen if remote healthcare services play an important role with the help of IoT. It is also to be noted that health data standards may be applicable now from a normal person's end, before he or she can be declared a patient.

In this complete scenario, ontology may play an important role as to help and support, as ontology can be reused, modified and enhanced with requirement. It can be processed and used by both computers and humans.

References

1. Babcock, S., Beverley, J., Cowell, L., Smith, B.: The infectious disease ontology in the age of covid-19 (05 2020). https://doi.org/10.31219/osf.io/az6u5
2. BioPortal: Covid-19 surveillance ontology (2020). https://bioportal.bioontology.org/ontologies/COVID19
3. BioPortal: Who covid-19 rapid version crf semantic data model (2020). https://bioportal.bioontology.org/ontologies/COVIDCRFRAPID
4. GeneOntology: Sars-cov-2 - coronavirus (2020). http://geneontology.org/covid-19.html
5. Groza, A.: Detecting fake news for the new coronavirus by reasoning on the covid-19 ontology (04 2020)
6. He, Y., Yu, H., Ong, E., Wang, Y., Liu, Y., Huffman, A., Huang, H.h., Beverley, J., Hur, J., Yang, X., Chen, L., Omenn, G., Athey, B., Smith, B.: Cido, a community-based ontology for coronavirus disease knowledge and data integration, sharing, and analysis. Scientific Data **7** (12 2020). https://doi.org/10.1038/s41597-020-0523-6
7. Mukherjee, N., Bhunia, S.S., Bose, S.: Virtual sensors in remote healthcare delivery: Some case studies. In: Proceedings of the 9th International Joint Conference on Biomedical Engineering Systems and Technologies - Volume 5: HEALTHINF, (BIOSTEC 2016). pp. 484–489. INSTICC, SciTePress (2016). https://doi.org/10.5220/0005823204840489

8. Olaide, O., Ezugwu, A.: Covid19: A natural language processing and ontology oriented temporal case-based framework for early detection and diagnosis of novel coronavirus (05 2020). https://doi.org/10.20944/preprints202005.0171.v2
9. Sen, P.S., Banerjee, S., Mukherjee, N.: Ontology-driven approach to health data management for remote healthcare delivery. In: Proceedings of the 7th ACM Workshop on ACM Mobile Health 2017. MobileHealth 17, Association for Computing Machinery, New York, NY, USA (2017). https://doi.org/10.1145/3084035.3084039
10. Shaikhani, D.M.: History taking (2008). https://www.slideshare.net/shaikhani/history-taking-presentation
11. who: Cancer screening 2020. https://www.who.int/cancer/prevention/diagnosis-screening/screening/en/
12. who: Coronavirus
13. who: Rolling updates on coronavirus disease (covid-19) pandemic
14. who: health workers: Chapter one a global profile (2006). https://www.who.int/whr/2006/06_chap1_en.pdf?ua=1#:~:text=Health%20workers%20are%20people%20whose,health%20of%20their%20com%2D%20munities.&text=It%20shows%20that%20there%20is,even%20across%20low%20income%20countries
15. World, R.W.: Types of medical sensors | functions of medical sensors. https://www.rfwireless-world.com/Articles/Medical-sensor-basics-and-medical-sensor-types.html (2012)

Framework for a Green Campus-LoRa™-Based Low-Power Smart Water Management for Campus

Saroja V. Siddamal, Vinod Talikoti, Preetham Umrani, Shambhavi S. Uppin, Vinuta Kulkarni, and Sushma Hiremath

Abstract In context to smart city initiatives in India, the paper proposes development and deployment of LoRa-based smart water management for campus, a framework for Green Campus. Water is a vital resource, and maintaining this is a key role. Present manual distributing of water in campus is leading to improper management and is time intense. The authors propose a LoRa-based system to automate water distribution in the campus. Lora is used to enable low data communication by sensors and gateway for M2M and IoT. The sensor unit is a low-power design consuming 4 μA in sleep mode and 16 mA in wakeup mode. The web-based solution takes care of water distribution based on the need. The system makes use of solenoid valves. The management of these valves is through cloud-based solution. The solution depicts the day-to-day consumption of water. This data is used in future for predicting water requirement using data analytical and for leakage management of water.

Keywords Lower power · LoRa · Green campus · M2M · IoT

1 Introduction

Water plays a very vital role in human existence but one of the limited recourses. For the proper utilization of this natural resource, usage and distribution of play an important role. With the knowledge of IoT, availability of sensor network can connect M2M and gather data for analysis of water distribution, and usage of the development of water distribution and monitoring system is possible. The authors propose to develop and deploy such a water distribution and monitoring system for campus. The rest of the paper is organized as follows. In Sect. 2, related work is discussed. Section 3 pays a way the available water distribution system and proposes

S. V. Siddamal (✉) · V. Talikoti · P. Umrani · S. S. Uppin · V. Kulkarni · S. Hiremath
KLE Technological University, Hubballi, India
e-mail: sarojavs@kletech.ac.in

P. Umrani
e-mail: preetham.umarani@kletech.ac.in

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_35

smart water distribution system for campus. An experimental result which includes deployment details data visualization graphs is discussed in Sect. 4 and finally the conclusion.

2 Related Work

Authors in paper [1] monitor campus-scale water distribution systems using off-the-shelf, low-cost sensors, electronics, communication links and cloud software stack. The deployment of the system is done in a medium-sized campus. In paper [2], the authors have focused on problem of water distribution and use as well as problems of managing technical infrastructure of water management in the City of Rybnik, Upper Silesia Region, Poland, since 2004. In another paper [3], authors discuss how wireless sensor networks can increase the spatial and temporal resolution of operational data from pipeline infrastructures and thus address the challenge of near real-time monitoring and eventually control. They focus on the use of WSNs for monitoring large diameter bulk-water transmission pipelines. The report [4] presents information on the water sector in an integrated, holistic manner. The report compiles data on the full range of water issues from water hydrogeology to resource use, water quality. The authors in paper [5] design an energy efficient and reliable IoT network for water management that connects sensor nodes to a gateway using very few hops inside the Indian Institute of Science (IISc) campus. The sub-GHz modules are tested in different real-world environments such as open ground, straight road, moderately and densely wooded area, inside a concrete building and on building roof-tops. The authors have compared the performances for characterization of the wireless channels in different environments. The authors in paper [6] proposed a method to watch and forecast the consumption of the water in the domestic pipeline through a web server. The authors in [7] demonstrate the development of sensor nodes for flow and pressure measurement. Using this sensor data, the consumer demands were predicted. The work involved analyzing the reliability of the sensors for long terms, which are used to measure parameters of water like pH, dissolved oxygen, etc. The main processor was powered by 12 V 33 Ah battery by 50 W solar panel. The water quality monitoring RANA kit developed to assist disaster hit location is discussed in [8]. A real-time water balance monitoring system deployed in IISc campus is shown in paper [9]. The real-time data analyzed is put on the server.

3 Water Distribution and Management in Campus

We are onboard working on a project to create product for smart water management system, with university campus as test bed. The objective of this project is to automate the water distribution across entire campus.

The campus water management system (CWMS) collects stores and distributes water from source to various buildings in the entire campus. It has ground reservoirs like bore wells, one over head main tank (OHMT), and each department with overhead tanks (OT) as shown in Fig. 1.

The water gets pumped to the OHMT from the bore wells. From the OHMT through three pipelines, the water gets distributed to the entire campus. The challenges with the CWMA are

- Poor water management
- Wastage of water
- Human dependent.

To address these challenges, authors propose smart water management (SWM) architecture as shown in Fig. 2. The proposed architecture is implemented for one

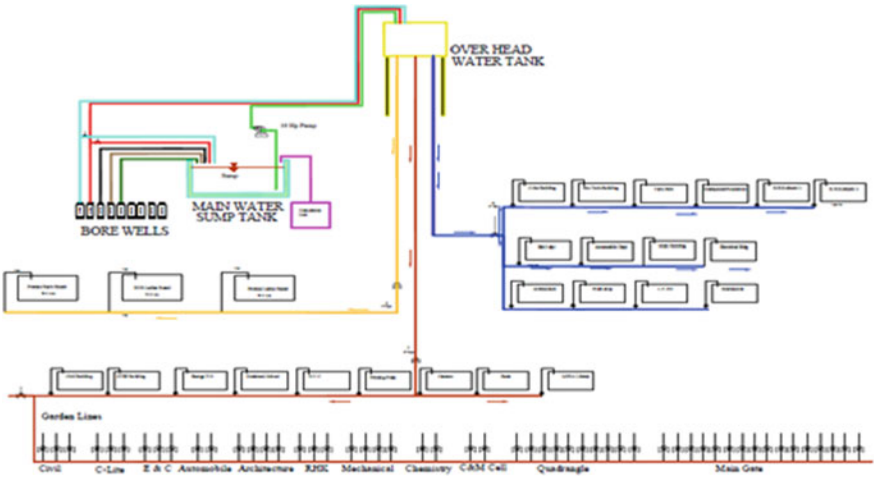


Fig. 1 Campus water distribution and management system

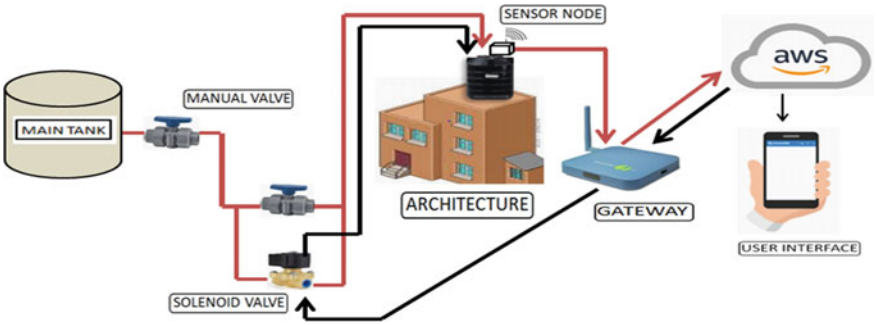


Fig. 2 Smart water management architecture

water pipeline which connects six departments. The practical challenges for such system are long-range communication. The campus is spread across 2 km with many obstacles such as tall buildings and trees. The network communication is to cover the entire campus. Battery-operated sensors and real-time monitoring are installed.

4 Proposed SWM Network

The proposed SWM network has two nodes:

Sensor nodes: The sensor nodes which act as end nodes are placed on the OHT of cluster of each department. These nodes wake up every 10 min, collect data from sensors and transmit to the gateway. Due to this periodic wakeup, low power is achieved. The water levels in the tanks are measured using distance measurement. This method is preferred as it is non-invasive. An industry grade waterproof ultrasonic sensor JSN-SRO4T measures the distance of water from tank top. A Lora™ RFM98W transceiver module is used to transmit the date to the gateway at 433 MHz frequency at the data rate of 300 kbps. This transceiver provides high interference immunity and consumes less current. The sensor unit is power up only during data acquisition. This transceiver solves the traditional compromise between range, immunity and energy consumption.

Gateway: The gateway uses a low-power ATmega328P microcontroller, and A Lora™ transceiver module is used to receive the date from the sensor nodes at 433 MHz frequency at the data rate of 300 kbps. It provides a communication range of 5 km (line of sight) with +14 dBm high efficiency PA.

Control Flow Design

To eliminate the ambiguity of water distribution when request is placed by more then one building, a trigger-based control system is designed as shown in Fig. 3. Control flow for water distribution works based on the below algorithm.

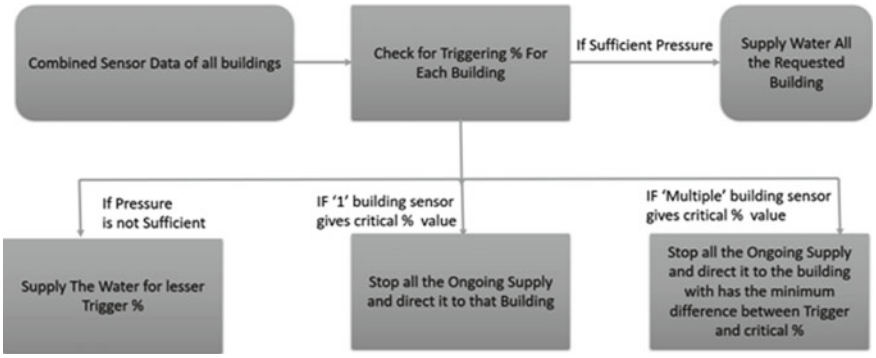


Fig. 3 Control flow for water distribution

Fig. 4 Gateway and sensor node deployed on campus

- Firstly, Check for trigger % for each building.
- Check for the availability of **Pressure** and serve all the requested building if sufficient.
- If less pressure serves for the building having lesser **Trigger %**.
- If building values are lesser than or equal to **Critical %**, water supply is diverted to the required building.
- If multiple building values are lesser than **critical %**, the building which has lesser difference between **Trigger %** and **Critical %** will be served.
- Scheduled filling of all the building in done sequential way after the college timings.
- Trigger % to be calculated based on the strength of the department.

System Architecture

The sensor node transmits the data to the gateway wirelessly as shown in Fig. 4. The network works in star topology. As the LoRa transceiver module is used to transmit the data which covers 2 km distance with obstacles, repeaters are not used in the architecture. AWS is used as cloud. The control flow is designed at the backend. Based on the water requirement, the triggering percentage the valves are controlled.

5 Experimental Results

The objective of the project is to develop CWMS, with university campus as test bed. Figure 4 shows map of the university campus where the gateway and sensors

are deployed. As an initial deployment, we have deployed across four departments with gateway placed in the center of the campus.

Each cluster has a sensor node deployed. It sends the status of water as distance measurement. This distance is later converted to litres at the backend. The installation for testing is as shown in Fig. 5. The proposed CWMS is compared with RANA kit [8] as shown in Table 1. The CWMA is a real-time monitoring device and with power optimization at the sensor node. The proposed device works with 433 kHz and is non-invasive. The usage of water is indicated for every 10 min. As shown in Table 1, the current consumption wakeup time is 19 mA, sleep time is 35 μA, and the battery life is approximately 1.5 years. The status of the battery is also indicated by the sensor node to know the battery health.

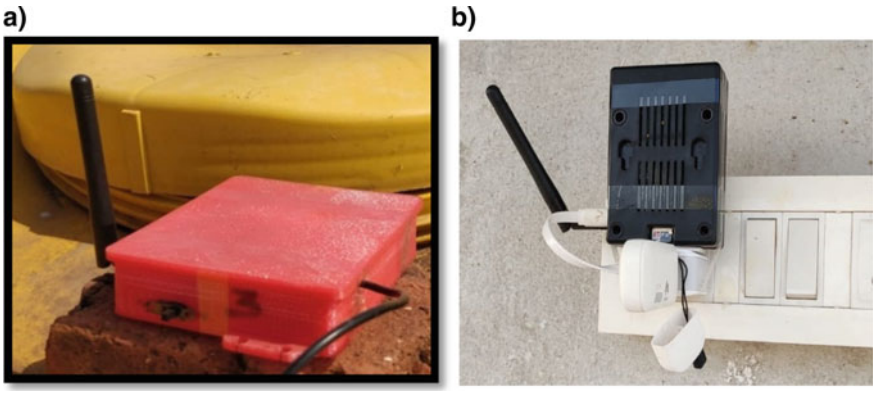


Fig. 5 Images of **a** sensor node and **b** gateway deployment

Table 1 Comparison of proposed CWMS with RANA kit [8]

Parameters	Proposed CWMS	RANA kit [8]
Battery life	Approximately 1.5 years	<then 72 h
Battery health	Indicates	No indication
Usage	Real-time water usage monitoring	Water quality monitoring for reservoirs of disaster hit areas
Communication	LoRa-433 kHz	900 MHz
Deployment	Non-invasive	Submersible
Duration	Long lasting	Deployment for a few minutes to maximum of three days
Monitored parameters	Water level	pH, turbidity, dissolved oxygen, temperature, total dissolved salts
Power mode	Deep sleep mode	–
Wakeup current	16 mA	–
Deep sleep current	4 μA	–

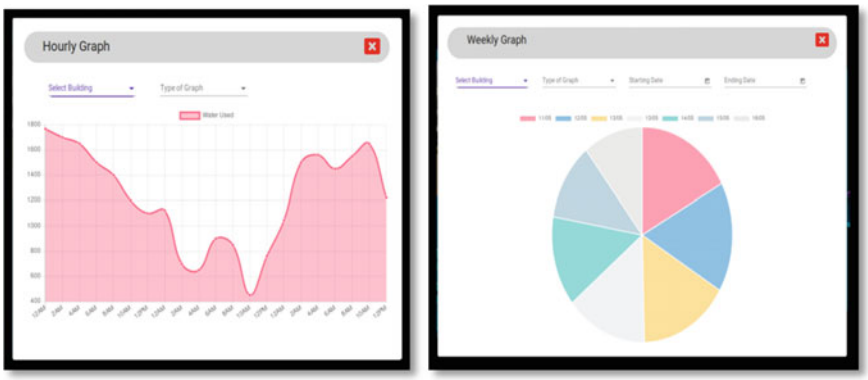


Fig. 6 Water level data for a week for sensor node placed in EC dept on cluster 1

The proposed device works with 433 kHz and is non-invasive. The usage of water is indicated for every 10 min. As shown in Table 1, the current consumption wakeup time is 19 mA, sleep time is 35 μA, and the battery life is approximately 1.5 years. The status of the battery is also indicated by the sensor node to know the battery health.

Data Visualization and Alert System
The data visualization is through a web page developed using angular ionic platform. The page has various displays like:

- Graph to indicate water level hourly, daily, weekly or monthly basis.
- Health of sensors and battery.

Figure 6 shows a sample graph showing water usage for a EC department hourly and weekly on cluster 1. The graphs show the water usage on non-working days is lesser.

The user dashboard is built to visualize various parameters as shown in Fig. 7. Water is pumped to the tanks every evening after college hours on regular basis. The graphs are also plotted to indicate water usage hourly basis for a day to understand the usage pattern. These graphs are used to study the usage pattern to make the system intelligent. This pattern will be studied for a year so the prediction algorithm will be built to make the CWMS intelligent. This will also be used to detect the water leakage.

6 Conclusion

The CWMS is deployed in KLE campus to monitor the water usage and distribution. The authors have designed a LoRa-based low-power sensor node and a gateway to

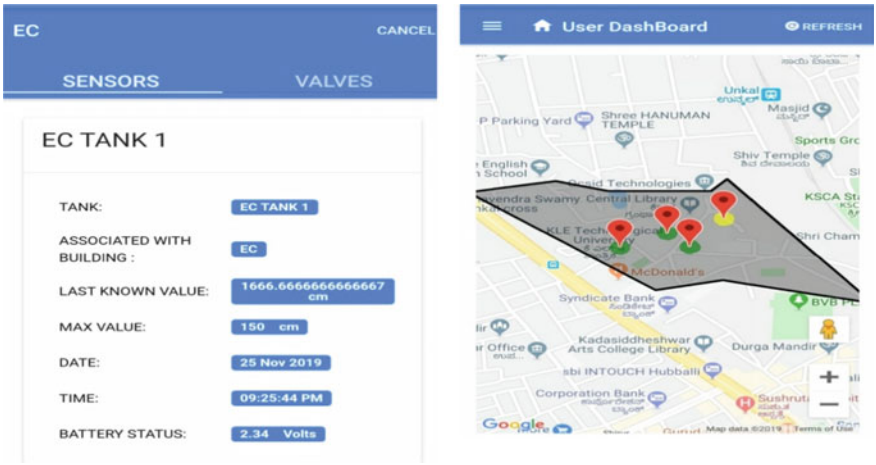


Fig. 7 Dashboards indicating tank level, battery status and time stamp

monitor the water distribution across the campus. The sensor unit is a low-power design consuming 35 μA in sleep mode and 19 mA in wakeup mode. The data is sent to the gateway every 10 min. The solution presents hourly, daily, weekly, and monthly usage of water. The analysis of usage pattern will be further used to develop prediction algorithms and leakage detection.

Acknowledgements The authors would like to thank VGST for funding this project under K-FIST Level 2.

References

1. Kudva, V., et al.: Towards a real-time campus-scale water balance monitoring system. In: 2015 28th International Conference on VLSI Design and 2015 14th International Conference on Embedded Systems
2. Karwot, J., et al.: Smart water in smart city. In: A case Study Conference: SGEM 16th International Scientific Conference on Earth & Geosciences, vol. 3/I, At Albena, Bulgaria, June 2016
3. Stoianov, I., Nachman, L., Madden, S.: Pipenet: A Wireless Sensor Network for Pipeline Monitoring, IPSN'07, 25–27 Apr 2007, Cambridge, Massachusetts, USA
4. Water in India: Situation and Prospects, UNICEF (2013)
5. Varma, P., et al.: Towards IoT based water management system. Performance Analysis of Wireless Devices for a Campus-Wide IoT Network. The 2015 International Workshop on Wireless Network Measurements and Experimentation, IEEE, pp. 84–89. 978-3-9018-8274-6/15/©2015
6. Gosavi, G., Gawde, G.: Smart water flow monitoring and forecasting system. In: 2017 2nd IEEE International Conference on Recent Trends in Electronics Information & Communication Technology (RTEICT), 19–20 May 2017, India, IEEE, pp. 1218–1222. 978-1-5090-3704-9/17/$31.00©2017

7. Whittle, A.J., Girod, L., Preis, A., Allen, M., Lim, H.B., Iqbal, M., Srirangarajan, S., Fu, C., Wong, K.J., Goldsmith, D.: WATERWISE@SG: a test bed for continuous monitoring of the water distribution system in Singapore. In: Water Distribution System Analysis (2010)
8. Angello, J.E., Corrigan, A.M., Garg, R.K., Hewitt, S.S., Hudgins, K.L., Lester, E.C., Sorensen, C.A.E., Wilson, M.R., Brinkman, B.M., Louis, G.E.: A rapid adaptive needs assessment kit for water quality monitoring in humanitarian assistance & disaster response applications. In: Systems and Information Engineering Design, Symposium (2012)
9. Kudva, V.D., et al.: Towards a real time campus scale water balance monitoring system. In: 28th International Conference on VLSI design and 14th International Conference on Embedded System, pp. 88–91. https://doi.org/10.1109/VLSIID

Rule-Based Recognition of Associated Entities in Hindi Text: A Domain Centric Approach

Arpana Prasad and Neeraj Sharma

Abstract There are several computing applications that process unstructured texts available in various natural languages on the Web. These applications provide worthwhile outcomes that are of interest to the end users, other application developers and researchers. Various approaches of Natural Language Processing (NLP) and related fields are used to develop these applications. Application's domain related lexical resources, in same language as the language of the text being processed, are considered useful and effective in improving performance of the related application. Several such resources are available in English, but there are certain languages that have scope for development and research. There is a need to develop the resources in those languages and bring them at par with their English counterparts. One such identified language is Hindi. Lexical resources in Hindi are developed in an ongoing research in the field of Opinion Mining. These are used in a rule based named entity recognition system developed as part of the research. The unstructured texts from Hindi Weblogs in the domain of home remedies are used in this research. The methodology used to develop these resources, outcomes achieved, experimental setup used for evaluation of the developed resources and the results obtained thereof, are discussed in this paper.

Keywords Food-health issue association recognition · Opinion mining · Rule based named entity recognition

1 Introduction

Currently, for knowledge generation and consumption a considerable section of society depends upon the digital media on the Web. In the current Corona virus

A. Prasad (✉) · N. Sharma
Department of Computer Science, Punjabi University, Patiala, Punjab, India
e-mail: arpanaprasad2013@gmail.com

N. Sharma
e-mail: sharma_neeraj@pbi.ac.in

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_36

(COVID-19) pandemic times, digital media and the Internet has almost become a lifeline for education, business, entertainment, medicine and almost every field one thinks of. The addition of huge digital contents in structured and unstructured forms, provide an arena of opportunities to researchers working in subject of Computer Science (CS). Methods from Natural Language Processing, a field of study in CS, are used for extraction and summarization of meaningful information from any unstructured text. Information Extraction (IE) and Opinion Mining (OM) are fields dedicated towards subtasks of information retrieval, classification and named entity recognition (NER) from any digital text. OM applications computationally derive author's sentiment towards an entity, aspect or an association. The sentiments are derived and classified into pre-determined polarity classes. Common polarity classes used by the application developers are; positive (1), neutral (0) and negative (−1). But there are no restrictions on the number and types of classes that an application may choose. NER is a task that is a part of IE, it is dedicated towards extracting word(s) and phrase(s) from text and classifying them into pre-determined categories like name(s), time-zone, location, food, health issue, plant, medicine, organ etc.

Derivatives from an ongoing research in the field of OM are discussed in this paper [1]. In this research an effort is made to derive author's opinion from Hindi text, in the domain of home remedies, available on Weblogs. Only Hindi sentences from Weblogs that demonstrate a relationship, as to how food/food combinations (F) affects health issue/body components (HI_BC) are chosen for the purpose of deriving opinions. For ease of reference the sentences that are chosen from Weblogs are referenced in the remaining part of this paper as 'Domain Specific Hindi Sentences from Weblogs (DSHSW)'. Samples of the sentences chosen for the purpose are presented in Table 1. In the remaining part of the paper wherever Hindi text is given, the translation of same is given within '{}'. There are two main components of the OM system being developed; (a) a computing mechanism to extract associated entities 'F' and 'HI_BC' (b) polarity classification of relationship of associated entities. The

Table 1 Sample sentences of 'DSHSW'

Examples of DSHSW
'हरे बैंगन में फाइबर पाया जाता है जिसको खाने से उदर भर जाता है और बहुत जल्दी भूख नहीं लगती , तो इस तरह वजन भी कम होता है ।' { 'Green eggplant contains fiber, which fills the stomach due to eating and does not cause hunger soon, so in such a way it also reduces weight.'}
'शोधकर्ताओं का कहना है कि चावलों को मांस या सोयाबीन के साथ लेने से खून में चीनी की मात्रा पर असर पड़ सकता है ।'{'Researchers say eating rice with meat or soybeans can affect blood sugar levels.'}
'दही के लगातार प्रयोग में लाने से आंतों के रोग और पेट की बीमारियां नहीं होती, और कई प्रकार के विटामिन भी बनने लगते हैं ।'{'Regular intake of yogurt does not cause intestinal diseases and stomach diseases, and causes a variety of vitamins.'}

work is novel for Hindi and outcomes are expected to be useful for summarization, handling home remedies related queries and subjecting the health beliefs to scientific verification. This paper presents the component 'a' of the OM system. Rest of the paper is organized as follows. An overview of related work in the area of the current research is given in Sect. 2. Work related to the development of rule-based recognition of related named entities is presented in Sect. 3. Experimental setup for evaluation of the developed resources and results obtained thereof is given in Sect. 4. Conclusions drawn are presented in Sect. 5.

2 Related Work

Work related to NER for an opinion mining system for Hindi text is presented in this paper; this section gives an overview of related work by other researchers. Approaches used for OM, domains explored and outcomes of different research are analyzed in survey [1–4]. Work related to OM for Indian languages are gaining momentum and some lexical resources are built for the same [5, 6]. Machine learning approach for NER is used in several researches. Term frequency and Inverted Document frequency approach is used for NER in a research [7]. Rule based approaches to Named Entity Recognition is used in several related studies [8–11]. Lexical knowledge is used for NER system for bio-medical text in a research [12]. Rule based system related to food identification from English text is presented in a research [13]. Analysis of various tools across several domain is presented in a study [14]. Corpus of news articles is used for recognizing temporal expression in Hindi Text. Approaches from machine learning and hand crafted twenty-five rules are also identified. Rule based system outperforms the machine learning approach when taken in isolation i.e. when a hybrid approach is not taken into account [15]. Rule based approach gives higher level of accuracy, compared to machine learning approach for NER task performed on health records in English [16].

Procedure adopted for dataset collection, data cleaning, semantic annotations of the datasets, syntactic annotations of the data sets for the current research are presented in a work contributed by the authors of this paper [1]. An overview of methodologies adopted to develop the lexical resources and rules formulated for identification of food disease are briefly presented in the work. The work also presents the polarity classification algorithm developed for the current research and evaluations of the same. In this paper an attempt is made to further elaborate the work done in the previous study [1]. Improvement in syntactical annotation of the datasets, algorithm developed and procedures involved in developing the lexical resources, and results thereof, which are not part of the existing study are elaborated in this paper.

3 Rule-Based Recognition of Related Named Entities

The sub-sections of the this section present; the objective of development of a system to identify related 'F' and 'HI_BC' entities from DSHSW, methodologies adopted towards development of lexical resources that support development of the required related NER system, rules developed for rule-based system under the current study. An effort is made to present here the improvements and extensions made to our previous study [1].

3.1 Objective of the Research

The core objective of development of the OM system under study, as mentioned in our previous study [1], is restated with more details as follows: Given finite sets 'F', 'A', 'FC', 'H', 'B' such that; $F = \{f1, \ldots, fn\}$, where 'fi' is a word or a phrase that is a name of a food item in Hindi, where $1 \leq i \leq n$ and n is a finite number greater than 0. An example of 'fi' is 'तुलसी' {Basil}. $A = \{a1, \ldots, am\}$, where 'ai' is a word or a phrase in Hindi which is used to describe a food item, where $1 \leq i \leq m$ and m is a finite number greater than 0. An example of 'ai' is 'काली' {black} which describes a type of food item 'तुलसी' {basil} commonly known as 'काली तुलसी' {black basil}. FC = {fc1…fcp}, where 'fci' is a word or phrase in Hindi which specifies the food component of the food item, where $1 \leq i \leq p$ and p is a finite number greater than 0. An example of 'fci' is 'पत्ते' {leaves} which describes a component of food item 'तुलसी' {basil} commonly referred to as 'तुलसी के पत्ते' {basil leaves}. $H = \{h1 \ldots hq\}$, where 'hi' is a word or a phrase that is a Hindi name of a health issue that a human being is subjected to, where $1 \leq i \leq q$ and q is a finite number greater than 0. An example of 'hi' is 'सिरदर्द' {headache}. $B = \{b1 \ldots br\}$, where 'bi' is a word or a phrase that is a Hindi name of a body component that is part of human body, where $1 \leq i \leq r$ and r is a finite number greater than 0. An example of 'bi' is 'सिर' {head}. Through this research we develop an algorithm 'A' such that the input to 'A' is a sentence from DSHSW. The algorithm determines; a set 'FOOD' from the input where FOOD = {set of single or multiple element from 'F' or combinations of elements from 'F'}, a set 'HI_BC' from the input where HI_BC = {{set of single or multiple element from set 'H'} UNION {set of single or multiple element from set 'BC'}}, and a polarity classification as an opinion number −1 or 1 i.e. whether the elements of set 'FOOD' have negative or positive impact on elements of 'HI_BC'. In this paper we limit the discussion to only the methodologies adopted, lexical resources developed and rules formulated for determining the sets 'FOOD' and 'HI_BC' for a sentence from DSHSW. Only extraction of related entities is the focus area of this study, and not the polarity classification of the association. An example to illustrate the input text and related expected output is as follows;

Input: चिकनाई वाली रोटी या खरबूजा, खीरा खाने के एकदम बाद पानी पीने से खांसी, जुकाम होने की संभावना ज्यादा है । {Drinking water with oily bread or musk melon or cucumber will lead to cold and cough.}

Output: The output is in the form of sets; FOOD = {'रोटी' {bread} Adjective: (चिकनाई) {oily} + 'पानी' {water}, 'खरबूजा' {musk melon} + 'पानी' {water}, 'खीरा' {cucumber} + 'पानी' {water}}. HI_BC = {'जुकाम' {cold}, 'खांसी' {cough}}.

Derived polarity classification: −1 (negative); means that the elements of FOOD are not good for consumption in combinations and it leads to the health issues mentioned in HI_BC. Please note that there is no 'body component' in the input text mentioned, hence they are not derived.

3.2 Lexical Resources Developed

As mentioned in Sect. 3.1 of this paper, the lexical resources identified as important inputs to the research are sets '*F*', '*A*', 'FC', '*H*', and '*B*'. Researchers of this research could not find an exhaustive list of Hindi lexicons for these sets. Lexical resources from food disease domain in English are available on the website of 'USDA National Nutrient Database for Standard Reference'[17], but translations are restricted. It does not cover the Hindi words and phrases as mentioned for different sets mentioned in Sect. 3.1 on Hindi Weblogs. Hence, the required lexical resources are developed as part of this research. The lexicons are also mapped to their respective synonym sets. Synonym sets (synsets) shortlisted and collected for each set are; 'F_Synset' for 'F', 'A_Synset' for 'A', 'FC_Synset' for 'FC', 'H_Synset' for 'H', and 'B_Synset' for 'B'. The lexicons are developed using semi-automated process. First the lexical sets are generated using a standard lexical resource in Hindi i.e. Hindi WordNet (HWN) [18]. This lexical resource has approximately 39,000 plus synonym sets along with their syntactic and semantic annotations. The generated lexical sets are put under manual verification to identify the relevant synonym set. As for several words in Hindi there are several synonym sets. For example the word ' मूली' (Transliteration: 'Mooli') synsets for food, plant, river and reptile lizard is generated. Hence manual word sense disambiguation is important. The procedure adopted to develop the synonym sets are given in the following steps; Step1: Manually generate a seed list of words or phrases for each set as mentioned above. The semantically annotated dataset developed in current research is used for the same, it is illustrated in our previous study [1]. Step 2: Python code is implemented and executed to search the HWN and generate all synonym sets for the seed list with semantic information like WordNet Id, Synonym Id, Synonym Set, Word Usage, and Word Meaning (all these are available in HWN). Step 3: The synonym sets thus generated are put to manual selection process using Word Usage and Word Meaning information. Step 4: From the Word Meaning and Word Usage stop words are removed. Step 5: Term frequency is determined for all remaining words in Word Meaning and Word Usage. Step 6: Top three words that

have maximum number of occurrences in Word Meaning (WM) and Word Usage (WU) are stored in sets 'WM' and 'WU' respectively. Example of one such word is 'रोग' in WM for the set 'H'. Step 7: Python code is implemented and executed to generate all synonym sets that have the words as in 'WM' and 'WU' in Word Meaning and Word Usage minus the synonym sets that are previously generated and verified for a particular set. Step 8: The newly generated synonym sets are put under manual verification.

Using the above mentioned process the number of words and the phrases generated for each set are approximately: F: 3128, A: 194, FC:92, H: 2800, B:1304, F_Synset: 303, A_Synset:38, FC_Synset:25, H_Synset:382, B_Synset:136: Total lexicons: 8400 approx.

3.3 Rules Identified for Related Entities Recognition

The rules identified for formation of conditional clauses, for the 'Food-Health Issue Association Identification' are given in subheading 'Rules' in this sub-section, they are extension and improvement to those mentioned in our previous work [1]. The base algorithm, on which the rules identified are used, is given under subheading Algorithm. Given any text, 'Input_Text' that belongs to DSHSW, this algorithm identifies the elements of the sets 'FOOD' and 'HI_BC', as mentioned in Sect. 3.1.

Detail of sub modules. The sub modules used in Algorithm are as follows. 'TextCleaning()' cleans the text of garbage values as mentioned in previous work. 'StopWordRemoval()' removes the stop words (commonly occurring words in a sentence e.g. 'a', 'an') from any text. 'POStagging()' performs syntactic tagging of each word of the sentence for its part of speech (POS). 'PerformStemming()' determines stem (root word e.g. root word for 'going' is 'go') of each word for a given text.

Algorithm Determine_FoodHI_BC(Input_Text)

{
Step1: Populate list variable Food_list from F, Adj_list from A, Component_list from FC, HealthIssue_list from H, BodyComp_list from BC.
Step 2: Use 'TextCleaning()' to clean all elements of each list obtained in Step 1.
Step 3: Use StopWordRemoval() to remove stop words from all elements of the lists obtained from Step 2.
Step 4: Use PerformStemming() to stem all elements of the lists obtained from Step3.
Step 5: CleanText = TextCleaning(Input_Text).
Step 6: POSTaggedText = POStagging(CleanText).
Step 7: StopWordRemovedPOSText = StopWordRemoval(POSTaggedText).
Step 8: StopWordRemovedText = StopWordRemoval(CleanText).

Step 9: StemText = PerformStemming(StopWordRemovedText).
Step 10: StemPOSText = PerformStemming(POSTaggedText).
Step 11: Apply all the rules mentioned under the heading 'Rules' in the current sub-section to determine 'FOOD' and 'HI_BC' sets using the lists generated after Step 4.
}.

Rules. *Rule 1*. To search for any value 'x' in any resource list 'Y'; in case 'x' has any symbol {':', '_', '-', '(',')'}, conduct search with and without these symbols and valid combination of these symbols.

Rule 2. Let 'x' be a unigram word from preprocessed text 'StopWordRemovedText'; element 'x' is added to 'FOOD', if (PerformStemming(x) ∈ 'F' or 'x' ∈ 'F', and x does not belong to 'FOOD'). Do so adhering Rule 1.

Rule 3. Let '$x1\_x2$' be a bigram word formed from preprocessed text 'StopWordRemovedText'; '$x1\_x2$' is added to 'FOOD', if (PerformStemming('$x1$') + '_' + PerformStemming($x2$)) ∈ 'F' or '$x1\_x2$' ∈ 'F' and '$x1\_x2$' does not belong to set 'FOOD'. If component unigrams of the bigram word belong to 'FOOD' they are removed from 'FOOD'. Rule 1 to be adhered also.

Rule 4. Let '$x1\_x2\_x3$' be a trigram word formed from preprocessed text 'StopWordRemovedText'; '$x1\_x2\_x3$' is added to 'FOOD', if (PerformStemming('$x1$') + '_' + PerformStemming($x2$) + '_' + PerformStemming($x3$)) ∈ 'F' or '$x1\_x2\_x3$' ∈ 'F' and '$x1\_x2\_x3$' does not belong to the set 'FOOD'. If component unigram and bigram words formed from '$x1\_x2\_x3$' belong to 'FOOD' then remove them. Rule 1 to be adhered in following this rule.

Rule 5. Let 'x' be a unigram word from preprocessed text 'StopWordRemovedText'; element 'x' is added to 'HI_BC' if PerformStemming(x) ∈ 'H' or 'x' ∈ 'H' or PerformStemming(x) ∈ 'B' or 'x' ∈ 'B', and x does not belong to 'HI_BC'. Do so adhering Rule 1.

Rule 6. Let '$x1\_x2$' be a bigram word formed from preprocessed text 'StopWordRemovedText'; '$x1\_x2$' is added to 'HI_BC', if (PerformStemming('$x1$') + '_' + PerformStemming($x2$)) ∈ 'H' or '$x1\_x2$' ∈ 'H' or (PerformStemming('$x1$') + '_' + PerformStemming ($x2$)) ∈ 'B' or '$x1\_x2$' ∈ 'B' and '$x1\_x2$' does not belong to set 'HI_BC'. In case component unigrams of the bigram word belong to 'HI_BC' then remove them. Rule 1 to be adhered.

Rule 7. Let '$x1\_x2\_x3$' be a trigram word formed from preprocessed text 'StopWordRemovedText'; '$x1\_x2\_x3$' is added to 'HI_BC', if (PerformStemming('$x1$') + '_' + PerformStemming($x2$) + '_' + PerformStemming($x3$)) ∈ 'H' or '$x1\_x2\_x3$' ∈ 'H' or (PerformStemming('$x1$') + '_' + PerformStemming($x2$) + '_' + PerformStemming($x3$)) ∈ 'B' or '$x1\_x2\_x3$' ∈ 'B' and '$x1\_x2\_x3$' does not belong to the set 'HI_BC'. In case components of trigram i.e. the unigrams or the bigrams formed from these trigrams are already in 'HI_BC' then remove them. Rule 1 to be strictly adhered in this rule.

Rule 8. All unigram words that are element of 'FOOD', 'HI_BC' are verified for POS from corresponding text in POSTaggedText. Only those words are retained that are having the part of speech as 'NN', 'NNP', 'JJ', 'VM' and 'VAUX'.

Rule 9. This rule is used to identify adjectives of the food Items. Take each element of set 'FOOD', search for the index at which that word occur in 'StopWordRemovedPOSText'. An example of 'StopWordRemovedPOSText' is 'गरम/JJ दूध/NN और/CC शहद/NN एक/QC साथ/NST लेने/VM मिक्रोबियल/NNP इन्फ़ैकशन/NNP इंफ्लेमेट्री/NN रोग/NN और/CC कफ/NN आराम/NN मिलता/VM'. For unigram words just find the index at which that word occurs in 'StopWordRemovedPOSText'. For bigram and trigram words of 'FOOD' find the index of first component word in 'StopWordRemovedPOSText'. Store the indexes obtained for each element in a temporary vector 'VECTOR'. To find the adjective of an element '*x*' of 'FOOD' the following process is to be followed. Find index of '*x*' from 'VECTOR' and store it in variable '*i*'. Find a number '*m*' that it is the greatest index stored in 'VECTOR' that is less than '*i*'. A word 'y' with tag 'JJ' or tag 'VM' and at index '*j*' in 'StopWordRemovedPOSText' is an adjective for '*x*' if $j < i$ and $j > m$ and '*y*' belongs to the set '*A*'. For bigram and trigram words in 'FOOD' extract the component word '*w*' from the element and if '*w*' belongs to the set '*A*' then '*w*' is the adjective which is part of the food name and may be ignored for any further processing.

Rule 10. This rule is used to identify components of the food items in the set 'FOOD'. For each element in 'FOOD' compute a vector 'VECTOR' that stores the index value of each element of 'FOOD'. For any unigram words in 'FOOD' find the index at which the word occurs in the text 'StopWordRemovedText' and store it in 'VECTOR'. For bigram and trigram words in the text 'StopWordRemovedText' find the index at which the first component word occurs in the text 'StopWordRemovedText'. Let '*x*' be a unigram word in set 'FOOD' and its index value in 'VECTOR' be '*i*'. Suppose '*s*' is the lowest index value in 'VECTOR' that is greater than '*i*'. A word '*y*' occurring at index value '*r*' in 'StopWordRemovedText' will be component for food item '*x*' if *y* belongs to set 'FC' and $i < y < s$.

Rule 11. Consider an example; ' शहद और नींबू के रस को एक साथ चाटने से खांसी में इममुनिटी बढ़ती है।' {licking honey and lemon juice together increases immunity in cough.} In this sentence it is expected that the algorithm derive elements of FOOD as {'शहद' {'honey'} + "नींबू' {lemon} ('रस' {juice})} and HI_BC as {'खांसी' {cough}}. The ' + ' symbol appearing in between the food elements signify that both the food items must be consumed in combination then it has positive effect on the health issue mentioned in HI_BC. The rule specified for identifying food combination is that, if two food items are to be derived to be taken in combination then phrases in as mentioned in a lexical resource 'Combination_specifier', must be present in the sentence. The distance of the phrase from food item(s) must be within the maximum distance window size mentioned in the 'Combination_specifier'. It may be noted that 'Combination_specifier' is a lexical resource in the form of record set developed as part of this research. This is developed using the semantically annotated dataset. Any record '*t*' of the 'Combination_specifier' is of the form *t* = {field1, field2}. Where 'field1' is a phrase that gives semantic orientation to the sentence that the food elements are to be consumed in combination. The 'field2' specifies the farthest distance of the phrase from the food items. If the phrase 'field1'

is suppose to be in between the food items then value of 'field2' is 0. If the phrase in 'field1' is before the two food items then the value of 'field2' is $-d$. If 'field1' must have right occurrence to the food item then the value of; 'field 2' is $+d$, where 'd' is the distance (i.e. no of words between a food item name and 'field1'). The rule is if in a sentence any phrase that is value of 'field1' from a record of 'Combination_Specifier' is present and the distance is less than the maximum distance of the phrase from food item as mentioned in 'field2' then it is derived that the food items are to be taken in combination. The records in 'Combination_specifier' for a sentence as mentioned in the example are: {'और', 0}, {'एक साथ', +2}, and {'साथ', +3}. Since the three phrases are present in the sentence and are within the maximum distance, so it is derived that the food mentioned in 'FOOD' are to be taken in combination.

Rule 12. For a complex sentence that illustrates two contrasting sentiments, break the sentence into two sentences such that, each sentence demonstrates one and only one sentiment. Determine separate sets 'FOOD' and 'HI_BC' for each newly formed sentences. If food items or health issues or organs do not occur in both newly formed sentences, then 'FOOD' and 'HI_BC' determined in one sentence is used for the other sentence too.

4 Experiment and Result

Algorithm, as mentioned in Sect. 3.3 is implemented using programming language, Python 3.5.3. The lexical resources as mentioned earlier are stored as record sets in a data base in SQLite (a relational database management system). The POS tagger used for this research, is developed by CFILT [18]. For a dataset of approximately 3500 sentences, consisting of approximately 30,000 plus word tokens, and approximately 4000 plus unique words, it identifies 22 valid unique POS tags. The different rules as mentioned in Sect. 3.3, are implemented using 'if and else' clause in Python. The experiment conducted extracted information from an input text belonging to 'DSHSW', and derived the sets 'FOOD' and 'HI_BC' defined in Sect. 3.1. The code was executed on datasets with 900 Hindi sentences from Hindi Weblogs in the domain of home remedies. The results were evaluated with respect to the manual semantic annotation done by two annotators for those 900 sentences. The results as obtained are presented in Table 2.

Table 2 Results obtained from the experiments

Parameter evaluated (Derived elements of sets)	True identification	False identification	Accuracy in percentage (%)
FOOD	810	90	90
HI_BC	855	45	95

5 Conclusion

The lexical resources developed as part of this research are semi automatically generated. Manual annotations and verification is a tedious task that is undertaken in this research. The part of speech tagger used in the research as mentioned in Sect. 4 is selected based on experimental comparison amongst three taggers. The two taggers that are not selected are; Hindi POS tagger available on a website [19] and POS tagger available for Hindi as part of Python NLTK package [20]. The first amongst them gave results close to the selected tagger, but approximately 75 word tokens from the datasets are marked with '?' meaning unknown tags. Hence it is dropped. The second tagger by NLTK, that is considered a good tagger for English text, did not give satisfactory results for the Hindi dataset used in this research. It tagged more than 1000 word tokens with 'UNK' tag i.e. unknown. The selected tagger which is used in this research did not tag even a single word token as unknown. The rules for extraction of associated entities are handpicked and are a result of tedious semi automated adopted procedures. The rules are formulated using datasets as mentioned in Sects. 3 and 4. As per studies rule based systems gives a competent rate of accuracy [11, 13]. Research is in progress to incorporate more rules to cover a larger dataset. Other approaches to NER are also under study for a comparative analysis of the outcomes obtained in this research.

This research is only dedicated towards determining author's opinion about polarity of the association between food and health issues for domain centric text from Hindi Weblogs. This study highlights the approach used in this research for recognition of associated entities, food and health issues. This research does not validate the derived associations.

References

1. Prasad, A., Sharma, N.: Lexicon based extraction and opinion classification of associations in text from Hindi weblogs. Int. J. Adv. Res. Eng. Technol. **11**(8), 439–451 (2020)
2. Mowlaei, M.E., Abadeh, M.S., Keshavarz, H.: Aspect-based sentiment analysis using adaptive aspect-based lexicons. Expert Syst. Appl. **148**(1), 11324 (2020)
3. Liu, B.: Many facets of sentiment analysis. In: Cambria, E., Das, D., Bandyopadhyay, S., Feraco, A. (eds.) A Practical Guide to Sentiment Analysis. Socio-Affective Computing, vol. 5. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-55394-8_2
4. Ravi, K., Ravi, V.: A survey on opinion mining and sentiment analysis: tasks, approaches and applications. Knowl.-Based Syst. **89**, 14–46 (2015)
5. Patra, B.G., Das, D., Das, A.: Sentiment analysis of code-mixed indian languages: an overview of SAIL code-mixed shared task. In: Banyopadhyay, S. (ed.) Proceedings of ICON-2017: 14th International Conference on Natural Language Processing. NLP Association of India (2017). https://www.aclweb.org/anthology/W/W17/W17-75
6. Joshi, A., Balamurali, A., Bhattacharyya, P.: A fall-back strategy for sentiment analysis in Hindi: a case study. In: Proceedings of the ICON 2010:8th International Conference on Natural Language Processing. NLP Association of India (2010). https://www.cse.iitb.ac.in/~adityaj/HindiSentiWordnet_AdityaJ.pdf. Accessed 2020/09/25

7. Upendraa, B., Babu, A.S.: KNN TFIDF based named entity recognition. Int. J. Sci. Dev. Res. **1**(12), 35–39 (2016)
8. Nguyen, L.H.B., Dinh, D., Tran, P.: An approach to construct a named entity annotated English-Vietnamese bilingual corpus. ACM Trans. Asian Low-Resour. Lang. Inf. Process. **16**(2) Article 9, 1–17 (2016)
9. Thao, P.T.X., Tri, T.Q., Dien, D., Collier, N.: Named entity recognition in Vietnamese using classifier voting. ACM Trans. Asian Low-Resour. Lang. Inf. Process. **6**(4) Article 14, 1–18 (2007)
10. Eftimov, T., Koroušić Seljak, B., Korošec, P.: A rule-based named-entity recognition method for knowledge extraction of evidence- based dietary recommendations. PLoS ONE **12**(6), 1–32 (2017). https://doi.org/10.1371/journal.pone.0179488
11. Zaghouani, W.: RENAR: a rule-based Arabic named entity recognition system. ACM Trans. Asian Low-Resour. Lang. Inf. Process. **11**(1) Article 2, 1–13 (2012)
12. Park, K.M., Kim, S.H., Rim, H.C., Hwang, Y.S.: ME-based biomedical named entity recognition using lexical knowledge. ACM Trans. Asian Low-Resour. Lang. Inf. Process. **5**(1), 4–21 (2006)
13. Popovski, G., Kochev, S., Seljak, B.K., Eftimov, T.: Foodie: a rule-based named-entity recognition method for food information extraction. In: Marsico, M.D., Baja, G.S.D, Fred, A. (eds.) ICPRAM 2019—Proceedings of 8th International Conference on Pattern Recognition Application Methods, pp. 915–922. Springer, Cham (2019)
14. Abdallah, Z.S., Carman, M., Haffari, G.: Multi-domain evaluation framework for named entity recognition tools. Comput. Speech Lang. **43**, 34–55 (2017)
15. Ramrakhiyani, N., Majumder, P.: Approaches to temporal expression recognition in Hindi. ACM Trans. Asian Low-Resour. Lang. Inf. Process. **14**(1) Article 2, 1–22 (2015)
16. Gorinski, P., Wu, H., et al.: Named entity recognition for electronic health records: a comparison of rule-based and machine learning approaches (2019). Available on https://www.researchgate.net/publication/331670597_Named_Entity_Recognition_for_Electronic_Health_Records_A_Comparison_of_Rule-based_and_Machine_Learning_Approaches. Accessed on 2020/9/25
17. USDA FoodData Central Homepage: https://ndb.nal.usda.gov/ndb/. Last accessed 2020/09/24
18. Computing for Indian language Technology (CFILT) Homepage: https://www.cfilt.iitb.ac.in. Last accessed 2020/09/24.
19. Part of Speech Tagger for Hindi Homepage: https://sivareddy.in/downloads. Last accessed 2020/09/24
20. Natural Language Toolkit—NLTK 3.5 documentation Homepage. https://www.nltk.org/. Last accessed 2020/09/05

Lung CT Image Segmentation: A Convolutional Neural Network Approach

Narendra Lalchand Lokhande and Tushar Hrishikesh Jaware

Abstract Among various types of cancer, lung cancer is the most severe which causes more deaths. Early disease assessment can increase survival rate. The computer-aided system for automatic detection of lung cancer requires accurate segmentation of lung region. So, it can be detected accurately for treatment. We know deep learning-based method performs very well compared to conventional methods of segmentation. In this work, we proposed convolutional neural network (CNN) with U-net model for segmentation of lung CT images. This network is trained with limited images with technique of data augmentation. Experimental results show training and validation accuracy of 0.9085 and 0.9161 with dice coefficient index, while corresponding loss of 0.0784 and 0.0816 reported for training and validation, respectively.

Keywords Convolutional neural network (CNN) · Lung cancer · CT image · ReLU · Segmentation

1 Introduction

A leading cause of universal cancer-related deaths is lung cancer [1]. The effective way to lower death rates is to recognize lung cancer in its initial stage [2]. However, it is difficult to find out the lung cancer in its initial stage and for which the physician asks the patients to undergo several regular intervals of computed tomography (CT) images [3]. Although CT imaging accurately captures the images of the lungs, the physician still finds it difficult to identify the nodules that are cancerous. The reason for this is due to the continuous cross-sectional images produced by the CT scanner and required for each cross section to be analyzed. This demand for extra effort put into detecting lung cancer by the radiologist, and therefore a high probability of error. Developing a computer-aided diagnostic (CAD) system can help the physician and radiologist interpret the CT images accurately to improve the accuracy [4].

N. L. Lokhande (✉) · T. H. Jaware
R C Patel Institute of Technology, Near Nimzari Naka, Shirpur Dhule, India

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_37

Subsequent automated research to diagnose lung cancer early includes significant steps such as preprocessing, segmentation, nodule identification, nodule recognition, and false positive reduction. So accurate segmentation of the lung in CT is one of the most relevant and important tasks to find out nodule in succeeded steps.

In order to take out lung region from CT scan image, this proposed work uses CNN.

The paper is structured as follows: It presents related work on lung segmentation in Sect. 2. Section 3 elaborates with the description of the model architecture, model configuration, and the proposed algorithm the methodology of the present work. Section 4 focuses on the results, and Sect. 5 concludes the paper and sets out directions for the future.

2 Related Work

Segmentation of lung is difficult because of the locational non-uniformity. Diligence of research work carried out in this field. Still segmenting lung portion is found to be a difficult task.

Conventional approaches for detection of lung region are mostly based on intensity, model, and morphological. Tong et al. [5] proposed adaptive threshold method for the detection of lung region and then using contour model lung vessel is separate out. Overall accuracy of 85% on the thoracic clinical dataset of 90 is achieved.

An incremental constrained non-negative matrix factorization for three-dimensional lung voxel by data cleaning with region map and data clustering is proposed by Hosseini et al. [6].

Sun et al. [7] presented an enhanced active type model for segmenting irregular lung area from forty pathological CT scanning images. Adaptive threshold-based method for segmentation of lung was proposed by B. Muthazhagan et al. [8]

Krizhevsky et al. trained a deep CNN model with five layers of convolution and followed by maxpooling layers and three fully linked layers [9]. This model is composed of about 60 million parameters.

Anthimopoulos et al. [10] come up with a CNN with five convolutional layers having leaky ReLU, method of average pooling with three fully connected layers. This model reported the accuracy of 85.5% for classification.

For automatic segmentation of lung CT scans a U-Net with eleven layered CNN was proposed by Shaziya et al. [11]. In the analysis of biomedical image, the absence of large annotated samples is constrained. U-Net, a specific type of CNN for analysis of medical images, was presented by the Ronneberger et al. [12]. Data augmentation is used to increase the data size in order to increase sufficient learning of U-Net. Imaging challenge for segmentation of biomedical image 2015 won by U-Net.

3 Methodology

A convolutional neural network (CNN) is bio inspired deep architecture. The standard CNN system consists of multiple convolutional, subsampling, and a fully connected layer. CNN model has been created from scratch and trained on lung CT images dataset on Kaggle [13]. The further explanation of all the stages in depth is described below.

3.1 CNN Architecture

CNN models are feed-forward networks with convolutional, pooling, flattening, and fully connected layers employing suitable activation functions.

Convolutional layer: It is the basic building block of the CNNs. In mathematics to perform polynomial multiplication, convolution operation is used. The function of convolution is to combine two mathematical entity. In the CNN models, the input image is first transformed in the form of a matrix. Convolution filter of size 3×3, 5×5 or 7×7 is used on the input image. The filter slides over the image, performing multiplication with each pixel and stores the sum. This creates a feature matrix. Preferably 3×3 filter is employed to create 2D feature maps. Several filters are operated with the input image to obtain more feature maps. This forms a feature map layer know as convolutional layer.

After the convolution layer, to deal with vanishing gradient an activation layer is added. Nonlinear activation function rectified linear unit (ReLU) [14] is most commonly used. This activation function gives output one when input is positive otherwise zero. Some variants of ReLU activation function are leaky ReLU, parametric ReLU, and concatenated ReLU [15]. The main advantage of using ReLU activation function is its sparsity, computationally simple, and converge fastly. In a proposed work, we used ReLU and sigmoid activation functions.

Pooling layer: Data dimensionality can reduce by pooling layers which may appears after convolutional layers. Pooling layers spatially reduces size. The pooling methods are max-pooling, min-pooling, average pooling, and global pooling. Two variations of pooling—general pooling and overlapping pooling—can also be used. The max pooling with dimension 2×2 is most commonly used to recognize important features in image. In a proposed work, max pooling is used.

Fully Connected Layer: Fully connected layer has full connection to all activations in the previous layer. So all the features calculated in previous layers are forwarded to all the neuron to this layer. Being fully connected layer, features from earlier node get combined with the present layer. From this, aggregated features are obtained. Based on extracted feature at fully connected layer prediction is to be made. A column vector is formed from image which is called flattening.

Fig. 1 Convolutional U-Net architecture for segmentation of Lung

3.2 Architecture Configuration

The algorithms used to segment lung region by CNN have been explained in Fig. 1. The epochs for the model presented in this work were fixed at 10. We have not trained various optimizer functions. The Adam optimizer function was chosen for training. The parameters for the network are specified as drop out of 0.25, learning rate of 0.8, and loss function as binary cross-entropy. Evaluation metric is accuracy in terms of dice coefficient.

The top layer is an input layer which takes an image of size 64 $\times$ 64 as an input. An original image size is of 512 $\times$ 512. The purpose of resizing is to reduce computational complexity. Configuration of CNN is presented in Table 1. It gives information about layer name, output shape, number of parameters trained, and the connection between the layers.

3.3 Data Augmentation

To improve the performance of computer vision, more and more data is required. For proper training of model, it is required to have sufficient large dataset. So, to increase data size data augmentation can be used. Common data augmentation operations are zooming, rotation, horizontal flip, vertical flip, shift, shearing, and cropping. Figure 2 shows CT images with augmentation operations described.

Table 1 Configuration of CNN

Layer used (type)	Shape	No. of Param	In connection with
in_1 (Input)	None, 64, 64, 1	0	–
con2d_1 (Con2D)	(None, 64, 64, 8)	208	in_1[0][0]
max_pool_1 (MaxPool2D)	(None, 32, 32, 8)	0	con2d_1[0][0]
con2d_3 (Con2D)	(None, 32, 32, 32)	6432	max_pool_1[0][0]
max_pool_2 (MaxPool2D)	(None, 16, 16, 32)	0	con2d_3[0][0]
drop_1 (Dropout)	(None, 16, 16, 32)	0	max_pool_2[0][0]
con2d_4 (Con2D)	(None, 16, 16, 64)	18496	drop_1[0][0]
max_pool_3 (MaxPool2D)	(None, 8, 8, 64)	0	con2d_4[0][0]
drop_2 (Dropout)	(None, 8, 8, 64)	0	max_pool_3[0][0]
con2d_5 (Con2D)	(None, 8, 8, 64)	4160	drop_2[0][0]
upsampling2d_1(UpSampling2D)	(None, 16, 16, 64)	0	con2d_5[0][0]
concat_1 (Concat)	(None, 16, 16, 128)	0	upsampling2d_1[0][0] con2d_4[0][0]
con2d_6 (Con2D)	(None, 16, 16, 32)	16416	concat_1[0][0]
upsampling2d_2(UpSampling2D)	(None, 32, 32, 32)	0	con2d_6[0][0]
concat _2 (Concat)	(None, 32, 32, 64)	0	upsampling2d_2[0][0] con2d_3[0][0]
con2d_7 (Con2D)	(None, 32, 32, 24)	6168	concat _2[0][0]
upsampling2d_3 (UpSampling2D)	(None, 64, 64, 24)	0	con2d_7[0][0]
concat_3 (Concat)	(None, 64, 64, 32)	0	upsampling2d_3[0][0] con2d_1[0][0]
con2d_8 (Con2D)	(None, 64, 64, 16)	2064	concat_3[0][0]
con2d_9 (Con2D)	(None, 64, 64, 64)	1088	con2d_8[0][0]
drop_3 (Dropout)	(None, 64, 64, 64)	0	con2d_9[0][0]
con2d_10 (Con2D)	(None, 64, 64, 1)	65	drop_3[0][0]

4 Results

On Windows 7, the proposed model was developed using python 3.7, Keras 2.3.1, tensorflow 2.0.

Dataset consists of 267 images with their corresponding masks [13]. The set of data is divided into the proportion of 30:70 as testing and training set. All images are in shade of 0 to 255 with spatial resolution of 512 × 512. In the preprocessing step, image dimensions are resized to 64 × 64 in order to train on low computational platform. A sample of CT scan image along with its ground truth is shown in Fig. 3.

Accuracy and loss plot for validation and training are shown in Fig. 4. The set of validation achieves accuracy of 0.9169 while training accuracy obtained for set is 0.9085. The training loss of 0.0784 and the validation loss is 0.0816. Figure 5 shows screen shot of original image at left, its ground truth in the middle and segmented lung region using proposed network at the left.

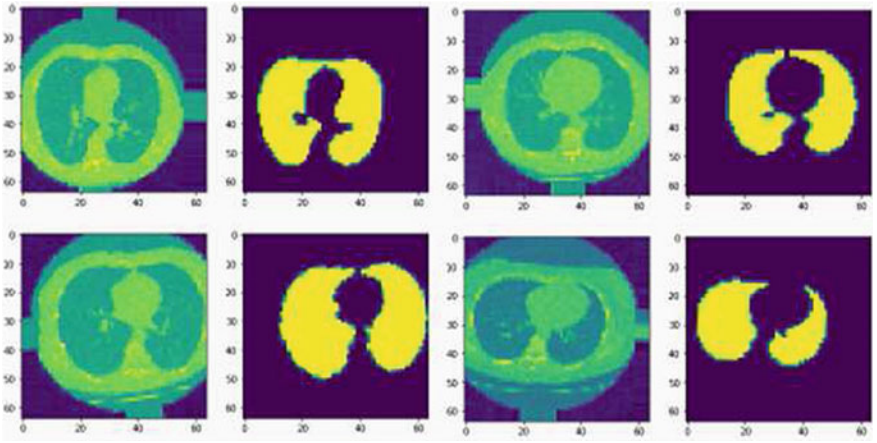


Fig. 2 Augmented lung dataset with ground truth

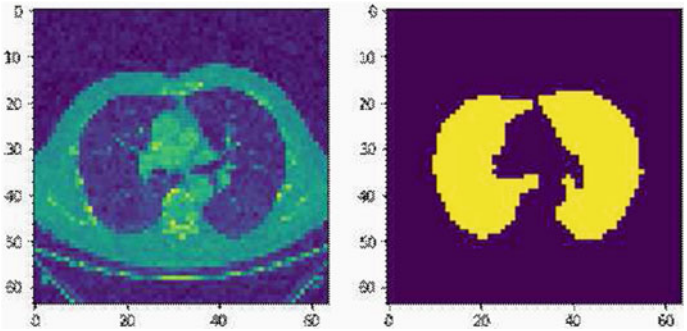


Fig. 3 Lung region and its ground truth

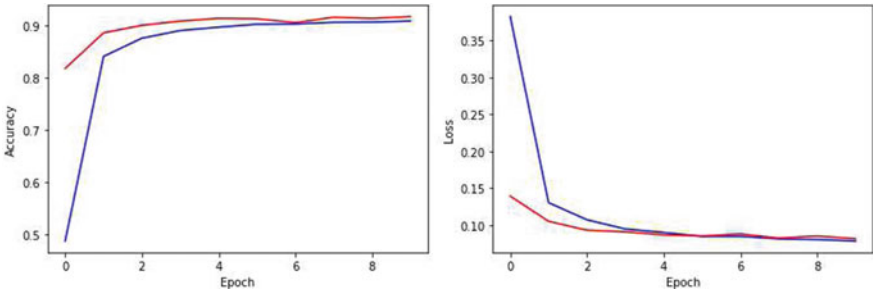


Fig. 4 Plot for accuracy and loss

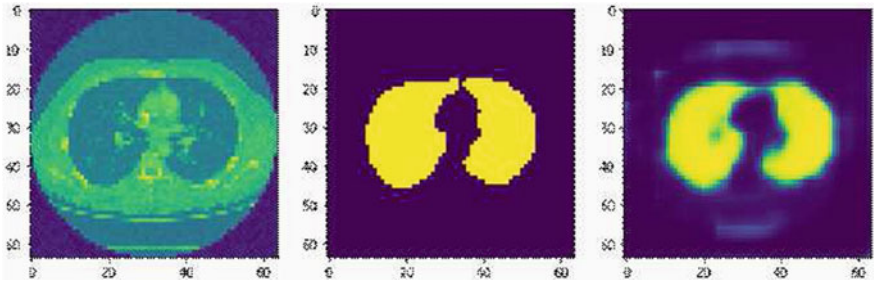


Fig. 5 Original lung image, ground truth, and segmented lung region

5 Conclusion and Future Work

Segmenting region of lung in CT image is one of the important steps in CAD system. Accurate segmentation of lung region is required. The proposed work uses rescaling of original images of size 512×512 to 64×64 in order to save computational complexity. Further, accuracy of segmentation can be increased by increasing larger training dataset with original dimensions. In the next phase of our work, we will try to increase segmentation accuracy by modifying model parameter or by increasing number of CNN layers.

References

1. Cancer.org. Lung Cancer Guide | What You Need To Know [online] (2020). Available at: https://www.cancer.org/cancer/lung-cancer.html. Accessed 5 February 2019
2. American Cancer Society Cancer Facts & Statistics: American Cancer Society | Cancer Facts & Statistics (2020) (online). Available at: https://cancerstatisticscenter.cancer.org/. Accessed 5 February 2019
3. Avinash, S., Manjunath, K., Senthilkumar, S.: Analysis and comparison of image enhancement techniques for the prediction of lung cancer. In: 2017 2nd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT) (2017)
4. Zhang, G., Jiang, S., Yang, Z., Gong, L., Ma, X., Zhou, Z., Bao, C., Liu, Q.: Automatic nodule detection for lung cancer in CT images: a review. Comput. Biol. Med. **103**, 287–300 (2018)
5. Tong, J., Ying, W., Dong, W.: A lung cancer lesions dectection scheme based on CT image. In: 2010 2nd International Conference on Signal Processing Systems (2010)
6. Hosseini-Asl, E., Zurada, J., Gimelfarb, G., El-Baz, A.: 3-D Lung segmentation by incremental constrained nonnegative matrix factorization. IEEE Trans. Biomed. Eng. **63**(5), 952–963 (2016)
7. Sun, S., Ren, H., Meng, F.: Abnormal lung regions segmentation method based on improved ASM. In: 2016 Chinese Control and Decision Conference (CCDC) (2016)
8. Muthazhagan, B., Ravi, T., Rajinigirinath, D.: An enhanced computer-assisted lung cancer detection method using content based image retrieval and data mining techniques. J. Ambient Intell. Human Comput. (2020). https://doi.org/10.1007/s12652-020-02123-7
9. Krizhevsky, A., Sutskever, I., Hinton, G.: ImageNet classification with deep convolutional neural networks. Commun. ACM **60**(6), 84–90 (2017)

10. Anthimopoulos, M., Christodoulidis, S., Ebner, L., Christe, A., Mougiakakou, S.: Lung pattern classification for interstitial Lung diseases using a deep convolutional neural network. IEEE Trans. Med. Imaging **35**(5), 1207–1216 (2016)
11. Shaziya, H., Shyamala, K., Zaheer, R.: Automatic Lung segmentation on thoracic CT scans using U-net convolutional network. In: 2018 International Conference on Communication and Signal Processing (ICCSP) (2018)
12. Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation. In: 2015 International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, Berlin, pp. 234–241 (2015)
13. Kaggle.com.: A Lung U-Net In Keras (online) (2020). Available at: https://www.kaggle.com/toregil/a-lung-u-net-in-keras/data. Accessed 9 March 2020
14. Eckle, K., Schmidt-Hieber, J.: A comparison of deep networks with ReLU activation function and linear spline-type methods. Neural Netw. **110**, 232–242 (2019)
15. Shang, W., Sohn, K., Almeida, D., Lee, H.: Understanding and Improving Convolutional Neural Networks via Concatenated Rectified Linear Units (2016). ArXiv abs/1603.05201

Overcoming the Complexities in Decision-Making for Enterprise Software Products: Influence of Technological Factors

Md. Hashmathur Rehman and M. Rajkumar

Abstract Technological aspects of an Information Technology (IT) innovation have a vital role in organization's decision for adopting the IT innovation. Due to innovations in distributed computing, customers has the provision to avail Software products in two delivery models namely On-Premise model and Cloud computing model. Each deployment model possesses advantages and disadvantages one over the other and at the same time organizations are posed with the challenge of adopting one among them. Relative advantage is one of the key parameters in decision making, it has an impact on the decision maker in terms of cost, performance, utilization etc. Complexity is how the user perceives the ease of use for an innovation. The other technical aspects are Compatibility with existing infrastructure, Data security, Data privacy, and Data backup. Observing the results of the software product under trial period and making comparisons between the two kind of delivery models helps in adopting the software product model in long term usage. In this research paper, we have determined and evaluated the technological aspects of enterprise software product which helps organization's in decision-making for adopting software product's delivery model. The Technological context comprises of technical characteristics of a software product which can ease the adoption of software model. Complexity, Compatibility, Relative advantage, Trialability, Observability, Data Privacy, Data Security, Data backup, Data Accessibility, Data Location are the major factors identified in this study. Focused group discussion and interviews were conducted with the organizations for finding as how various technological factors has an impact on the decision of organization's in the adoption of software products in the form of on-premise model or the cloud computing model.

Keywords Technology-organization-environment framework · Diffusion of innovation · DEMATEL · On-premise · Cloud computing · Software products · Adoption

Md. H. Rehman (✉) · M. Rajkumar
ICFAI University Jharkhand, Ranchi, India
e-mail: rehman.mh@iujharkhand.edu.in

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_38

1 Introduction

Software refers to the programs, instruction or data that are used for operating and executing the specific tasks in computers [1]. A Software Product is the application made available to customers with a specific order number or a distinct name [2]. With the day-to-day upgradation in the field of software delivery models, software products are now available to customers in two kinds of delivery models and they are On-Premise and Cloud Computing.

1.1 On-premise Delivery Model of Software Product

Software products that are available physically in the customers premises and that are installed on computers or appliances are termed as On-Premise software products. They are available to end customers via downloads, ISOs, DVDs, CDs, etc. The end user needs license key in order to use this [3].

1.2 Cloud Computing Delivery Model of Software Product

In this model, The Software product is installed in the vendors data center of vendor and the vendor is the actual manufacturer of software product. In this model, through internet the software product and its services are made available to the customer [4]. Cloud computing delivery model of software product is available as Platform-as-a-service, Infrastructure-as-a-service, and Software-as-a-service.

Software-as-a-service is defined as a cloud computing service model that supports virtualization of software applications, [5]. Ex: Google Apps, Office 365.
Platform-as-a-service is defined as a cloud computing service model that provides virtualization of Operating systems, [6], e.g., Windows, Linux.
Infrastructure-as-a-service is defined as a model that provides virtualization of computing resources, [7]. Ex: Networking hardwares and Servers.

2 Theoretical Background

Research studies indicate that organizations will perform better when technological innovations are adopted and may succeed over their competitors [8]. The Technological adoption models which are considered as theoretical reference in this research study are Technology-Organization-Environment, Diffusion of Innovation, DEMATEL.

2.1 Technology-Organization-Environment Framework (TOE)

The Predominantly used framework towards the adoption of innovations by organizations is Technology-Organization-Environment developed by Tornatzky and Fleischer [9]. According to TOE framework, an innovation in communication as well as information technology is influenced by environmental, organizational, and the technological factors [10]. Wen and Chen [11], in their research study states that TOE can be used independent of type of industry and size of organization.

2.1.1 Technology Context

Technological context of an innovation describes its technical characteristics. There are two types of technological factors namely external and internal factors which influences the organizations to adopt IT innovation. Internal technological factors are existing technologies in the organization and the external technological factors are those which are available in the marketplace [12]. Among the technological factors, the most influencing factors in organization's decision for adopting innovations are Observability, Trialability, Relative advantage, Complexity, and Compatibility [10, 13–16].

2.1.2 Organizational Context

Organizational context of an IT innovation describes the traits, attributes, and characteristics of an organizations which influences adoption of IT innovation by the firms. They are top management support, readiness of organization, size of an organization, Level of awareness among its IT employees [10, 17–19]. For example, Support from the top management acts as an enabler for adoption of cloud-based technologies, Managers with good experience of Cloud computing will adopt it [19].

2.1.3 Environmental Context

For an organization, the environmental situation refers to the surroundings in which it operates. It comprises of several stakeholders such as business competitors, government, customers, governing board members, and suppliers, etc. These stakeholders are responsible for influencing the decision to adopt an innovation by the organization [20]. The most common factors in the context of environment are Business partner pressure, Industry type, Competitive pressure, Support from the Government, External support, Advice of IT Specialist, and Service Level Agreement [10, 20, 21].

2.2 Diffusion of Innovation

Diffusion of Innovation states that Diffusion is a systematic approach using which a new idea or process is adopted and conveyed through various channels to its stake holders over a period. An innovation is defined as a new object or process or idea which is perceived as new by an individual or unit of adoption. Communication is a process of creating and sharing the information among its members till a mutual understanding is reached [22]. The five characteristics of innovation are observability, compatibility, complexity, relative advantage, observability, and trialability were identified by Rogers and found helpful in the assessment of adoption rate [22, 23].

2.3 DEMATEL

Decision Making Trial and Evaluation Laboratory (DEMATEL) is a method of dividing factors into cause and effect groups with a causal diagram [24, 25]. Wu et al. [26], has emphasized that trust is the key factor of success in adoption of any type of e-commerce i, e.. when there is more than one choice of e-commerce products are available. Deutsch [27] states that trust will get established when perceived benefits surpasses perceived risks. The eight perceived benefit factors identified are easy to deploy to end-users, pay as you go, monthly payments, requires less in-house IT staff and costs, seems like the way of future encourages standard systems, always offers latest functionally, sharing systems with partners is simpler [28] and the seven perceived risk factors identified are network and web application security, data access and backup, data integrity and segregation, authentication and authorization, identity management, data locality and security, virtualization vulnerability, and sign-on process [29].

3 Review of Literature

Cloud Computing is found to be the remarkable innovation in IT-related software products. The keywords used for searching in Google Scholar are “adoption”, “cloud”, “PaaS”, “IaaS”, “SaaS”, “TOE”, “Diffusion of Innovation” and “DEMATEL”. It is evident that majority of research scholars carried out their research on adoption factors for SaaS service model of cloud computing. This is in line with the approach followed by Van de Weerd et al. [19]. Al-Hujran et al. [12] has done his research using TOE framework limited to Jordanian companies. The cloud adoption in Jordanian organizations is very low but they found it very useful. The results of the research reveal that cloud computing delivery model makes sense to SMEs. At the

same time, issues were found relating to TOE (Technical, Organization, and Environmental) context which needs to be addressed before using the cloud computing services effectively by organizations in Jordan. The technological factors identified are privacy concerns, compatibility, security, and trust. On the other hand, the organizational factors identified are organizational culture, characteristics of CEOs and top management support. Finally, the critical factors that is hindering adoption of cloud services by organizations in Jordan from situational standpoint is the need for SLAs contractual agreements, and regulatory framework. Hsu [15] has also adopted the TOE framework for inspecting the cloud adoption in TAIWAN and stated that cloud adoption is significantly influenced positively by factors such as "IT capability" and "Perceived Benefits" and negatively by "Business Concerns" towards cloud computing adoption. According to [13, 17], organizational factors exhibit a major role in cloud adoption. They did their research study on Indonesian organizations and noted "top management support as the enabler for cloud adoption" [19]. Mujinga and Chipangura [21] stated that in cases of privacy and security breaches, the major concerns were business continuity and service availability. Security issues need to be taken seriously by small business operating in the developing economies when they decide on cloud computing and while planning to migrate to cloud. Wu et al. [28] has used DEMATEL framework finding the factors influencing cloud adoption. Seven perceived risk factors and Eight perceived benefit factors were found to influence the decision of adoption of cloud as listed under critical factors.

3.1 Technological Factors Influencing Organizations to Adopt On-cloud or On-premises Computing Delivery Model of Enterprise Software Product

After integration of three frameworks TOE, DOI, and DEMATEL and with the help of literature review, we now define the technological factors used in this research study.

Relative advantage is defined as "The extent to which a new idea or process is observed as better than the technology, or product or service it replaces, [9, 22]".
Compatibility is defined as "The extent to which a new idea or process is perceived as consistent with existing values, adopter needs and past experiences, [9, 22]".
Complexity is defined as "The extent to which a new idea or process is perceived as relatively difficult to understand and use, [9, 22]".
Trialability is defined as "The extent to which an IT innovation may be evaluated on a limited basis, [9, 22]".
Observability is defined as "The extent to which a customer can observe the innovation and its positive effects is known as observability, [9, 22]".
Data Availability is defined as, "It is the responsibility of providers to make data available for 24*7 days without delay, [29]".

Table 1 Overview of organizations

Organizational characteristics						Respondent characteristics	
#	Industry type	Employees	IT staff	Total assets (Dollars)	Migration status	Position	Education
O1	Banking	97,535	2710	1.72 Trillion	Migrated	Assistant Vice President	IT
O2	Banking	85,000	2360	2.374 Trillion	Migrated	IT Staff	IT
O3	Financial	10,000	300	3.15 Billion	Migrated	Manager	IT
O4	Education	1200	50	10 Million	No	IT Staff	IT
O5	Pharmaceutical	126,000	2500	133 Billion	No	IT Staff	IT
O6	Transportation	12,000	400	15.3 Billion	Migrated	Manager	IT
O7	IT	380,300	250,000	125.35 Billion	Migrated	Sr. Manager	IT
O8	IT	49,000	40,000	70 Billion	Migrated	Director	IT
O9	IT	4500	3500	1 Billion	Migrated	IT Staff	IT
O10	IT	160,000	90,000	11.68 Billion	Migrated	IT Staff	IT
O11	IT	8071	6000	5482 Million	No	IT Staff	IT
O12	IT	120	100	20 Million	No	Manager	IT
O13	IT	124,000	80,000	241 Billion	No	IT Staff	IT
O14	IT	500	350	500 Million	No	IT Staff	IT
O15	IT	450	300	200 Million	No	IT Staff	IT
O16	IT	4712	3000	5 Billion	No	IT Staff	IT
O17	IT	1200	1000	1 Billion	No	Director	IT
O18	IT	1200	1000	1 Billion	No	Sr. Manager	IT
O19	IT	1100	1000	1 Billion	No	IT Staff	IT
O20	On-line shopping	566,000	28,385	131 Billion	No	Sr. Manager	IT

Data Accessibility is defined as "It is the control mechanism of data access which confirms secure authorized access to data and prevent un-authorized access of data, [29]".
Data Security is defined as "Data is secured from natural and man in the middle attacks, [29]".
Data Integrity is defined as "protection of data from unauthorized access, modification and deletion from unauthorized user, hackers, intruders. Integrity ensures that data is not tampered during its journey from source to destination. Service provider must guarantee that data will be transmitted in a secure channel without getting tampered, [30, 31]".
Data Privacy is defined as "The provider must guarantee that data is segregated at physical layer for each user. The service provider must have intelligence to isolate data among different users in its services, [29]".
Data backup is defined as "Provider must ensure to its customers that the backup of stored data is taken at regular interval and is available for quick restore. If disasters like data loss or fire, Server crash happens then data should be made available from disaster recovery center of data [32]".
Data Locality is defined as follows, "Service provider must ensure that storing and processing of data will be in jurisdiction limits and the service level agreement will comply to the policies of judiciary systems, security managements, and data privacy, [32]".
Virtualization vulnerability is defined as "Virtualization provisions a user to create, share, copy, roll back, migrate virtual machines. It also allows them to run different types of software applications, [33]". Virtual machine technology does not offer perfect isolation of data and machine. Research on Virtualization security and management of virtual machine's environment is still in progress. Researchers are doing continuous work for enhancing security measures and performance of virtual machines [29].
Identity Management is a process of assigning Identity (ID) to an individual or process. The process then verifies the credentials and gives access to resources in the system. The process also enforces restrictions on already established identities [29].
Single Sign-on process is defined as, "Authentication to one application gives authorization to other applications. SAML (Security Assertion Markup Language) and WS-Fed (Web Service-Federation) technologies are widely used for single sign-on process. The alternative to WS-Fed and SAML is VPN SSO which is a Single Sign-on solution implemented using a Virtual Private Network tunnel, [34]". An end user proves his identity to one SAML or WS-Fed supported software application and the same user can access other software application enabled with SAML, WS-Fed without proving his identity again.

4 Research Methodology

Multiple case study method was adopted in this research work for investigating the influence of technological factor in the adoption of delivery model of software product through cloud computing. Yin [35] has detailed on multiple case studies, with this approach we can investigate the adoption of enterprise software products delivery models through cloud by the organizations in real-world context. Eisenhardt [36], detailed on using the research design and with the help of the suggested design, the researchers compared and contrasted the findings systematically for all the cases in the data analysis. With the research method, we intended to find how organizations decide whether to adopt on-premise delivery model or cloud computing delivery model of enterprise software product.

4.1 Case Selection

In this section, we first discussed how we have selected case organizations for this research study. All case organizations are multinational companies operating in INDIA. In our previous studies, we have examined how organizational and environmental factors influences the organization's decision to adopt on-premise or cloud service model of delivery for enterprise software product [18, 20, 37–39]. We have identified the case companies who are using Mobile Device Management, Salesforce, On-Premise email servers, Office 365, G-Suite, Mobile Threat defense, SAP CRM software products. These software products are available in both On-premise and cloud computing delivery model. "There are atleast two different cases needed within each subgroup in order to supplement the theoretical replications across the groups with literal replication within each group, [35]".

4.2 Data Collection

Yin [35], has described the framework on data collection using focused group discussions and interviews. For primary data collection, we have used similar techniques of focused group discussions and interviews which is in line to the work of [35]. Pare [40] has described the process of obtaining information from the participants of focused group discussions and interviews. We have used the similar technique in determining and examining the factors influencing organizations to adopt delivery model of enterprise software product. An interview guide is prepared and used for all interviews. Once an input is received, the feedback in incorporated in the guide. The period of data collection is from December 2019 to July 2020. The interviews and group discussions were recorded using Audio–video conferencing tool Cisco Webex meetings. The participants comprises of decision makers from different levels of

organizations and they are IT Staff, Senior Managers, Directors, Vice Presidents and this is presented in Table 1 also. Before the start of focused group discussion or interview, the interviewer described the factors with suitable examples to the participants. The participants were assured that the information will be kept confidential and organizations are represented as O1–O20 in Table 1.

5 Data Analysis and Findings

5.1 Data Analysis

For analysis, we have used fifteen codes to organize the data: Relative advantage (C1), Compatibility (C2), Complexity (C3), Trialability (C4), Observability (C5), Data Availability (C6), Data Accessibility (C7), Data Security (C8), Data Integrity (C9), Data Privacy (C10), Data backup (C11), Data Locality (C12), Virtualization vulnerability (C13), Identity Management (C14), Single Sign-on process (C15). The code C1 comprises of responses from participant about the reasons of relative advantages be-tween two technologies in terms of performance, cost, deployment, ease of use, and so on. The data analysis is done in multiple iterations which are within-case analysis, qualitative comparative analysis and Across-case analysis. For within-case analysis, we have examined each individual case separately with the help of technological factors and outcome variable adopt/migrate, not-adopt/not-migrate. Field notes have been used to refine the findings of individual case analysis. With this approach, several influencing factors are found for adoption of enterprise software product delivery model. Qualitative comparative analysis has been used in an informal way for processing the outcome of within-case analyses. "Schulze-Bentrop [41], has described qualitative comparative analysis as a process to discover the group of conditions which can be used in explaining a specific outcome". "Rihoux and Ragin [42], has described that qualitative comparative analysis of data should be used for within-case analysis and Across-case analysis of data. All of these cases should be converted to configuration sets for comparisons". These configurations are various influencing factors in this research study. The findings are concluded with Across-case analysis section.

5.2 Data Validity

Yin [35], has defined that there are four types of data validation such as reliability, external validation, internal validation, and construct validation that reveals the quality of the research. Construct validation provides correct measures to be used for a research study. The sources of construct validation used in this research study are (1) Group discussions, documentation, interviews, and notes (2) Building continuous

evidences during interviews, and (3) Summarization of results of individual cases for feedback. Internal validation defines and evaluates the relationships discovered in the research study [35]. We have used explanation-building procedure in our study that strengthens the internal validation of data. Generalization of findings is known as external validation [35]. We used replication logic of multiple case design to strengthen generalization of findings in our research study. In the last, the reliability of the research is ensured using database for each case study. This guarantees the analysis and data collection could be repeated [35].

5.3 Findings

The findings from data analysis are presented in separate section each section like within-case analysis, qualitative comparative analysis, across-case analysis. In within case analysis, we have discussed and elaborated the results of each case organization. During qualitative comparative analysis, we have shown how various case organizations responded for each influencing factor relating to outcome variable adoption or migration to cloud services and not adopt or not migrate to cloud services. In across-case analysis section, we have shown the similar patterns found across different case organizations. We have used Nvivo software for analyzing the data. This is inline with [43], "The Nvivo Qualitative Project Book" and also used the quotes given by the interviewees to show the findings of the study.

5.3.1 Within-Case Analysis

We have examined each individual case based on fifteen variables and also the outcome variables adopt/migration or not-adopt/not-migrated. For evaluating, measurable values were assigned to the variables and the responses were obtained through group discussion or interview. For example, the factor relative advantage in terms of performance is calculated by considering the configuration of computing resources, the performance of application is assigned a number of user check-in requests per second from their mobile devices and we acquired its value from the participants during interviews. For respondent O8, "The number of user check-in attempts for On-premise server is 40 per second whereas it is approximately 25 in cloud for same amount of computing resources." For the variable compatibility, we present our analysis from O10. The informant, in this case, states that "we migrated to cloud servers for email software from Exchange on-premise to Exchange on-line. The cloud server system is compatible with existing on-premise user authentication server and is supported by vendor". For the variable complexity, we present our analysis using case O15, the respondent, in this case, states that "we directly adopted cloud solution for email system like office 365, It is easy to deploy and configure for end user".

Data Availability was represented from case organization O9. It was reported that "Cloud providers ensure that data is available for 24*7 days to users". For the variable data accessibility, the respondent O20 states that "Multi-Factor Authentication (MFA) is being used for allowing only legitimate users to access the data". In case of variable data Security, the respondent O3 states that "Data stored in cloud servers are secured before getting saved on them. Stronger encryption technology like AES is used for securely transmitting the data to users from cloud servers". For the variable data integrity, respondent O6 states that "Data integrity is guaranteed by Integrity algorithms like Message-Digest5 and Secure Hash. The providers support both algorithms which ensures that data in not tampered when traveling between cloud servers and end user devices". Data privacy represented in O9 revealed that cloud products are micro services-based models and providers ensures that data for each customer is maintained separately in tenants. For the variable data backup, the respondent O20 states that, "Data backup is planned and taken on a regular basis. The disaster recovery setup ensures that the data will be available to users in case of natural disasters or unexpected failures". For the variable data locality, the respondent O01 which is a banking organization states that "provider has adhered the data retention laws of the government. To comply the data retention laws we are storing data in on-premise servers and application in cloud servers".

5.3.2 Qualitative Comparative Analysis

Qualitative comparative analysis comprises of classifying each case organizations according to variables. In line with the approach explained by Rihoux and Ragin [42], we have used the informal qualitative comparative analysis to illustrate our findings. We have assigned values to each individual variable with either 1 or 0. If the value is 1 then there is the presence of an outcome or condition and the absence of the same is represented as 0. This is in line with the approach mentioned by Rihoux and Ragin [42]. Table 2 represents the values assigned to the variables.

The findings are represented in Table 2 which is also the database set of all organizations. A table known as truth table is extracted from Table 2 and it summarizes all configurations of fifteen conditions that had an impact on cloud computation adoption/migration decision of the organizations.

Table 3 represents the possible combination of all configuration sets. The configuration sets which are relevant to the study are displayed and the configurations which are not relevant to the study are not presented in the table. All the configuration sets are found in our case organization database set. It can be noted that the configuration sets A, B, C are leading to adoption/migration of cloud service model where as, the configurations sets D, E, F are not leading to adoption/migration of cloud model of software product, which means these organizations have adopted or remained with on-premise delivery model of software product.

Table 2 Database set of cases

#	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15	Adopted/Migrated
O1	0	0	0	0	0	1	1	1	1	1	0	1	1	1	1	0
O2	0	0	0	0	0	1	1	1	1	1	0	1	1	1	1	0
O3	0	0	0	0	0	1	1	1	1	1	0	1	1	1	1	0
O4	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
O5	0	0	0	0	0	0	0	1	1	1	0	0	1	1	1	0
O6	0	0	0	0	0	0	0	1	1	1	0	0	1	1	1	0
O7	0	0	0	0	0	0	1	1	1	1	0	0	1	1	1	0
08	0	0	0	0	0	0	1	1	1	1	0	0	1	1	1	0
O9	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	1
O10	0	0	0	0	0	0	1	1	1	1	0	0	1	1	1	0
O11	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
O12	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	1
O13	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
O14	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	1
O15	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	1
O16	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	1
O17	1	1	1	1	1	1	1	1	1	1	0	0	0	1	1	1
O18	1	1	1	1	1	1	1	1	1	1	0	0	0	1	1	1
O19	1	1	1	1	1	1	1	1	1	1	0	0	0	1	1	1
O20	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

Table 3 Truth table

Configuration	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15	Adopt/Migrate	Not Adopt/Not Migrate
A: 11111111100000	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	5	
B: 11111111100011	1	1	1	1	1	1	1	1	1	1	0	0	0	1	1	3	
C: 11111111111111	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	4	
D: 00000111101111	0	0	0	0	0	1	1	1	1	1	0	1	1	1	1		3
E: 00000001100111	0	0	0	0	0	0	0	1	1	1	0	0	1	1	1		2
F: 00000011100111	0	0	0	0	0	0	1	1	1	1	0	0	1	1	1		3

5.3.3 Across-Case Analysis

In this section, the output from the interviews conducted and the focus group discussion were represented. An identical patterns that explaining the adoption/migration decision of organizations towards cloud computing is mentioned.

Relative advantage has an influence on Twelve of the organizations selected as the case in the study. All (O4, O9, O11, O12–O20) of them have adopted/migrated to cloud service model of delivery for enterprise software product. All the above-mentioned case organizations irrespective of their industry type are influenced by this reason and adopted/migrated to cloud service delivery model of software product where as case organizations O1, O2, O3, O5, O6, O7, O8, O10, adopted or remained with on-premise model.
Compatibility has an influence on Twelve of the organizations selected as the case in the study. All (O4, O9, O11, O12–O20) of them have adopted/migrated to cloud service model of delivery for enterprise software product where as case organizations O1, O2, O3, O5, O6, O7, O8, O10, adopted or remained with on-premise model.
Complexity has an influence on Twelve of the organizations selected as the case in the study. All (O4, O9, O11, O12–O20) of them have adopted/migrated to cloud service model of delivery for enterprise software product where as case organizations O1, O2, O3, O5, O6, O7, O8, O10, adopted or remained with on-premise model.
Trialability has an influence on Twelve of the organizations selected as the case in the study. All (O4, O9, O11, O12–O20) of them have adopted/migrated to cloud service model of delivery for enterprise software product where as case organizations O1, O2, O3, O5, O6, O7, O8, O10, adopted or remained with on-premise model.
Observability has an influence on Twelve of the organizations selected as the case in the study. All (O4, O9, O11, O12–O20) of them have adopted/migrated to cloud service model of delivery for enterprise software product where as case organizations O1, O2, O3, O5, O6, O7, O8, O10, adopted or remained with on-premise model.
Data Availability has an influence on fifteen of the organizations selected as the case in the study. Case organizations O4, O9, O11–O20 have adopted/migrated to cloud service model of delivery for enterprise software product where as case organizations O1, O2, O3, O5, O6, O7, O8, O10, adopted or remained with on-premise model.
Data Accessibility has an influence on eighteen of the organizations selected as the case in the study. Case organizations O4, O9, O11–O20 have adopted/migrated to cloud service model of delivery for enterprise software product where as case organizations O1, O2, O3, O5, O6, O7, O8, O10, adopted or remained with on-premise model.
Data Security has an influence on all the organizations selected as the case in the study. Case organizations O4, O9, O11, O12–O20 have adopted/migrated to

cloud service model of delivery for enterprise software product where as case organizations O1, O2, O3, O5, O6, O7, O8, O10, adopted or remained with on-premise model.

Data Integrity has an influence on all the organizations selected as the case in the study. Case organizations O4, O9, O11, O12–O20 have adopted/migrated to cloud service model of delivery for enterprise software product where as case organizations O1, O2, O3, O5, O6, O7, O8, O10, adopted or remained with on-premise model.

Data Privacy has an influence on Twelve of the organizations selected as the case in the study. Case organizations O4, O9, O11, O12–O20 have adopted/migrated to cloud service model of delivery for enterprise software product where as case organizations O1, O2, O3, O5, O6, O7, O8, O10, adopted or remained with on-premise model.

Data backup has an influence on four of the organizations selected as the case in the study. All of them (O4, O11, O13, O20) have adopted/migrated to cloud service model of delivery for enterprise software product where as case organizations O1, O2, O3, O5, O6, O7, O8, O10, adopted or remained with on-premise model.

Data Locality has an influence on seven of the organizations selected as the case in the study. Case organizations O4, O11, O13, O20 have adopted/migrated to cloud service model of delivery for enterprise software product where as case organizations O1, O2, O3, O5, O6, O7, O8, O10, adopted or remained with on-premise model.

Virtualization vulnerability has an influence on Twelve of the organizations selected as the case in the study. Case organizations O4, O11, O13, O20 have adopted/migrated to cloud service model of delivery for enterprise software product where as O1, O2, O3, O5, O6, O7, O8, O10 adopted or remained with on-premise model.

Identity Management has an influence on eighteen of the organizations selected as the case in the study. Case organizations O4, O11, O13, O17, O18, O19, O20 of them have adopted/migrated to cloud service model of delivery for enterprise software product where as O1, O2, O3, O5, O6, O7, O8, O10 adopted or remained with on-premise model.

Single Sign-on process has an influence on fifteen of the organizations selected as the case in the study. Case organizations O4, O11, O13, O17, O18, O19, O20 of them have adopted/migrated to cloud service model of delivery for enterprise software product where as case organizations O1, O2, O3, O5, O6, O7, O8, O10, O14, O15, O16 adopted or remained with on-premise model.

6 Conclusion

This research was made in an objective to explore and evaluate the impact of technological factors on the decision of the organization's towards adoption/migration and

also non-adoption/non-migration towards cloud computing model of software products. It was noted from the study that these factors influence the decision of the organizations. This paper makes various contributions that has both theoretical and practical implications. First with regard to theoretical implications, we have explored and investigated the factors which influences the organizations decision towards adoption/migration or non-adoption/non-migration towards cloud computing model of software product. Factors like Complexity, Compatibility, Trialability, Observability, and Relative advantage affected the decision. Whereas Banking organizations, Financial institutions are not-adopting/not-migrating due to factors like Data Integrity, Data Security, Virtualization vulnerability, Data Locality and Identity Management. This research paper not only covers theoretical implications but practical implications also. Government regulations for data retention in the geography of the country has large impact on the decision of organizations to adopt/migrate back to on-premise technologies. Apart from this, banking and financial institutions are adoption hybrid model of software product which means that applications will be running in cloud whereas data will be stored in on-premise infrastructure. With this approach, they are in line with government regulations and comply to industry acts like FATCA. Finally, for SMEs cost saving is the main concern, the factor Relative advantage in terms of cost is the main reason to adopt/migrate to cloud services model of delivery. The limitations of this research study are as follows. The scope of this study is majorly IT industry, though we have other types of industries like Banking, Pharmaceutical, On-line shopping, Transportation, the number of case organizations are also limited to 20 covering different type of industries. Although we have used replication to strengthen the reliability of our research study, several other types of industries are not included. To solve the issue of small amount of cases, one can use survey-based data collection method so that a greater number of organizations can be covered. The other limitation in Interviews based data collection method is if an interesting issue rises then the follow-up question cannot be asked.

References

1. Rouse, M.: Software (2019). Retrieved from https://searchapparchitecture.techtarget.com/definition/software
2. Rouse, M.: Product (2005). Retrieved from https://whatis.techtarget.com/definition/product
3. Sweeney, M.: The Difference Between SaaS Applications and On-premises (2014). Retrieved from https://clearcode.cc/blog/saas-applications-vs-on-premises/
4. Rouse, M.: Cloud Computing (2020). Retrieved from https://searchcloudcomputing.techtarget.com/definition/cloud-computing
5. Rouse, M.: Software as a Service (2019). Retrieved from https://searchcloudcomputing.techtarget.com/definition/Software-as-a-Service
6. Rouse, M.: Platform as a Service (2019). Retrieved from https://searchcloudcomputing.techtarget.com/definition/Platform-as-a-Service-PaaS
7. Rouse, M.: Infrastructure as a Service (2018). Retrieved from https://searchcloudcomputing.techtarget.com/definition/Infrastructure-as-a-Service-IaaS
8. Geroski, P., Machin, S., Van Reenen, J.: The profitability of innovating firms. Rand J. Econ. **24**, 198–211 (1993). https://doi.org/10.2307/2555757

9. Tornatzky, L., Fleischer, M.: The Process of Technology Innovation. Lexington Books, Lexington, MA (1990)
10. Oliveira, T., Thomas, M., Espadanal, M.: Assessing the determinants of cloud computing adoption: an analysis of the manufacturing and services sectors. Inf. Manage. **51**(5), 497–510 (2014)
11. Wen, K.W., Chen, Y.: E-business value creation in small and medium enterprises: a US study using the TOE framework. Int. J. Electron. Bus. **8**(1), 80–100 (2010). https://doi.org/10.1504/IJEB.2010.030717
12. Al-Hujran, O., Al-Lozi, E.M., Al-Debei, M.M., Maqableh, M.: Challenges of cloud computing adoption from the TOE framework perspective. Int. J. E-Bus. Res. **14**(3) (2018)
13. Mangula, I.S., Van de Weerd, I., Brinkkemper: the adoption of software-as-a- service: an Indonesian case study. In: Proceeding of PACIS (2014)
14. Low, C., Chen, Y., Wu, M.: Understanding the determinants of cloud computing adoption. J. Ind. Manage. Data Syst. **111**(7), 1006–1023 (2011)
15. Hsu, P.F.: Cloud computing adoption in Taiwan: an empirical study. In: 2013 International DSI and Asia Pacific DSI Conference (2013)
16. Alshamaila, Y., Papagiannidis, S., Li, F.: Cloud computing adoption by SMEs in the North East of England. J. Enterp. Inf. Manage. **26**(3), 250–275 (2012)
17. Mangula, I.S., Van de Weerd, I., Brinkkemper, S.: Adoption of the cloud business model in Indonesia: triggers, benefits, and challenges. In: Proceeding of the 14th International Conference on Information Integration Web-based Applications & Services (IIWAS), ACM Digital Library (2012)
18. Rehman, M.H., Majumdar, S., Krishna, C.Y.S.: Adoption of infrastructure as a service (IaaS) in organizations in Bengaluru, India: a study with respect to organizational factors for mobile device management software products. Int. J. Appl. Bus. Econ. Res. **15**(Part-II), 497–509 (2017)
19. Van de Weerd, I., Mangula, I.S., Brinkkemper, S.: Adoption of software as a service in Indonesia: examining the influence of organizational factors. Inf. Manage. **53**(2016), 915–928 (2016)
20. Rehman, Md., Rajkumar: Buying behavior of organizations for software products: influence of environmental factors. Restaurant Bus. **118**, 252–271 (2019). https://doi.org/10.26643/rb.v118i10.9321
21. Mujinga, M., Chipangura, B.: Cloud computing concerns in developing economies. In: Proceedings of the 9th Australian Information Security Management Conference, Edith Cowan University, 5th–7th December, Perth Western Australia (2011)
22. Rogers, E.M.: The Diffusion of Innovations, 3rd edn. Free Press, New York (2003)
23. Bensley, R.J., Brookins-Fisher, J.: Community Health Education Methods: A Practical Guide, 2nd edn. Jones & Bartlett Learning, New York (2003)
24. Gabus, A., Fontela, E.: Perceptions of the world problematique: communication procedure. In: Communicating with Those Bearing Collective Responsibility (DEMATEL Report No. 1). BATTELLE Institute, Geneva Research Centre, Geneva, Switzerland (1973)
25. Gabus, A., Fontela, E.: World Problems, an Invitation to Further Thought Within the Framework of DEMATEL. BATTELLE Institute, Geneva Research Centre, Geneva, Switzerland (1972)
26. Wu W.-W., Lan L.W., Lee Y.-T.: Exploring decisive factors affecting an organization's SaaS adoption: a case study. Int. J. Infor. Manage. https://doi.org/10.1016/j.ijinfomgt.2011.02.007.
27. Deutsch, M.: Cooperation and trust: some theoretical notes. In: Jones, M.R. (ed.) Nebrasca Symposium on Motivation, pp. 275–318. University Press, Nebrasca (1962)
28. Wu, H.H., Chen, H.K., Shieh, J.I.: Evaluating performance criteria of employment service outreach program personnel by DEMATEL method. Expert Syst. Appl. **37**(7), 5219–5223 (2010)
29. Subashini, S., Kavitha, V.: A survey on security issues in service delivery models of cloud computing. J. Netw. Comput. Appl. (2010). https://doi.org/10.1016/j.jnca.2010.07.006
30. Linthicum, D.S.: Approaching SaaS integration with data integration best practices and technology. White Paper (2009)

31. Kandukuri, B.R., Paturi, V.R., Rakshit, A.: Cloud security issues. In: IEEE International Conference on Services Computing, Bangalore, 21–25 September 2009
32. Chen, D., Zhao, H.: Data security and privacy protection issues in cloud computing. In: 2012 International Conference on Computer Science and Electronics Engineering. IEEE (2012). https://doi.org/10.1109/ICCSEE.2012.193
33. Catteddu, D.: Cloud computing: benefits, risks and recommendations for information security. In: European Network and Information Security. Springer, Berlin (2009). https://doi.org/10.1007/978-3-642-16120-9_9
34. Jansen, W., Grance, T.: Guidelines on security and privacy in public cloud computing. NIST Special Publication 800 (144) (2011)
35. Yin, R.K.: Case Study Research: Design and Methods, 5th edn. Sage Publications, Thousand Oaks, CA (2014)
36. Eisenhardt, K.M.: Building theories from case study research. Acad. Manage. Rev. **14**(4), 532–550 (1989)
37. Rehman, M.H., Majumdar, S., Rajkumar, M.: Benefit and risk factors influencing organizations to migrate from on-premise to cloud computing model of software product, smart intelligent computing and applications. In: Third International Conference on Smart Computing and Informatics, vol. 2, pp. 185–202. Springer, Berlin (2019)
38. Rehman, M.H., Majumdar, S., Krishna, C.Y.S.: Assessment of factors impacting customer buying behaviour in software product companies. IUJ J. Manage. **5**, 21–25 (2017)
39. Rehman, M.H., Majumdar, S., Rajkumar, M.: Examining the influence of decisive factors on organizations: migration to on-premise software product from cloud computing. Test Eng. Manage. **83**, 24178–24196 (2020)
40. Pare, G.: Investigating information systems with positivist case research. Commun. Assoc. Inf. Syst. **13**(1) (Article 8) (2004)
41. Schulze-Bentrop, C.: Qualitative Comparative Analysis (QCA) and Configurational Thinking in Management Studies. PL Academic Research, Frankfurt (2013)
42. Rihoux, B., Ragin, C.C.: Configurational Comparative Methods: Qualitative Comparative Analysis (QCA) and Related Techniques. Sage Publications, Thousands Oaks, CA (2009)
43. Richards, L., Bazeley, P.: The Nvivo Qualitative Project Book. Sage Publications, London (2000)

Analysis of Encrypted ECG Signal in Steganography Using Wavelet Transforms

Prashant Patil and S. A. More

Abstract The proposed strategy gives the improved verification to secrete data communication through encoded data mask in ECG signals. The Implemented encryption strategy used to scramble the classified information into mixed-up structure and not just upgrades the wellbeing of anonymous carrier data by creating the data distant to some intruder with an random technique. After encryption of message data, the data hider will conceal the mystery data into the ECG signal coefficients. Despite the method, that encryption achieves certain security effects; it marks the mysterious messages indiscernible and unnatural or floating. This framework is yet improved with encode messages utilizing chaos crypto framework. Thereby arising the need for another security approach called reversible data hiding. It is the specialties of secreting the presence of information in another transmission medium to accomplish secrete data correspondence. It does not replace cryptography but instead supports the security utilizing its lack of definition highlights. Here the discrete wavelet transform is utilized to break down an ECG sign to various recurrence subbands. The data concealing method utilizes the LSB substitution algorithm for hiding the mystery message bits into the high recurrence coefficients. In the information extraction module, the secret information will be separated by utilizing pertinent key for selection of the important information. By utilizing the unscrambling keys, separated content information will be decoded from encryption to get the original data.

Keywords Chaotic encryption LSB substitution · ECG signals · Inverse transform · Performance measurements · Wavelet decomposition

1 Introduction

Evolution in Point of Care systems and wearable biomedical devices allow to record bio-physiological parameters and these can be communicated over internet along with required personal information straightening remote healthcare and reduce

P. Patil (✉) · S. A. More
Department of Electronics and Telecommunication Engineering, R. C. Patel Institute of Technology, Shirpur, Dhule, India

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_39

increasing patient traffic at hospitals. Here arises the need of protection to privacy and security from unauthenticated sources to patient data in transmission techniques. These days, cryptography strategies have shown up at a level of refinement with the ultimate objective that fittingly scrambled interchanges, which can be standard secured, as well as past the significant presence of the information which is to be communicated. In fact, it is all encompassing that the greatest estimations using multi kilobit key lengths could not be contained through computation power, whether or not all the figuring power worldwide for the accompanying 20 years was based on the attack form the attackers [1]. Obviously, the possibility exists that shortcomings could be found, or enlisting power forward jumps could occur, or registering power forward leaps could happen, yet for many applications, current crypto graphical procedures are commonly adequate (Fig. 1).

"Security through uncertain quality" is not generally a terrible thing, given that it is not the fundamental security instrument used. Steganography grants us to conceal scram-drained messages in mediums more opposed to stick out. The essential information in the photographs is as yet mixed, which attracts far less thought being flowed in the picture then it would something different. The technique is contrasted and past strategies as applied with virtual and unfortunate limits including mean square mistake just as pinnacle sign to commotion extent will be resolved for execution evaluation.

LSB [2, 3] is one of the most punctual and easiest techniques for concealing data in sound signals. It is the regularly utilized procedure for sound steganography. In LSB encoding, the least enormous pieces of the spread media/unique sound is modified to incorporate the secrete data. Despite the fact that this is a basic technique, an assailant can undoubtedly remove the secrete data from the stegano based object.

The process of Parity coding [4, 5] works on a gathering of tests rather than singular patterns. Here individual examples are gathered and parity of each cluster is determined. For inserting message bit individually, check the parity piece-bit of a group of tests. It does nothing if the event that the parity bit and message bit matches. In any case, change the LSB's of any of the individual samples in that cluster to make the parity bit equivalent to the message bit.

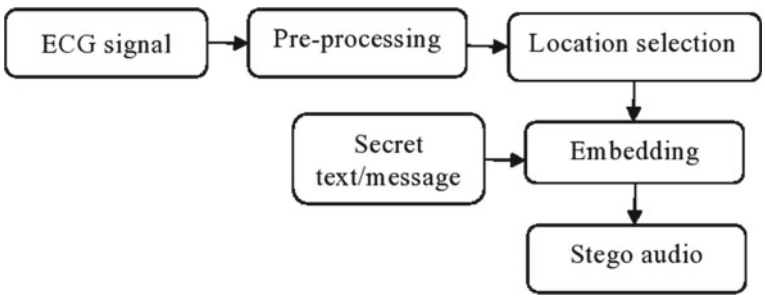


Fig. 1 Tentative model

In echo covering up [6] technique, information is fixed in the echo part of the host sound signal. The echo is a reverberation added to the host signal and subsequently the issue with the added substance noise is kept away from here. While utilizing echo concealing, three boundaries are to be thought of: they are same beginning amplitudes, counterbalanced (offset delay), and rot rate (decay), with the goal that echo is not perceptible. The fundamental burden of this strategy is tolerant discovery and low identification proportion. Because of its low inserting rate and low security no investigates are going on echo concealing procedure.

2 Proposed Method

At first, the message needs to scrambled with some ordinary encryption scheme and ECG signal has decomposed using wavelet transform with multiple levels which can improve the security of secret data. Embedding of encrypted data to ECG decomposed signal is archived by Adaptive LSB technique, same inverse algorithm has implemented at Receiver side to extract the data. To identify the efficiency of our system we have calculated PSNR, MSE, and Correlation Coefficient.

2.1 Algorithm for Encoding

Input: ECG signal is used as medium carrier and the encrypted secret message to hide.

Output: Stego ECG signal.

The stages are as per the following:

Stage1: The emit message must be scrambled utilizing chaotic encryption. Consider the binary-twofold of the code text of the secrete characters to be covered up. On the off chance that the secrete message is in text, at that point convert it into the individual ASCII [7] esteem and after that it will be changed over into binary-twofold shape of pattern.

Stage2: Peruse a secrete bit from the arrangement to stow away.

Stage3: Apply 5level wavelet decomposition on Input ECG signal.

Stage4: For each Wavelet decomposition.

- Combine all high-frequency component of Wavelet decomposition.
- Insert a secret bit into high-frequency component of ECG signal.
- Rehash the steps until all the secrete bit esteems are supplanted.

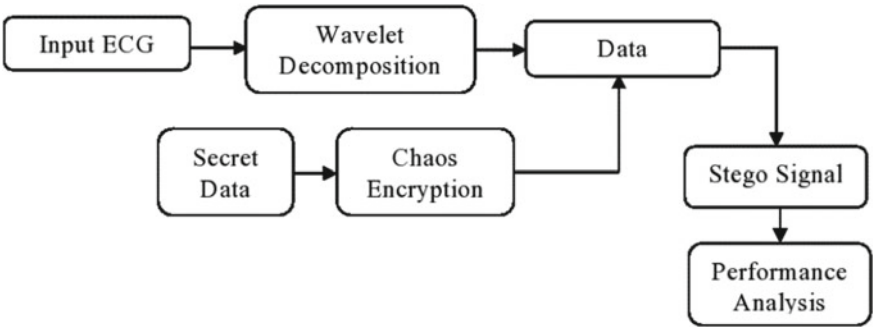
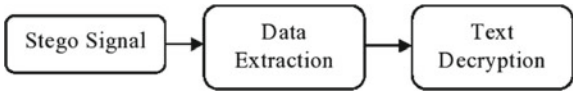


Fig. 2 Embedding process

Fig. 3 Extraction process



2.2 Embedding

See Figs. 2 and 3.

2.3 Chaotic Encryption Scheme

The encryption strategy of broad chaos is the least complex procedure to scramble video information or message by chaotic condition. This technique can encourage to find some basic data and set up the significant phase of security. The benefit of chaotic encryption is its level of security. The encryption is accomplished by cycles if iteration. No alternate routes are accessible with this methodology. Though the necessity of enormous code stockpiling and delayed in speed are viewed as the significant detriments. The properties of disorder are somewhat creating a few changes in the whole cryptography (Fig. 4).

The transitivity and the sensitivity are the main assets for the proposed methodology. In an underlying condition, disorderly is consistently sensitive. Thus it will deliver a slight contrast in direction for all values. Indistinguishable direction just can create similar qualities. The geography transitivity characterizes that the state focuses bounced in a limited space state.

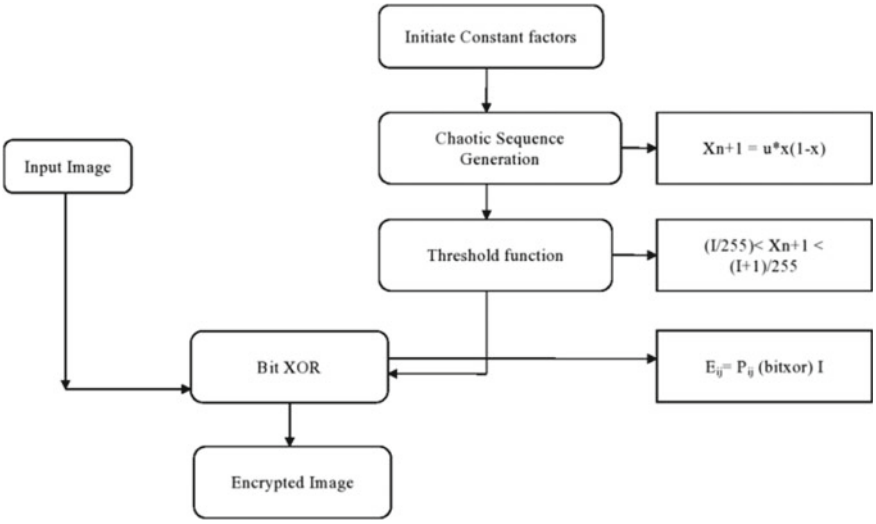


Fig. 4 Secret data encryption process

2.4 Discrete Wavelet Transform

In this methodology, wavelet transform is utilized. The wavelet investigation is likely the latest resolution for defeat the deficiencies of the Fourier transform. In wavelet investigation the utilization of a completely window adjusted window takes care of the signal cut issues. The window is moved along the signal and for each position the range is determined. Window is moved along the sign and for each position, the range is determined. Wavelet transform breaks down a signal into coefficients indicating to recurrence segment of sign at specific moment of time. Numerous applications utilize discrete signals accordingly DWT is correct decision to utilize. Channel groups are utilized to sum up conduct of consistent wavelet change and DWT decay can be accomplished by applying wavelet change to the sign utilizing band channels. Channel filters are utilized to deducting the computational efforts of the constant wavelet transform. The signal is decayed with a high pass filter and a low pass filter. The coefficients of these filters are processed for utilizing numerical investigation and made accessible (Fig. 5).

Where, LPd denotes Decomposition of Low Pass Filter, HPd denotes Decomposition of High Pass Filter, LPr denotes Reconstruction of Low Pass Filter and HPr denotes Reconstruction High Pass Filter.

In wavelet, breaking down of a picture, the disintegration is done line by line and afterward segment by segment. For example, here is the technique for a $N \times M$ picture. Filtering each line and afterward down-example to get two $N \times (M/2)$ pictures. At that point of filter, every section and subsample the filter yield to acquire four $(N/2) \times (M/2)$ picture, the one got by low-pass separating the lines and segments are alluded to as the LL picture.

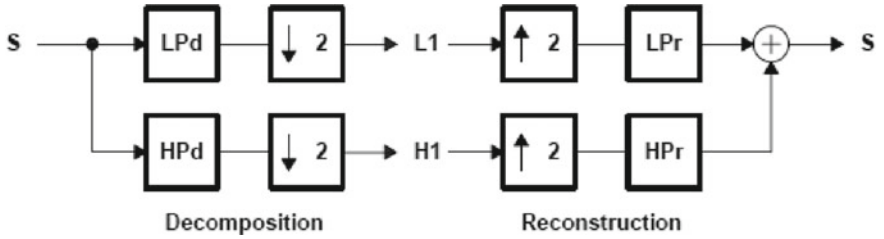


Fig. 5 Discrete wavelet transform process

The one acquired by low-pass filter the lines and high-pass separating the segments is indicated to as the LH pictures. The one acquired by high-pass separating the lines and low-pass sifting the segments is known as the HL picture. The subpicture got by high-pass sifting the lines and segments is indicated to as the HH picture. Every one of the sub pictures got in this design would then be able to be separated and subtested to acquire four more subpictures. This cycle can be proceeded until the ideal subband structure is acquired.

3 The Superiority Measures for Picture and Results

The Quality of the reconstructed picture is estimated regarding mean square error (MSE) as given in (1) and peak signal to noise (PSNR) proportion. The MSE is regularly called remaking blunder difference σq^2. The MSE between the first picture f and the recreated picture g at decoder is characterized as in (2)

$$\text{Mean Square Error}\,\sigma q^2 = \frac{1}{2}\sum\nolimits_{l,m}(f[l,m] - g[l,m]) \tag{1}$$

where l and m means the entirety over all pixels in the picture and N is the quantity of pixels in each picture. From that the peak signal-to-commotion proportion is characterized as the proportion between signal difference and remaking blunder fluctuation. The PSNR between two pictures having 8 pieces for every pixel regarding decibels (dBs).

$$\text{Peak Signal-To-Noise Proportion} = 10\log_{10}\left(\frac{255^2}{\text{MSE}}\right) \tag{2}$$

For the most part when PSNR is 40 dB or more prominent, at that point the first and the reproduced pictures are for all intents and purposes indistinct by natural eyes. Relationship coefficient is utilized to discover the comparability between two distinct pictures with their powers. It will be portrayed by (3),

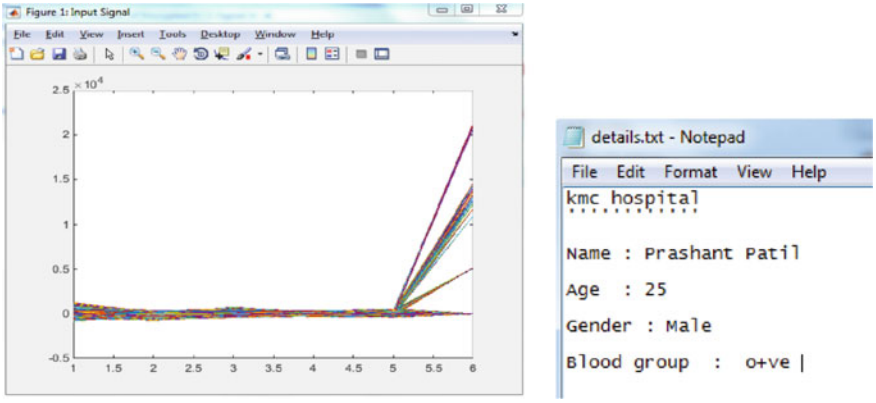


Fig. 6 Input ECG signal and secrete message data

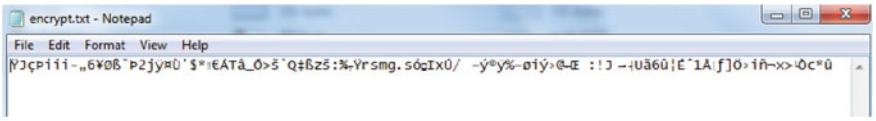


Fig. 7 Encrypted data

$$\textbf{Rel\_coef} = \sum[\sum(u1 * u2)/\sqrt{\sum u1^2 * \sum u2^2}] \tag{3}$$

Following validation has been succeeded with the proposed method including Root Mean Square Error: 80.382391, Peak Signal to Noise ratio (dB): 43.065404 Correlation Coefficient: 0.944545 (Figs. 6, 7, 8 and 9).

4 Conclusion

In the executed strategy, the message to be secretive will be inserted at particular situations inside the ECG carter additionally we had utilized encryption strategy to expand the security (the specific situations to be created by wavelet deterioration measure), it very well may be considered as better and effective technique for concealing the information. This executed framework will not change the size of the record even in the wake of inserting and furthermore appropriate for an ECG document design. Likewise, the encoding and deciphering strategies are like be executed. Along with MSE, PSNR between input signal and reconstructed signal, correlation coefficient is calculated and shown better results in signal decomposition and reconstruction. In spite of the fact that it is an all-around manufactured framework, it has been restricted

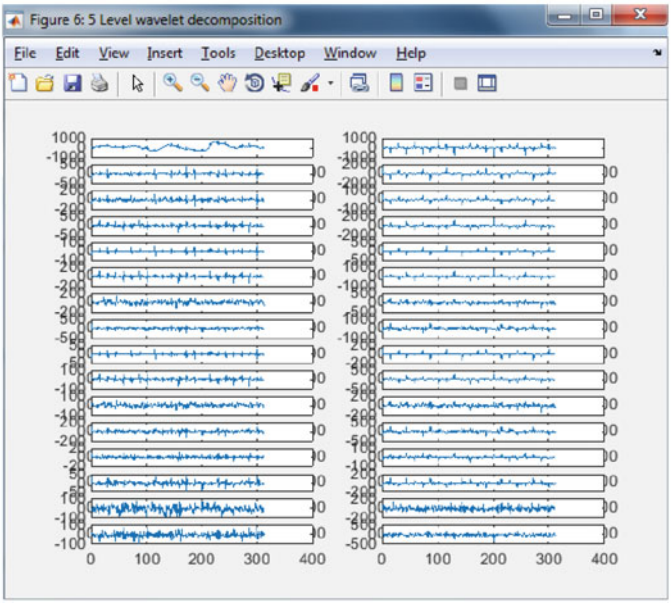


Fig. 8 5 level wavelet decomposition

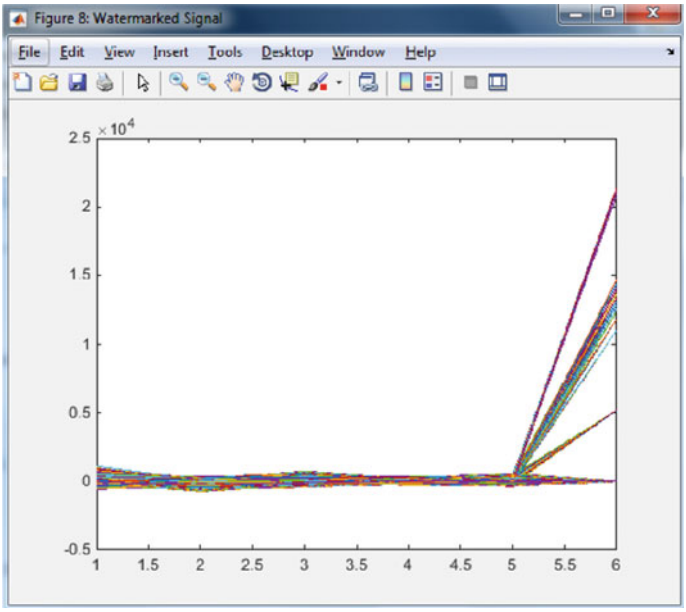


Fig. 9 Setgo ECG signal

to certain limitations. Performance and data accuracy have been improved in discrete wavelet transform while comparing to some of other transforms.

References

1. Yahya, Z., Hassan, M.: Probabilistic analysis of targeted attacks using transform-domain adversarial examples. IEEE Access **8**, 33855–33869 (2020)
2. Anwar, M., Sarosa, M., Rohadi, E.: Audio steganography using lifting wavelet transform and dynamic key. In: International Conference of Artificial Intelligence and Information Technology (ICAIIT), Yogyakarta, Indonesia, pp. 133–137 (2019)
3. Simmons, G.: The prisoners' problem and the subliminal channel. In: Advances in Cryptology (CRYPTO '83), pp. 51–67, California (1985)
4. Wang, H., Peng, D.: Resource-aware secure ECG healthcare monitoring through body sensor networks. IEEE Wirel. Commun. **17**(1), 12–19 (2010)
5. Kaur, N., Kaur, U.: An audio watermarking using Arnold transformation with discrete wavelet transform (DWT) and discrete cosine transform (DCT). Int. J. Comput. Sci. Eng. **2**(06), 286–294 (2013)
6. Dey, N., Mukhopadhyay, S., Das, A., Chaudhuri, S.S.: Analysis of P-QRS-T components modified by blind watermarking technique within the ECG signal for authentication in wireless telecardiology using DWT. Int. J. Image Graphics Signal Process. **4**(7) (2012)
7. Cvejic, N., Seppänen, T.: Reduced distortion bit-modification for LSB audio steganography. In: ICSP Proceedings. IEEE (2004)

The Canny Blocks for Generation of Secured Epass in Pandemic COVID-19

R. Durga and E. Poovammal

Abstract The SARS COVID-19 has immensely affected human life across the world. The transport which completely blocked was released by issuing epass to the travelers with valid emergency. The purpose of issuing the epass is not only to make the restriction of frequent travel but also to control the traffic and to save the livelihood of people in local zones. There are distinct challenges in epass generation and issuance. The existing system uses the traditional cloud database for storing the user-sensitive data who applied for epass. It may result in data leakage and security threats due to centralized server. The unique technology which meets the current requirement of highly secured sensitive data available to all authorized person is Blockchain Technology. Blockchain is expounded as distributed decentralized ledger, maintains the non-modifiable ledger at all nodes in the network. In this paper, we proposed blockchain technology-based framework that can be deployed for generating and issuing the epass to improve the availability of sensitive data to the authorized person for effective controlled movement between the country. This decentralized distributed blockchain-based epass maintains the history of the user applied for the epass properly in the non-modifiable ledger properly.

Keywords Blockchain technology · Cloud database · Consensus agreement · Decentralized ledger · Distributed data · Mining · Privacy · Proof of work · Traffic · Threats

R. Durga (✉) · E. Poovammal
Department of Computer Science and Engineering, SRM Institute of Science and Technology, Kattankulathur, Tamil Nadu 603203, India
e-mail: dr8529@srmist.edu.in

E. Poovammal
e-mail: poovamme@srmist.edu.in

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_40

1 Introduction

The COVID-19 pandemic in India is the wedge of worldwide coronavirus disease. In January 2020, the very first person affected by COVID-19 in India [1]. The Prime Minister Narendra Modi announced fourteen hours voluntary curfew on 22 March, followed by national wide lockdown till May 31, 2020. The government announced unlocking from June first week except on the containment zones. During the unlocking phase, all state government refused interstate traffic without epass. There are many online applications adopted by various state government to progress the epass as per the guidelines of Union ministry [1]. The purpose of issuing the epass is not only to make the restriction of frequent travel but also to control the traffic and to save the livelihood of people in local zones. The epass applied through the portal developed by National Informatics Center under the ministry of Electronics of Information Technology. It is used by 14 states of India to manage the epass movement request for emergency, marriage, attending funeral. The COVID-19 virus destroys the roots of essential services by means of travel restrictions due to unrivalled transmission of virus. It rises hordes of challenges to the human civilization. These challenges are met by all countries need to be resolved soon by technological innovation. There are many challenges faced by government in all countries are Social Distancing, essential government services on transport, Data Sharing, Education, Supply Chain and Logistics [2].

The existing epass generation technique is consisting of traditional database storage. The user needs to register the mobile number and personal id proof like Aadhaar, passport, vehicle ID for registration phase. The user data are stored in the cloud database. Once the user completes the registration form, it sends the one-time password (OTP) to the registered user mobile number. Then the next step is submitting the OTP on the form depicted on the screen. The user will see the login page to apply for epass. Login details includes sensitive data like profile details such as user image, name, gender, email ID, photo proof ID like PAN, Aadhaar, voter ID, vehicle details, place of visit, purpose. The registered user after logging in can fill the application. Once the application submitted successfully, the confirmation message is viewed by the user on the screen. Then the submitted request is moved to the corresponding district authority for the approval. Once district-level authority approved, the user received the message through the registered mobile number. Then the epass can be downloaded from the link given in the message.

The traditional database which stores all sensitive data are in the risk mode since the attackers can hack the centralized database. The storage of data can be incurred with different attacks, latency problems, data leakage and privacy-preserving issues [3]. The single point of failure and Unauthorized third-party access control, power consumption are the major issues with the traditional cloud-based database system.

The European Parliamentary Research [4] proposed ten identical technologies for the challenges raised by COVID-19 such as Blockchain Technology, Robotics and Unmanned Aerial Vehicles, 3D Printing, 5G Communication, Cloud and Edge Computing, Artificial Intelligence and Machine Learning, Internet of Things. Don

Tapscott [5], famous author, and chief executive in Canadian business confined in role of advanced technology and in social cells, pronounced that "Blockchain is the revolutionary second generation of Interconnected networks" [6].

Blockchain is an eccentric technology among pop ups in 2008 as non-modifiable decentralized application with disintermediation between the promoter and the end user. The blockchain had some reasons to reach the various sectors beyond the boundaries of cryptocurrencies. In prevailing, there are essential features such as personalized transparency, disintermediation, non-modifiable ledgers, consensus mechanism, and cryptographical lock version.

- Personalized Transparency: The ability to maintain the information, which is visible to all members connected in the sequence.
- Disintermediation: There is no intercessor between the transceiver and receiver.
- Non-Modifiable Ledger: The impotence to alter or remove the information from the block.
- Consensus-based Agreement Mechanism: The technique is employed to distribute the agreement among the nodes to validate the information based on Proof of Work (PoW), Proof of Stake (PoS), Proof of Authority (PoA), Practical Byzantine Fault Tolerance (PBFT).
- Cryptographical Lock Version: Each block holds an identical cryptographical key to lock the information is verified and validated through consensus mechanism.

By combining these characteristics such as the blockchain is able to do traceability, decentralization, smart contracts. The advantages are highly increased collaboration, increment in sharing of promised information, effective efficiency, low cost, reduced risk and innovative models.

The paper is organized as Sect. 2 describes the art of blockchain technology and the methodology used to develop the framework, Sect. 3 drives on the literature review, Sect. 4 describes how epass is generated based on blockchain technology, and Sect. 5 concludes with the key features.

2 The Art of Blockchain Technology

Every precinct of economic, social, and technical spheres of research could transform over on nearby future due to emerging second-generation called blockchain. The blockchain consists of nodes initiates the process of transaction by using its private key infrastructure employed with digital signature method. The transaction can be validated by the peers based on preset mining conditions. The peers will verify and validate the transaction by verifying that the initiator node had significant bitcoins for transaction and also check with the past entries update the data in the wallet.

Let us consider a typical scenario of a blockchain with five nodes namely B1, B2, B3, B4 and B5 in which B1 is trying to transfer an amount of Rs. 500 to B3. B1 initiates the transaction of transferring Rs. 500 to B3. Any node out of B1 to B5 performs the preset mining task. The node which completes the mining task

successfully will verify whether B1 has enough amount to initiate the transaction and also check the prehistory of the B1 transaction. After the successful verification, the transaction is validated and the amount is transferred from B1 to B3.

The verified and validated transaction are stored in the blocks. Miners solve the computational puzzle or preset mining condition to publish the block added to the network. Miners are credited with the small amount of incentives. Every user can act as a miner in the network to verify and validate the user using the consensus mechanism. The distributed ledger maintains the copy of the added block in all peer ledger. The newly added block linked by using a hash function. Miners use different types of consensus mechanism to validate the block such as Proof of Work (PoW), Proof of Stake (PoS), Practical Byzantine Fault Tolerance (PBFT).

- Proof of Work is a proof-based consensus mechanism basically works on the concept of consensus method to determine that particular node will have the right to add the new block to the existing blockchain. The advantage of proof of work clears the confusion that every user in the network will broad cast their nodes for verified transaction, thereby leads to the duplicated transaction in the ledger.
- Proof of Stake is an energy-efficient algorithm than proof of work. Instead of using more computational resources for solving the puzzle, it relies on the process of creating the block in the network. The validator is selected based on the stake, so the validator with wealthiest stake can get more chances for validating the blocks.
- Practical Byzantine Fault Tolerance is a high-performance consensus mechanism used for communicating safely between the nodes in the network. In PBFT the nodes are queued in sequence manner based on the voting procedure, the leader node was selected. The remaining nodes in the network are acting as backups. The leader node receives the request and forwarded it to the backups to process the request.

Blockchain is generally categorized into three kinds. They are (i) Public blockchain, (ii) Private blockchain, (iii) Hybrid blockchain.

As the name implies, in the public blockchain the data are visible to the public. Private blockchain involved with the particular organization, where every member known about others or their personal identity is validated or to meet the certain criteria to be the part of the member.

In public blockchain, the control can take over by everyone in the network but in the private blockchain the control is spread among the some of the members and no data are visible to the public. Blockchain can be designed to match the requirement of the user story. The epass generator system uses the public blockchain.

The healthcare collaborators need the COVID-19 patient data to be shared during this pandemic period [7]. The COVID-19 patient datasets play the vital role in the research of COVID-19. Since the patient data had access control policies that impulses from the national and international bodies like HIPAA [4]. The patient data should be secured is the major concern. The disintermediation storage and non-modifiable ledger help in improve the privacy of the patient data [2]. Like traditional system, all data are in the form of files but blockchain enables the data sharing in real time.

The information regarding the heart rates, oxygen levels, blood pressure can be deployed through Internet of Things (IOT) devices. Medical IoT [8] devices help to integrate all the data regarding COVID patient directly into the blockchain excludes the forgery and mutation of the data. The smart contract in the blockchain technology makes the process easier across the Medicare providers in the various countries around the world. The personalized transparency will integrate all patient data using medical IoT devices can store and share the data will create the trust between patient and stake holders.

3 Related Work

In the year 2017, $3.7 bn of funding on startups those using the blockchain in the transport sectors. In 2018, it is efficiently increased with doubly $6.5 bn and $130 bn at the April end. The department of UK transport in 2017 discusses about the transportation by passengers in England during the high traffic is 14 million [6]. The blockchain needs 20,000 per second. The University of Sydney's Red belly blockchain generates 400k transaction per second comparatively Bigchain DB uses PoS is computationally less expensive [9].

Gartner prognosticated that the overall global business value per annum for the blockchain will target as $176 billion by 2025 and increases doubly by $3.1 billion in transport industry supports micro payment instinctively recorded on the blockchain for every transaction [10]. The disintermediation excludes the third party to superintend the data. The user data should be protected and stored in the decentralized-based server to improve the security.

Li et al. identified the blockchain framework for encrypting the user data by inherit the attribute-based encryption technique [11]. The cloud storage which had duplication of data there by increase the storage space and cost in cloud server. Zhang et al. proposed blockchain-based hashing technique reduces the duplication of data [12]. Dagher et al. identified the hash values for the blocks which is resilient to DoS attacks happen in the blockchain [13]. Lin et al. [14] proposed a smartphone-based ticketing system inherits the property of smart contract and multisignature in blockchain technology. The enhanced multi signature process corroborates the authenticity and security of buying the tickets. He concludes that blockchain-based ticketing system implores at very nominal cost. CATAPULT is the transport system that explores blockchain that can be used in the future transport sector. Many use cases are discussed about implications of cybersecurity in the blockchain technology [15].

There are some drawbacks associated with the traditional cloud database system that can be resolved by means of cryptographically lock version of blockchain technology [7], there by excludes the centralization concept and can maintain the user privacy.

4 Blockchain-Based Epass Generation

The proposed blockchain-based framework for generation of epass is developed with three major objectives on user data stored in the block. They are:

- Highly secured transmission of user data.
- Prevention of data leakage.
- Reduction in latency compared to traditional method.

The user-sensitive data is secured and it is available to all the authorized nodes in the blockchain. The data leakage is prevented by generating the secured hashes. Hence the security and verification features of every transaction are achieved by implementing consensus algorithm.

In traditional method, the epass is generated after the approval of the two authorities of source and destination. So, the time lapse is unpredictable which may or may not be reliable. Since epass approval process relies on working style of the induvial, the automated, decentralized epass generation in blockchain enhances the process of epass issuance. The preset mining task can be light weighted, so that the generation of blockchain epass outperforms the traditional epass.

The proposed architecture consists of four phases such as Registration phase, Generation of hashes, Mining phase, Epass Generation phase.

4.1 Registration Phase

The user needs to register their personal details such as Aadhaar number, mobile number, gender, age, address, and results of COVID test. Once the details are filled, the mining of data into the block is initiated. In traditional epass generation, there is no entry of the COVID-19 test report results. Each and every user applied for epass is registered in the non-modifiable ledger in the sequence.

4.2 Generation of Hashes

The solidarity of the user data is attained by the hash value of the block. Each block attached to another block like chain structure. The hash value connects the chain with one another. The hash value of the block is distinctive of the particular block. The cryptographical lock version helps to accept the input of any size and gives the desired hash as the output.

4.3 Mining

The miner uses the cryptographical lock of the hash value to mine the data directly inside the block. All the blocks are mined with the timestamp. So, the blocks are arranged in sequence based on chronological order. The user can benefit, from this nontampering proof of blocks. The user who applied first will gain the preference. No forgery can be done by the authorities who involved in the epass generation.

The mining of the user data along with the COVID test is maintained in the record of non-modifiable ledgers. The intruder cannot alter the information. If any altered, then the hash generated becomes invalid. The mining of the user data done inside the block in one second into the blockchain successfully in our proposed system. The mining time is calculated based on the CPU time.

4.4 Generation of Epass

The automate process of the blockchain technology helps in the generation of epass easier. The rules used in the work is, if particular users COVID-19 test results in positive, then the epass will not generated. The agreement was verified with the user data that he had COVID-19 test report was negative. Then the epass was generated using the decentralized-based blockchain technology.

5 Results and Discussion

The contributed work is experimented on Spyder (Anaconda 3), 64-bit processor, and implemented in python. The epass generation registration filled form is depicted in Fig. 1. The epass registration form consists of Aadhaar card, phone number, gender, age, address and COVID-19 test results. All the fields are mandatory. The user should verify all the details before submission of the registration form. Once the user submitted the details, fields are validated in the respective column. After submission of the mandatory data. the request is sent to mine the user data inside the block. The mining phase is initiated. The blockchain verifies the user is the authenticated user by means of consensus mechanism. It validates the address that no person can register more than once. Then the address was saved as smart contract to avoid multiple addresses. The smart contract uses the agreement was verified with the user data that he had COVID-19 test report was negative. If the COVID test result was negative, the user data was mined successfully is depicted in Fig. 2. The epass generated message was displayed on the user screen is depicted in Fig. 3.

After the successful mining of data is done. The user sensitive data was mined in the blockchain is timestamped. The blocks are mined sequentially in timestamp order. The mining time of the user data in to the block done in one second.

Fig. 1 Data filled in the registration form

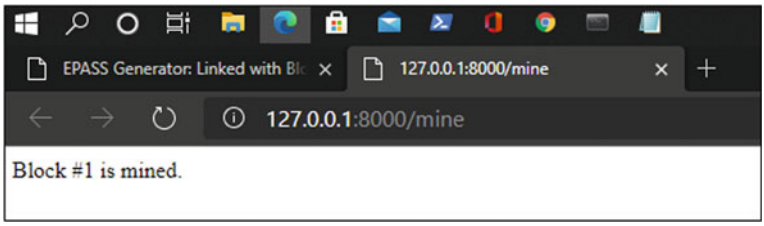


Fig. 2 Mining of block

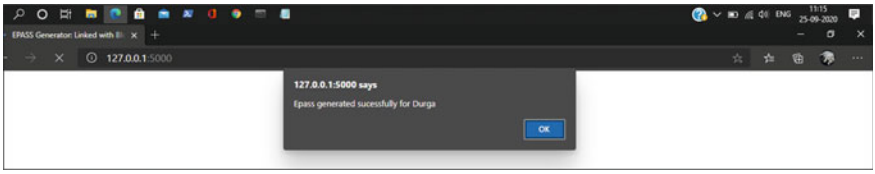


Fig. 3 Notification of epass generation

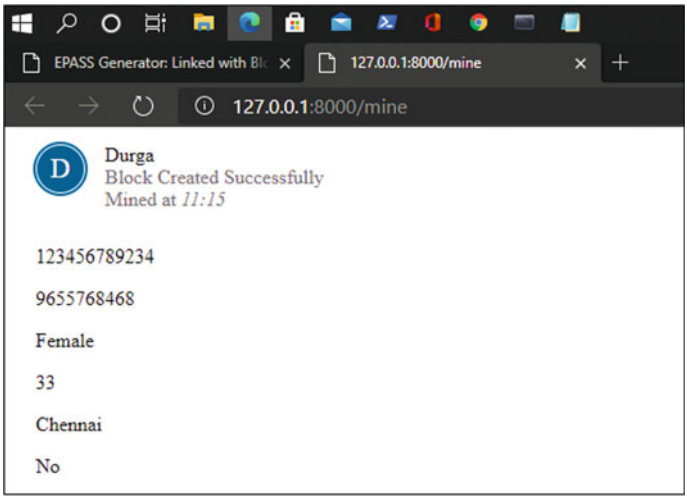


Fig. 4 Generation of epass

It avoids the forgery done by the third party. Thus, the epass was generated successfully for the user is depicted in Fig. 4.

6 Conclusion

The proposed blockchain-based framework specially with decentralized nature provides the generation of epass in pandemic period provides authenticated epass, highly secured of data with non-modifiable ledger, prevents the data leakage, low latency in storing and mining the data. The automate process in blockchain eliminates the third party by prosecution of smart contract to generate the epass. All the data are depicted on the non-modifiable ledger are traceable, thereby privacy of the data are preserved and enhances the trust with the users.

References

1. https://en.wikipedia.org/wiki/Coronavirus_disease_2019
2. Yang, K.: Unprecedented challenges, familiar paradoxes: covid-19 and governance in a new normal state of risks. Public Adm. Rev. **80**(4), 657–664
3. Jose, L., Malathy, C., Poovammmal, E.: A survey on privacy preservation of healthcare data using cryptographic techniques. Int. J. Pharm. Technol. **8**(4), 22322–22329 (2016)
4. Kritikos, M.: Ten technologies to fight coronavirus. European Parliamentary Research Service (ERPS). Accessed on 04.06.2020 [Online]. Available: https://www.europarl.europa.eu/RegData/etudes/IDAN/2020/641543/EPRSIDA(2020)641543EN.pdf

5. Tapscott, D., Tapscott, A.: The Blockchain Corridor: Building an Innovation Economy in the 2nd Era of the Internet (2017)
6. https://www.gov.uk/government/statistical-data-sets/nts05-trips
7. Bansal, A., Garg, C., Padappayil, R.P.: Optimizing the implementation of COVID-19 immunity certificates using blockchain. J. Med. Syst. **44**(9), 140 (2020)
8. Yeole, A.S., Kalbande, D.: Use of internet of things (iot) in healthcare: a survey. In: Proceedings of the ACM Symposium on Women in Research 2016, pp. 71–76 (2016)
9. https://redbellyblockchain.io/
10. https://www.gartner.com/imagesrv/media-products/pdf/Talari-Networks/talari-networks-1-4A5JDFW.pdf
11. Li, M., Yu, S., Zheng, Y., Ren, K., Lou, W.: Scalable and secure sharing of personal health records in cloud computing using attribute-based encryption. IEEE Trans. Parallel Distrib. Syst. **24**(1), 131–143 (2013)
12. Zhang, Y., Zheng, D., Deng., R.H.: Security and privacy in smart health: efficient policy-hiding attribute-based access control. IEEE Internet Things **5**(3), 2130–2145 (2018)
13. Dagher, G.G., Mohler, J., Milojkovic, M., Marella, P.B.: Ancile: privacy-preserving framework for access control and interoperability of electronic health records using blockchain technology. Sustain. Cities Soc. **39**, 283–297 (2018)
14. Lin, K., Chang, Y., Wei, Z., Shen, C., Chang, M.: A smart contract-based mobile ticketing system with multi-signature and blockchain. In: 2019 IEEE 8th Global Conference on Consumer Electronics (GCCE), Osaka, Japan, pp. 231–232 (2019)
15. The Case for Government to Incentivise Data Sharing in the Intelligent Mobility Sector. Transport Systems Catapult (2017)

Behavioral Changes in Strength of Concrete Reinforced by Waste Plastic Fibers

Nirav M. Patel and M. N. Patel

Abstract The quality of environment has been deteriorated in last few centuries, with increase in intervention by humans. Because of its low production cost, durability, and availability, plastic has become one of the highly preferred man-made material. This paper portrays experimental outcomes of strength the performance of concrete toughened with waste plastic fibers. Polypropylene terephthalate (PET) waste bottles used for the experiments. PET waste relatively unfit for the reuse and also its recycling process is energy consuming, thus, it contributes littering and growing leftover which cause environmental menaces. In this research, by accessing compressive, slump, flexural and splitting tensile strength, it help us to find out the constructiveness of PET waste fibers as bracing element for concrete. PET fibers shredded into three different length of 10 mm, 15 mm and 20 mm and mixed in concrete by its magnitude ranging between 0 and 2%. The investigation outcomes represented strength performance of concrete due to the introduction of PET fibers. The flexural and splitting tensile strength was increased with usage of PET fibers up to 1.5% by volume of concrete, whereas the compressive strength showed slight decrement with increase of PET fibers from 1% onwards. The presence of waste fibers makes structure durable and lightweight.

Keywords Green concrete · Sustainable concrete · Plastic waste PET fibers · Compressive strength · Splitting tensile strength · Crack resistance

N. M. Patel (✉)
Civil Engineering Department, PIET, Parul University, Vadodara, Gujarat, India
e-mail: nirav.patel8986@paruluniversity.ac.in

M. N. Patel
Parul University, Vadodara, Gujarat, India
e-mail: provost@paruluniversity.ac.in

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_41

1 Introduction

The plastic derives from the Greek word "plastikos," meaning "being capable to shaped or molded" [1]. Plastic has low density, lightweight with good strength, long life, user-friendly designs, fabrication capabilities, and low cost are the major factors behind the extreme high growth of plastic products, due to these major benefits plastics have become an integral and inseparable part of a human lives [2]. Being the world's second-largest populated country, India brings about 5.6 million tons of plastic debris per annum and India's per capita consumption is 11 kg, according to Federation of Indian Chambers of Commerce and Industry (FICCI) in 2017. As per the Indian officials, plastics make about eight percentage of total solid waste in India. Figure 1 shows the plastic waste produced and its % contribution in solid waste for the major Indian cities.

At the present Indian, construction industries consuming about 328 million metric tons of cement annually and it is anticipated to touch about 580 million metric tons per

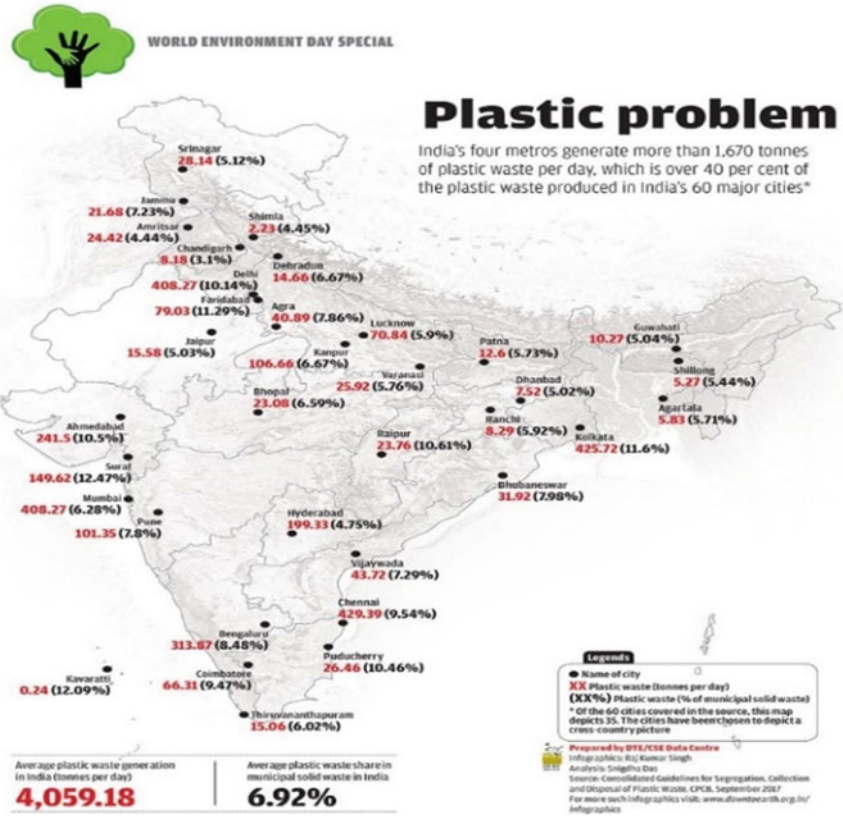


Fig. 1 Plastic waste generated by major Indian cities [3]

annum in 2022. Concrete is made up of natural materials and the re-establishment of natural assets beyond the control of humanity. As the requirement of concrete increases, serious question lifted up the rapidly dissipated precious natural assets [4]. So now it becomes essential to use various alternative materials other than natural along with traditional materials to reduce the rapid and immense usage of natural origins. For counteracting these issues, a new concept was coined—green concrete. It is equipped by using various industrial waste product or waste materials along with conventional natural materials. Number of various such wastes are already added and tested for different inspections and their influences on various aspect of concrete characteristics [5].

Concrete has certain limitation to developing the microcracks and insufficiency to water resistance and chemical permeability. As this microcracks starts to develop in concrete, which leads to concrete distress by chemical attack and impact resistance. Using of waste plastic directly within the need of concrete industry can be considered as a novel approach [6]. The most commonly found material in plastic water-soda bottle and similar packaging is polyethylene terephthalate (PET). It is generally preferred to use and throw PET material as the recycling may cause an adverse effect on the health of the consumer.

1.1 Related Study

By inculcating polyethylene terephthalate (PET) fibers as reinforcement for concrete to improve its tensile strength and compressive strength [7]. A greater compressive and tensile strength limit the occurrence of macro or micro shrinkage cracks that ultimately results in visible cracks. High percentage of fibers in spite of improving concrete behavior cannot be recommended as the concrete become less workable. This happens due to the circular fibers and their lesser adherence with concrete.

The decomposition rate of PET is slower compared to conventional reinforced concrete beams. The suggested procedure enhances durability of concrete elements by applying concept of sustainability [8]. It is best when the waste plastic is recycled by using other materials to evolve products with improved characteristics. Plastic such as PET is used pretty much everywhere in a use and throw concept. The no corrosive nature of waste plastic results in more amount of solid waste accumulation than the actual recycling.

1.2 Research Significance

The outcome of the experimental investigations is a comprehensive effort to answer the questions given below related to usage of the PET in concrete.

- What should be ideal amount of PET fibers in concrete?

- The amount of the changes in workability and strength by doing modification in the concrete by adding PET?
- How the standard material will behave in process of axial compression and flexure loading by PET to concrete?

The experiment results in particular tensile capacity, compressive and flexural strength, and workability performance on concrete with PET fibers have been successfully able to answer the above questions.

2 Materials and Their Mix Proportions

2.1 Concrete Mix

Cement is used for the experiment work is grade 53—ordinary Portland cement (OPC), produced by UltraTech cement organization from Gujarat possessing specific gravity of 3.15 g/cm^3. Coarse aggregate of sizes 20 mm and 10 mm locally available near Vadodara city was used. Sand from Aursang River near Vadodara city used for the test performance. The fineness modulus of the sand and coarse aggregate was 2.95 and 4.78, respectively. The specific gravity value for FA and CA was 2.62 and 2.70 having 0.78% and 0.50% water absorption, respectively.

2.2 PET Plastic Waste Fiber

PET is a part of polyester family which is basically a thermoplastic lacquer collected of phthalates. It has low tensile strength in compared to common reinforcement steel but clearly it excessive than the very moderate tensile capacity of the concrete. PET fibers were acquired by cutting wastewater and soda bottles as shown in Fig. 2. The experiments were carried out with varying length of fibers 10 mm, 15 mm and 20 mm at constant width of 2 mm and thickness of 0.25 mm and denoted by Type P1, Type P2 and Type P3. The fibers were included in concrete in differing fractions ranging from 0 to 2% by mix volume. 0% of PET fibers was considered as a reference concrete. PET has tensile strength of 55–75 MPa. The elongation at break is on a higher side up to 150%. Modulus of elasticity ranging between 2800 and 3100 MPa. Density of the PET about 1.37 g/cm^3 with water absorption of merely 0.1 [9].

2.3 Specimens and Mix Proportions

Total 15 numbers of batches for concrete were made of M25 grade concrete with fixed 0.47 W/C ratio. The particular mix proportions and batch details are mentioned

Fig. 2 Waste plastic bottles and PET fibers [10]

Table 1 Mix proportion of concrete constituents (per m^3)

Cement (in kg)	Coarse aggregates (in kg)	Fine aggregates (in kg)	Sand (in kg)	W/C ratio	Water (L)
336	888	225	834	0.47	158

in Tables 1 and 2, respectively. For each batch was prepared three specimens to find average value of compression and splitting tensile capacity and one specimen for

Table 2 Batch allocations of concrete mixes

PET type	Batch	PET (%)	PET wt (in kg/m^3)	PET size (W × L × T) (in mm)
Type P1	CM1	0	–	2 mm × 10 mm × 0.25 mm
	CM2	0.5	6.85	
	CM3	1	13.70	
	CM4	1.5	20.55	
	CM5	2	27.40	
Type P2	CM1	0	–	2 mm × 15 mm × 0.25 mm
	CM2	0.5	6.85	
	CM3	1	13.70	
	CM4	1.5	20.55	
	CM5	2	27.40	
Type P3	CM1	0	–	2 mm × 20 mm × 0.25 mm
	CM2	0.5	6.85	
	CM3	1	13.70	
	CM4	1.5	20.55	
	CM5	2	27.40	

flexural strength measurement. Total 105 members of elements were built including 45 cubes, 45 cylinders, and 15 beams.

3 Experimental Program and Tests

3.1 Workability Test of Fresh Concrete

The slump test was performed to monitor the consistency of freshly mixed concrete. Results of the slumps compared with standard slump values to identify low, medium, and high workability of the concrete mix. Generally, slump value between 50 and 100 mm suggested for average workability of mix and suitable for plain and reinforced concrete.

3.2 Harden Concrete Tests

Compression tests were performed on CTM machine of 2000 kN load caring capacity with constant loading rate at 14 N/cm^2/min. Load was continuously applied on the cubes of size (150 × 150 × 150) mm examined on the machine until the final failure occurs. Splitting tensile capacity helps to study the indirect tensile capacity of the concrete. Cylindrical specimen of size (150 dia. × 300 height) mm used for the test. The specimens were subjected to compression load in the CTM machine. The crack pattern was observed in the axial directions. The flexural test was carried out with plain concrete beam of size (150 × 150 × 700) mm and applied three-point load in UTM machine having 500 kN capacity. The outcomes were exhibited as load versus deflection graph. Deflection values are taken into count until final failure of the specimen occurs.

4 Results

The following graphs depict the findings of the experiments (Figs. 3, 4, 5, 6 and 7).

5 Conclusion

The experimental end results on concrete comprising altering percentages and types of PET allocated with the following judgments;

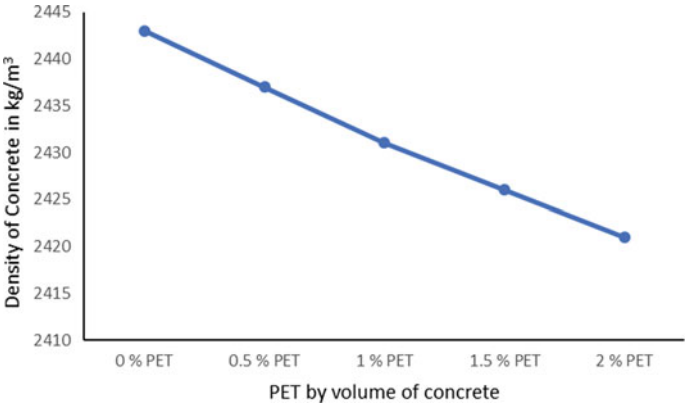


Fig. 3 Relationship between density and % PET alteration

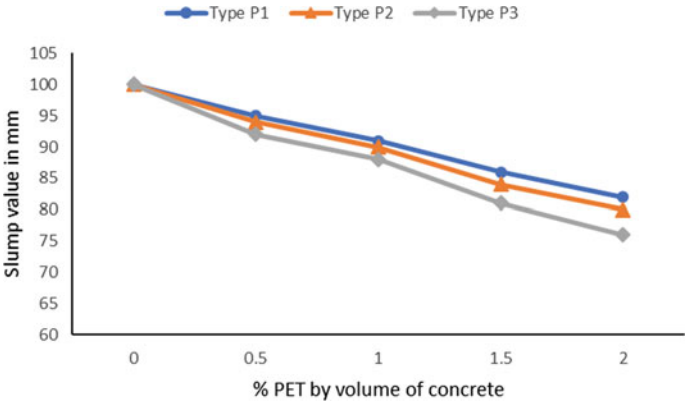


Fig. 4 Correlation between slump values and % PET alteration

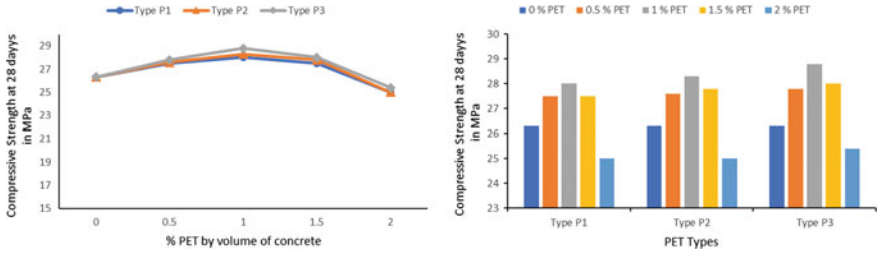


Fig. 5 Modification of compressive capacity with altering % PET fiber and types

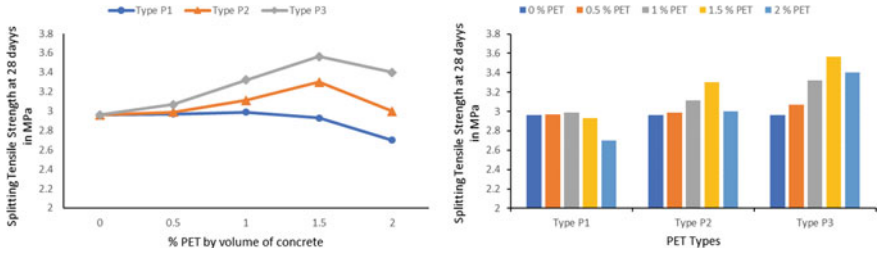


Fig. 6 Modification of splitting tensile capacity with altering % PET fiber and types

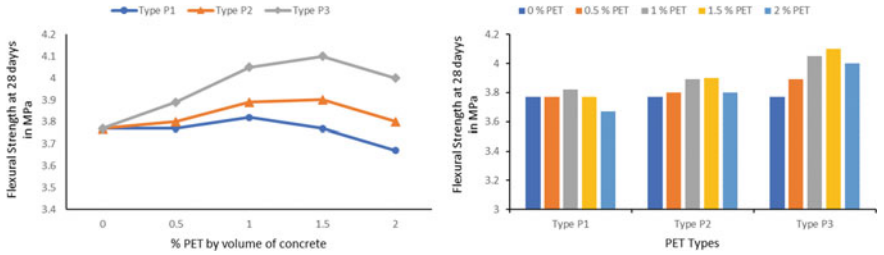


Fig. 7 Modification of flexural capacity with altering % PET fiber and types

- Workability of concrete mixes decreased and became more stiffer at a higher percentage, with the increased of waste PET fibers. The main reason behind this, the addition of macro fibers influences the viscosity of the mix and fibers restrained the stability of the concrete mixes.
- Result indicates lengthy fibers reduce the slump better than the tiny fibers at a same percentage inclusion. Fibers with shorten length get mixed properly with other materials. However, at an extensive percentage of tiny fibers reduced slump also.
- The crack resilience capacity of standard concrete was increased by inclusion of PET fibers. The generation or extension of minute (micro) cracks of harden mass and ductility of cement paste against the brittle failure was improved in the existence of PET fibers. PET fibers hold concrete properly, which help to reduce cracks.
- The existence of PET fibers in concrete did not directly affect the strengths.
- The flexural and splitting tensile strength was increased with usage of PET fibers up to 1.5% by volume of concrete, whereas the compressive strength showed slight decrement with increase of PET fibers from 1% onwards.
- Type P1 and P2 revealed appropriate adaption of concrete characteristics at an inferior amount up to 1% for all test situations. Type P3 with lengthy fibers displayed outstanding crack-resisting efficiency and provided to the ductile reaction of concrete.

- An increment in waste PET fiber percentage leads to a decrement in the density of concrete mixes with an enhancement in the strength of the mixes, which make structure more durable and lightweight.
- The task implemented to acquire the response of concrete reinforced with PET fibers revealed probable usage of waste PET water and soda bottle fibers as a concrete essential component in a dependable way.
- The PET concrete is applicable near water bodies and chemical prone areas as it provides water and chemical tightening.
- Practically, it produces lesser heat of hydration as compared to conventional concrete, which leads to reduction in temperature rise in mass concrete, and this is a major advantage of PET concrete.

Acknowledgements Authors are obliged to Mr. Gaurav Saraiya, Project Manager of Shree Balaji Wind, Vadodara for providing supply of all raw materials required and friendly support for the performance of experiments.

References

1. Pacheco-Torgal, F.: Introduction to the use of recycled plastics in eco-efficient concrete. In: Use of Recycled Plastics in Eco-efficient Concrete, pp. 1–8. Elsevier, Amsterdam (2019)
2. Siddique, R., Khatib, J., Kaur, I.J.: Use of recycled plastic in concrete: a review. Waste Manag. **28**(10), 1835–1852 (2008)
3. https://www.downtoearth.org.in/news/waste/an-indian-consumes-11-kg-plastic-every-year-and-an-average-american-109-kg-60745
4. Bhogayata, A., et al.: Performance of concrete by using non-recyclable plastic wastes as concrete constituent. Int. J. Eng. Res. Technol. **1**(4), 1–3 (2012)
5. Bhogayata, A., Arora, N.K.: Green concrete from the post-consumer plastic wastes: Indian scenario, pp. 437–440 (2011)
6. Jain, A., et al.: Fresh, strength, durability and microstructural properties of shredded waste plastic concrete. Iran. J. Sci. Technol. Trans. Civ. Eng. **43**(1), 455–465 (2019)
7. Al-Hadithi, A.I., Abbas, M.A.: Innovative technique of using carbon fibre reinforced polymer strips for shear reinforcement of reinforced concrete beams with waste plastic fibres. Eur. J. Environ. Civ. Eng. 1–22 (2019)
8. Lakhiar, M.T., et al.: Flexural performance of concrete reinforced by plastic fibers. Eng. Technol. Appl. Sci. Res. **8**(3), 3041–3043 (2018)
9. Parra, C., et al.: Recycled plastic and cork waste for structural lightweight concrete production. Sustainability **11**(7), 1876 (2019)
10. Al-Hadithi AI, Abbas MA.: Innovative technique of using carbon fibre reinforced polymer strips for shear reinforcement of reinforced concrete beams with waste plastic fibres. Eur J Environ Civ Eng **25**(3), 516–537 (2021)

Efficient Drip Irrigation System Using Soil Moisture Sensor with IoT Sync to Webserver

M. Srikar, N. Samanth, and S. Sampreeth

Abstract This work focuses on efficiently managing an agricultural field without human intervention by making the use of an IoT-based system. The principle concept used here is, electrical conductivity of the moisture is present in the soil. This setup includes a servomotor which is used to regulate the water flow, based on the input from the option selected by the user on the webpage pertaining to the type of soil and the corresponding crop. The input is generated from the webserver hosted on the microprocessor which is further then processed to the microcontroller based on the temperature of the surroundings. The PIR sensor is used to ensure the security of the agricultural field.

Keywords Drip irrigation · Soil moisture sensor · Servomotor · DHT11 sensor · PIR sensor · Motion sensor · Efficient irrigation · Water conservation · Smart agriculture · Embedded systems · IoT sync · Serial communication

1 Introduction

Agriculture is the backbone of the country; it is also undoubtedly the major contributor to the GDP or gross domestic product of India. This study was motivated by the desire for the enhancement of the use of technology in the field of agriculture. This study contributes to enhancing the life of a farmer by using the appropriate amount of water required for the crop.

Now the following questions arise—how do we simplify the life of a farmer using technology? What would be the ideal and efficient way to regulate the water consumption?

This paper provides the solutions for the above-raised questions.

M. Srikar (✉) · N. Samanth · S. Sampreeth
Department of Electronics and Communication Engineering, B.N.M Institute of Technology, Bengaluru, Karnataka, India

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_42

1.1 Moisture Value and IoT Sync

To obtain an ideal yield each crop requires a stipulated amount of water [1], the soil moisture sensor [2] records the active moisture value of the soil. The servomotor regulates the water flow based on the concurrent reading of the soil moisture sensor. The concurrent reading of the soil moisture sensor is compared to the formatted input value received from the webserver provided by the user.

The formatted input value is calculated based on the annual consumption of water [1] and also taking into consideration the temperature of the surroundings. The webserver is hosted on the microprocessor making use of web sockets to emit data. This data is logged to the terminal which is then sent serially to an external microcontroller [3]. This webserver serves as an interface to communicate between the user and the setup.

1.2 PIR Sensor

PIR or the passive infrared sensor is used to enhance the security of the field. Upon detecting motion, a light system is triggered to indicate activity in the field. In certain cases, the user cannot depend upon human intervention to look after the security of the field, hence, using such a setup helps in automating the process of securing the field.

2 Literature Survey

There have been implementations of efficient irrigating system and controlling the same with the help of a microcontroller, but these setups have the capability of measuring the soil moisture value and regulating the flow of water, in addition to this the following setup helps in regulating the water flow based on the temperature of the surrounding environment and also ensuring the security of the field.

The following setup is modular and hence can be installed as per the need of the user even in operating agricultural fields and thus increasing the customizability of the field.

3 Experimental Setup

To prove that this study is positive and provides a considerable change in the amount of water required, we constructed a small-scale drip irrigation setup [4] using drippers as shown in the results section. The setup includes PIR or passive infrared sensor,

servomotor, LED strip, DHT11 sensor connected to digital pins 3, 10, 7, 5, respectively. The soil moisture sensor is connected to analog pin A0 of the microcontroller. The microcontroller and the microprocessor are connected through an USB cable for serial communication.

4 Interfacing Block Diagram

The microprocessor receives input from the webserver as depicted in Fig. 1. The soil moisture sensor [2, 5] possesses the ability to output both analog and digital values, and this setup requires the soil moisture sensor to return an analog output which is received at an analog pin on the microcontroller. The values of the soil moisture sensor range from 0 to 1023, where zero indicates high moisture content and 1023 indicates low moisture content in the soil. The PIR sensor returns digital output and hence connected to a digital pin on the microcontroller. When the output from the PIR/motion sensor is high, that is; motion is detected, the light system is triggered,

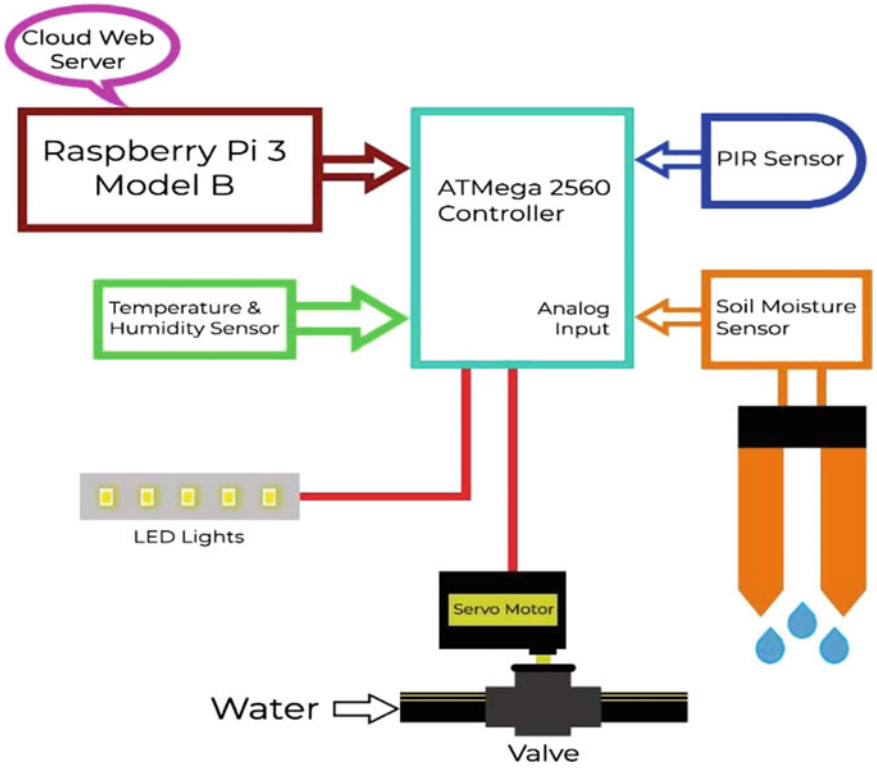


Fig. 1 Interfacing block diagram

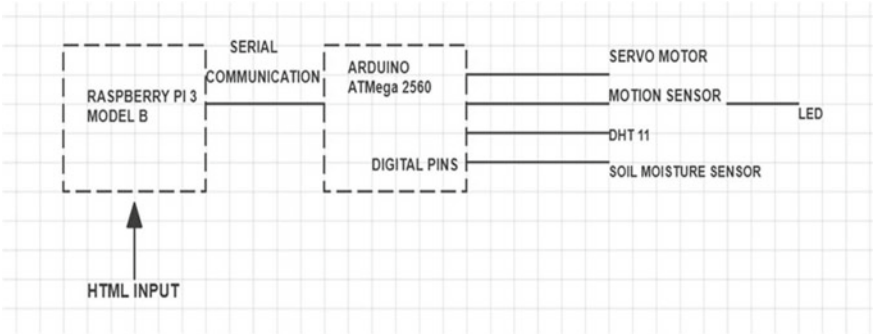


Fig. 2 Signal conditioning circuit [1]

consequently, when no motion is detected, this implies that the PIR/motion sensor returns a low output.

The temperature sensor plays a vital role as the value from the webserver is formatted and returned by the DHT11/temperature sensor [2]. The servomotor is explicitly used in the setup to obtain accurate regulation of the water during drip irrigation. The servomotor possesses the capability of turning between 0 and 1800. Thus, ensuring the smooth movement of the valve [4]. The microcontroller and the microprocessor are connected serially and the data communication takes place through port communication.

4.1 Signal Conditioning Circuit

As depicted in Fig. 2 that the input received from the webserver is transferred to the microcontroller from the microprocessor through the serial communication established between the two. The power for the smooth functioning of the microcontroller is supplied from the microprocessor. On receiving the input, the microcontroller formats the input to correspond to the pertaining surrounding temperature.

The microcontroller compares the soil moisture value to the input value and the servomotor is triggered if the required conditions are not met, hence, regulating the flow of water to the crops.

The PIR/motion sensor detects movements on the field and triggers a light system which alarms the user.

Table 1 depicts the moisture value required for a crop in a particular soil. These values were obtained after comparing soil moisture sensor values with that of the annual consumption of water by the given crop.

Table 1 Pertaining moisture values for some crops [1]

Soil type	Crop name	Moisture value
Alluvial	Rice	542
Black	Cotton	736
Laterite	Rubber	103
Mountainous	Maize	635
Arid	Barley	910
Saline	Sunflower	900
Marshy	Rice	500

5 The Algorithm

The algorithm for the setup is as shown in Fig. 3. It depicts the sequential workflow.

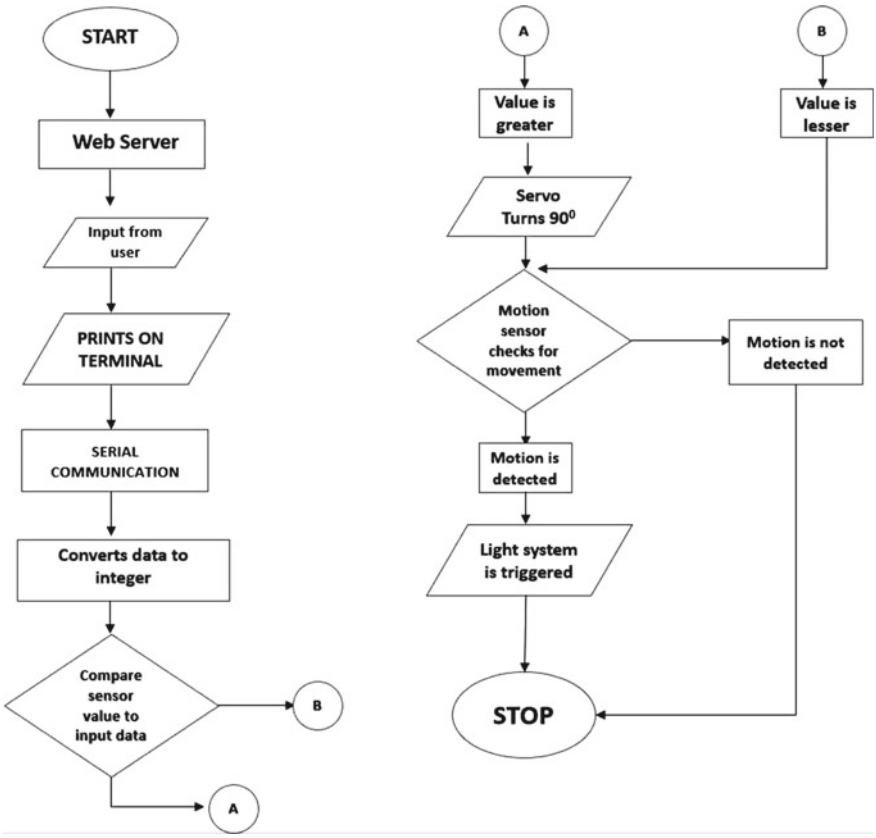


Fig. 3 Algorithm

6 Webserver

The webserver is written in node js using an express [6] framework to host static HTML and CSS files. The bi-directional communication is established using web sockets which logs the information obtained from the webserver to the terminal. The inbuilt Python function subprocess is used to start the webserver while setting up serial communication parallelly. The baud rate and the port for serial communication are also specified. The communication is started using the HTTP protocol via the port specified to the corresponding IP address of the Raspberry Pi.

There is a script which dynamically generates options using the virtual DOM object property of HTML. This script populates the entries of the second field based on the selected option of the first field. The script is used to emit web socket events which is caught by the webserver and transferred over to the microcontroller.

7 Results

The soil moisture sensor is implanted in the soil. Further, the user inputs the desired value from the webpage, and hence on comparison of the formatted value with that of the sensor value, the servomotor regulates the valve controlling the flow of water. To ensure the security of the field and reduce human intervention, the PIR/motion sensor triggers a light system which alarms the user from visible distance.

Figure 4 depicts that the LED is turned OFF as no motion has been detected, this is communicated to the microprocessor terminal as shown in Fig. 8. This message is also displayed on the serial monitor of the microcontroller as shown in Fig. 11. The webserver is successfully connected and the port runs on 3831 which is seen in Fig. 8. The interface between the user and the setup is depicted as shown in Fig. 10. Once the user inputs the desired value from the webpage, the value is emitted through web sockets and displayed on the terminal of the microprocessor as depicted in Fig. 9, and this is further communicated to the microcontroller using serial communication. The microcontroller initially senses the soil moisture value and as it is indicated that the soil is dry, and the valve is regulated as shown in Fig. 11 and the water starts to drip as shown in Fig. 6. Once motion is detected, the LED switches on as seen

Fig. 4 LED is OFF

Fig. 5 LED is ON

Fig. 6 Ripples are visible

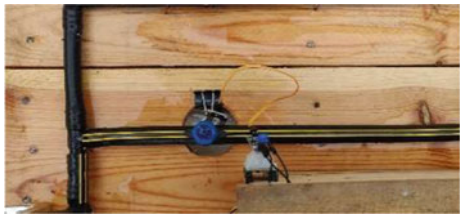


Fig. 7 Ripples disappear

in Fig. 5, once the soil moisture increases up to the desired value the valve closes and the ripples disappear as shown in Fig. 7 and this is communicated on the serial monitor of the microcontroller which is seen in Fig. 12. The valve which is regulated by the servomotor remains open until the set moisture value for the pertaining crop is met.

8 Conclusion

Thus, this setup has described and implemented a real-time efficient drip irrigation setup. The data has been successfully synced via Wi-Fi to a remote webserver. Also, the user-friendly interface has been implemented so that the user can focus to improve the yield of crops in the farm with least amount of human interaction. Each crop in

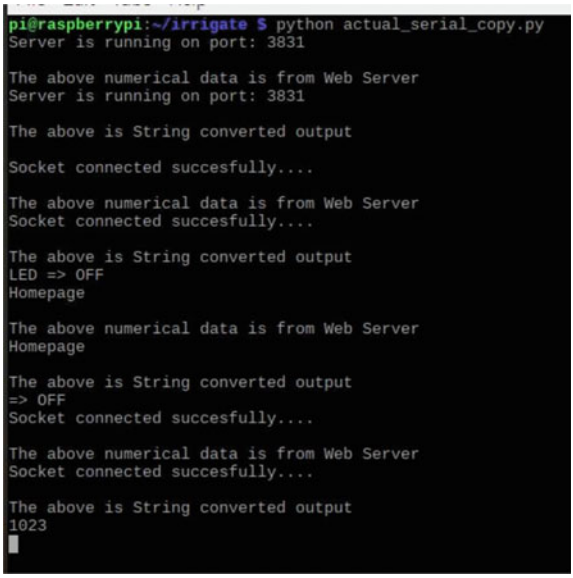


Fig. 8 Webserver loads

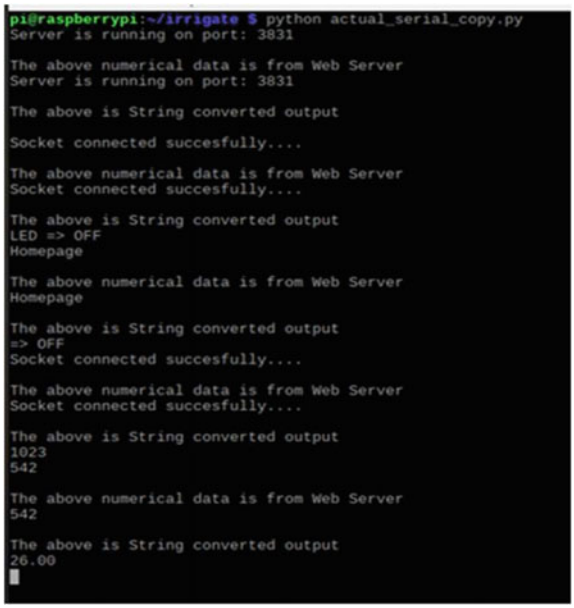


Fig. 9 Input from webpage

Fig. 10 Interface/webpage of the setup

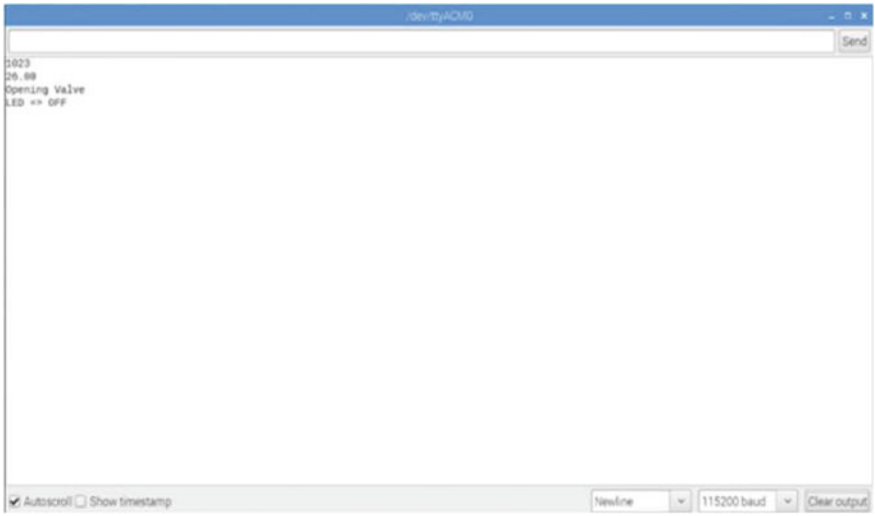


Fig. 11 O/p on serial monitor

each soil has been given pre-set moisture values which enhance the yield. Further, the developments can be done for the supply of power. The self-made modular setup is cost effective, and thus the overall design is an effective low-cost device.

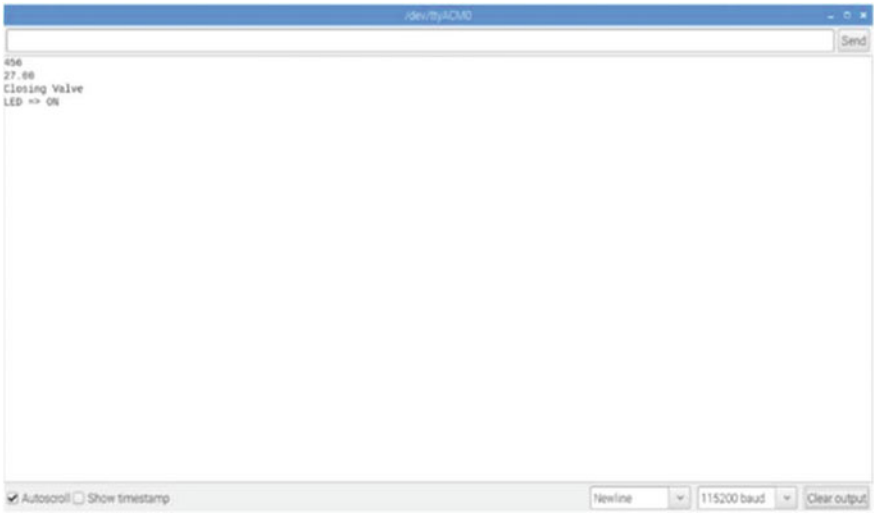


Fig. 12 O/p on serial monitor

References

1. AgroPediaHomepage: https://agropedia.iitk.ac.in/content/water-requirement-different-crops
2. Baseline Systems DHT11 Datasheet, Components 101
3. Atmega2560 Datasheet
4. Venot, J.P.: Drip Irrigation for Agriculture: Untold Stories of Efficiency, Innovation and Development. Taylor & Francis, UK (2017)
5. Ramon, M.C.: Creating a soil moisture sensor. In: Intel® Galileo and Intel® Galileo Gen 2. Apress, Berkeley, CA (2014). https://doi.org/10.1007/978-1-4302-6838-3_8
6. Express Node Package Module: https://www.npmjs.com/package/express

Early Childhood Anxiety and Depression Detection Based on Speech Using Machine Learning Analysis

Suvarna Nandyal and P. S. Swathi

Abstract Uneasiness and wretchedness in youngsters habitually go undiscovered. On the off chance that these conditions are not treated at the ideal time, generally known as disguising issues, they bring about long-haul negative results like medication misuse and expanded self-destruction hazard. This paper gives an appropriate technique to perceive small kids with incapacities being disguised utilizing a three-minute discourse task. We present that AI examination of sound information from the errand can be utilized to characterize youngsters with an incapacity that disguises. The most significant attributes of disguising issue are concentrated inside and out, indicating that influenced youngsters have low-pitch voices, with repeatable discourse emphases and content, and a shrill reaction to startling control improvements. The current procedure relies upon clinician's audit of the patient. These techniques are abstract, done on meet and rely upon reports by the patient. With increment in the downturn, some programmed and dependable methods for gloom acknowledgment are required. Endeavors are being made to survey and distinguish despondency through PC vision and AI. Subsequently, in this paper, we present our techniques for text, sound, and video highlight extraction, trailed by choice level combination which helps in recognizing discouraged kids with exactness.

Keywords Anxiety · Speech · Facial expression · Depression · Machine learning

1 Introduction

Melancholy is a significant psychological wellness issue that is quickly influencing lives of kids, youthful, and old with no age obstruction. Discouragement impacts passionate as well as physical and mental condition of the individual [1]. The space for dysfunctional behavior location through quick, minimal effort, and possible advancements can possibly offer voice to populaces who have difficulties understanding and communicating their own pain and looking for suitable assistance. Around one of

S. Nandyal (✉) · P. S. Swathi
Department of Computer Science and Engineering, PDA College of Engineering, Kalaburgi, Karnataka, India

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_43

every five kids experience the ill effects of uneasiness and melancholy, altogether known as disguising disorders. But in light of the fact that kids younger than eight cannot dependably communicate their enthusiastic misery, grown-ups should have the option to comprehend their psychological state and perceive potential emotional wellness issues. Side effects of despondency remember the absence of enthusiasm for every day exercises, feeling low, uneasiness, dissatisfaction, loss of weight, and in any event, sentiment of self-loathing. Three principle highlights demonstrated to show these sicknesses: low pitch, rehashed discourse expressions and subject, and a higher pitch when shocked. As the pace of sorrow is expanding step by step, some programmed and dependable methods for wretchedness acknowledgment are required. This instrumented mindset acceptance task gives a target proportion of kid movement without the restrictions of social coding and, when taken with these primer outcomes, can possibly be utilized as a screening device for youth disguising disorders.

To identify children with an internalizing disorder [2] using information from such errands, a model of the unpredictable connection between target voice sound sign attributes and conclusion must be set up. An information-driven methodology, similar to AI, is unmistakably appropriate for such undertaking. Actually, our past work has investigated the utilization of wearable inertial sensors [3]. These wearable sensors are utilized for distinguishing youngsters with disguising psychopathology dependent on AI investigation of their development designs during a dread acceptance task. The space for psychological instability is recognized by means of quick, minimal effort, and possible advances that can possibly offer voice to populaces who have difficulties understanding and communicating their own trouble and looking for fitting assistance. In grown-ups, this is remarkable for people with madness and self-destructive goal. Nonetheless, these techniques can likewise serve any kid younger than eight who has more trouble comprehension and imparting their theoretical feelings. While some work has been done in distinguishing influence in youngsters between the ages of 10 and 13 [4] and in catching conversational correspondence in kids with chemical imbalance, the network presently cannot seem to utilize the examination of vocal highlights inside kid discourse ages for recognizing kids with hidden psychopathology. In this, we utilize discourse examination of youngster voice chronicles during a three-minute discourse undertaking and AI to identify clinically inferred nervousness and burdensome judgments in kids between the ages of 3 and 8 years of age. In this paper, we portray the trial which incorporates information preparing, model turn of events and assessment draws near, present execution qualities of a typical parent-report survey, and the AI models for recognizing kids with disguising issues.

2 Related Work

Environmentally legitimate long-haul temperament observing of people with bipolar turmoil utilizing discourse speech designs is adjusted by the passionate and neurophysiological condition of the speaker. There exists a developing collection of work that computationally analyzes this tweak in patients experiencing gloom, mental imbalance, and post-horrible pressure issue. In any case, most of the work here spotlights on the examination of organized discourse gathered in controlled conditions. Here, we develop the current writing by analyzing bipolar confusion (BP) [5].

A diagram of flow mechanical improvements which are encouraging exploration in the field of self-destruction anticipation including different methods of screening, for example, network examination of cell phone gathered availability information, programmed recognition of suicidality from online media substance, and emergency location from acoustic inconstancy in discourse designs is presented [6, 7]. The current field of mhealth applications for self-destruction counteraction is evaluated, and an inventive application for an indigenous populace is introduced. From this review, future difficulties—specialized and moral—are talked about.

Another methodology for diagnosing tension and sadness in small kids is depicted. At present, analysis in this populace requires long periods of organized clinical meetings spread over days and weeks. They analyze the presentation of an assortment of capabilities and displaying ways to deal with recognize the best performing strategic relapse that gives an indicative precision of 80%. This precision is equivalent to existing demonstrative procedures, however at a little portion of the time and cost as of now required. These outcomes highlight the future utilization of this methodology in a clinical setting for diagnosing youngsters with disguising issues [8].

The paper presents another methodology for diagnosing tension and discouragement in little youngsters. It proposes the utilization of a 90-s dread enlistment task during which time member movement is observing and utilizing an industrially accessible wearable sensor. AI and information extricated from the most clinically achievable 20-s period of the undertaking are utilized to anticipate analysis in an example of kids with and without a disguising determination. They likewise look at the presentation of an assortment of capabilities and displaying ways to deal with recognize the best performing strategic relapse that gives a demonstrative precision of 80% [9].

3 Proposed System

State of mind enlistment assignments have been utilized often in research settings to show on edge, baffling, happy, amazing, or disheartening effect. A kids physiological and social reaction are recorded to these errands utilizing an assortment of advances like video-recorded and coded practices and legitimately estimated

cortisol, pulse inconstancy, electrodermal action, development, physically handled and concentrated according to hypothesis-driven desires.

The Trier Social Stress Task (TSST) is one such mindset enlistment undertaking to incite execution tension by having a member give a short, improvisational discourse to a confederate crowd claiming to be altogether exhausted and basic. Conduct coding and physiological measures have been related with task influence and psychological instability (uneasiness, misery, and bipolar issue) [10]. While useful for the research, results from objective assessment of this task are not considered have an immediate clinical utility due to long and costly process which requires specialized equipment. The factors considered while selecting studies for review include structured diagnostic interviewing including non-depressed to ensure construct validity, comparing the samples of both child and adults to get better accuracy, examining physical activities, and comparing cortisol responses to psychological stressors between depressed and non-depressed, Finally, effect sizes must have been calculable for both baseline and stress time periods. The current methods rely upon clinician's audit of the patient. These techniques are abstract, done on meet, and rely upon reports put together by the patient.

With increment in the pace of gloom, some programmed and solid methods for wretchedness acknowledgment are required. As shown in Fig. 1, a training database is designed, and the input audio signals are pre-processed, and the features are extracted. A SVM modeling classifier classifies the data into depressed or non-depressed and testing is carried out (Fig. 2).

The input layer receives normalized images with the same sizes, and then, the image is convolved with multiple kernels using shared weights. CNN is composed of a stack of convolutional layers; a convolutional layer is parameterized by the

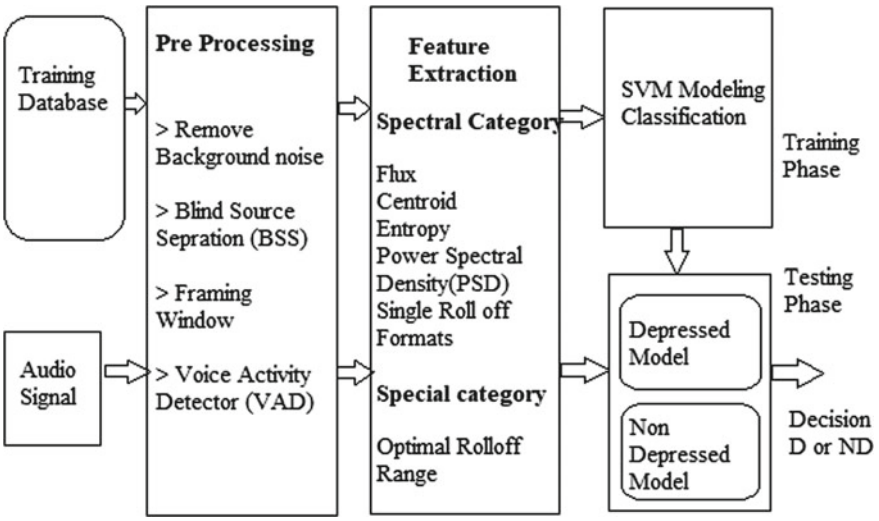


Fig. 1 Procedural flow of acoustic depression detection

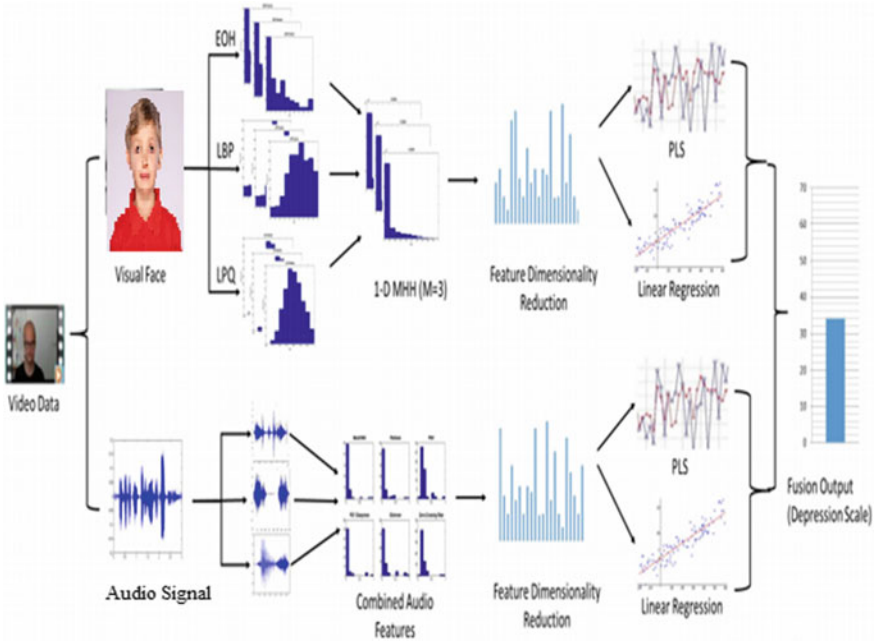


Fig. 2 Facial feature recognition

number of maps. CNNs apply many filters to the raw pixel data of an image to extract and learn features, which the model can then be used for classification. The feature map of each pixel is calculated as follows:

$$Cn = f(p * W + b) \tag{1}$$

where " $*$ " indicates the convolution computation, n is the pixel in the feature map, p is the pixel value, W is the convolution kernel, and b is bias and f is the convolution kernel.

After this convolution layer, we apply a nonlinear layer (activation layer) because the network is able to train faster

$$f(p) = \max(0,\ p) \tag{2}$$

This function is applied to all of the values in the input data to introduce nonlinearities into the model.

Algorithm 1: General Algorithm

Input: X: Training set. δ: Threshold. **Output:** X_R: $X_R \subset X$ s.t. $|X_R| \ll |X|$. **Begin** Train a decision tree $\mathcal{T}$; // X_R Begins empty $X_R \leftarrow NULL$
For each leaf $\mathcal{L}_i$ of $\mathcal{T}$ **do**
  for each opposite class neighbor $\mathcal{L}_j$ **do**
  if entropy of $\mathcal{L}_j$ is low **then**
  //Select closest examples
    Use $\mathcal{L}_i$ and $\mathcal{L}_j$ to build X^+;
    Compute ω (Eq. (12));
    Add $x_i \in \mathcal{L}_j$ to X_R according to (12);
    end for
  else
  //Add all the elements in $\mathcal{L}_j$ to X_R.
    $X_R \leftarrow X_R \cup \mathcal{L}_j$;
  end if **end for return** X_R **End**

Fig. 3 SVM algorithm

3.1 Algorithms and Methods

Supporting vector machines (SVMs, additionally supporting vector organizations) in AI are managed learning models with related learning calculations which can investigate information utilized for order and relapse examination. A support vector machine (SVM) is an officially developed discriminative classifier with a different hyperplane (Fig. 3).

4 Experimental Results

4.1 Audio Segmentation

Sound division (regularly called sound order) is a preprocessing step in sound investigation. It isolates various sorts of sound, for instance, discourse, music, natural sounds, quiet, and blends of these sounds. Division is a significant pre-handling stage for the vast majority of sound investigation applications. The principle objective is to part a continuous sound sign into homogeneous portions. Division can either be directed or unaided (Figs. 4 and 5).

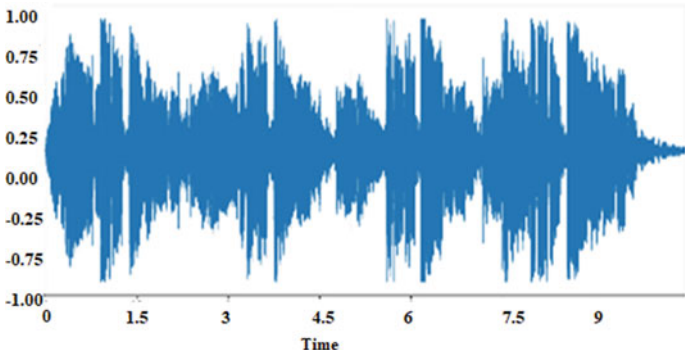


Fig. 4 Child audio file array

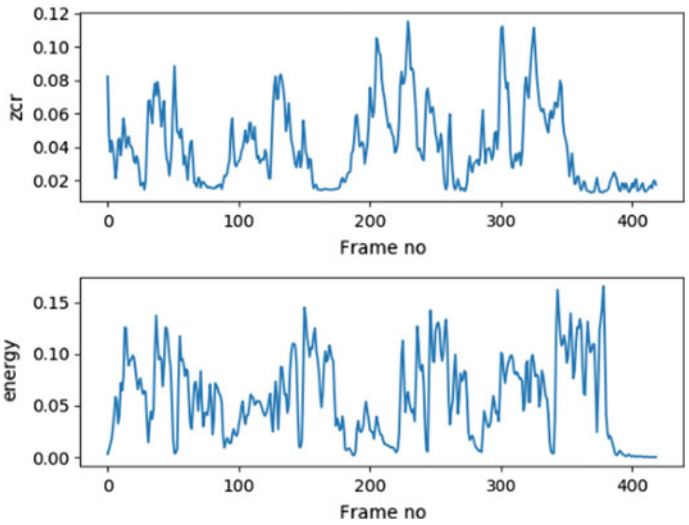


Fig. 5 Audio file frames

4.2 Feature Extraction

There are a few highlights to any sound and video signal. We should consequently remove the qualities pertinent to the issue that we are attempting to unravel (Figs. 6, 7, and 8).

Here, we explore the use of audio data from a three-minute mood induction task and machine learning for identifying young children with internalizing psychopathology. We examine the features (high-pitch tone, low-pitch tone, chaotic speech, repetitive speech, breathy, light timbre, and full timbre) selected by the feature selection process to lend clinical interpretability to the model. Logistic regression

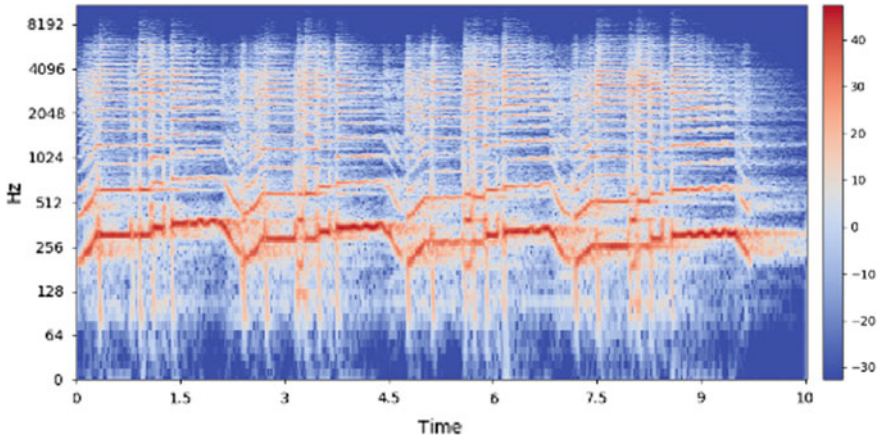


Fig. 6 Spectrogram

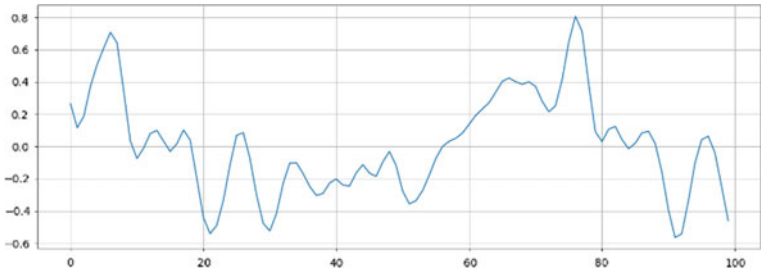


Fig. 7 Zero crossing rate

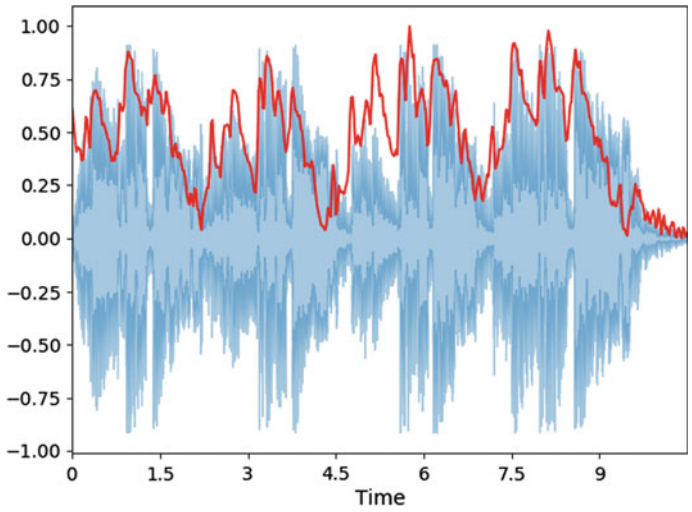


Fig. 8 Spectral centroid. The center of gravity of spectrum

and support vector machine models were able to identify children with an internalizing disorder with 80% accuracy. Children with an internalizing disorder generally exhibits low-pitch voices with repeatable speech inflections and content and high-pitched response to surprising stimuli.

The facial expression is extracted by converting the input colored image obtained during three-minute mood induction task into gray image, represents the binary gradient mask for the image for edge detection, dilates the image to represent the image more specifically, erodes the image (filling holes), segments it and shows image with outlines of features. Then, the data is compared with training datasets and classifies whether the external features depict happy, sad, anger, anxiety, fear, etc. Children with internalizing disorder generally exhibit sad, anger disgust, and anxiety.

5 Conclusion

The outcomes introduced in this paper propose that an AI investigation of youngster talking designs during a short uneasiness acceptance task can distinguish kids with disguising psychopathology. Here, we are using .wav audio file. The length of an audio file varies from about 1 to 2 s. Bitrate is 256 kbps. The facial emotion-based stress recognition is done with the help of jpg images. Based on classes such as angry, disgust, fear, happy, neutral, sad, and surprise, the images are classified. This factual characterization model beats clinical edges on parent-detailed kid side effects gathered with CBCL, demonstrating its potential as a target screening instrument in this populace. A nitty-gritty examination of sound highlights and outward appearances chose for this characterization demonstrates that influenced youngsters display low-pitch voices, with dull intonation, shrill reaction to amazing boosts and grinning, pitiful or astonishing countenances. In this paper, we presented detection of depression or non-depression through visual and sound highlights utilizing SVM and convolutional neural networks as classifier.

References

1. Chansky, T.E., Kendall, P.C.: Social expectancies and self-perceptions in anxiety-disordered children. J. Anxiety Disord. **11**(4), 347–363 (1997)
2. Egger, H.L., Angold, A.: Common emotional and behavioral disorders in preschool children: presentation, nosology, and epidemiology. J. Child Psychol. Psychiatry **47**(3–4), 313–337 (2006)
3. McGinnis, R.S., et al.: Wearable sensors and machine learning diagnose anxiety and depression in young children. In: 2018 IEEE International Conference on Biomedical and Health Informatics (BHI), Las Vegas, NV (2018)
4. Schuller, B., Steidl, S., Batliner, A.: The INTERSPEECH 2009 Emotion Challenge, p. 4

5. Karam, Z.N., et al.: Ecologically valid long-term mood monitoring of individuals with bipolar disorder using speech. In: 2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 4858–4862 (2014)
6. Elenko, E., Underwood, L., Zohar, D.: Defining digital medicine. Nat. Biotechnol. [Online] (2015). Available: https://www.nature.com/articles/nbt.3222 [Accessed: 21-May-2018]
7. Achenbach, T.M., Howell, C.T., Quay, H.C., Conners, C.K.: National survey of problems and competencies among four- to sixteen-year-olds: parents' reports for normative and clinical samples. Monogr. Soc. Res. Child Dev. **56**(3), v–120 (1991)
8. Topol, E.J., Steinhubl, S.R., Torkamani, A.: Digital medical tools and sensors. JAMA **313**(4), 353–354 (2015). Calkins, D., Dedmon, S.E., Gill, K.L., Lomax, L.E., Johnson, L.M.: Frustration in infancy: implications for emotion regulation, physiological processes, and temperament. Infancy **3**(2), 175–197 (2002)
9. Lopez-Duran, N.L., McGinnis, E., Kuhlman, K., Geiss, E., Vargas, I., Mayer, S.: HPA-axis stress reactivity in youth depression: evidence of impaired regulatory processes in depressed boys. Stress **18**(5), 545–553 (2015)
10. Buske-Kirschbaum, A., Jobst, S., Wustmans, A., Kirschbaum, C., Rauh, W., Hellhammer, D.: Attenuated free cortisol response to psychosocial stress in children with atopic dermatitis. Psychosom. Med. **59**(4), 419–426 (1997)

Effect of Dynamic Process Load Variation on Relative Deadlines in Multitasking Embedded Systems

Sandeep S. Chapalkar and K. Karibasappa

Abstract The real-time systems have a set of processes continuously running in the system and will also have a set of processes triggered on events or user interaction. In such systems, the number of processes in scheduling queue varies based on frequency of trigger activation. The hard real-time systems are designed to meet the deadline of all processes under peak load of processes. Hence, the execution speed of the processing unit is chosen much higher than the required speed for average load. The DVFS technique addresses the issue based on process load; however, it does not take deadlines into account while scaling the processors operating frequency.

Keywords Process load · Frequency scaling · Deadline analysis · Real-time system · Scheduling

1 Introduction

The embedded applications are designed to finish the assigned job within an estimated time and failing to do so will be considered as system failure. The time constraint of the real-time systems is one of the most important features which decides the criticality of the system and also affects the design issues. The absolute deadlines of processes can be optimized during software development, but the relative deadlines depend on process load and operating conditions in real time. The hard real-time systems need to be designed to meet deadline requirement of all processes under peak process load, which leads to high-frequency processor, additional thermal stability features and hence increase in overall cost of development. Therefore, processor utilization factor also degrades as system runs with optimal number of processes most of the time and the high-performance processing units may complete the task well within the deadline and stay idle for rest of the time. The full capacity of

S. S. Chapalkar (✉)
Dayananda Sagar College of Engineering, Bengaluru, India

K. Karibasappa
Dayananda Sagar Academy of Technology and Management, Bengaluru, India

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_44

processing unit is used only occasionally when the system has maximum expected processes in the queue. But at a same time, hard real-time applications cannot afford to miss deadline under any circumstances. Hence, it is important to devise a solution improve processor utilization factor as well as meet all process deadlines under various conditions.

The study in this paper analyzes the issue by considering relative deadlines of processes and calculates the scaling factor of the processor operating frequency dynamically. The number of processes in execution queue is presumed to be changing dynamically which in turn affects the relative deadlines of other processes depending on their priorities. The analysis is done considering the system with N number of processes with "EDF" and RM scheduling algorithms.

2 Related Work

The real-time system correctness is decided not only by the output of the system but also by how much time a system took to produce results. A hard real-time system has stringent time constraints and it is necessary to complete its operations within all its absolute deadlines. For a given set of tasks, it is helpful to understand what modifications can be made to a tasks of the system that achieves border line scheduling. It is useful in a technical paradigm to estimate the least required frequency of a processing unit in this regards wherein implementation of different scheduling algorithms is prominently considered for the efficient utilization of processors [1].

These constraints make embedded system software different than generic software and they usually contribute to around 70% cost total development cost [2]. The widely used scheduling policies in real-time systems are earliest deadline first ("EDF") for dynamic priority allocation and rate monotonic (RM) for fixed or static priority allocation [3].

The periodic hard deadline and aperiodic soft deadline tasks are scheduled together in this paper based on critical task indicator (CTI). Minimum response time is achieved if both sets of processes are scheduled using "EDF" algorithm [4]. Brinkley Sprunt, Lui Sha and John Lehoczky propose that an individual sporadic server can be dedicated to hard deadline aperiodic tasks to achieve the required deadline [5].

As the embedded systems even though there can be combination of periodic, sporadic and aperiodic tasks, the system tasks are reasonable have security issues as they become predictable over period of its operation. Therefore, a dynamic priority allocation and obfuscation makes it difficult for unauthorized intruder in real-time systems [6].

The hardware accelerators are used to increase the speed of the critical sections of the operations and enhance the total efficiency and performance. The hardware accelerators operate as isolated processors which reduces load on primary processing unit by taking over the specific processes as and when required. In unified design paradigm, software implementation is recommended to support scalability,

upgradation and easy improvization with cost effectiveness [7]. However, the hardware realization is important in realization of the critical section of task and operating frequently executed processes to enhance the efficiency of operation. "EDF" scheduling policy is a preemptive algorithm that allows the preemption if process with sooner deadline compared to the process in execution state is arrived. The reference model considered is a uniprocessor "real-time kernel" implementing "EDF" scheduling policy with integrated support for coprocessor. The tasks can block coprocessor once per job for feasibility analysis. Depending on the updated scheduling policy proposed, in addition to existing feasibility analysis method, it identifies a critical sections accounting for reduction of process execution period.

The "EDF" algorithm is more efficient if the preemption is allowed compared to non-preemptive systems. The "EDF" scheduling is similar to FCFS method if the deadlines are deterministic [8]. The feasibility analysis in "EDF" is most often based on processor demand. Once the processes are scheduled according to deadlines, then at end of every deadline, processor demand of all processes in queue needs to be checked again to identify any new instances [9].

In many multitasking systems, the number of processes to be scheduled may change continuously. However, the hard real-time systems are to be designed to satisfy the deadline requirement under worst-case execution time [1]. Preemptive systems increase the CPU utilization factor as scheduler need not wait for lower priority process to complete its execution [10]. During operation every time, a new process gets triggered the schedulability has to be recalculated and processes accordingly [11]. This process adds as overhead to the overall performance of the system.

3 Comparison of RM and "EDF" Scheduling Policies

A real-time system with maximum of N number of processes having average execution time of T_{exe} is considered for analyzing relative deadline variation under different process load conditions. In operating conditions, the number of processes can vary between standby mode processes to peak load processes. Hard real-time systems are designed to operate at its maximum process load, but the operating frequency can be scaled down just to meet the required deadlines [1]. This technique reduces the power consumption of the system and improves the operating life of the system as well.

The RM and "EDF" are compared systematically based on important properties below.

1. Implementation complexity favors RM as many kernels support this algorithm due to its implementation simplicity compared to "EDF" [12].
2. Runtime overhead calculations show that the number of context switches required for RM is significantly high compared to "EDF" [13].

3. Schedulability analysis shows that the CPU utilization in RM ranges from 69 to 91% and "EDF" utilization factor is around 80% [14].
4. Under permanent overload condition, "EDF" scales periods and executes tasks [15] whereas RM completely blocks lower priority tasks.
5. In case of transient overload condition, both "EDF" and RM performance is almost similar [14].
6. Jitters and latency performance is also similar in both scheduling methods [14].

4 Relative Deadline Analysis

Considering the hard RTS, for which the full process load or worst-case situation is used for the design, i.e., the system is designed to satisfy the absolute deadline requirements even if all the processes are ready for execution in scheduler queue. In such system design usually, there will not be any restriction on resources and developer will have flexibility to acquiring higher-order processing units and peripherals for intended performance. As a result, it is necessary to enforce the design with highest standard of quality assurance. However, in case of firm or soft RTS, the design is constrained by cost, power and resources which might compromise performance under some limitations.

The effect of processes on relative deadline in preemptive system can be illustrated considering N interfaces. If T_{NExe} is the time required to execute the Nth process, then the absolute deadlines (T_{AD}) of all processes having higher priority "interrupt service routine (ISRs)" will govern the relative deadline of Nth process. The absolute deadline T_{AD} of a process is the period required for the process starting from the time when process is added in scheduling queue, i.e., arrival time T_{A} to termination of the process execution, which also encapsulates waiting period T_{W} and execution period T_{NExe}.

$$T_{\mathrm{AD}} = T_{\mathrm{W}} + T_{\mathrm{EXE}} \tag{1}$$

The execution times of the processes are assumed to be considerably large compare to context switching time and hence ignored in this analysis. However, in hard real-time systems and systems in which the process execution time is small enough, the context switching time plays an important part in total turnaround time. Such system is discussed further in the following sections.

The waiting period T_{w} is execution time remaining for previous process scheduled in non-preemptive method. Whereas if it is preemptive method, then there arise two conditions based on whether the executing process is of lower priority than new process or higher priority. If the scheduled process is of higher priority, then T_{w} will be same as that of non-preemptive method, and if scheduled process is of lower priority, then it is immediately preempted; in this case, T_{w} is equal to context switching time.

The relative deadline of process T_{RD} is the addition of T_{AD} of all processes with higher priority currently present in the scheduling queue. As the number of processes in the queue changes, then relative deadlines also change, respectively.

$$T_{\mathrm{RD}} = \sum T_{\mathrm{iAD}} \quad \text{for } i = 0 \text{ to } X - 1 \tag{2}$$

Using Eq. 1, Eq. 2 can be written as:

$$T_{\mathrm{RD}} = \sum T_{\mathrm{iW}} + T_{\mathrm{iEXE}} \quad \text{for } i = 0 \text{ to } X - 1 \tag{3}$$

Average relative deadline T_{RD} is the arithmetic mean of relative deadlines of all the processes.

$$T_{\mathrm{RDavg}} = \sum T_{\mathrm{iW}} + T_{\mathrm{iEXE}} \quad \text{for } i = 0 \text{ to } X - 1 \tag{4}$$

where "X" is the total processes count having priority higher than the current process in the scheduled queue. Therefore, to determine relative deadline, the average of the higher priority scheduled processes is included. This average value can be used to predict the probability of deadline miss by the process. If the absolute deadline defined for process is less than the average value, then the probability of missing deadline is very high. Similarly, if absolute deadlines for a process are more than the average value, then probability of missing the deadline reduces proportionately to low value. Based on this, it can be concluded that all critical processes should have absolute deadline more than average value for safe operation of system.

Worst case: The worst-case scenario is when all processes scheduled by the application of the system initiate the service request and scheduled for execution. Hard real-time systems are designed to execute all its processes within the deadline under this condition. The processing unit operates at its highest frequency to satisfy the application necessities and meet the required deadlines of every process. The execution speed necessary to satisfy the deadline conditions of all processes depends on the execution time of individual process in queue and their relative deadlines. The total turnaround time T_{T} for least priority process is calculated by cumulative addition of turnaround time T_{Ti} of all processes.

$$T_{\mathrm{T}} = \sum T_{\mathrm{Ti}} \quad \text{where } i = 0 \text{ to max} - 1 \tag{5}$$

In case, the calculated turnaround time T_{T} is greater than the absolute deadline of the least priority process, then the operating speed of the processing unit required to be scaled up. The scaling factor in the processor is proportional to the difference of total execution time and absolute deadline. Therefore, the algorithm should produce worst-case delay as minimum as possible.

$$\mathrm{F_{exe}} \propto \sum T_{\mathrm{Ti}} \quad \text{where } i = 0 \text{ to } N - 1 \tag{6}$$

Average case: The system process load is considered to be average if the number of processes raising the request for service in the application of the system is nearly half the total number of processes. That implies that only some periodic processes are triggered; hence, preemptions probability and number of context switching are moderately less in average load scenario. Hence, the operating speed of the processor need not to be scaled up in order to satisfy the processes deadlines, as the CPU selected during the design will be capable of handling the process load. In this condition, the relative deadlines for most of the processes are less than the absolute deadline.

Best case: The best-case scenario is when there is only one process or priory known background processes are scheduled. Therefore, the total turnaround time of any process will almost be same as its running or execution time. Therefore, both relative and absolute deadlines of such processes will be almost equal. However, in this situation, the utilization factor of the processor might decrease compared to previous cases considered. Such type of scenarios usually arises during sleep mode or standby of multitasking embedded systems. Hence, the operating frequency can be scaled down to save power. Some processing units provide inbuilt option to switch to power-saving mode.

As the objective is to optimize power consumption of the system by scaling the operating frequency in accordance with relative deadlines of processes, the execution time of each process is calculated with standard operating frequency of the processing unit based on which absolute deadline can be estimated. Also execution time can be estimated using a known host and then appropriate scaling factor can be used to estimate execution times on target hardware. Then depending on the number of processes scheduled at any instant of time, a relative deadline can be derived, and operating frequency of the system can be scaled accordingly to meet the defined deadlines for the processes in execution queue.

If embedded system has N number of processes having execution time of T_{iExe} of process X, then the absolute deadlines of all higher priority processes available in execution queue will collectively decide the relative deadline of the process. The time taken by the respective process from the time it arrived in the process queue to completion of execution of process including waiting period T_{W}, IO operations, context switching time T_{CWT} is termed as total turnaround time T_{T} of the process. If the processes are strictly CPU bound without any IO operations, then T_{T} is defined by the following equation.

$$T_{\text{T}} = T_{\text{iExe}} + T_{\text{W}} + T_{\text{CWT}} \tag{7}$$

If all the processes are simultaneously triggered and process queue is at its maximum length, then the highest number of context switching is possible. Even under such situations if process execution time T_{iExe} is very large compared to the total context switching time $T_{\text{CWT,}}$ then T_{CWT} can be ignored for estimating turnaround time.

$$T_{\text{T}} = T_{\text{iExe}} + T_{\text{W}} + T_{\text{CWT}} \tag{8}$$

$$T_{\mathrm{AD}} = T_{\mathrm{W}} + T_{\mathrm{iExe}} \tag{9}$$

The relative deadline T_{RD} is the sum of T_{AD} of every process having priority more than the scheduled process in the queue at any given time.

$$T_{\mathrm{RD}} = \sum T_{\mathrm{iAD}} \quad \text{for } i = 0 \text{ to } N - (X - 1) \tag{10}$$

$$T_{\mathrm{RD}} = \sum (T_{\mathrm{iW}} + T_{\mathrm{iExe}}) \quad \text{for } i = 0 \text{ to } N - (X - 1) \tag{11}$$

Average relative deadline T_{RD} is the arithmetic average of relative deadlines of every process.

$$T_{\mathrm{RDavg}} = \frac{(\sum (N - (X - 1)) + T_{\mathrm{iExe}})}{(N - (X - 1))} \quad \text{for } i = 0 \text{ to } X - 1 \tag{12}$$

where $N - (X - 1)$ are the total number of processes with higher priority in the scheduled queue of processes. Hence, relative deadline calculation is done by taking the arithmetic mean or average of only the processes which are scheduled are considered.

The absolute deadlines are defined for all the processes in the system, and every time the new process enters the execution queue, the relative deadlines of every lower priority process are recalculated. In real-time embedded systems, all processes execution time and resource requirements are priory known based on which the initial static priorities are assigned. These priorities are then changed dynamically depending on the process density and their absolute deadlines.

The frequency scaling technique is used to modify the operating clock period of CPU to achieve relative deadlines less than the absolute deadlines of the respective process. The frequency is scaled up if any process is missing the deadlines to meet the system requirement and similarly frequency is scaled down if processes are expected to complete the job much earlier than required to optimize the power consumption.

4.1 Relative Deadline Equation

Considering N number of processes in system with predefined priorities P_{N} for both scheduling policies discussed in Chap. 1 viz "EDF" and RM, the relative deadline for process X in an execution queue is given by Eq. 13.

$$T_{\mathrm{RD}} = \sum T_{\mathrm{iW}} + T_{\mathrm{iEXE}} \quad \text{for } i = 0 \text{ to } X - 1 \tag{13}$$

where T_{iW} includes context switching time and actual waiting period of process.

4.2 Peak Process Load Condition

A real-time embedded system having maximum of N number of processes having average execution time of T_{exe} is considered for analyzing relative deadline variation under different process load conditions. In operating conditions, the number of processes can vary between standby mode processes to peak load processes. Hard real-time systems are designed to operate at its maximum process load, but the operating frequency can be scaled down just to meet the required deadlines. This technique reduces the power consumption of the system and improves the operating life of the system as well.

The worst-case condition is considered when maximum expected number of processes are available in the process queue for scheduling, i.e., peak load of the system. System is expected to be in such situation occasionally as most systems operate with average load condition or standby mode in some cases. Hence, the operating frequency of the processing unit is scaled to its maximum level to meet the absolute deadlines T_{AD} of each process. The relative deadline T_{RD} of all processes should be less than their respective absolute deadlines.

$$T_{iRD} < T_{iAD} \quad \text{for } i = 1 \text{ to } N \tag{14}$$

The absolute deadline for any process should be defined less than sum of its actual execution time T_{EXE} and its waiting period T_W and context switching time T_C. The context switching time can be ignored if it is very less compared to average execution time of processes T_{EXEAvg} and when a number of processes in the system are considerably less.

$$T_{AD} = T_W + T_{iExe} + T_C \tag{15}$$

But if the process is of much lower priority, then context switching time should be considered as probability of its preemption increases. Also, the number of preemptions will be considerably high to ignore.

The relative deadline T_{RD} of process Pi is based on number of processes with higher priority present in the process queue at time of consideration. T_{RD} is directly proportional to the product of probability of preemption P_P and sum of the execution time of higher priority processes.

$$T_{PiRD} \propto \Pi\left(P_p * \Sigma T_{jexe}\right) \quad \text{for } j = i + 1 \text{ to } N \tag{16}$$

The proportionality constant Δ can be defined as the ratio of execution time of ith process to the sum of the execution time of all higher probability processes.

$$\Delta = T_{iExe} / \Sigma T_{jExe} \quad \text{for } j = i + 1 \text{ to } N \tag{17}$$

Therefore, the relative deadline for any process in the system can be derived as

$$T_{\mathrm{RD}} = \Delta * \Pi\left(P_{\mathrm{p}} * \Sigma T_{\mathrm{jexe}}\right) \quad \text{for } j = i + 1 \text{ to } N \tag{18}$$

The absolute deadline must be compared with relative deadline every time new process enters the queue. The total turnaround time T_{iT} of any process i should always be maintained less than its absolute deadline T_{iAD}, i.e., $T_{\mathrm{iT}} < T_{\mathrm{iAD}}$ which is possible only if relative deadline T_{iRD} is less than sum of T_{iAD} and process load time T_{iL}. The process load time is period between the event trigger of process and process gets listed in process queue, i.e., $T_{\mathrm{iRD}} < T_{\mathrm{iAD}} + T_{\mathrm{iL}}$.

If the execution time of a process is T_{Exe}, then total turnaround time of any process X is given by.

$$T_{\mathrm{T}} = T_{\mathrm{XExe}}(X) + \sum (T_{\mathrm{iExe}} + T_{\mathrm{iW}} + T_{\mathrm{iC}}) \quad \text{for } i = 0 \text{ to } N - (X - 1) \tag{19}$$

4.3 Frequency Scaling Factor

If in any case, T_{RD} of process exceeds its absolute deadline T_{AD}, then frequency of operation must be scaled up, and if T_{RD} is much lesser than T_{AD}, then operating frequency needs to be scaled down. The frequency scaling is proportional to difference between T_{RD} and T_{D} of last process in the process queue.

$$f \propto (T_{\mathrm{RD}} - T_{\mathrm{AD}}) \tag{20}$$

If $(T_{\mathrm{RD}} - T_{\mathrm{AD}})$ is positive, then frequency is scaled up, and if it is negative, then frequency is scaled down. The scaling factor is calculated as the ratio of difference between the relative and absolute deadline to relative deadline achieved. In multitasking, the process having highest scaling factor is considered.

$$k = (T_{\mathrm{D}} - T_{\mathrm{RD}})/T_{\mathrm{RD}} \tag{21}$$

Hence, the new required frequency of operation f_{C} required to meet the deadline is given by Eq. 22.

$$f_{\mathrm{C}} = k * f \tag{22}$$

The frequency scaling algorithm involves following steps.

1. Identify the processes completing the execution within scheduling period but missing the deadline.

2. If all scheduled processes are meeting the deadline and some processes are left out without getting even single slot, then frequency is arbitrarily scaled up by 25% and checked for updated status.
3. The frequency scaling factor is calculated for all processes completing the execution but missing the deadline.
4. Select the highest factor and scale the frequency accordingly.
5. Check if all the processes are meeting the deadline using revised frequency.
6. If "YES," then same frequency is maintained for application; otherwise, the cycle repeats from step i till all processes meet the deadline requirements.
7. The lowest operating frequency is computed such that at least one process violates the deadline below that frequency.

The system requires an additional programs to monitor the change in turnaround time due to dynamic variation in number of processes and to calculate the scaling proportion of operating frequency. These processes are to be active throughout the operating period of the system. The additional overhead introduced due to these background process to monitor this operation also should be considered during design cycle.

5 Conclusion

The proposed technique is suitable for systems with sufficiently large set of processes. Hence, this technique is efficient for system with relatively large number of processes which usually the case in hard real-time systems. The power consumption of the system also reduces to great extent which in turn provides longer life for the system and better thermal stability. As the frequency scaling is decided based on deadline analysis, the system can operate at much lower frequencies if deadlines are defined sufficiently large. This method is very useful in defense, industrial critical systems and space applications where deadlines are stringent and system is designed for their peak load. Whenever system is operating with fewer processes, automatically frequency can be scaled down for effective utilization of resources.

The background processes may have negative impact on performance of system if system is having very few set of processes. Also, additional design challenge is faced in scaling the frequency of operation of processing unit.

References

1. Zhang, F., Burns, A., Baruah, S.: Sensitivity Analysis of Arbitrary Deadline Real-Time Systems with "EDF" Scheduling. Springer Science Business Media, LLC, Berlin (2011)
2. Ra, O., Torbjo: Trends and Challenges in Embedded Systems—CoDeVer and HiBu Experiences. Corpus ID: 15066429 (2003)

3. Balbastre, P., Ripoll, I., Crespo, A.: Minimum deadline calculation for periodic real-time tasks in dynamic priority systems. IEEE Trans. Comput. **57**(1), 96–109 (2007)
4. Lee, S., Kim, H., Lee, J.: Soft Aperiodic Task Scheduling Algorithm in Dynamic-Priority Systems. KOSEF under Grant No. 941-0900-046-2. IEEE (1995)
5. Spruntlui, B., Sha, L., Lehoczky, J.: Aperiodic task scheduling for hard-real-time systems. J. Real-Time Syst. **1**, 27–60 (1989)
6. Chen, C.-Y., Hasan, M., Ghassami, A.E., Mohan, S., Kiyavash, N.: Securing Dynamic-Priority Real-Time Systems Using Schedule Obfuscation. arXiv: 1806.01393v1 [cs.CR] (2018)
7. Morton, A., Loucks, W.M.: "EDF" Feasibility Analysis of Accelerated Tasks. University of Waterloo, Waterloo, Ontario, Canada. IEEE (2007)
8. Kargahi, M., Movaghar, A.: A method for performance analysis of earliest-deadline-first scheduling policy. In: Proceedings of the 2004 International Conference on Dependable Systems and Networks (DSN'04) 0-7695-2052-9/04 © 2004 IEEE
9. Morton, A., Loucks, W.M.: "EDF" Feasibility Analysis of Accelerated Tasks. 0840-7789/07©2007 IEEE
10. Short, M.: Improved Schedulability analysis of implicit deadline tasks under limited preemption "EDF" scheduling. In: IEEE ETFA'2011
11. Zhang, F.: Schedulability analysis for real-time systems with "EDF" scheduling. IEEE Trans. Comput. **58** (2009)
12. Lehoczky, J.P., Sha, L., Ding, Y.: The rate-monotonic scheduling algorithm: exact characterization and average case behavior. In: Proceedings of the IEEE Real-Time Systems Symposium, pp. 166–171 (1989)
13. Bini, E., Buttazzo, G.C., Buttazzo, G.M.: A hyperbolic bound for the rate monotonic algorithm. In: 2001 Proceedings of the 13th Euromicro Conference on Real-Time Systems. Delft, the Netherlands, pp. 59–66
14. Buttazzo, G.C.: Rate Monotonic vs. "EDF": Judgment Day. Springer Science Business Media, Inc., The Netherlands (2005)
15. Cervin, A., Eker, J., Bernhardsson, B., Arzen, K.-E.: Feedback-feedforward scheduling of control tasks. Real-Time Syst. **23**(1), 25–53 (2002)

Influence of Customer Relationship Management for the Success of E-Business

Amala Siby and Jossy P. George

Abstract Customer relationship management has recently been one of the key factors in the success of many organizations. Organizations have realized the importance of customer satisfaction and are integrating their operations with that of customer relationship to serve the customers in a better way. This paper seeks to understand the importance of CRM in e-business. It also talks about the importance of customer relationship for an organization in its growth. Relationship marketing has been studied to show how customer relationship management software can be made use of for the benefit of an organization.

Keywords E-commerce · Retailing · Online marketing · Digital advertisements · B2B commerce

1 Introduction

The emerging trends in the way business is done these days are vast and changing every day. E-commerce is being picked up by almost every organization that start their business [1]. E-commerce helps in establishing relationships between the customers and the organization, and the most commonly used application to do so is the customer relationship management software. It is a software that helps an organization to make effective decisions regarding their business, and it is based on the customer responses that are collected and analyzed through detailed consideration. It helps an organization to understand what its customers are expecting from the company, so that they can make changes accordingly and improve their current services. Electronic CRM plays a vital role in helping an organization to understand the tastes and preferences of the customers.

A. Siby (✉) · J. P. George
CHRIST (Deemed to be University), Bengaluru, India
e-mail: amala.siby@res.christuniversity.in

J. P. George
e-mail: frjossy@christuniversity.in

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_45

Customer relationship management influences the success of e-business in various ways. Customers are the primary reason for the success of any organization; an organization must manage the relationship with its customers in the best way. Customer satisfaction is essential for an organization to grow. If the customers are unhappy, it can lead to various problems for an organization. The factors that can influence customer relationship is the way the organization approaches its customers in terms of the kinds of service they offer, which in turn reflects on the success of the e-business. Studies have shown that organizations which have good customer relationship are more successful in the present scenario. These organizations tend to retain their customers and not lose them to other organizations. ECRM helps in building customer loyalty, which is very difficult in the current market with a lot of competition [2].

2 Literature Review

Reviewing the articles and journal papers written by several well-established authors regarding e-business and its influence on the customer relationship management, this helps understanding the different perspectives of different authors on the subject of how customer relationship management has an impact on e-business in various areas.

2.1 E-Business and CRM

Business architecture has an impact on customer relationship management. Small businesses can have an advantage over their large competitors by making use of the business architecture where customer relationship is given almost attention. In the case of larger enterprises, the amount of attention given to customers is minimal, and hence, being a small-scale organization can prove to be an advantage during the growth process of any organization which will benefit in the long run. This research helps in understanding the networking aspect of an organization through an e-business platform [3].

2.2 E-commerce Opportunities and Challenges Using CRM Software

The concept of e-commerce and how a business is carried out in the digital platform where customers are exposed to goods at a lesser price with much efficient customer handling mechanisms is the greatest opportunity for e-commerce. There are various opportunities available for organizations getting into the internet mode of operating

a business, provided that they are ready to face challenges in the process of shifting from their traditional way of doing business. With the changing trends in technology, the brick and mortar way or the traditional way of conducting business will soon come to an end. CRM software manages the customers with much efficiency in the shifting stages of business from traditional platforms to online platforms [4]. Customer relationship has an impact on the development of an organization. The customers are the main reason behind the success of any organization without whom the business would not go further; the customer relationship management proves to be an important tool that helps an organization to understand customer tastes and preferences through lots of quantitative and qualitative data that helps in the process of decision making. With the increasing number of Internet users every day, the e–marketplace has a lot of scope and opportunities for the customers who engage in e-commerce [5].

2.3 Successful CRM Systems

The launching of customer relationship management software is significant as it has a massive influence on the shareholders' value and the customers' value. Most of the organizations are implementing CRM applications in their business; an organization needs to check the way the application is responding toward the industry by checking the effectiveness of the CRM software deployed by doing a compatibility check. The factors that contributed towards the success of the CRM software are reduced chaos, business goals drive functionality and alignment of all department strategies, etc. [6].

The customer behavior in an online platform can be traced using CRM software, which can further be used as data sources for the companies to build in their customer base [7].

3 Relationship Marketing

The integration between various elements of retail chain is one important characteristic of relationship marketing. The different elements of supply chain can efficiently be integrated to bring the maximum efficiency, and this can be achieved through the integrated functioning as shown in Fig. 1.

E-business in the present scenario operates on the same nature such as this to manage its operations on a digital platform. E-business employs a customer relationship software which will help it in having a perpetual relationship with its customer. Figure 1 talks about how an organization integrates its business with its operators that includes the manufacturers, the resellers and the consumers to manage relations so that it can earn maximum revenue from their integrated functioning.

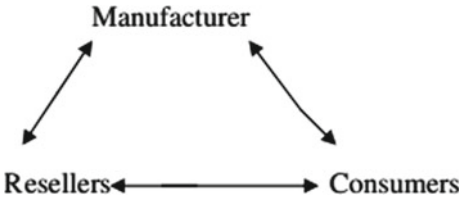


Fig. 1 Relationship marketing integration

4 Methodology

The qualitative research focuses majorly on the literature review understanding of CRM use for e-business. The rapid increase of CRM consultants and literature reviews on the increased use of CRM make business houses run into the ill-considered malfunctioning CRM programs [8].

4.1 Objectives of Customer Relationship Management

- To know the customers better and hence increasing the margin of sales.
- To enjoy the benefits of higher customer satisfaction by dealing in with more customer-specific needs.
- Reduction in general sales and other administrative costs.
- Improving the process of communication with right consumers.
- Creating customer value through customer retention and building of virtual community.

Since the research is mostly done on a qualitative scale where quantitative data has not been used to make any inferences regarding the subject matter of study at hand. The qualitative data has helped in understanding consumer behavior through customer relationship management, and how it has helped build the e-business on a digital platform. The research also revolves around an exploratory styles information in this paper which is gathered by scrutiny by referring works of different authors on the same study that is the influence of customer relationship management for the success of e-business. Since the CRM and its influence is vast on the e-business no clear study can be made about the subject, and hence, exploratory style has been used to understand this concept to a certain extent possible.

5 Customer Relationship Management Software

Customer relationship management is a software which manages relationships with present and prospective future customers by organizing and coordinating the sales

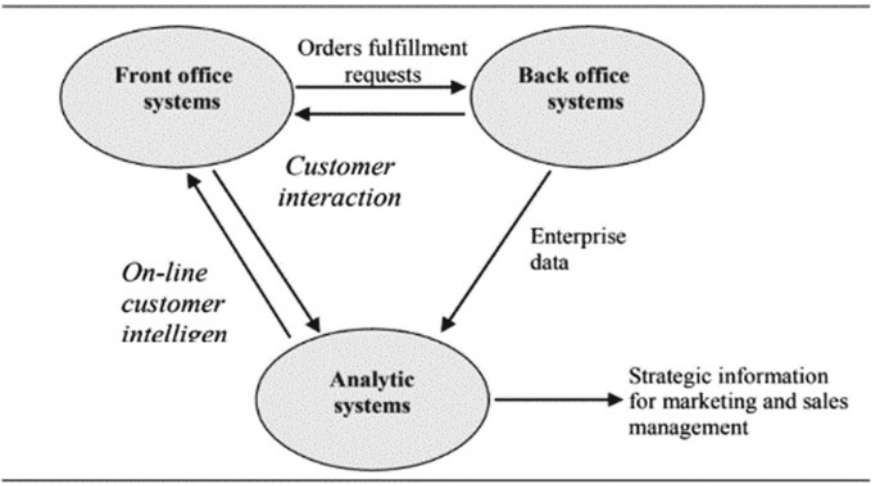


Fig. 2 Virtuous triangle of CRM

and the marketing activities of the organization providing the best of customer and technical support to its consumers [9].

The CRM software integrates the multiple channels in which the customers are dealt with by allowing a constant flow of information through the analytical softwares [10].

As shown in Fig. 2, the integration between back office, front office and system is dealt smoothly with the help of CRM software. It helps in identifying the patterns in which tactical information can be passed on to determine future policies of business with the customer interactions [11].

5.1 Scope of Customer Relationship Management

The wider scope of customer relationship management is explained in the below diagram.

Figure 3 elaborates on the wider scope of CRM into multidisciplinary aspects. The advantages that CRM offers to people are wide ranged from offering a better customer service and increase the revenues of the customers, finding new customers for the business, helping staff for better deal management, simplification of overhead costs of business. These do not imply that the software is flawless. The affordability of software with regard to cost of training and overhead costs is still an issue of consideration [12].

Fig. 3 Scope of customer relationship management

6 Case Study on CRM Practices of Leading Organizations

CRM influence on the success of e-business in organizations such as Apple, Amazon and BMW is discussed below:

6.1 CRM in Amazon

Amazon is an online e-commerce giant with world-class CRM practices. Any irregular customer also returns to the alluring world of Amazon because Amazon knows its customers more than anybody else knows [13].

The integration of CRM activities in Amazon happens in three levels: personal data storage, user interface and customer interaction all with the help of its CRM software. The neatly organized user interface and adequate, appropriate, customized and timely search results also play a huge role in the success of CRM in Amazon. Customer interaction and expanding the horizons of its products to live streaming, payment and media services catering to the needs of customers are a major reason behind the success of Amazon [14].

6.2 CRM in BMW

BMW as a brand has established its existence through a premium, luxury car with high quality. The advantage BMW has created over its automobile competitors is the customized quality and unique designing of its products. The exclusive communication strategy adopted by them in communicating all the new products to its existing consumer base is a commendable factor. The customer retention policy of BMW by way of managing the need of customers by offering services like opportunity to change or sell old BMW with a new one has attracted a wide range of customers and helped in retaining the existing ones too [15].

6.3 CRM in Apple

The tech giant has always kept its customer first. Steve Jobs, the founder, had his vision set upright 'start working with the customer experience and work backwards to the technology.' The targeted marketing strategy along with the self-renewing limitless data sets give an upper hand for apple to customize its products and services to its customers [16].

The high-end CRM strategies are achieved by Apple in the following manner:

1. Knowing its Customers
   The knowledge Apple has about its customers is commendable. Respecting privacy of its customers, Apple knows what are the individual expectations of each and every consumer are. Knowing the real needs of customers even before they realize, it is the sophistication in technology which Apple has achieved as a brand.
2. Anticipate Customer Needs
   Innovation is the strategic advantage which Apple has over its competitors. Products that were much ahead of times were always Apple trademarked. Know your customer before they know thyself is the famous slogan which Apple always upheld.
3. Stores Focused on Customer Experience
   The store design and customer service provided by Apple stores are irreplaceable. A casual visitor to the store also will never forget the impact an Apple store would have created in his mind. By foregoing the conventional store concept, Apple would by then have penetrated into the customer minds [17].

Branding that lasts for generation is Apple's uniqueness. Owning an Apple product is a pride for many of its users and potential customers. This is an achievement which the brand has created using its extensive management of its CRM.

7 Findings

The organizations that were taken for study here have similar objectives with respect to their CRM objectives. The most major objectives include cost savings, revenue enhancement and strategic impact. Also, the literature reviews emphasized on the fact that companies do not wish to outsource their CRM, instead they like to retain it as internal operation only.

The variables that were identified through extensive literature survey in managing the customer relations in the business through CRM software's are as follows: a) top management commitment, integration of systems in CRM, b) customer behavior prediction and c) event-based marketing strategies.

The variables that were found relevant to the organization in order to evaluate the effectiveness of CRM in the success of e-business included as follows: (a) Recording customer information, (b) analysis of customer information using software, (c) integration of business process, (d) formulating customer retention policies and (e) extending support services.

8 Conclusion

The usage of CRM for organizations is the best way to improve their business to make maximum revenue. It is found that organizations with CRM applications implemented have a better chance at perpetual succession compared to the ones that do not maintain a steady customer relationship which was clearly understood from the case studies of Amazon, Apple and BMW. The importance of customer relationships in the present market dynamics is highly relevant as there is a lot of competition between the organizations to be the market leader. In the case of Apple, Amazon and BMW, we understood how these companies employ CRM for their benefit to maintain a steady and good relationship with its customers for the purpose of enhancing sales that will invariably help in the growth of the organization in the long run. Scope of further study can be extended to the variables contributing toward the success of CRM from a consumer perspective. Analytical studies can be conducted in order to substantiate the variables identifies in the study.

References

1. Anshari, M., Almunawar, M.N., Lim, S.A. and Al-Mudimigh, A.: Customer relationship management and big data enabled: personalization & customization of services. Appl. Comput. Inform. 94–101 (2019)
2. Bannan, K.J.: Customer relationship management, 1 July 2001. (Online). Available: https://www.pcmag.com/article2/0,2817,17174,00.asp. Accessed 26 Jan 2018
3. Dhillon, R., Kaur, S.: Review paper on e-business & CRM. Int. J. Comput. Bus. Res.

4. Amith Vikram, M.A.: E-commerce: opportunities and challenges. In: Bangalore Conference (2012)
5. Sayed, S., Khan, A., Syed, N.: CRM-centric E-commerce web application. Int. J. Eng. Res. Gener. Sci. (2015)
6. Wilson, H., Daniel, E., McDonald, M.: Factors for success in customer relationship management (CRM) systems, 2002. (Online). Available: https://www.tandfonline.com/doi/abs/10.1362/0267257022775918. Accessed 30 Jan 2018
7. Usman, U.M.Z., Jalal, A.N., Musa, M.A.: The impact of electronic customer relationship. Int. J. Adv. Eng. Technol. (2012)
8. Kanagal, N.: Role of relationship marketing in competitive marketing strategy. J. Manage. Market. Res. 2
9. Tutorials Point: Customer relationship management, (Online). Available: https://www.tutorialspoint.com/management_information_system/customer_relationship_management.htm. Accessed 21 Jan 2018
10. M.J. Kaur, Customer relationship management: a study of. Glob. J. Financ. Manage. 155–158 (2016)
11. David Bradshaw, C.B.: Managing customer relationships in the E-business world: how to personalise computer relationships for increased profitability. Int. J. Retail Distrib. Manage. **29**(12), 520–530 (2001)
12. Bashar Shahir Ahmed, F.A.M.B.M.: The intelligence of E-CRM applications and approaches on online shopping industry. In: Advanced Methodologies and Technologies in Digital Marketing and Entrepreneurship, pp. 70–82 (2019)
13. Niharika Mishra, S.M.: Effect of artificial intelligence on customer relationship management of amazon in Bangalore. Int. J. Manage. **10**(4), 168–172 (2020)
14. Wadhwa Bharti, W.B.P.N.: Business model of amazon India—a case study. J. Market. Manage. Res. **10**(1), 32–40 (2020)
15. David Cosgrave, M.O.: Ethical standards and perceptions of CRM among millennial consumers. Int. Market. Rev. (2020)
16. Tawfik Jelassi, F.J.M.-L.: Overview of the e-Business strategy framework. In: Strategies for E-Business, pp. 35–48 (2020)
17. Hollebeek, L.D., Srivastava, R.K., Chen, T.: S-D logic–informed customer engagement: integrative framework, revised fundamental propositions, and application to CRM. J. Acad. Market. Sci. 161–185 (2019)

Implementation of Different Methods for Decomposing the Rhythms of EEG Signal

Snehal Dongare and Dinesh Padole

Abstract Brain signals reflect the different psychological states and conditions of the human being. These signals can be recorded using the sensing mechanisms like EEG. EEG is the group of frequencies that ranges from 0.1 to 64 Hz. These bands indicate different mental states and activities. Hence, on separating these bands, particular signal pattern can be identified for the selected activity. Also, the band separation can be used for removing the noise from EEG signals. This paper focuses on different methods for separating EEG signal rhythms. In this paper, temporal filtering, wavelet-based filtering and empirical mode decompositions are used to realize the separation of EEG signal bands. From the selected methods, it is found that EMD is more promising.

Keywords EEG signal · Frequency bands · Power in the EEG bands

1 Introduction

The electroencephalograph signals are the mostly used signals for evaluation, detection and categorization of different conditions of brain. These EEG signals are non-stationary and nonlinear in nature. Its band division is given as—Delta (0.1–3 Hz), Theta (3–7 Hz), Alpha (7–12 Hz), Beta (12–30 Hz) and Gamma (30–64 Hz) [1–3]. Different mental activities carried out by the particular user can be represented with the help of these frequency bands [2, 3]. Human eyes cannot directly analyze the hidden information embedded inside EEG signals. Hence, in order to understand the brain functionality of the human, it is required to find the changes in these signals [1]. These bands indicate different mental states and activities. Hence, it is essential to separate out each of the band. Significance of EEG signal bands is shown in Table 1 [3].

To reveal the information out of EEG signal, it is needed to separate bands in the signal and then focus on the particular band of the interest. This information can

S. Dongare (✉) · D. Padole
G H Raisoni College of Engineering, Nagpur, India

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_46

Table 1 Details of EEG signal bands

Brain waves	Frequency range (Hz)	Mental activities
Delta	0.1–3	Defines the depth of sleep
Theta	3–7	Generally during dreams
Alpha	7–12	Alert and relaxed
Low beta	12–15	Musing, integrated
Mid beta	16–20	Thinking
High beta	21–30	Thoughts, alertness
Gamma	30–64	Motor functions, higher mental activity

further be used to perform some dedicated task so as to assist the human being. The goal is to separate the bands of EEG signals [1, 4]. Different methods are available for separation of bands in the EEG signal. This article proposes comparative study on of various techniques that can be used for separating the different EEG bands like temporal filtering, empirical mode decomposition and discrete wavelet transform.

This paper is framed as follows: In Sect. 2, basic steps for analyzing EEG signal are explained. Section 3 gives the details of methods considered for decomposition of EEG signal. Section 4 gives the details of experimentation and results. Conclusion is provided in Sect. 5.

2 Basic Steps in EEG Signal Analysis

For the analysis of the EEG signal, basic steps involve acquisition of the EEG, decomposition of the EEG, then features extraction and the classification. Proper analysis of the EEG depends on the band selected for the feature extraction, since different bands convey the different information. Hence, in order to extract the particular information, particular band of frequency is important. This will give us the more accurate information regarding the brain activity. So, we should focus on well-defined method for decomposing the EEG signal bands. Various methods are available for the decomposition of the signal bands. Some of the methods used for study are given in the following section (Fig. 1).

3 Methods Considered for Decomposition of EEG Signal Bands

EEG signal is non-stationary and dynamic in nature. It contains lots of information in time, frequency and spatial domain. In order to separate the EEG signal bands, the

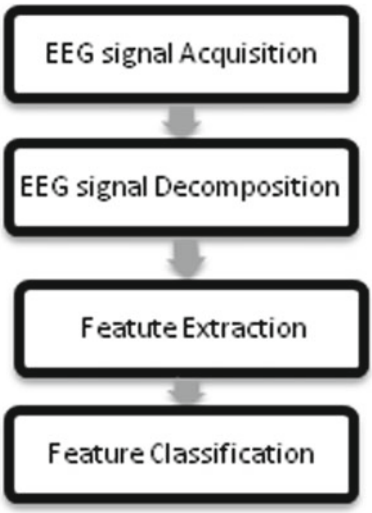


Fig. 1 Basic steps for EEG analysis

methods considered are temporal filtering, discrete wavelet transform and empirical mode decomposition (EMD).

3.1 Temporal Filtering

Temporal filtering is the method to remove the unwanted frequencies within the raw EEG signal [5]. Frequency bands in the EEG signal are predefined. Hence, filtering method can be directly applied on the EEG signals. Bandpass filters can be used effectively for filtering the signal in the desired range. Butterworth bandpass filter is best suited for selecting the frequencies of the interest [5]. It only passes the frequencies of interest and attenuates the unwanted frequencies. The frequency ranges for various bands of the EEG signal are known, and hence, pass band frequency and stopband frequencies can be assigned to select the desired band [5]. Design flow used for temporal filtering is shown in Fig. 2.

3.2 Discrete Wavelet Transforms (DWT)

EEG signals are non-stationary and carry the information in both the frequency and time domain. Wavelet transform is the one of the methods to achieve this kind of separation [6]. In order to process the signals, Daubechies 4 (db4) wavelet can be used as the mother wavelet. Depending on the sampling rate of the input signal, different level of the decomposition can be applied to the EEG signal to grab the required

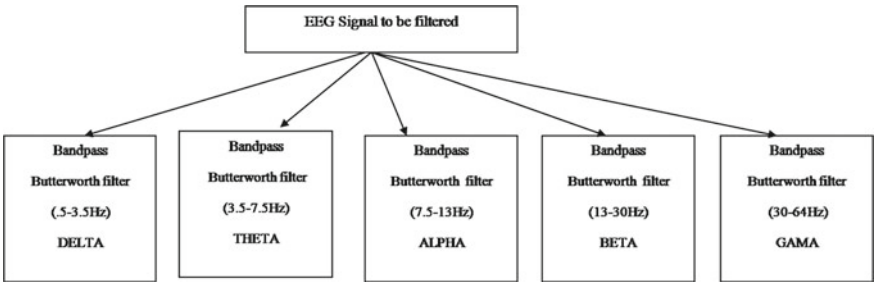


Fig. 2 Band separation using temporal filtering

Table 2 Detail coefficients after eight-level wavelet decomposition

Coefficients of wavelet	Frequency band (Hz)	Signal information
D8	0–4	Delta
D4	32–63	Gamma
D5	16–32	Beta
D6	8–16	Alpha
D7	4–8	Theta
D1	250–500	Noise
D2	125–250	Noise
D3	63–125	Noise

range of EEG signal between 0 and 60 Hz. The input signal used is sampled at the frequency of 512 Hz. In order to cover the entire frequency band of 0–60 Hz and to be separated up to the lowest delta frequency band, eight-level signal decomposition has to be used. It is shown in Table 2. Coefficients D1, D2 and D3 can be removed as it is out of the band [6].

DWT can be used to acquire both the temporal and frequential information in the EEG signal.

3.3 Empirical Mode Decomposition (EMD)

Signal decomposition explained above is all based on linear and stationary assumptions. EEG signal is nonlinear and non-stationary, and moreover, it is susceptible to the noise due artifacts. Using EMD algorithm, the original EEG signal can be decomposed into the series of intrinsic mode function. It also removes the noise, and information is represented by the different IMF [6]. The dynamic EEG signal can be effectively represented by EMD since EMD has adaptive basis function. The signal is decomposed into the IMF from which we can find the most significant band.

Hilbert transform can be applied on the intrinsic mode function to obtain instantaneous frequency. It shows better temporal resolution [7]. In EMD, the original signal $x(t)$ can be expressed as summation of an IMF and the residue as shown below [7, 8]

$$\mathrm{x}(t) = \sum_{i=1}^{i=n} ci(t) + rn(t) \tag{1}$$

where

n = number of IMF

$ci(t)$ = IMF

$rn(t)$ = Residual.

4 Experimentation and Results

The experimentation is carried out on the mentioned dataset using Open-Vibe and EEGLAB platform.

4.1 Dataset Used for Experimentation

The experiments are carried out on the dataset available at the Open-Vibe platform [9]. It consists of recordings regarding their imagination of the left- and right-hand movements. It is recorded by Mind media NeXus32B amplifier at the sampling frequency of 512 Hz [3, 6]. Signal pattern of the above dataset at a particular moment is shown in Fig. 3.

4.2 Scenario and Output Windows for Band Separation

Temporal Filtering
The EEG Signal is decomposed into various frequency bands using temporal filtering. Butterworth bandpass filter of the order 4 is used for experimentation. The Open-Vibe scenario for extracting band is shown in Fig. 4.

The various EEG band extracted along with the original EEG signal at that moment is shown in the following window (Fig. 5).

Wavelet Decomposition
The scenario for EEG signal band decomposition using discrete wavelet transform is shown in Fig. 6.

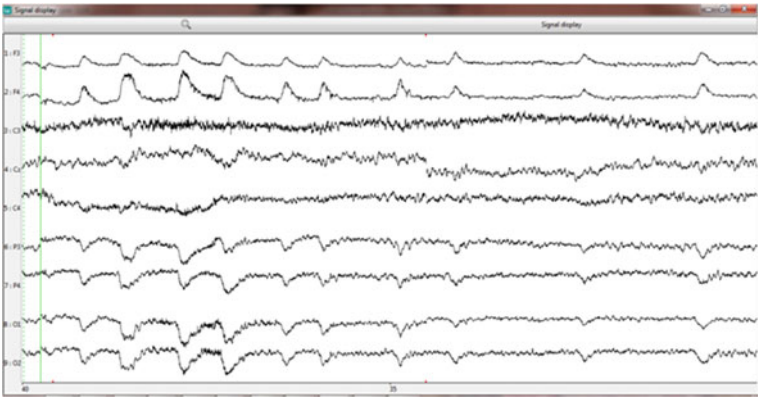


Fig. 3 Output window of original EEG signal

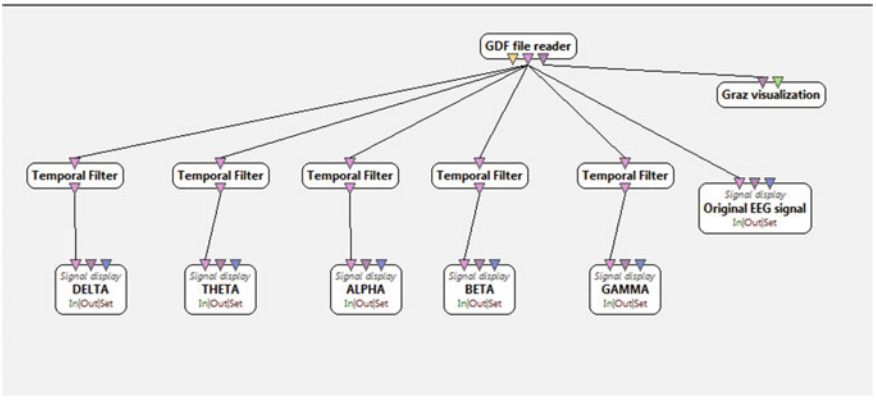


Fig. 4 Band separation scenario of temporal filtering

The output of the decomposition is shown in the following window (Fig. 7).

Empirical Mode Decomposition

The empirical mode decomposition is implemented in the EEGLAB on the above-mentioned dataset. The original signal is decomposed into following seven IMF which corresponds to the different frequency bands. First IMF can be neglected since it contains noise and beyond the required band (0.1–64 Hz) (Fig. 8).

5 Conclusion

This paper explores the different methods for separating the EEG rhythms. Proposed methods automatically decompose the EEG signals into various sub-bands. EMD

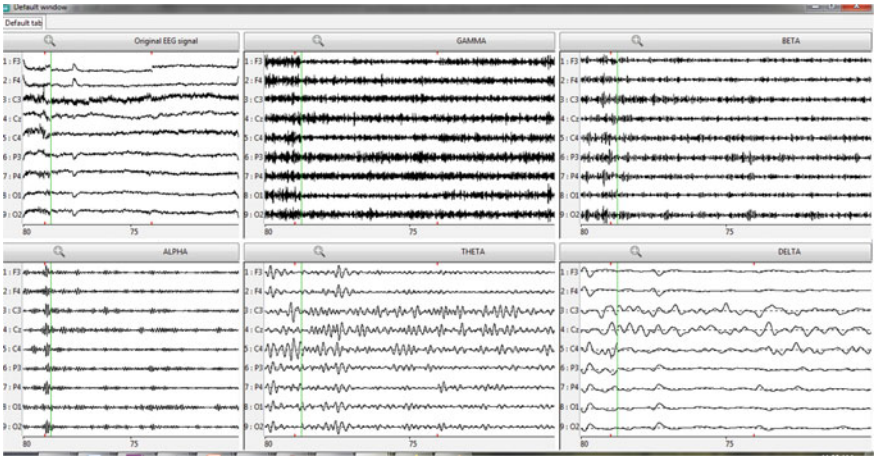


Fig. 5 Band separation using temporal filtering

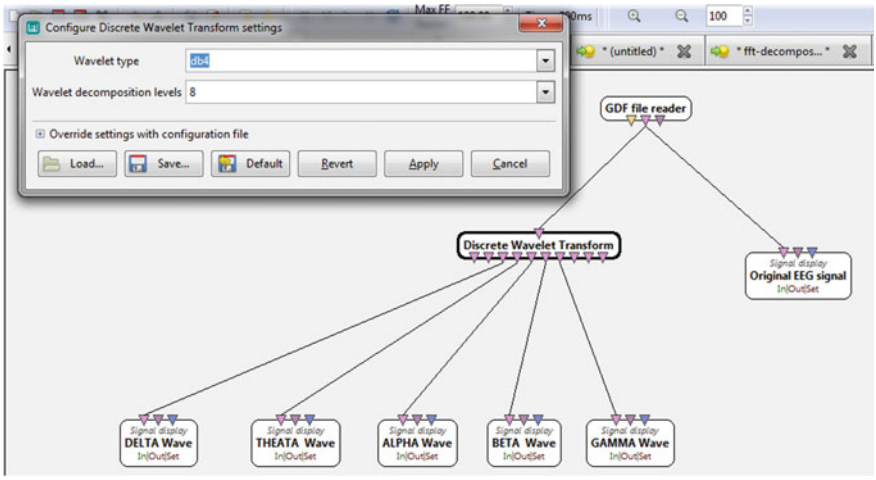


Fig. 6 Band separation scenario of wavelet transform

method preserves the nonlinear ad non-stationary features of the EEG signal while removing the noise due to the artifacts. Band power from each of these methods can be calculated to find the best suited band for particular activity. Hence, the detection of particular activity can be improved by selecting the desired EEG band.

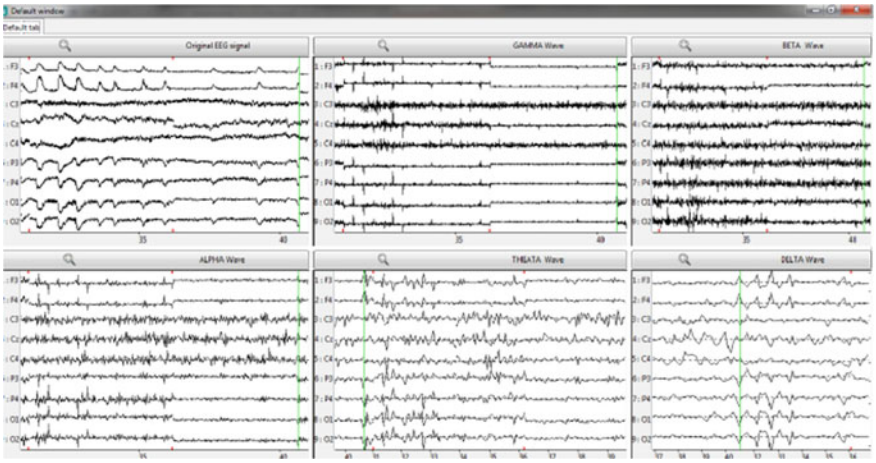


Fig. 7 Band separation using wavelet transform

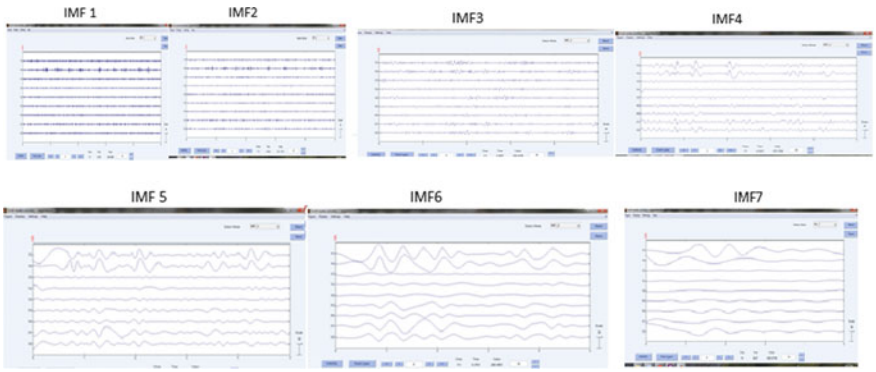


Fig. 8 Signal decomposition into the IMF

References

1. Bajaj, V., Pachori, R.B.: Separation of rhythms of EEG signal based on Hilbert Haung transform with the application to Seizure Detection. In: ICHIT2012, LNCS 7425, pp. 493–500
2. Dongare, S., Padole, D.: Development of feature-based hybrid method for brain signature Identification. In: 9th International Conference on Emerging Trends in Engineering and Technology, 2019. https://doi.org/10.1109/ICETET-SIP-194685.2019.9091985
3. Telphan, M.: Fundamentals of EEG measurement. Meas. Sci. Rev. **2**(2) (2002)
4. Zhong, J., Shuren, Q., Chenglin, P.: Study on separation for the frequency bands of EEG signal and frequency band relative intensity analysis based upon EMD. In: International Conference on Signal Processing, Robotics and Automation (ISPRA '08), 20–22 Feb 2008. ISSN: 1790-5117 151, ISBN: 978-960-6766-44-2

5. Singla, E.M., Singh Guru, M.H.: Paper on frequency based audio noise reduction using butter worth, Chebyshev & Elliptical filters. Int. J. Recent Innov. Trends Comput. Commun. **3**(10), 5989–5995. ISSN: 2321-8169
6. Cheong, L.C., Sudirman, R., Hussin, S.S.: Feature extraction of EEG signal using wavelet transform for autism classification. ARPN J. Eng. Appl. Sci. **10**(19) (2015)
7. Zhang, J., Wei, J., Liu, X., Wu, C., Wang, Y.: A novel application of empirical mode decomposition (EMD) to feature extraction of epileptic EEG. 1473–8031. https://doi.org/10.5013/IJSSST.a.17.29.39. ISSN: 1473-804x
8. López, M.B., Giraldo, E., Molinas, M.: Analysis of neural activity from EEG data based on EMD frequency bands. https://doi.org/10.1109/ICECS.2017.8292116
9. https://openvibe.inria.fr/datasets-downloads/, 2019

Policy-Approximation Based Deep Reinforcement Learning Techniques: An Overview

Mohit Sewak, Sanjay K. Sahay, and Hemant Rathore

Abstract Until recently, Deep Reinforcement Learning was restricted to innovations in games like Atari, Dota2. Despite surpassing the benchmarks established by their human counterparts in multiple games, these methods could not scale to real-life and industrial automation tasks. The main reason for this was the essential requirement of complex and continuous action control and sophisticated physics of the domain involved in these tasks. Because of these reasons, most of the incumbent solutions for such applications involved the invent of custom planning algorithms. The design of such sophisticated custom solutions required complete knowledge to the dynamics of the domain and its derivatives and hence were not scalable. Policy-based DRL has democratized this space, as now deep reinforcement learning agents could be trained to learn similar sophisticated policies just by learning from the data generated by interacting with these systems or their respective simulations. This has led to significant innovations in real-life and high-value control automation applications like autonomous vehicles, drones, and industrial robots. Therefore, in this paper, we present an overview of different types of policy-approximation based technique in Deep Reinforcement Learning that are the basis of many advanced control automation systems.

Keywords Deep reinforcement learning · Artificial intelligence · Deep learning · Policy approximation

M. Sewak (✉) · S. K. Sahay · H. Rathore
Department of CS&IS, BITS Pilani, Goa, India
e-mail: p20150023@goa.bits-pilani.ac.in

S. K. Sahay
e-mail: ssahay@goa.bits-pilani.ac.in

H. Rathore
e-mail: hemantr@goa.bits-pilani.ac.in

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_47

1 Introduction

The objective of any Reinforcement Learning (RL) agent is to learn an optimal action policy to maximize reward against a particular environment that represents a Markov Decision Process (MDP) [1]. The environment represents multiple states to the agents in sequence which may at least be partially influenced by the actions taken by the agent. This could be a challenging problem depending upon the complexity involved in the state representation or the type of control required. Of late, RL was used mostly for problems which deals with simple state that could be easily represented in simple vectors. But with the invent of algorithms like the Deep Q Networks (DQN), [2] this changed. In DQN, RL was combined with Deep Learning (DL) [3]-based value approximators-based upon Convolutional Neural Networks (CNN) [4] that could directly work on image and video frames to learn complex action policies. Another application can be malware detection [5–7] where same mechanism can be visualized and used for achieving better results. Such algorithms and techniques that combine DL with RL comes are collectively known as Deep Reinforcement Learning (DRL) [1]. Though DQN could solve the problems related to complex state, it could not work well under conditions that in addition also required complex tasks involving large action space or continuous action control. There are multiple variants of the DL-based Q learning algorithm [8, 9], each fall under the family of algorithms known as value approximation-based DRL [10]. All the algorithms under the value approximation-based DRL suffer from similar drawbacks. This is where a more recent advancements called the policy-approximation based DRL shines.

In *Industry 4.0*, many processes require complex control over large and often continuous action spaces. For such types of Reinforcement Learning, policy approximation-based approaches provide better results [11–13], to empower real-life applications like autonomous vehicle [14] to complex robotic automation [15–17]. Until recently, DRL was restricted only to innovations in games (like Google's Alpha-Go [18] and Open AI's Dota 2 [19]); and despite surpassing the benchmarks established by human counterparts in multiple games, these methods could not scaled to real-life and industrial automation tasks because of the inherent drawbacks. In this policy-based Reinforcement Learning, instead of learning an approximator for the value function (so as to indirectly determine an optimal policy), the policy is learnt/approximated directly [1]. With the combined capabilities of deep learning and policy approximation-based techniques, the resulting agent could learn sophisticated policies from nonlinear observations [20]. Therefore, recently, such approaches have become very popular in designing automated control solutions for complex applications in Industry 4.0.

In DRL, value-based RL techniques fail to efficiently learn effective action policies for complex control scenarios in sophisticated domains that require continuous action control [11]. This drawback becomes significant in many real-life application ranging from autonomous vehicles to complex robotic automation where understanding and control of complex domain mechanics involves sophisticated physics [21, 22]. In the recent past, any effort to implement a decent control in such areas invariably required designing a custom planning/optimization solutions [23].

To design such bespoke solution, an intricate understanding of the physics and domain of each of its control and their respective interactions is required. This approach is obviously not scalable, and hence, the application of such techniques was restricted to a few high-value control automation problems [24]. With the advancements in policy-based RL techniques, this space has recently been democratized, as these techniques could automatically learn complex control policies using the data generated from interaction with the system/environment and obviate the need to invent complex custom algorithms after thoroughly understanding the complete domain and application intricacies [25]. The recent advancements in this field have also led to the invent of the algorithms that could guarantee monotonic improvements in the learnt policy with increasing data [12] and have also made the learning mechanism simple to implement and highly sample efficient to learn [13]. Therefore, in this paper, we cover these techniques as these have potential for application into many real-life sophisticated applications.

In Sect. 2, we cover the policy approximation methods and the sub-categorization of the different agents that use policy approximation methods. We also cover the details of the Stochastic and Deterministic policy gradient in this section. Next, in Sect. 3, we cover the different agents that are based upon the Stochastic Policy Gradient mechanism. Then, in Sect. 2.3, we cover the different agents that are based upon Deterministic Policy Gradient mechanism. Finally, in Sect. 5, we conclude the paper.

2 Policy Approximation Methods

The ultimate goal of an RL agent is to find an optimal (action) policy. Therefore, taking an indirect route of optimizing value estimates [2, 8, 9] to determine an optimal policy may be sub-optimal. But the policy of an agent is not often a continuous and differentiable function, and hence, it is difficult to optimize it mathematically. In many value approximation/estimation methods, although the value estimation for different state action may be stochastic, since the policy is applied over an $\arg\max \text{value}_{\text{state}-\text{action}}$ function of it, the policy itself is deterministic (i.e., we have a single action output based on the value estimates). This approach although works for Reinforcement Learning (RL) problems with small or limited cardinality action space but is not computationally inefficient for RL problems with large action space cardinality, or continuous action space (limiting condition with $\lim_{|\text{a:PermissibleAction(a)}|\to\infty}$).

To optimize/maximize a function, we have to first define it mathematically. Therefore, we quantify the utility of a given policy by a policy utility function denoted by symbol J [26], which could be mathematically defined as below:

$$J_{(\theta)} = \mathbb{E}\sum_{t\geqslant 0}[\gamma^t r_t|\pi_\theta] \tag{1}$$

In Eq. 1, θ is the vector of policy parameters that need to be trained to optimize the the policy utility function π. γ is the discounting factors for (future) rewards r received over different time steps $t \in [0, t]$.

2.1 Categorization in Policy-Based RL

The gradient of the Policy Utility function 1 needs to be obtained so as to find its maximum. Since this is not as simple, therefore, we need to have assumptions for the policy gradient functions. The different agents under the policy approximation DRL could be further categorized based on the fact whether they have a stochastic assumption of the policy (gradient) [27] or a deterministic assumption of the policy (gradient) [28]. Another criterion of categorizing these DRL agents is based on their action vs behavior policy similarity, i.e., whether they follow on policy [29] or off policy [30] behavior approach. All the above criteria of categorizations RL agents could be combined to obtain finer agent categorizations as shown in Fig. 1.

Since the gradient assumption-based difference is salient to the policy-based RL, we will go into more detail of the two sub-types in this area. Next, we will cover the (Stochastic) Policy Gradient theorem, which provides the mathematical basis to estimate the gradient of a policy (utility function) and hence to optimize/maximize the expectancy of the policy utility function. Then, we will cover the Deterministic Policy Gradient (DPG). Until recently, the deterministic policy gradient was assumed to be infeasible [28]. But recently, this has also been mathematically approximated using the deterministic policy gradient theorem [28]. Interestingly, optimizing the

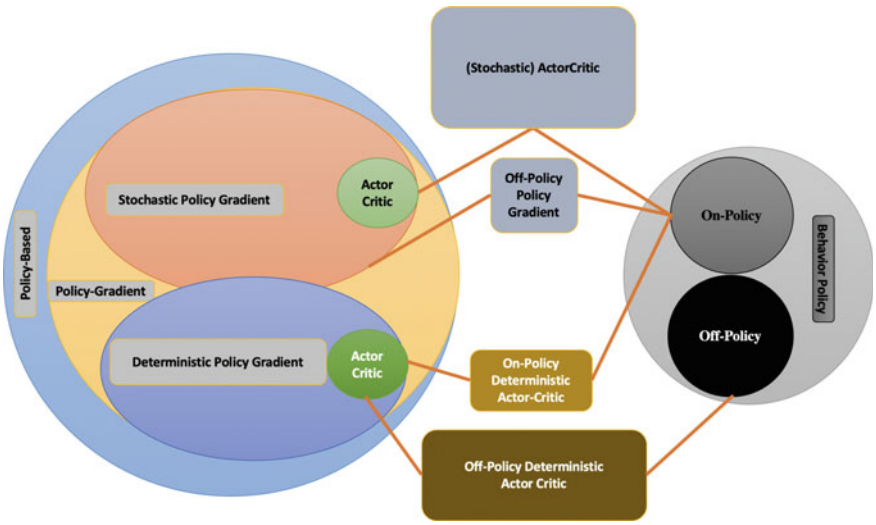


Fig. 1 Categorization of policy-based RL agents

deterministic policy gradient is found to be more efficient than its stochastic counterpart. Therefore, much of the recent development is focused toward agents that are based on deterministic policy gradient approach [11–13].

2.2 Policy Gradient

The mathematical gradient of the policy utility function (J) is intractable. This could be further explained by giving the mathematical function that needs to be optimize using the policy gradient approach as in the below equation:

$$\theta^* = \arg\max_{\theta} J_{(\theta)} \tag{2}$$

The policy utility function is dependant upon the state-transition probability function across each subsequent state ($s \in [s_0, s_n]$) in the sequence steps $[0, n]$ and hence the trajectory τ, which denotes the sequence of states transitioned, conditioned on action a in the previous sequence and the resultant reward r received.

where $\tau = [(s_0, a_0, r_0); (s_1, a_1, r_1); \ldots; (s_n, a_n, r_n)]$.

In terms of the trajectory, τ, Eq. 2 could be expressed as given below:

$$J_{(\theta)} = \mathbb{E}_{\tau} \sum \mathbb{P}(\tau; \theta) r_{(\tau)} \tag{3}$$

Next, Eq. 3 could be expressed in calculus form to show the differentiation required to obtain its gradient as below.

$$\Delta_{\theta} J_{(\theta)} = \int_{\tau} r_{(\tau)} \Delta_{\theta} \mathbb{P}(\tau; \theta) d\tau \tag{4}$$

To alleviate these challenges, the policy gradient theorem provides the necessary mathematical simplifications to compute the policy gradient. Until Deterministic Policy Gradient was not proposed, the Stochastic Policy Gradient was assumed to be the only Policy Gradient, and hence, still, sometimes it is referred to without the '*Stochastic*' prefix. We cover the (Stochastic) Policy Gradient next.

Stochastic Policy Gradient Theorem: Eq. 4 is differentiated (to obtain the gradient), over the parameter θ to give the function $\mathbb{P}(\tau; \theta)$. But the function itself is dependant upon the parameter θ. This is what causes the circular dependency and hence intractability in computing the policy gradient which is required so as to optimize the policy value function. This is where the Policy Gradient Theorem helps in alleviating the problem by assuming some simplifications.

As per the policy gradient theorem, the function $\mathbb{P}(\tau; \theta)$ could be given as equation below:

$$\mathbb{P}(\tau; \theta) = \prod_{t \geqslant 0} \Pr(s_{t+1} \mid s_t, a_t) \pi_{\theta}(a_t | s_t) \tag{5}$$

The state-transition probability from a given state is conditioned on all the previous states visited in the trajectory τ. The conditional probabilities for sequence of mutually exclusive events are a product of the individual event probabilities. A log of the joint probability of such mutually exclusive events would give us a sum of the log of individual probabilities. Therefore, if the visited states could be considered as mutually exclusive, then taking the log of the product of probability of individual elements in the trajectory, and it could be expressed as in equation below:

$$\log \mathbb{P}(\tau; \theta) = \sum_{t \geqslant 0} \log \mathbb{P}(s_{t+1} \mid s_t, a_t) + \log \pi_\theta(a_t \mid s_t) \tag{6}$$

It should be noted that in Eq. 6, the function $\mathbb{P}(\tau; \theta)$, does not depend upon the trajectory probability $\sim$, but only a series of independent state-transition probabilities. Therefore, Eq. 6 could be simplified as below as per the Policy Gradient Theorem:

$$\Delta_\theta \log \mathbb{P}(\tau; \theta) = \sum_{t \geqslant 0} \Delta_\theta \log \pi_\theta(a_t \mid s_t) \tag{7}$$

This finally given us the gradient of the policy value function as per the Policy Gradient Theorem as Eq. 8. This function could be maximized to learn an optimal policy (the objective of any DRL/RL the agent).

$$\Delta_\theta J_{(\theta)} \approx \sum_{t \geqslant 0} r_{(\tau)} \Delta_\theta \log \pi_\theta(a_t \mid s_t) \tag{8}$$

2.3 Deterministic Policy Gradient

A Deterministic Policy Gradient (**DPG**) [28] is the gradient of a (utility function of a) deterministic policy. A deterministic policy, as the name suggests, is non-stochastic. Since the stochastic policy theorem does not apply to a deterministic policy, it was believed that the deterministic policy gradient could not be easily computed, and hence, it is infeasible. But like the (stochastic) policy gradient theorem, now there is also a deterministic policy gradient theorem, which powers some of the most powerful agents based on deterministic policy gradient approach as is describe next mathematically.

A deterministic policy μ parameterized on weight vector θ maps the possible state set S to possible action set A and could be mathematically expressed as below:

$$\mu_\theta : \{s : S(s)\} \rightarrow \{a : A(a)\} \tag{9}$$

The policy utility function (also called **Performance Function** and denoted by symbol '**J**') under a state-transition probability $p(s \rightarrow s', t, \mu)$ could be then given as below:

$$J(\mu_\theta) = \mathbb{E}[r_1^\gamma \mid \mu] \tag{10}$$

The performance function as in Eq. 10 could be expressed in expectancy format as below:

$$J(\mu_\theta) = \mathbb{E}_{s\sim\rho_\mu}[r(s, \mu_\theta(s)] \tag{11}$$

Deterministic Policy Gradient Theorem: The Deterministic Policy Gradient ***Theorem -1*** states that if $\Delta_\theta \mu_\theta(s)$, and $\Delta_a Q_\mu(s, a)$ exists; and also the Deterministic Policy Gradient exists, then the gradient of the Performance Function could be expressed as below:

$$\Delta_\theta J(\mu_\theta) = \mathbb{E}[\Delta_\theta \mu_\theta(s) \Delta_a Q_\mu(s, a; \theta) \mid_{a=\mu_\theta(s)} \tag{12}$$

The Deterministic Policy Gradient ***Theorem-2*** states that under certain limiting conditions such as $\lim_{Var[\Delta_\theta J_{(\theta)}]\to 0}$, the Deterministic Policy Gradient is the same as the Stochastic Policy Gradient. Intuitively also, as the variance of any stochastic process tends to lim $\to$ 0 (i.e., $\lim_{Var[\Delta_\theta J_{(\theta)}]\to 0}$), and the given stochastic process is no longer stochastic and transform into a deterministic process, and hence, it is the case when the variance of the Stochastic Policy tends to 0, and it starts behaving like a Deterministic Policy.

3 DRL Agents Based on Stochastic Policy Gradient

In this section, we cover the agents that are based on the (stochastic) policy gradient theorem. The actor-critic [1] mechanism (Sec. 3.1) form the basis of these. The actor-critic algorithm uses a concept of baseline, and based upon the selection of baseline criteria, there could be multiple forms of actor-critic algorithms. The simplest actor-critic algorithm could use the estimates of required value functions as the baseline. The version of actor-critic that is more common is the advantage actor-critic (A2C), which use advantage estimates instead of value estimates, and asynchronous action advantage actor-critic (A3C) [31] that is a distributed variants with asynchronous updates (section: 3.3). Despite using advantage estimates, these do fall under policy-based methods as the associated DL network output is not the intermediate (advantage) estimates themselves as is in the case of value approximation-based techniques [10], but instead, the (stochastic) action probability for different actions when in a given state. Since the policy generated by this technique is probabilistic/ stochastic, it has exploration inbuilt within the action policy itself, and hence, a separate behavior policy is not required to mandate effective exploration of the space. Therefore, these techniques fall under on policy methods of DRL (Sect. 1).

3.1 Actor-Critic Agents

The Actor-Critic algorithm [27, 32] is a combination of value estimation and the policy approximation techniques. In the actor-critic algorithm, the agent is divided into two sub-components, namely the actor and the critic (hence the name actor-critic). The actor encapsulates the policy of the agent and takes the action, and the critic provides the advantage/value estimation to refine the actor's policy [33].

During every step, the state received from the environment is passed to both the actor and the critic. But the reward received from the environment is passed only to the critic. The critic uses the received state and rewards to compute the (state) value estimate like any value approximation-based agent. The Temporal Difference Error (*TD Error*) [34] between the reward and critic's estimates are used to update the actor's policy approximation function and the critic's value approximation function. This mechanism is as shown in Fig. 2.

The TD error at any step (δ_t) could be given as below:

$$\delta_t = R_t + \gamma V(S_{t+1}; W_t) - V(S_t; W_t) \tag{13}$$

In each step, the critic receives the TD error as shown in Eq. 13 and updates the weights vector (W_t) of the critic's value approximation function. All the other symbols have their standard reference as defined in earlier sections. The weights updates are moderated weighted using the learning rate α_w as below:

$$W_{t+1} = W_t + \alpha_w \delta_t \Delta_w V(S; W_t) \tag{14}$$

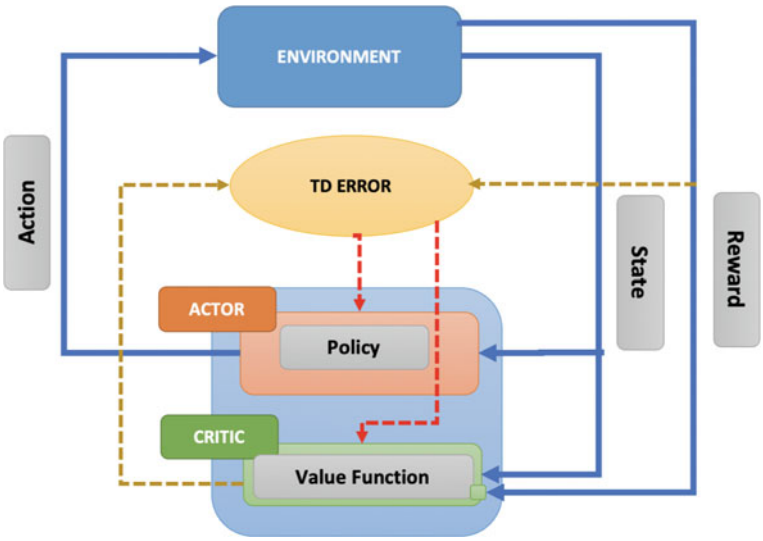


Fig. 2 Working of the actor-critic agent

The symbols A, α, π have the standard references:

$$\theta_{t+1} = \theta_t + \alpha_\theta \delta_t \Delta_\theta \log \pi(A \mid S; \theta_t) \tag{15}$$

3.2 Advantage Actor-Critic

The advantage actor-critic combines the ideas of Dueling DQN ([9]) and that of the actor-critic agent [35]. In this, the critic instead of estimating (absolute) state value baseline uses the *advantage* (the incremental utility over a default) baseline. By doing so, the agent becomes more stable, as could also be seen in the REINFORCE algorithm [1] which is based on the (Stochastic) policy gradient theorem (2.2). This stability enhancement is due to the reduced variance in the actor's policy approximation function.

Another inspiration that the Advantage Actor-Critic takes from the Dueling DQN [9] is the use of branched sub-network architecture of the CNN network. Like Dueling Q networks, the Advantage Actor-Critic agent's (CNN/) DL network also branches into two sub-networks after the convolutional layers, where one of the sub-networks serves as critic and estimates the advantage, the other sub-network serves as actor and estimates the action (policy).

3.3 Distributed Actor-Critic Agents

The stochastic policy gradient-based techniques mostly require second-order optimization which is computationally expensive. This combined with the additional computational cost of DL makes the training and convergence challenging. The Actor-Critic Agent (3.1) and the Advantage Actor-Critic Agents (3.2) use this mechanism and hence are computationally very expensive. Also, since these algorithms do not work in parallel, the high computational complexity directly translates into a higher clock times as well. The Asynchronous Advantage Actor-Critic Agent (A3C) [31] and the (Synchronous) Advantage Actor-Critic Agent (A2C) attempts to alleviate this problem by making the agent's working parallel and distributing the computation [1].

The Asynchronous Advantage Actor Critic In the Asynchronous Advantage Actor-Critic (A3C), there exists a master/central copy of the agent (combination of actor's and critic's weight vectors). Many individual instances of the worker agents could spawn in parallel and distribute across different General-Purpose GPU (GP-GPU) machines. Each such agent instantiates with a copy of the master agent's weight during. Each agent next trains against a randomized individual instance of the environment (in isolation), and thereupon merges its weight updates (learning) with the centralized agent's weights (collective learning). This sync happens either

after a pre-determined number of states or on reaching terminal state, whichever is earlier. Besides expediting the learning process, such distributed setup also reduces the collinearity in the weight updates as would have otherwise generated if a single agent had been learning alone sequentially. This aspect, besides expediting the training also makes the training more stable.

In the A3C, one drawback exists that is of noise updates. Since the worker agents could update the master agent's weight asynchronously and do not coordinate with each other, a worker agent could end up updating the master's agent after many other agents have updated the central weights before this agent. Since the weight update proposal of this agent is based upon the initial weight that it copied from the master agent's weight, and by the time, this worker agent returns to update the weight, the master agent's weight may drift quite a lot since this agent's initial weights were copied from the master, and the resultant updates computed on that copy. This results into noisy updates of the master's weights.

Advantage Actor Critic The Advantage Actor-Critic (A2C) follows a similar philosophy of training multiple worker agents in parallel and having a central/master agent's weight. But unlike A3C, the worker agents do not directly update the master agent's weight anytime asynchronously. Instead, the A2C has a coordinator which updates the master/central agent's weights, and no other worker is allowed to update the master agent's weight directly. Also in A2C, the instantiating and update of weights from all the worker agent's are synchronized, i.e., they start simultaneously using the same copy of initial weights, and the coordinator waits until it receives the individual weights updates from each worker. Since all the workers have iterated the same number of steps each, a group update is synonymous to an average of update of all individual weight updates, making the weight updates simpler and more efficient.

4 DRL Agents Based on Deterministic Policy Gradient

The Deterministic Policy Gradient-based approaches are increasingly becoming more popular. One reason for this is simplicity of the gradient mathematics as is evident from the equation in Theorem-1 and Theorem-2 in this section. Another very important reason for this is the computational simplicity as the agents based on DPG are computationally less expensive than agents for the same RL task and that could claim similar total reward that are based on Stochastic Policy Gradient. Therefore, we will next cover some popular DRL agents based on the DPG.

4.1 Deep Deterministic Policy Gradient

The original DPG paper [28] besides covering the DPG theorem also proposed the DPG algorithm. The DPG agent as proposed was based on the DPG Theorem and

used linear approximation functions. Even with linear approximation functions, the DPG algorithm [28] claimed to outperform the DQN algorithm [2] on 20 different continuous action control-based standardized RL problems. Much of the credit for a linear approximation function-based algorithm outperforming a Deep Learning approximation function-based agent algorithms. This could be due to the effectiveness and efficiency that the DPG-based approach has overvalue estimation (or even stochastic policy gradient-based approach) for RL problems involving continuous action control [36].

The Deep Deterministic Policy Gradient (DDPG) algorithm [11] is an enhancement of the original proposed DPG algorithm with Deep Learning approximation function. DDPG would also require the enhancements that DQN required to stabilize while moving from a linear approximation function to a DL approximation function. These enhancements involve the using *experience-replays* from a *memory-buffer*, in which the experience tuples are stored. DQN also used a secondary offline Q network as the target Q network to overcome the instability caused in training due to the estimates from a given network being used to update itself. The Double-DQN algorithm [8] improves upon this by using an offline and an online Q network for the update and estimation, respectively. Even in Double DQN, there are drawbacks due to infrequent updates of the Actor-Critic [9]. The actor-critic algorithm is sensitive to such infrequent use the Dueling DQN [9] style sub-network architecture. Since DDPG dates ahead of both DPG, DQN, Double DQN, Duelling DQN, and Policy Gradient-based Actor-Critic, it takes the best of all approaches and uses the Double-DQN style twin Q-Network architecture for implementing the actor-critic style agent on DPG-based approach. To overcome the infrequent and sudden update that the Double DQN faces, the DDPG uses a different mechanism for the weight updates of the offline target network.

In DDPG, the target network parameter vector θ is updated using a mechanism called 'Soft Update.' The 'Soft Update' follows an exponential smoothing function (with exponential smoothing coefficient τ) to smooth-en the updates. This is as shown below:

$$\theta_{t+1} = \tau\theta_t + (1 - \tau)\theta_{t+1} \tag{16}$$

Since the DDPG is based on Deterministic Policy, hence it misses the Stochastic component in the policy essential for an on-policy agent. Hence, DDPG is essentially an *off-policy* learning algorithm and requires an external *Behaviour-Policy* to optimize on the exploration part. The Behaviour Policy used for the DDPG is as given below:

$$\mu_b(s_t) = \mu_a(s_t \mid \theta_t) + \mathcal{N} \tag{17}$$

In Eq. 17, μ_a is the action policy (parameterized by weight vector θ) as coming from the DDPG algorithm, and μ_b is the Behaviour Policy, which is arrived by adding a noise function $\mathcal{N}$ to the action policy to ensure optimal exploration. 's' is the current state and 't' the current time step.

4.2 Trust Region and Proximal Policy Optimization

In this section, we cover two more deterministic policy gradient-based DRL algorithms which are enhancements of the DDPG.

Trust Region Policy Optimization Deterministic Policy Gradient (2.3)-based deep reinforcement learning approaches like the 'Deep Deterministic Policy Gradient (DDPG)' (4.1) claimed to deliver the best in class performance on large, even continuous action space-based reinforcement learning tasks [11]. The problem with such approaches like the DDPG is that their line search-based policy gradient update (used during optimization) either proves too big for updated involving nonlinear trajectory making the update go beyond the target or makes the learning too slow. Since in the deep reinforcement paradigm, nonlinear gradients are quite common so algorithms-based online search gradient update are not very robust and cannot provide guarantees of near monotonic policy improvements.

This is where the 'Trust Region Policy Optimization' (TRPO) [12] an algorithm based on Trust Region-based policy update helps. The TRPO algorithm use the 'Minorize Maximization' (MM) (second order) Gradient update and claims to solve the above problem with the DDPG like algorithms and provide guarantee for near monotonic general (stochastic) policy improvement even for nonlinear policies like that approximated by (deep) neural networks.

Additionally, TRPO uses a mechanism called *Importance Sampling* to compute the expectancy of the policy from previous trajectories instead of only the current trajectory to stabilize the policy gradient. This method has an underlying assumption that the previous trajectory's distribution ($Q(x)$) is not very different from the current trajectory's distribution ($P(x)$).

The policy gradient for a Stochastic Policy Gradient [37] method and associated algorithms like actor-critic [31] looks like below:

$$\nabla_\theta(J_\theta) = \mathbb{E}_{\tau\sim\pi_\theta(\tau)}[\nabla_\theta \log \pi_\theta(\tau) r(\tau)] \tag{18}$$

In Eq. 18 above, the trajectory τ over which the samples for computing expectancy is gathered (to update the gradient ∇ of the policy value function J) is the same (current) trajectory of the policy as used in the policy π (parameterized over θ).

But, in the case of TRPO that uses importance sampling and the past trajectory for sampling, this policy value function update looks as below:

$$\nabla_{\theta\prime}(J_{\theta\prime}) = \mathbb{E}_{\tau\sim\pi_\theta(\tau)}[\sum_{t=1}^{T} \nabla_{\theta\prime} \log \pi_{\theta\prime}(\prod_{t\prime=1}^{t} \frac{\pi_{\theta\prime}}{\pi_\theta})(\sum_{t\prime=t}^{T} r)] \tag{19}$$

Proximal Policy Optimization This is where the 'Proximal Policy Optimization' (PPO) [13] algorithm helps. The PPO algorithm works similar to TRPO but is computationally more efficient. This is because instead of the computationally expensive second-order gradient optimization techniques; the PPO algorithm uses a linear vari-

ant of the gradient update called the 'Fisher Information Matrix' (FIM). In Eq. 19 for TRPO, the trajectory is being sampled from the policy previous time $= t(\pi_\theta = Q(x))$ and the expectancy over such collected samples are used to update the policy at next time step $(\pi_{\theta'} = P(x))$. When the ratio of expectancy over the two trajectory distributions varies significantly $\frac{P(x)}{Q(x)}$ as in the case of linear-gradient update in PPO, the previously stated assumption may not hold, leading to high variance in policy updates. To avoid this, there are two methods that the PPO algorithm recommends. The first one uses a *Adaptive KL (kulback-liebler) Penalty*, and the second one uses *Objective Clipping*. As per the original PPO paper [13], the *Objective Clipping* variant, with the clipping factor $\epsilon = 0.2$ provided the best result.

5 Discussion and Conclusion

Policy-based DRL techniques have revolutionized the domain of control automation with inventions ranging from autonomous drones to robotic industries. Until recently, DRL was restricted only to innovations in games; and despite surpassing the benchmarks established by their human counterparts in multiple games, these methods could not scale to real-life and industrial automation tasks. The main reason for this was the essential requirement of continuous action control and complex physics of the domain involved in these tasks. Because of these reasons, most of the earlier solutions for such applications involved the invent of custom planning algorithms. The design of such sophisticated custom solutions required complete knowledge to the dynamics of the domain and its derivatives and hence were not scalable. Policy-based DRL has democratized this space as now DRL agents could be trained to learn similar sophisticated policies just by learning from the data generated by interacting with these systems. This has led to significant innovations in real-life and high-return applications like autonomous vehicles, drones, and industrial robots and automation.

Although the policy-approximation based DRL agents are computationally more expensive than their value approximation, but they also offer excellent results on many complex RL tasks. For complex RL tasks involving large and even continuous action space, these easily surpass the performance of value estimation-based DRL agent algorithms like the Deep Q Networks and its variants. Moreover, the policy-based DRL agents also prove to be more sample efficient for such complex RL tasks. With recent advancements in policy-based DRL, a lot of these drawbacks have been addressed, and these techniques are now more sample efficient and offer better monotonic improvement guarantees. Therefore, in this paper, we discussed different policy-based DRL techniques and the mathematical advancements empowering it to understand its applicability for different applications to achieve better results over the existing ones.

References

1. Sewak, M.: Deep Reinforcement Learning: Frontiers of Artificial Intelligence. Springer (2019)
2. Mnih, V., Kavukcuoglu, K., Silver, D., Graves, A., Antonoglou, I., Wierstra, D., Riedmiller, M.: Playing atari with deep reinforcement learning. arXiv preprint arXiv:1312.5602 (2013)
3. Sewak, M., Sahay, S.K., Rathore, H.: An overview of deep learning architecture of deep neural networks and autoencoders. J. Comput. Theoret. Nanosci. **17**(1), 182–188 (2020)
4. Sewak, M., Karim, M.R., Pujari, P.: Practical Convolutional Neural Networks: Implement Advanced Deep Learning Models Using Python. Packt Publishing Ltd (2018)
5. Sahay, S.K., Sharma, A., Rathore, H.: Evolution of malware and its detection techniques. In: Information and Communication Technology for Sustainable Development, pp. 139–150. Springer (2020)
6. Rathore, H., Sahay, S.K., Chaturvedi, P., Sewak, M.: Android malicious application classification using clustering. In: International Conference on Intelligent Systems Design and Applications, pp. 659–667. Springer (2018)
7. Rathore, H., Agarwal, S., Sahay, S.K., Sewak, M.: Malware detection using machine learning and deep learning. In: International Conference on Big Data Analytics, pp. 402–411. Springer (2018)
8. Hasselt, H.v., Guez, A., Silver, D.: Deep reinforcement learning with double q-learning. In: Thirtieth AAAI Conference on Artificial Intelligence, pp. 2094–2100 (2016)
9. Wang, Z., Schaul, T., Hessel, M., Hasselt, H., Lanctot, M., Freitas, N.: Dueling network architectures for deep reinforcement learning. In: International conference on machine learning, pp. 1995–2003 (2016)
10. Sewak, M., Sahay, S.K., Rathore, H.: Value-approximation based deep reinforcement learning techniques: an overview. In: 5th IEEE International Conference on Computing, Communication and Automation (ICCCA 2020) (2020)
11. Lillicrap, T.P., Hunt, J.J., Pritzel, A., Heess, N., Erez, T., Tassa, Y., Silver, D., Wierstra, D.: Continuous control with deep reinforcement learning. In: ICLR (Poster) (2016)
12. Schulman, J., Levine, S., Abbeel, P., Jordan, M., Moritz, P.: Trust region policy optimization. In: International conference on machine learning, pp. 1889–1897 (2015)
13. Schulman, J., Wolski, F., Dhariwal, P., Radford, A., Klimov, O.: Proximal policy optimization algorithms. CoRR arxiv: abs/1707.06347 (2017)
14. Zhu, M., Wang, X., Wang, Y.: Human-like autonomous car-following model with deep reinforcement learning. Transp. Res. part C: Emerg. Technol. **97**, 348–368 (2018)
15. Kohl, N., Stone, P.: Policy gradient reinforcement learning for fast quadrupedal locomotion. In: IEEE International Conference on Robotics and Automation, 2004. Proceedings. ICRA '04. 2004, vol. 3, pp. 2619–2624 (2004)
16. Sewak, M., Sahay, S.K., Rathore, H.: Comparison of deep learning and the classical machine learning algorithm for the malware detection. In: 2018 19th IEEE/ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD), pp. 293–296. IEEE (2018)
17. Sewak, M., Sahay, S.K., Rathore, H.: An investigation of a deep learning based malware detection system. In: Proceedings of the 13th International Conference on Availability, Reliability and Security, pp. 1–5 (2018)
18. Silver, D., Huang, A., Maddison, C.J., Guez, A., Sifre, L., Van Den Driessche, G., Schrittwieser, J., Antonoglou, I., Panneershelvam, V., Lanctot, M., et al.: Mastering the game of go with deep neural networks and tree search. Nature **529**(7587), 484–489 (2016)
19. Christopher, B.: Dota 2 with large scale deep reinforcement learning. arXiv preprint arXiv:1912.06680 (2019)
20. Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A.A., Veness, J., Bellemare, M.G., Graves, A., Riedmiller, M., Fidjeland, A.K., Ostrovski, G., et al.: Human-level control through deep reinforcement learning. Nature **518**(7540), 529–533 (2015)

21. Smart, W.D., Pack Kaelbling, L.: Effective reinforcement learning for mobile robots. In: Proceedings 2002 IEEE International Conference on Robotics and Automation (Cat. No.02CH37292), vol. 4, pp. 3404–3410 (2002)
22. Rajeswaran, A., Kumar, V., Gupta, A., Schulman, J., Todorov, E., Levine, S.: Learning complex dexterous manipulation with deep reinforcement learning and demonstrations. CoRR arxiv: abs/1709.10087 (2017)
23. Poznyak, A.S.: Advanced Mathematical Tools for Automatic Control Engineers: Stochastic Techniques. Elsevier (2009)
24. Socha, K., Dorigo, M.: Ant colony optimization for continuous domains. Eur. J. Oper. Res. **185**(3), 1155–1173 (2008)
25. Finn, C., Levine, S., Abbeel, P.: Guided cost learning: deep inverse optimal control via policy optimization, pp. 49–58. PMLR (2016)
26. Sutton, R.S., McAllester, D.A., Singh, S.P., Mansour, Y.: Policy gradient methods for reinforcement learning with function approximation. In: Advances in Neural Information Processing Systems, pp. 1057–1063 (2000)
27. Bhatnagar, S., Ghavamzadeh, M., Lee, M., Sutton, R.S.: Incremental natural actor-critic algorithms. In: Advances in Neural Information Processing Systems, pp. 105–112 (2008)
28. Silver, D., Lever, G., Heess, N., Degris, T., Wierstra, D., Riedmiller, M.: Deterministic policy gradient algorithms. In: 31st International Conference on International Conference on Machine Learning, pp. I–387 (2014)
29. Singh, S., Jaakkola, T., Littman, M.L., Szepesvári, C.: Convergence results for single-step on-policy reinforcement-learning algorithms. Mach. Learn. **38**(3), 287–308 (2000)
30. Munos, R., Stepleton, T., Harutyunyan, A., Bellemare, M.: Safe and efficient off-policy reinforcement learning. In: Advances in Neural Information Processing Systems, pp. 1054–1062 (2016)
31. Mnih, V., Badia, A.P., Mirza, M., Graves, A., Lillicrap, T., Harley, T., Silver, D., Kavukcuoglu, K.: Asynchronous methods for deep reinforcement learning. In: International Conference on Machine Learning, pp. 1928–1937 (2016)
32. Degris, T., Pilarski, P.M., Sutton, R.S.: Model-free reinforcement learning with continuous action in practice. In: American Control Conference, pp. 2177–2182. IEEE (2012)
33. Sutton, R.S., Barto, A.G.: Introduction to Reinforcement Learning, 1st edn. MIT Press, Cambridge, MA, USA (1998)
34. Sutton, R.S.: Learning to predict by the methods of temporal differences. Mach. Learn. **3**(1), 9–44 (1988)
35. Parisotto, E., Ba, L.J., Salakhutdinov, R.: Actor-mimic: deep multitask and transfer reinforcement learning. CoRR arxiv: abs/1511.06342 (2015)
36. Duan, Y., Chen, X., Houthooft, R., Schulman, J., Abbeel, P.: Benchmarking deep reinforcement learning for continuous control. In: International Conference on Machine Learning, pp. 1329–1338 (2016)
37. Sutton, R.S., McAllester, D., Singh, S., Mansour, Y.: Policy gradient methods for reinforcement learning with function approximation. In: 12th International Conference on Neural Information Processing Systems, pp. 1057–1063 (1999)

Offline Character Recognition of Handwritten MODI Script Using Wavelet Transform and Decision Tree Classifier

Solley Joseph and Jossy P. George

Abstract MODI script is derived from the Nāgari family of scripts, and it was used for writing Marathi until twentieth century. Though currently not used as an official script, it has historical importance, as a large volume of manuscripts are preserved at various libraries across India. With the use of an appropriate recognition system, the handwritten documents can be transferred into digital media, so that it can be conveniently viewed, edited, or transliterated to other scripts. The research on MODI script is still in the initial stages, and there is a considerable demand for more research in this field. An implementation of wavelet transform-based feature extraction for MODI script's character recognition is discussed in this paper. The experiment is performed using Daubechies, Haar, and Symlet wavelets, and performance comparison between these different mother wavelets is carried out. Decision tree classifier is used for the classification process, and the results indicate that the feature extraction using Daubechies wavelet yielded better character recognition result.

Keywords Handwritten character recognition · MODI script · Decision tree classifier · Wavelet transform · Pattern recognition

1 Introduction

Character recognition is a well-researched area and has accomplished remarkable recognition accuracy in the case of printed text compared to handwritten text. Though the character recognition systems of non-Indic scripts such as English, Chinese, Japanese, Korean, and German have already matured, the character recognition of regional languages in India is still a challenging task. MODI script is an example of the kind of script in which very little work is done on character recognition of the

S. Joseph (✉) · J. P. George
CHRIST (Deemed to be University), Bengaluru, India
e-mail: solley.joseph@res.christuniversity.in

S. Joseph
Carmel College of Arts, Science and Commerce for Women, Goa, India

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_48

script [1–3]. The aim of our study is the wavelet transform-based feature extraction method in the recognition of the MODI script.

It was in twelfth century that MODI script was developed, and until 1950, it was the official script in use in Maharashtra [4, 5]. During this time, large volumes of documents and correspondence were written in MODI scripts [6]. These collections are preserved in the libraries and temples in various parts of India. Some of these collections are archived in libraries outside India as well [2, 6]. The feature extraction for MODI character recognition using the method of wavelet transform is discussed in this paper. The decision tree-based method is used at the classification phase of the character recognition process.

Handwritten character recognition of cursive script is generally difficult; for MODI script, it is even more complex due to various reasons such as the presence of similar-looking characters and the absence of word demarcation in MODI documents [7]. Another challenge is in the segmentation stage of the recognition process. The segmentation process is difficult due to various factors such as degradation of text, the presence of multiple skews, uneven alignment, slanting lines, overlapping lines, and touching lines [8, 9].

The MODI script has a basic character set of forty-six characters, which includes thirty-six consonants and ten vowels. In addition to Marathi, some other languages like Gujarathi, Hindi, Tamil, and Urdu also used MODI script [10]. With no symbol for demarcation of words or sentences, the script adopted the usage of cursive "shorthand." The absence of a demarcation symbol makes it challenging to perform word segmentation on MODI documents [11]. Figure 1 shows the MODI character set.

The organization of the remaining contents of the paper are the related literature listed in Sect. 2, and in Sect. 3, the methodology is described. The experimental study and the results are given in Sect. 4, and Sect. 5 contributes to the conclusion of the paper.

Fig. 1 Basic MODI script characters

2 Literature Review

Feature extraction using methods based on wavelet energy has been experimented on various pattern recognition applications. A combination of wavelet energy feature extraction was used to get accuracy of 95.59% for character recognition of Malayalam scripts [12]. Wavelet transform has also been used for English character recognition, and a 97% accuracy was achieved [13]. An analysis and comparison of wavelet families for image compression was performed [14]. The study reported that the biorthogonal wavelet provided the best result for image compression compared to other wavelet families such as Daubechies, Coiflet, and Symlet. A method based on wavelet energy was implemented for finding the word boundary, and it was reported that the performance was better compared to other detection methods [15]. Wavelet transform-based feature extractions have successfully been implemented in various other applications as well. Tajane et al. have carried out a comparative analysis of different mother wavelets suitable for the electrocardiographic (ECG) signal [16]. Cristina et al. experimented with the wavelet-based method for traffic forecasting, and the results indicate that Haar wavelet yields better accuracy [17].

Moment-based feature extraction has been used for MODI script recognition. An implementation of Hu's moment-based feature extraction for MODI characters recognition resulted in an accuracy of 82.6% [18]. An experiment using Otsu's binarization was employed on the MODI script for character recognition, which yielded an accuracy of 72.6% [19]. A feature extraction method using empirically determined heuristics was implemented on MODI script and achieved a recognition rate of 91.2% [20]. The representation learning-based method was used as a feature extractor for handwritten multilingual numeral recognition, and the accuracy achieved was 96.7% [21].

3 Methodology

The focus of our study is the comparison of the performance of various mother wavelets in the process of MODI script's character recognition. The wavelet transform-based method is implemented in the feature extraction stage of the process. At the stage of classification, the extracted features are classified using a decision tree classifier.

Wavelet transform has diverse applications. Computer scientists use it in applications like signal processing, speech recognition, image processing, computer graphics, and Internet traffic description [22, 23]. It has applications in the fields of biology and geology also [12, 24]. The proposed methodology is depicted in Fig. 2.

Feature extraction using the wavelet transform is performed on the input image of 60 * 60 pixel size. The experiment is conducted using Daubechies, Haar, and

Fig. 2 Proposed methodology

Symlet wavelet families. The extracted features are then subjected to the decision tree classifier for character recognition.

3.1 Feature Extraction

The feature extraction stage can be considered as the most important stage in the process of character recognition as the accuracy of the process is highly dependent on the selected features. Wavelet transform is an effective tool for multi-resolution feature extraction, and by using wavelets, the image can be transformed into approximate and detail (diagonal, vertical, and horizontal) sub-signals, as shown in Fig. 3. These sub-signals can be efficiently used for feature extraction in pattern recognition and image processing applications [25]. It is possible to set these details to zero for smaller values, as it will not make any major changes in the image. Hence, filtering and compression can be achieved [12, 26].

For formulating wavelet energy, in the given 'n' decomposition from the three sub-signals, wavelet coefficients are used [27]. The horizontal wavelet energy $e_n^H(x, y)$, $e_n^V(x, y)$ as vertical, and $e_n^D(x, y)$ as diagonal, detail sub-signals, at the pixel position

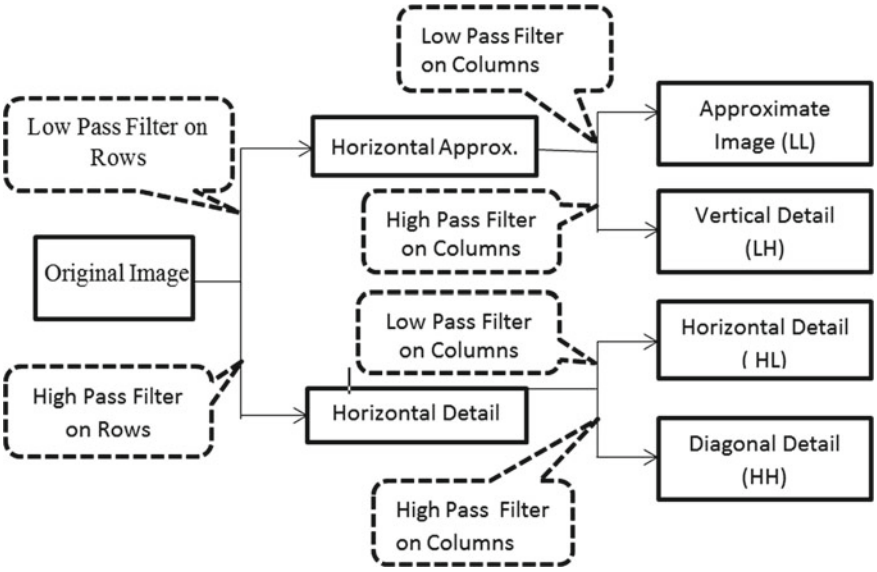


Fig. 3 Discrete wavelet transform

(x, y) at each of the level are given below

$$e_n^H(x, y) = \sum_{r,s \in Z} (D_H)^2 G(x - r, y - s)$$
$$e_n^V(x, y) = \sum_{r,s \in Z} (D_V)^2 G(x - r, y - s)$$
$$e_n^D(x, y) = \sum_{r,s \in Z} (D_D)^2 G(x - r, y - s) \quad (1)$$

The horizontal, vertical, and diagonal detail sub-signals are D_H, D_V, and D_D, which are the wavelet coefficients at the pixel position (x, y), respectively. Z refers to the set of integers, with r, and s are the x-coordinate and y-coordinate of the neighborhood pixel, respectively. $G(x, y)$ is the two-dimensional Gaussian kernel function given by

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{\left(\frac{-(x^2+y^2)}{2\sigma^2}\right)} \quad (2)$$

where σ is the SD of the kernel.

The family of wavelets consists of Haar, Daubechies, Symlets, Coiflets, Biorthogonal, Meyer, Gaussian, Morlet, etc. For the study, we have performed single-level 2D decomposition using Daubechies, Haar, and Symlet, and a short description of them is given below.

The Daubechies wavelets are grounded on the work of Ingrid Daubechies. They are orthogonal wavelets. The names and wavelet functions ψ, of the nine members of the family, are shown in Fig. 4. In our experiment, we have used db2, db3, and db4.

The Haar wavelet is similar to a step function, and it is the most simple wavelet compared to the others. The wavelet functions ψ of the Haar wavelet family is shown in Fig. 5.

The Symlets get that name as they are almost symmetrical wavelet, and is suggested by Daubechies as modifications to 'db' family. Both wavelets have similar properties. The wavelet functions ψ for Symlet family which we have used in our experiment are shown in Fig. 6.

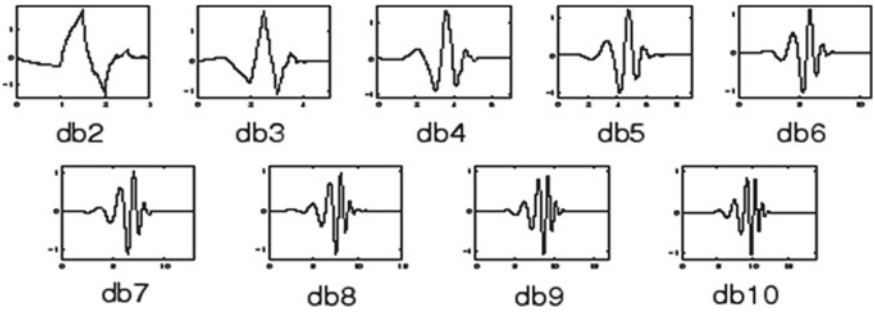
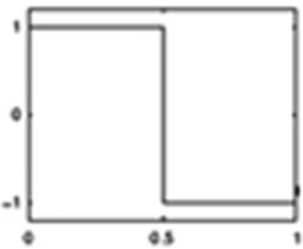


Fig. 4 Daubechies wavelet family

Fig. 5 Haar wavelet family



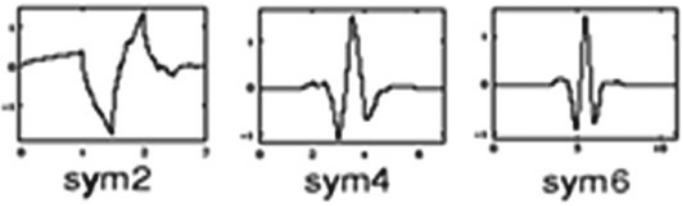


Fig. 6 Symlet wavelet family

3.2 Classification

The extracted features are then subjected to classification. A decision tree classifier is implemented in this experiment. It is a supervised model that can be effectively used for classification and prediction. Each internal node flowchart like a tree structure denotes a test on an attribute, an outcome of each branch represents the test, and each leaf node (terminal node) holds a class label [28]. The classification is performed by recursively partitioning the source set into subsets based on an attribute value.

4 Experimental Results

For the experiment, the data set consists of 4600 MODI characters, which are divided for training and testing purposes. Seventy percent of the total samples (3220) are used for training and for testing, and the rest of 1380 samples are used. For extracting the desirable features, a single-level discrete wavelet transform (DWT) is used. (The approximate sub_signal is used in our study). The experiment is conducted using four wavelets families such as Daubechies, Haar, and Symlet wavelets.

Each input image of 60 * 60 size is subjected to feature extraction using the 'dwt' function (Wavelet transform function). Table 1 describes the feature size of different mother wavelets for single-level decomposition.

The output of this level (the extracted features) is used for the classification. At the end of the classification stage, the accuracy is computed. A comparative study is performed by recording the accuracy rate as well as the absolute time is taken for the process. A consolidated version of the comparison report is given in Table 2.

The performance comparison of various mother wavelets given in Table 2 indicates that Daubechies wavelet performs better in the feature extraction and for character recognition. The highest accuracy of 91.75% was achieved using the Daubechies wavelet (db2).

Table 1 Size feature vector extracted using different mother wavelets

S. No.	Mother wavelets		Extracted feature size
1	Daubechies wavelet family	db2	17 × 17
		db3	20 × 20
		db4	23 × 23
2	Haar wavelet family	Haar	17 × 17
3	Symlet wavelet family	sym2	17 × 17
		sym4	20 × 20
		sym6	23 × 23

Table 2 Comparison of accuracy of character recognition, using different types of wavelets

S. No.	Wavelet filter	Accuracy (%)	Absolute time in seconds
1	Daubechies (db2)	91.75	56.18
2	Daubechies (db3)	87.31	59.59
3	Daubechies (db4)	85.07	63.97
4	Haar	86.67	34.10
5	Symlet (Sym2)	87.76	41.30
6	Symlet (Sym3)	86.59	46.05
7	Symlet (Sym4)	84.49	72.41

5 Conclusion

An algorithm for the optical character recognition of handwritten MODI script, based on the wavelet transform feature extraction method and decision tree classifier, was implemented. Multiple mother wavelets are used to carry out the experiment. Daubechies, Haar, and Symlets wavelets are implemented, and the output of this level is subjected to classification using the decision tree classification method. The performance of the Daubechies wavelet is better compared to the rest of the wavelet families. The highest accuracy achieved is 91.75% using Daubechies wavelet (Db2), and we can conclude that the Daubechies wavelet family is better compared to Haar and Symlet in MODI script character recognition.

References

1. Joseph, S., George, J.: Feature extraction and classification techniques of modi script character recognition. Pertanika J. Sci. Technol. **27**(4), 1649–1669 (2019)
2. Khillari, R.: History of MODI Script. Shri Modi Vibhav, 2008. (Online). Available: http://modi-script.blogspot.com/2008/05/history-of-modi-script.html
3. Joseph, S., George, J.: Efficient handwritten character recognition of MODI script using wavelet transform and SVD. In: Data Science and Security. Lecture Notes in Networks and Systems, vol. 132. Springer, Singapore (2021)
4. Ramraje, R.A.: History of MODI script in Maharashtra. Glob. Online Electron. Int. Interdiscip. Res. J. (GOEIIRJ) **2**(1) (2013)
5. Joseph, S., George, J.: Handwritten Character Recognition of MODI Script using Convolutional Neural Network Based Feature Extraction Method and Support Vector Machine Classifier, pp. 32–36. IEEE (2020)
6. Joseph, S., George, J.P., Gaikwad, S.: Character recognition of MODI script using distance classifier algorithms. In: Fong, S., Dey, N., Joshi, A. (eds.) ICT Anal. Appl. Lect. Notes Networks Syst., vol. 93. Springer, Singapore (2020)
7. Joseph, S., George, J.: Data Augmentation for Handwritten Character Recognition of MODI Script Using Deep Learning Method. In: Senjyu, T., Mahalle, P.N., Perumal, T., Inf, Joshi A., Commun. Technol. Intell. Syst. ICTIS, (eds.) Smart Innov. Syst. Technol, vol. 196, p. 2021. Springer, Singapore (2020)

8. Kulkarni, S., Borde, P., Manza, R., Yannawar, P.: Text line segmentation of handwritten historical MODI documents. In: International Conference on Information, Communication and Computing Technology (ICICCT2017), pp. 1–12 (2017)
9. Gharde, S.S., Ramteke, R.J.: Recognition of characters in Indian MODI script. In: Proceedings of International Conference on Global Trends in Signal Processing, Information Computing and Communication, pp. 236–240 (2016)
10. Pandey, A.: Final Proposal to Encode the MODI Script in ISO/IEC 10646, pp. 26–27 (2011)
11. Joseph, S., Datta, A., Anto, O., Philip, S., George, J.: "OCR System Framework for MODI Scripts using Data Augmentation and Convolutional Neural Network", in Data Science and Security. Lecture Notes in Networks and Systems, Springer, Singapore. **132**, 201–209 (2021)
12. Chacko, B.P., Krishnan, V.V., Raju, G., Anto, B.P.: Handwritten character recognition using wavelet energy and extreme learning machine. Int. J. Mach. Learn. Cybern. **3**(2), 149–161 (2012)
13. Shelke, S.: Handwritten character recognition using wavelet transform for feature extraction, Mar 2014, pp. 3–7 (2016)
14. Kumari, S., Vijay, R.: Analysis of orthogonal and biorthogonal wavelet filters for image compression. Int. J. Comput. Appl. **21**(5), 17–19 (2011)
15. Juang, C.F., Cheng, C.N., Tu, C.C.: Wavelet energy-based support vector machine for noisy word boundary detection with speech recognition application. In: Proceedings of 21st Conference on Computational Linguistics and Speech Processing, ROCLING 2009, pp. 21–29 (2009)
16. Kapil Tajane, D.J.U., Pitale, R.: Review paper: comparative analysis of mother wavelet functions with the ECG signals. Int. J. Eng. Res. Appl. **4**(1) (2014)
17. Stolojescu, C., Railean, I., Moga, S., Isar, A.: Comparison of wavelet families with application to WiMAX traffic forecasting. In: Proceedings of International Conference on Optimization of Electrical and Electronic Equipment, OPTIM, pp. 932–937 (2010)
18. Borde, P.L., Manza, R.R., Kulkarni, S., Borde, P.L., Manza, R.R., Yannawar, P.: Analysis of orthogonal moments for recognition of handwritten MODI numerals. J. Sci. Technol. **4**(1), 36–43 (2015)
19. Anam, S.: An Approach for Recognizing Modi Lipi using Otsu's Binarization Algorithm and Kohenen Neural Network. Int. J. Comput. Appl. **111**(2), 28–34 (2015)
20. Maurya, R.K., Maurya, S.R.: Recognition of a medieval Indic- ' Modi' script using empirically determined heuristics in hybrid feature space, no. 2 (2018)
21. Solley, T.: A study of representation learning for handwritten numeral recognition of multilingual data set. Lect. Notes Networks Syst. **10**, 475–482 (2018)
22. Sifuzzaman, M., Islam, M.R., Ali, M.Z.: Application of wavelet transform and its advantages compared to Fourier transform. J. Phys. Sci. **13**, 121–134 (2009)
23. Juang, C.-N.-C., Chan, T.-M.: Speech detection in noisy environments by wavelet energy-based recurrent neural fuzzy network. Expert Syst. Appl. 36 (2009)
24. Li, T., He, H., Yi, X.: Multi-component seismic response analysis of offshore platform by wavelet energy principle. Coast. Eng. **56**(8), 810–830 (2009)
25. Mamatha, B., Valli Kumar, V.: ISAR image classification with wavelet and watershed transforms. Int. J. Electr. Comput. Eng. **6**(6), 3087–3093 (2016)
26. Wu, X.Q., Wang, K.Q., Zhang, D.: Wavelet energy feature extraction and matching for palmprint recognition. J. Comput. Sci. Technol. **20**(3), 411–418 (2005)
27. Achuthan, A., Rajeswari, M., Ramachandram, D., Aziz, M.E., Shuaib, I.L.: Wavelet energy-guided level set-based active contour: a segmentation method to segment highly similar regions. Comput. Biol. Med. **40**(7), 608–620 (2010)
28. Akouaydi, H., Abdelhedi, S., Njah, S., Zaied, M., Alimi, A.M.: Decision trees based on perceptual codes for on-line Arabic character recognition, pp. 153–157 (2017)

Applications of Artificial Intelligence in the Context of Challenges Amid COVID-19

Chintal Upendra Raval, Priyalba Vaghela, Aishwariya Budhrani, and Rakshil Vijaybhai Kevadiya

Abstract *Objectives*: The aim is to study the potential contribution of AI technologies and tools to all aspects of the COVID-19 crisis response. *Methods*: The research puts together the combined perspectives of rapid analysis of the literature conducted on the Google Scholar, Scopus, and PubMed database using the artificial intelligence or AI and COVID-19 or Coronavirus keyword. *Findings*: In this research, we have gathered the latest AI information for COVID-19 and then evaluated it to identify its possible applications to highlight substantial opportunities, challenges and research directions agendas resulting from the rapid introduction of AI in a number of areas: health care, psychiatry, education, packaging industry, data and network security, and Business-to-Business (B2B) Marketing Industry during COVID-19.

Keywords AI application · Artificial intelligence (AI) · AI technologies · B2B Marketing · COVID-19 · Data and network security · Education · Health care · Packaging industry · Psychiatry · AI solutions · Survey on AI tools

1 Introduction

Artificial intelligence (AI) has been a breakthrough technology in recent times. We have started to learn ways of communicating with the devices. We have been successful in guiding machines to perform a set of specific tasks using their intelligence. Machines have recently been doing a great deal of intelligent activities using cognitive intelligence as referred to as the natural intelligence (NI) that humans and other animals have shown [1].

AI will play an important role in shaping core-sector growth in emerging economies such as India. For a positive result, we are adopting new technologies and conducting experiments with lots of new things. AI technology will be the exquisite technology to give core sectors a raise and help us digitize rapidly [1].

C. U. Raval (✉) · P. Vaghela · A. Budhrani · R. V. Kevadiya
Devang Patel Institute of Advanced Technology and Research (DEPSTAR), Faculty of Technology & Engineering (FTE), Charotar University of Science and Technology (CHARUSAT), Anand, India

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_49

Many sectors have grown markedly and have contributed more to the economy. The result would be more optimistic by adding technology to those core sectors. Sectors like infrastructure, financial services, technology, automotive, and health care have grown rapidly all throughout the country. The research puts together the combined perspectives of rapid analysis of the literature conducted on the Google Scholar, Scopus, and PubMed database using the artificial intelligence or AI and COVID-19 or Coronavirus keyword [2].

The remaining sections of this article are organized according to the following: Sect. 2 presents many of the challenges and applications in various domains. The study is concluded in Sect. 3.

2 Application of Artificial Intelligence in the Context of Challenges During COVID-19 for Various Disciplines

2.1 Health Care

As with almost every aspect of the healthcare environment, the COVID-19 pandemic has restructured applications of artificial intelligence. AI has numerous healthcare applications. Whether used for discovering ties between genetic codes, controlling surgical robots or even optimizing hospital efficiency, artificial intelligence reinvents—and revitalizes—modern health care through machines capable of predicting, learning, and reinvigorating it [3]

1. **Using artificial intelligence for COVID-19 early detection, response and recovery**: AI tools and techniques can help to detect epidemiological patterns by mining and diagnose the virus. It helps in predicting the evolution of virus, helps to prevent or identify virus transmission chains by monitoring and contact tracing of individuals, respond to the health crisis by learning personalized information and monitoring the treatment and enhancing early warning systems.
2. **Case Projection and Mortality**: AI technology can monitor and forecast against the threats and the possible spread of the infection and the existence of the virus. In addition, one can estimate the number of cases that resulted in positive and mortality in any region. AI could help to classify the regions, the people and the country most vulnerable and take action accordingly.
3. **Drug Discovery and Vaccine Development**: Analysis of available data for COVID-19 by AI is used for drug research. This seems to be important for drug delivery design and development. This technology is used to accelerate drug testing in actual environments, where regular testing takes a great deal of time and therefore benefits significantly from accelerating this cycle, which may not be possible for humans. AI can help to determine useful medicinal products for treating patients with COVID-19. This has become an effective tool for designing diagnostic tests and for creating vaccines. In addition, AI is useful for clinical trials throughout vaccine development.

4. **Reducing healthcare employees' workload**: In healthcare facilities, semi-autonomous robots and drones are configured to respond to urgent needs such as supplying food and medications, sanitation and sterilization, assisting medical professionals, and distributing equipment.

2.2 AI in Psychiatry

In the midst of the COVID-19 pandemic, developments in artificial intelligence (AI) innovations have shown new healthcare capabilities and opportunities. Whilst the detection, likely therapy, and mitigation are commonly spoken of, its implications for mental health are a rarely addressed topic. Many were plague by the psychiatric condition of being unable to experience a normal life during lockdowns [4]. Additionally, the extra stress of cryptic terms of employment and an approaching recession has induced psychiatric disorders and human psychological issues. In addition, with strict social distance rules in place and undermining current healthcare facilities, individuals are searching for options that will still be open, removing the need to schedule appointments early.

1. **Detecting mental illness with artificial intelligence**: Modern streams of ample data mean that data-driven AI methodologies in terms of prediction and detection can be used to develop a detection/prediction models for psychological disorders. In particular, an individual's "digital exhaust" may be mined for behavioral or psychological wellness information, and the data obtained from their various personal digital platforms and connections with social media.
2. **Reporting and socially distanced mental health monitoring with AI**: Premised on language and other digital information which we as individuals create in our everyday routines, there is immense potential to use language technology and AI to help track psychological health shifts sensory perceptions. Such data may be posts on social media or digital platforms; we may exchange messages or any written text or spoken content that we generate. Individual sensing or online phenotyping, and the processing in a natural language of clinical content and social media information [5].
3. **AI Chabot's**: Chabots are based on artificial intelligence and assess questions, make understanding of them and answer in tones that are clinically safe.

2.3 Education

C-19 compelled universities across India, and indeed the world, to abandon physical classrooms and switch on to online classrooms. While for most private universities in India, this transformation has been smooth, and the public ones are still adapting. There have also been debates about the existence of courses and the future of assessment and evaluation—whether or not they should be done online. Artificial intelligence can have a serious influence on our educational future.

1. **Differentiated and customized learning**: Using machine learning, an AI branch that considers patterns in the data, teachers can deduce valuable insights from student academic performance and make informed, efficient decisions that help guide individuals in the finest potential direction. By collecting the data, ML algorithms can determine where a student has more difficulties and support them by providing them with handcrafted material, assignments and lessons which can help them bridge the gap in learning [6]. ML algorithms can classify appropriate learning pathways for students by collecting and analyzing large-scale student data and ensuring that they face the least difficulty on their way to professional development.
2. **Innovative tutoring**: AI helps connect the best personnel and strengthen the tutoring experience through digital platforms that communicate and focus on solving homework and assignments for millions of students. The platform now lightens the burden of moderators by identifying and filtering out spam as well as low-quality questions and answers owing to AI algorithms, helping the staff focus on delivering quality services.
3. **AI and Teacher collaboration**: AI will drive productivity, personalization and optimize admin tasks to empower teachers to have the time and freedom to provide understanding and adaptability—uniquely human skills where machines will combat.
4. **Universal access to all the students**: Artificial intelligence technologies may help make global classrooms accessible to all, including those who speak different languages or who may have visual or hearing impairments. This also brings new opportunities for students who may not be able to attend school due to an illness or may need to study at a different level or on a subject that is not accessible in their own school [7].

2.4 Transport and Logistics

COVID-19 and the initiatives aimed at controlling its spread have produced unprecedented conditions for transportation and supply chain networks across the nation and the world. Existing networks are strained by the rapid movement of medical supplies and essential goods to where needed. In many other nations, delays in ocean shipping, airfreight, and manufacturing challenge the timely supply of essential commodities [8].

With huge disruptions in today's supply and demand chains, precise, connected inventory and order, data are essential to driving business decision making.

1. **Logistics optimization using historical and real-time data**: Do smarter route planning by taking into account historical and real-time road conditions, weather, traffic, waiting times, maintenance stops, and more.
2. **Robotic transportation and Delivery solutions**: In the present times where human-to-human contact can be fatal in many areas and everything depends on one-touch away services via mobile devices, service providers need to promote

the best service and delivery experiences for customers. Drones (unmanned air vehicles) carry goods and passengers. AI in drone transport is a solution to possible human errors, in all modes of transport, including aviation, which is the safest mode of transport, and innovative research is undertaken to improve safety standards by removing errors in the human factor by using emerging technologies such as artificial intelligence, machine learning and advanced sensors [9].

3. **Optimization of the supply chain by data-driven applications**: With minimum manual effort, control the workflow more effectively. Empower the supply chain experts with smart tools trained in big data to manage supply and demand, scrub staff schedules and automate Chabot's customer support [10].
4. **Route planner with traffic prediction and real-time updates**: Provide traffic congestion and incidents, adjust the route on the go, and allocate extra time when traffic jams are inevitable.

2.5 Packaging Industry

The packaging is vital to pharmaceuticals, FMCG, food, and a whole host of categories that come under the purview of the critical maintenance act of services. Many packaging plants are now operational alongside food and pharmaceutical plants across the country. Still, a countrywide lockdown has resulted in a shortage of raw materials, lack of workforce, restrictions on transportation, and more [11].

Artificial intelligence has immense potential to offer the packaging industry innovative experiences. With AI being enforced in the packaging sector in wider processes, it is probable to expect the technology to leave its imprint across the entire supply chain [12].

1. **AI-enabled Vision Inspection Systems**: AI can also help with intelligent product testing to ensure that shipments are consistent on what is shipped, to whom, and in what quantity. Using AI-powered tools such as AI-powered computer vision can carry out quality inspection, identification, and monitoring in real time, ensuring minimum human interaction [13].
2. **Date Labeling**: By applying machine learning techniques in processes such as date labeling, businesses not only benefit from better-standardized procedures but also from customer dissatisfaction. Incorrectly, labeled products can result in inspection failures, dissatisfied customers and even potentially costing companies their hard-earned revenues. Using machine learning in the packaging industries, date labeling can become a standard, ensuring fewer human intervention and greater process efficiencies [14].
3. **Artificial Intelligence for warehouse automation**: It is a complex process when it comes to warehouse management and comes with its unique set of challenges. There is strong interconnectedness between warehouse and automation systems the more efficient and connected the results are to other tools and processes. Intelligent warehouses are warehouses operated by the Internet

of things (IoT), artificial intelligence, and other intelligent technologies such as radio frequency identification (RFID) that incorporate data from different sources in the warehouse to develop and automate processes on their own, without programming or human involvement.

4. **Sorting Recycled Goods**: Sorting goods is essential to ensure the effective use of recycled products and by-products. However, this tedious, laborious, and time-consuming work gives humans little benefit. The use of artificial intelligence and robotics should provide a decent fit for that reason. Robots are equipped with artificial intelligence to execute such activities, with the dawn of intelligent automation upon us. They are quicker, more efficient and are able to work more hours a day.

2.6 Data and Network Security

Employees and other stakeholders will begin remote work in their response to the Coronavirus outbreak, and there is increased pressure on the management of cyber-security risk in an organization. The IT function, too, is under enormous pressure. In certain companies, IT professionals need to expand remote work capability to workers who had not previously worked from home. That includes, in some cases, their service partners.

1. **Fraud detection**: Decision intelligence, an AI-based fraud detection, uses predictable consumer behavior-based algorithms. It evaluates the typical spending habits of customers, the vendor, the location of the purchase, and a variety of other sophisticated algorithms, to evaluate whether a purchase is out of the ordinary.
2. **Mitigating malware**: AI systems by using complex algorithms can help to detect viruses and malware, so that AI can then run pattern recognition in applications. AI systems can be trained before joining the network to recognize even the smallest behaviors of ransomware and malware attacks and then isolate them from that network. They can also use predictive functions that go beyond conventional approach speeds.
3. **Security and Crime Prevention**: Systems running on AI unlock the capacity for natural language processing that automatically collects information by combing posts, news, and cybe- threat studies. This knowledge will provide insight into trends, cyber-attacks, and methods for prevention.
4. **Spam filter applications**: Gmail detects and blocks unwanted spam and fraudulent emails using AI. Gmail's AI has been trained by the billions of active Gmail users—whenever you click on an email with "Spam" or "Not Spam", you are actually helping to train the AI to recognize spam. Thus, the AI has developed so that it can detect even the sneakiest of spam emails that attempt to go undetected as "regular" emails.

2.7 Business to Business Marketing Industry

In contexts of marketing lead generation, AI can also be extremely beneficial for companies, so that while people are working remotely, your business is not influenced. AI interference will boost the company with usages such as enhanced marketing strategies and increased sales using the right product lead. AI can help process huge fractions of data in split seconds, which is unparalleled with the amount of the time it takes for an individual to complete the very same task [15].

1. **Sales prediction using AI-enabled Marketing and Sales package**: Marketing Software powered by artificial intelligence is the finest business tool to boom. Brands can now run numerous different campaigns across a unified platform, easily. The networks are designed to combine and incorporate marketing campaigns with Twitter, Facebook, and Instagram. Such suites are responsible for the time-consuming 24×7 tasks from automating the ad-buying process to optimizing promotions.
2. **Account-Based Marketing**: AI will help analyze data to help advertisers notify about which accounts have the greatest inclination to buy. In addition, the ability of AI to recognize behavior patterns can also create relationships, allowing marketers to better understand human motives, thus allowing to produce better messages.
3. **Streamlining Communications**: Chabot's are really a game-changer. Numerous Web sites use Chabot's because they are wonderful to answer general user-requested FAQs without the need for personnel. Chabot's can have an impact on the customer experience and are of great help to employers with a short staff. Intelligent Chabot's are getting all the rage. They are driven by AI and can interact with humans in real time. Using predefined keywords, questions and answer, classifications for your Web site, SMS, Facebook, and WhatsApp profile, these custom bots can be developed to generate automated responses.
4. **AI-powered business marketing**: It has indeed now become task with the entire country under complete lockdown knowing whether there is any large-scale grocery store or pharmacy near you. This is where you can market your business on Google Search or Maps using AI-enabled tools.

3 Conclusion

The route to increasing AI applications has the potential for changing and influencing many aspects of human life throughout. This study presents consolidated yet possible applications to highlight substantial opportunities, challenges, and research directions agendas resulting from the rapid introduction of AI in a number of areas: health care, psychiatry, education, packaging industry, data and network security, and B2B Marketing Industry during COVID-19. COVID-19 can be mitigated by equipping ourselves with AI and using it as a powerful weapon in case of need.

References

1. Technologies, L.: Learnitude Technologies (Internet). Learntechx.com (cited 13 June 2020) (2020). Available from: https://learntechx.com/blog/applications-of-artificial-intelligence-in-top-10-areas
2. Ahuja, A., Reddy, V., Marques, O.: Artificial intelligence and COVID-19: a multidisciplinary approach. Integr. Med. Res. 100434 (2020)
3. Thomas, A.: Significance of AI-based mental healthcare systems Amid Covid-19. Analytics India Magazine (2020). https://analyticsindiamag.com/significance-of-ai-based-mental-healthcare-systems-amid-social-distancing-measures/. Last accessed: 23/09/2020
4. https://www.healtheuropa.eu/ai-in-psychiatry-detecting-mental-illness-with-artificial-intelligence/95028/. Last accessed: 23/09/2020
5. D'Alfonso, S.: AI in Mental Health. Current Opinion in Psychology (2020)
6. Editorial, G.: AI applications to intelligent vehicles for advancing intelligent transport systems. IET Intel. Transp. Syst. **14**(5), 267–269 (2020)
7. Dickson, B.: How artificial intelligence is changing education-TechTalks. https://bdtechtalks.com/2017/03/09/artificial-intelligence-education-edtech/. Last accessed: 23/09/2020
8. Editorial, G.: Big data analytics and artificial intelligence (AI) applications for smart transportation-selected papers from world transportation congress (WTC) 2018. IET Intel. Transp. Syst. **13**(3), 425–426 (2019)
9. Kurpjuhn, T.: Demystifying the role of AI for better network security. Netw. Secur. **2019**(8), 14–17 (2019)
10. Alshanbari, R., Khan, S., El-Atab, N., Mustafa Hussain, M.: AI powered unmanned aerial vehicle for payload transport application. In: 2019 IEEE National Aerospace and Electronics Conference (NAECON) (2019)
11. AI in Packaging industry a Smart Solution for the Future-. Day One (2020). https://www.day1tech.com/ai-in-packaging-industry-a-smart-solution-for-the-future-day-one/. Last accessed: 23/09/2020
12. Miklosik, A., Kuchta, M., Evans, N., Zak, S.: Towards the adoption of machine learning-based analytical tools in digital marketing. IEEE Access **7**, 85705–85718 (2019)
13. https://www.packaginginsights.com/news/ai-in-packaging-how-artificial-intelligence-is-driving-the-packaging-industry-forward.html. Last accessed: 23/09/2020
14. Hollstein, K., Weide-Zaage, K.: Advances in packaging for emerging technologies. 2020 Pan Pacific Microelectronics Symposium (Pan Pacific) (2020)
15. How COVID-19 Could Accelerate AI Adoption in the B2B Marketing (Internet). AiThority (cited 13 June 2020) (2020). Available from: https://www.aithority.com/guest-authors/how-covid-19-could-accelerate-ai-adoption-in-the-b2b-marketing-industry/
16. Vaishya, R., Javaid, M., Khan, I., Haleem, A.: Artificial intelligence (AI) applications for COVID-19 pandemic. Diabetes Metab. Syndr. **14**(4), 337–339 (2020)

Hardware Implementation of English to Regional Language Translator Using Arduino

Navneet Kaur Brar, Ajay Kumar, Manu Bansal, and Alpana Agarwal

Abstract According to statistics, only 20% of Indian population can understand English; others generally communicate and understand their regional languages only. Therefore, there is a need to demonstrate multilingual unique secure language translator in a country like India. In this paper, hardware implementation of a system is shown that converts English input into Hindi or Punjabi. It is implemented on Arduino UNO Atmega232 microcontroller. The experimental results have shown that this model occupies minimum memory and power to which system it will be attached. Moreover, the system is economical and robust.

Keywords Arduino · Language translator · Serial peripheral interface (SPI) · Integrated development environment (IDE)

1 Introduction

In this multilingual world, many a times, language barrier becomes a menace to understand the language of other people. India is a country, where people usually understand their regional languages rather than the English that has become a global language. They face problem where instructions are given in English at public places such as railway station or airport. Therefore, to overcome this problem, a system, as a helpdesk for them, is required that can help the layman to understand English language which could be converted into their native language. This model must be

N. K. Brar (✉) · A. Kumar · M. Bansal · A. Agarwal
Thapar Institute of Engineering and Technology, Patiala, India

A. Kumar
e-mail: ajay.kumar@thapar.edu

M. Bansal
e-mail: mbansal@thapar.edu

A. Agarwal
e-mail: alpana@thapar.edu

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_50

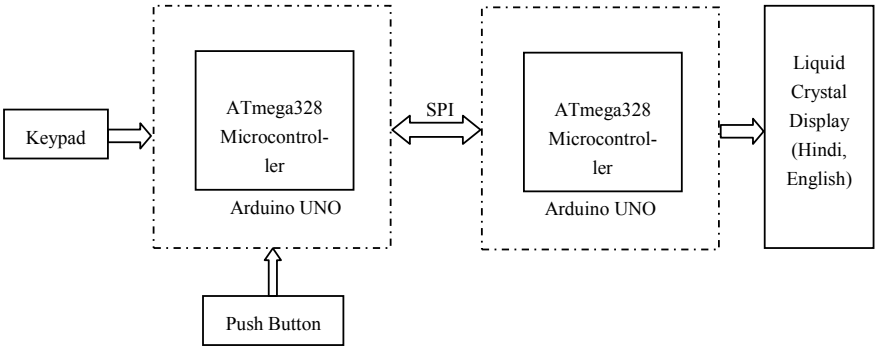


Fig. 1 Block diagram of system

economical and durable but not at the cost of memory and power to which system it will be connected.

The researchers have made various models to make such a reliable language translator. Much work has been done in sign language translator [1, 2]. However, there are only a few researchers who have designed a model that converts English to other languages [3]. In this paper, a language translator is designed that converts the English input into other regional languages, and over here Hindi and Punjabi are used as an example. The Arduino UNO is selected for the controller owing to its monetary value and has open software that is supported by most of the operating systems [4]. For the interface, serial peripheral interface (SPI) is used which is full duplex and works synchronously. The detail of the SPI can be studied from [5–7]. In the designed system, two Arduino boards are used. The first Arduino works as master as per SPI protocol which is connected to keypad and push button. The other Arduino acts as a slave in which language conversion is done and displayed on LCD. The block diagram is shown in Fig. 1.

The main contributions of the designed system are as follows:

- It uses negligible resources such as power or memory to which it is attached.
- In this system, more than one language can be displayed on LCD, which is prior selected by the user via a push button.
- It is robust, cost effective and as well as reliable.

The remainder of the paper is designed as follows. In Sect. 2, hardware description is discussed in detail. Section 3 demonstrates the proposed methodology. Further, in Sect. 4, results are shown. Finally, the conclusion is drawn in Sect. 5.

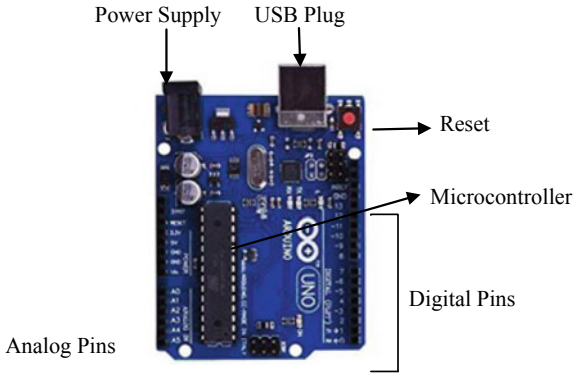


Fig. 2 Arduino UNO board

2 Hardware Description

In this section, the basics of the Arduino board and serial peripheral interface (SPI) is discussed in detail. Section 2.1 presents the overview of Arduino board. Section 2.2 explains the SPI interface protocol between master and slave.

2.1 Overview of Arduino Board

Arduino is a board that contains ATmega microcontroller, SRAM memory, flash and EEPROM storage [8]. There are different types of Arduino boards such as Arduino UNO, Mega, UNO, Pro mico and RS232. They are differentiated based on their memory, processor speed and number of analog and digital pins. The schematic of Arduino UNO that contains Atmega 328 microcontroller is shown in Fig. 2. The Arduino can be easily interfaced with other devices as libraries are available for various protocols and peripherals such I2C, SPI, UART and LCD, Keypad, servomotors etcetra respectively. In addition to it, Arduino can be used in various applications such as obtain input data from sensors, output physical, visual and audio output [9].

2.2 Interface Protocol

The interface protocol is a way to communicated data between two devices. Nowadays, many interface protocols such as universal asynchronous receiver–transmitter (UART), serial peripheral interface (SPI) and inter-integrated circuit (I2C) are used. However, among all of them, SPI is widely used due to its fast speed and synchronicity.

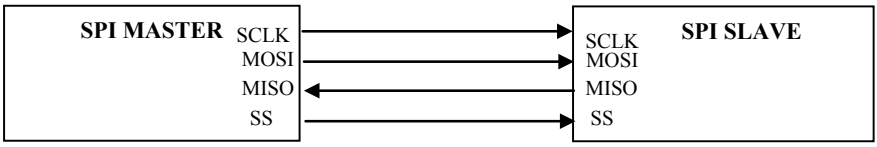


Fig. 3 Master–slave configuration in SPI

SPI is a serial communication protocol that works in full duplex mode. In SPI interface protocol, the communication is between a master and slave. The SPI interface uses four pins generally called as slave select (SS), master output slave input (MOSI), master input slave output (MISO) and synchronous clock (SCK) as shown in Fig. 3 [5]. In Arduino, digital pins are fixed for SPI, i.e. 10 for SS, 11 for MOSI, 12 for MISO and 13 for SCLK [7]. In SPI, there is only one master that can transmit and receive data to a single or multiple slaves. The three pins SCLK, MOSI, MISO are common for all slaves, but chip select is different for every slave. The slave is configured by chip select in master [10].

3 Proposed Methodology

The language translator, which is both robust in hardware and economical at the same time, has become a problem. So, to meet the cost and durability constraints, this system is proposed. The technique is implemented on Arduino UNO board at hardware level, whereas at software level, the Arduino integrated development environment (IDE) platform 1.8.9 version is used for the coding purposes. The code is written in C language. In this section, the proposed methodology for language translator and the hardware connections is explained in detail.

3.1 Proposed Design

The entire system can be divided into three steps: (1) Selecting the language (2) Keying the input data through a keypad and (3) Display of output on LCD. The flow chart is shown in Fig. 4.

Step I: Selecting the language: The push button is used to decide the language, which is to be displayed on LCD. If the push button is pressed once, then the Hindi language is selected. Similarly, if the push button is pressed twice, Punjabi language is selected. This selection key goes along with the data to another Arduino via SPI.

Step II: Keying the input data from Keypad: The user gives the input in English from alphanumeric keypad, i.e. 4×4. In the code, alphkey() library is added in the Arduino from Sketch $\rightarrow$ Include library $\rightarrow$ Alphkey. In this keypad from each

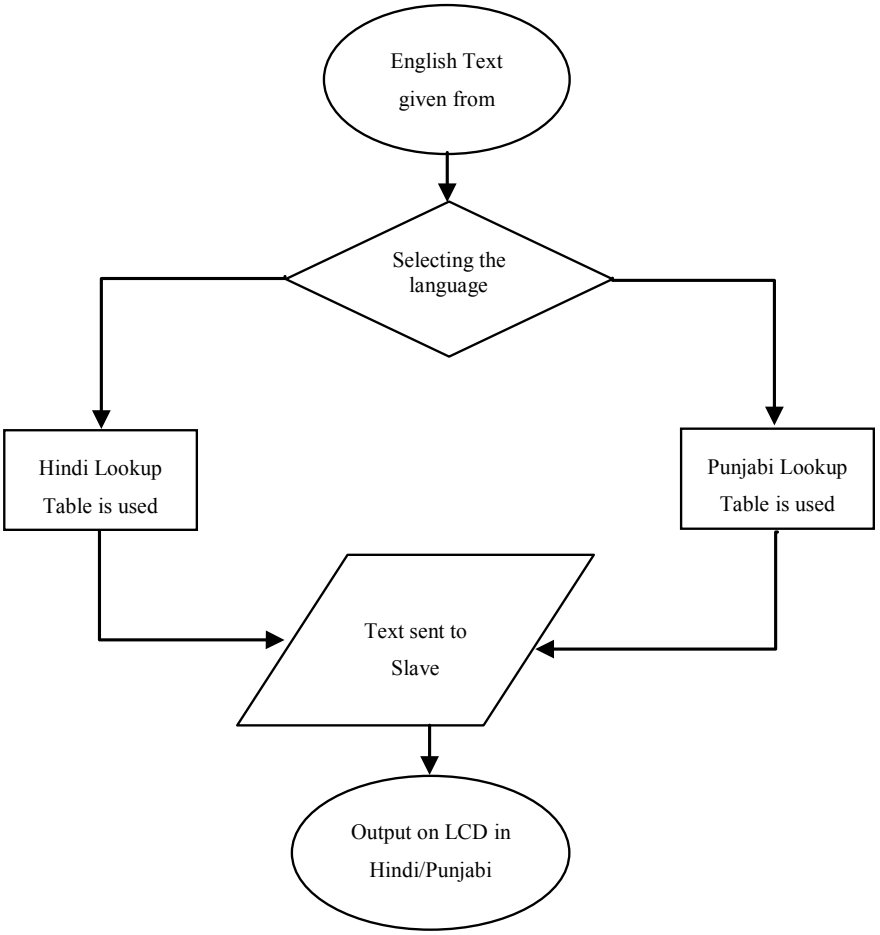


Fig. 4 Flow chart of system

key, three English characters input can be given. If the same key is pressed once, then the first character is selected, and if the same key is pressed within 0.5 s, then second character is selected. This message in English is stored in buffer, and along with that, language key is transmitted from Arduino 1 to Arduino 2 via SPI. In Arduino 2, received data is in ASCII characters corresponding to which Hindi or Punjabi characters are assigned from lookup table. If Hindi language was selected, then characters from Hindi lookup table are assigned and vice versa. The lookup table uses only 7% of the memory. Few of them have been shown in Table 1.

Step III: Display of output on LCD: After SPI transfer, according to the language selector key, the corresponding lookup table is selected. For example 1, key selected was 1, and then, the lookup table of Hindi characters corresponding to English

Table 1 Lookup table of characters

Keypad digit	English character	Hindi character	Punjabi character
1	A	क	ਪ
1	B	ख	ਮ
1	C	न	ਸ
2	D	आ	ਤ
2	E	त	ਜ
2	F	म	ਕ

characters will be displayed on LCD and vice versa for Punjabi language if key selector key is 2.

3.2 Hardware Connections

The hardware connections for the system are done in three parts. Initially, the connections of Arduino 1 to push button and keypad are made as per the reference manual of PmodKYPD [11]. Further, the Arduino 2 connections with LCD are done as per reference manual of Digilent LCD [12]. The default SPI connections of Arduino 1 to Arduino 2. The association between the various devices is listed in Tables 2 and 3.

Table 2 Pin assignment of Arduino 2

Hardware	Pin assignment	
Device	Arduino 2	Device
LCD	D9	RS
LCD	D8	E
LCD	GND	R/W
LCD	D7	DB3
LCD	D6	DB2
LCD	D5	DB1
LCD	D4	DB0
Arduino 1	D10	P10
Arduino 1	D11	P11
Arduino 1	P12	P12
Arduino 1	P13	P13

Table 3 Pin assignment of Arduino 1

Hardware	Pin assignment	
Device	Arduino 1	Device
Keypad	D5	C4
Keypad	D4	C3
Keypad	D3	C2
Keypad	D2	C1
Keypad	D9	R4
Keypad	D8	R3
Keypad	D7	R2
Keypad	D6	R1
Push Button	D1	P1
Arduino 2	D10	P10
Arduino 2	D11	P11
Arduino 2	D12	P12
Arduino 2	D13	P13
LCD	D5	DB1
LCD	D4	DB0
Arduino 1	D10	P10
Arduino 1	D11	P11

4 Results and Discussions

In this section, the results of the proposed system are discussed. The scheme is implemented on Arduino UNO that comprises ATmeg328 controller. The code is simulated on integrated development environment (IDE) Arduino 1.8.9, i.e. an open source platform at 115,200 baud rate. The Digilent 16 × 2 Liquid Crystal Display (LCD) and 4 × 4 Hex Keypad are used for the display and input data, respectively. The output at LCD in Hindi as well in Punjabi is shown in Figs. 5 and 6, respectively.

The performance of the system has been estimated on three parameters, i.e. overall cost, memory usage and execution time. The overall cost includes the cost of two Arduino, LCD, keypad and push button. In memory, both the program storage space and global variables are included. The execution time is the time of output display in another language to that of input given in English. The performance parameters are shown in Table 4.

5 Conclusion

In this paper, a language translator system is implemented at hardware level. The main advantage of this system is that it is cost effective; the components used are

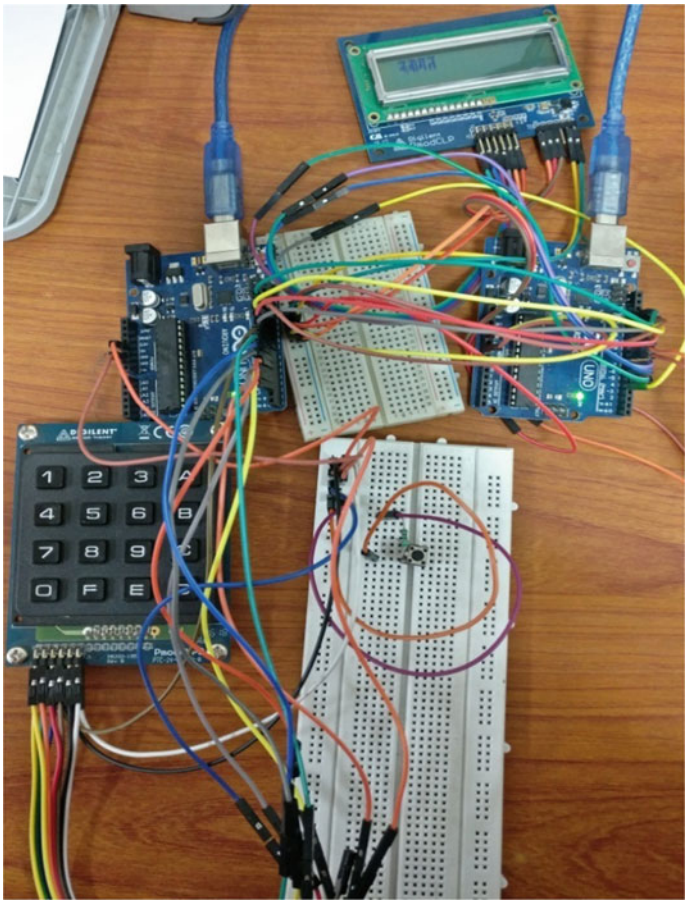


Fig. 5 Hindi text display on LCD

economical, i.e. less than INR2000 and are easily available. Moreover, it consumes less memory space approximately 42,000 bytes. The SPI protocol is used for the communication. The system is expandable to other multiple regional languages. In addition to it, cloud-based application language translator can also be designed. So that, it can provide security to the input data as well.

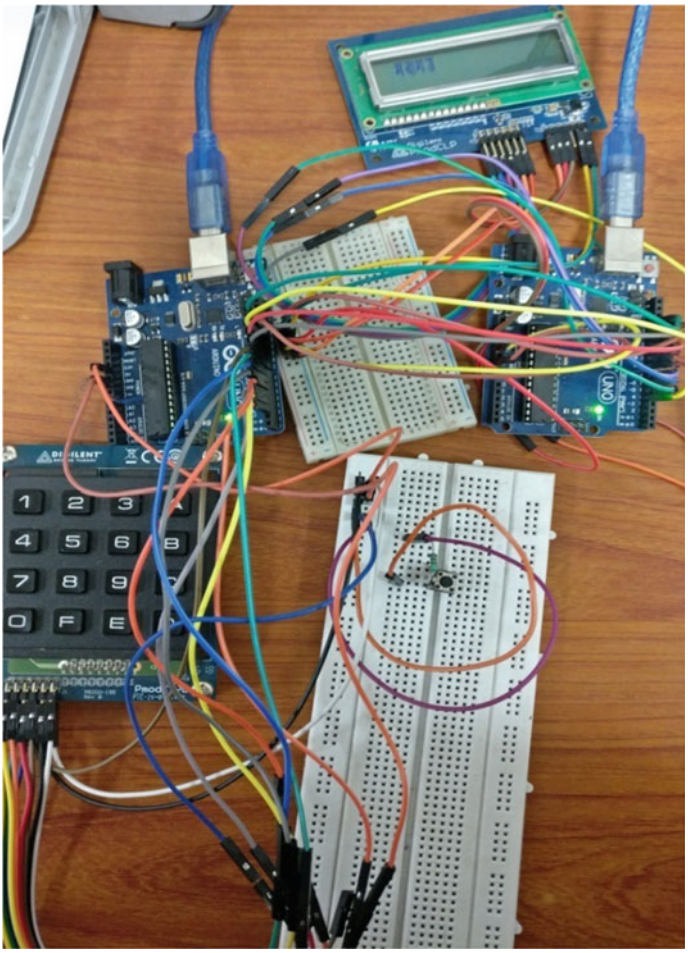


Fig. 6 Punjabi text display on LCD

Table 4 Performance parameters

Constraints	Value
Overall cost	Rs. 1830/- (per system)
Memory	42,256 Bytes
Execution time	0.325 μsec

Acknowledgements The authors are grateful to Ministry of Electronics and Information Technology (MeitY) for the financial support through 'SMDP Chip to System Design' Project. The authors would also like to thank Director, Thapar Institute of Engineering and Technology, Patiala, Punjab, for providing necessary resources to carry out this research.

References

1. Sapkota, B., Gurung, M.K., Mali, P., Gupta, R.: Smart glove for sign language translation using Arduino. In: KEC Conference 2018, Kantipur Engineering College, Dhapakhel, Lalitpur, pp. 5–11 (2018). https://kec.edu.np/wp-content/uploads/2018/10/2.pdf
2. Quiapo, C.E.A., Ramos, K.N.M.: Development of a sign language translator using simplified tilt, flex and contact sensor modules. In: 10 Conference (TENCON) 2016 IEEE Region, Singapore, pp. 1759–1763 (2016). https://doi.org/10.1109/TENCON.2016.7848321
3. Johnson, M., Schuster, M., Le, Q.V., Krikun, M., Wu, Y., Chen, Z., Thorat, N., Viégas, F., Wattenberg, M., Corrado, G., Hughes, M.: Google's multilingual neural machine translation system: enabling zero-shot translation. Trans. Assoc. Comput. Linguist. **5**, 339–351 (2017). https://doi.org/10.1162/tacl_a_00065
4. Heera, S.Y., Murthy, M.K., Sravanti, V.S., Salvi, S.: Talking hands—An Indian sign language to speech translating gloves. In: International Conference on Innovative Mechanisms for Industry Applications (ICIMIA), pp. 746–751. IEEE, Bangalore (2017). https://doi.org/10.1109/ICIMIA.2017.7975564
5. Wootton, C.: Serial Peripheral Interface (SPI). In: Samsung ARTIK Reference. Apress, Berkeley (2016). https://doi.org/10.1007/978-1-4842-2322-2_21
6. Anand, N., Joseph, G., Oommen, S.S., Dhanabal, R.: Design and implementation of a high-speed serial peripheral interface. In: International Conference on Advances in Electrical Engineering (ICAEE). IEEE, Vellore, pp. 1–3 (2014). https://doi.org/10.1109/ICAEE.2014.6838431
7. Qiang, J., Gu, Y., Chen, G.: FPGA Implementation of SPI bus communication based on state machine method. J. Phys. Conf. Ser. 1449, 012027 (2020). (IOP Publishing Ltd.). https://doi.org/10.1088/1742-6596/1449/1/012027
8. Turner, R.: Arduino Programming 2 Books in 1, the Ultimate Beginner's and Intermediate Guide to Learn Arduino Programming Step by Step, Ryan Turner (2019)
9. Margolis, M., Jepson, B., Weldin, N.R.: Arduino Cookbook: Recipes to Begin, Expand, and Enhance Your Projects. O'Reilly Media, Sebastopol (2020)
10. Wang, S., Shi, Y., Han, Z., Kong, L., Wang, N.: Dialog Semiconductor UK Ltd., 2020. Serial Communication Protocol. U.S. Patent Application 16/568,840
11. DIGILENT A National Instrument Company. https://reference.digilentinc.com/reference/pmod/pmodkypd/reference-manual
12. DIGILENT A National Instrument Company: https://reference.digilentinc.com/reference/pmod/pmodclp/reference-manual

A Comparison Between Power Spectral Density and Wavelet Transform for EEG-Based Sleep Onset Detection

K. P. Jayalakshmi, Y. Mahesha, and Seema Miranda Priya

Abstract In many neurological studies, the study on brain activity with respect to the continuous evaluation on the work of safe driving has received noticeable attention. The electroencephalogram (EEG) is a non-invasive physiological device that measures brain activity. Brain signals analyzed using electroencephalography can help in the prediction of the drowsiness in drivers. The proposed work is to establish a simple and precise sleep onset prediction method based on EEG signals acquired from a single-channel electrode. These EEG signals are processed using two different methods, fast Fourier transform and discrete wavelet transform. Distinct features are extracted from both processing methods separately and are used as training data set to train the classification models in MATLAB classification learner. A comparison has been made between the processing methods to evaluate which of the two methods give better accuracy. According to the obtained results, models trained using wavelet transform features gave an accuracy of 98%. Among the models trained, decision tree is the most accurate model having the fastest prediction speed, approximately up to 30,000 observations/second.

Keywords Electroencephalogram · Power spectral density · Fast Fourier transform · Wavelet transform · MATLAB · Sleep onset detection

1 Introduction

Sleep is a key function of the human brain, and it is important for a person to get decent amount of sleep. Lack of sleep can have a negative impact on aspects of understanding, reasoning, and vigilant attention. Driver drowsiness has been one of

K. P. Jayalakshmi (✉) · Y. Mahesha · S. M. Priya
St Joseph Engineering College, Mangaluru, Karnataka, India
e-mail: jayalakshmi@sjec.ac.in

Y. Mahesha
e-mail: maheshy@sjec.ac.in

S. M. Priya
e-mail: priyam@sjec.ac.in

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_51

the main causes of road accidents worldwide in recent years, leading to significant physical injuries, casualties, and substantial economic losses. Many of these road accidents and deaths could be avoided if the driver's drowsiness could be carefully monitored and drivers get early warnings. To make commutes safer for all, it is also necessary to establish a system for early detection of drowsiness. Many techniques have been proposed, including vehicle-based techniques such as the lane departure warning system and the steering wheel movement system, to allow both simple and cheap detection of driver drowsiness. Ingre et al. [1–4] video-based techniques such as the identification of the degree of eyelid closure [5–9] over the pupils in terms of time-based physiological signal methods and electroencephalography [10, 11]. The most effective means of detection amongst these methods are considered physiological signal-based methods, and the EEG, which is a non-invasive physiological means of monitoring and measuring the brain activity, is considered to have the closest relationship with drowsiness compared to other physiological signals. Swarnkar et al. [11–14]. Since the early 1930s, the electroencephalogram (EEG) signal has been used to describe the phases of sleep. In the standard specifications of Rechtschaffen and Kales, the determination of sleep stage includes stage awake, 1, 2, 3, 4, and rapid eye movement (REM) sleep [15]. Stage I is a drowsiness phase and is considered to be a transition between wakefulness and sleep, during which a person may not be aware of the fact that they were sleeping and can be awakened easily. Here, EEG signals are of low amplitude and high frequency, and the most dominant wave is the alpha wave, whereas gamma waves and beta waves are dominant during wake state. In second sleep stage, temperature of one's body reduces, and the heart rate slows down. Stages three and four are deeper stages of sleep. During REM sleep, dreaming occurs and brain activity increases. Throughout the sleeping phase, the sleeper continuously cycles through each of these stages. With regard to the age of the sleeper, these cycles differ [16, 17]. Human brain produces brainwaves depending on the state of a person, which is classified into delta (0–4 Hz), theta (4–8 Hz), alpha (8–12 Hz), beta (15–30 Hz), and gamma (greater 30 Hz) [15, 18]. Our work focuses on beta and alpha waves for sleep detection. The block diagram for sleep onset detection is shown in Fig. 1, and EEG signals from the Pz-Oz channel are selected in our approach as the theta- and alpha-based features in Pz-Oz EEG channel have more accuracy than the Fpz-Cz channel [19, 20]. In MATLAB, before signal processing, the sampling frequency of the signal is resized 100, 128 Hz; this is performed to make further computations easier. Signal processing and feature extraction are performed using two different methods, fast Fourier transform and wavelet transform. Features for training the classifier model are extracted, and the datasets are fed into the MATLAB classification learner. A comparison between the methods for extracting the features of EEG signal for sleep detection is performed to find out which method yields the most accurate classification result.

Fig. 1 Block diagram representation of the proposed method

Table 1 30 s epoch of wake and stage 1 sleep

Total subjects	Wake	Sleep Stage1
Six	500 epochs	316 epochs

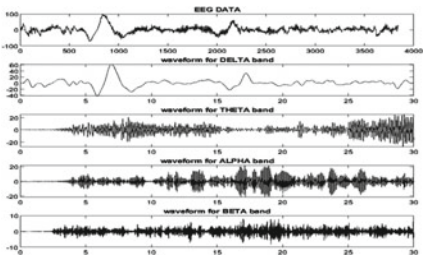
2 EEG Dataset

The dataset (Sleep-EDF [Expanded] Database) used in this work is available online from PhysioNet.Org. The dataset record consists of a series of 61 males and females sampled 100 Hz for almost 24 h. These signals are recorded in a 30 s epoch pattern.

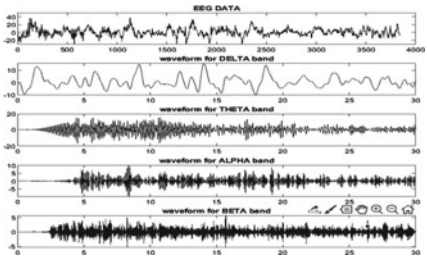
Table 1 shows 30 s epoch of wake and sleep samples of the EEG signal from Pz-Oz channel (for wake and stage one) which are used as inputs to the designed filter.

3 Signal Processing

EEG signals from Pz-Oz channel of six subjects were converted from EDF format into a text file which was loaded into MATLAB. Only the wake and stage one sleep state, signals are considered for further processing. Brainwaves pattern in these two stages are studied in order to extract classification features. A digital band-pass filter was designed to classify the signals into delta, theta, alpha, and beta. Signal of 30 ss epoch was given as the input to the band-pass filter. Figures 2a and 2b show the output of the digital Bandpass filter, with the wake and sleep stage 1 signal being filtered out to four sub-bands, respectively.



(a) Output of Bandpass filter during Wake stage



(b) Output of Bandpass filter during sleep stage

Fig. 2 Outputs of bandpass filter

4 Feature Extraction

The features extracted after PSD and DWT are exported to a table as set one and set two, respectively.

4.1 Power Spectral Density

Power spectral densities (PSD) have been determined from the performance of these filters. Power spectral density measures the signal intensity in the domain of frequency. PSD is very useful to know the frequency and amplitude of signal given in time series. Features like energy, entropy, and mean of power spectrum of the signal were extracted. Total of 816×12 training datasets were extracted from the following features:

- Energy: It represents the capability of the signal by determining the range under the curve of power at any period of time.

$$\text{energy} = \sum_{n=1}^{N} |x_i[n]|^2 \tag{1}$$

  where x is the EEG signal for each sub-band ranging from $i4 = 1$ to 4, and $N4$ is equal to the total number of data samples.
- Entropy: EEG signal is complex, nonlinear, non-stationary, and random in nature. Entropy is used to measure the level uncertainty. Entropy contains useful information and features that are exclusive to individuals.

$$\text{entropy} = \sum_{j=1}^{N} \left(x_{ij}^2\right) \log \left(x_{ij}^2\right) \tag{2}$$

  here, x is the EEG signal for each sub-band ranging from $i = 1$ to 4, and N is equal to the total number of data samples.
- Mean of PSD: Mean of the PSD of each sub-band is computed

4.2 Discrete Wavelet Transform

Discrete wavelet transform (DWT) is used in our paper. The method of calculating a DWT for a signal is by passing it through a series of filters. WT decomposes a signal into its sub-bands. By computing the energy of sub-bands, only desired bands are used for reconstruction which eliminates noise or unwanted signals. Since several mother

Fig. 3 Frequency decomposition using wavelet transform for Fs = 128 Hz

wavelet functions exist, the one with the highest similarity (in terms of frequency bands) to the EEG character is selected. According to the study [19], the order two (db2) smoothing function of the Daubechies wavelet makes it more acceptable for detecting changes in the observed signal. Depending on the dominant frequency variable of the EEG signal, the number of decomposition levels is selected. The categorization of decomposed EEG signal into four levels with matching frequency distribution at sampling frequency 128 Hz is shown in Fig. 3. Each stage output provides detail of the signal "D" and an approximation of the signal "A", where the output of the previous stage becomes the input for the next step. Here, D2 is considered as the beta sub-band, D3 as the alpha sub-band, D4 as the theta sub-band, and A4 as the delta sub-band. Decomposition coefficients of each sub-bands are processed to extract the features for the training data set. Total of 816×8 training data set were extracted from the following features:

- Finite summation of decomposed coefficients: sum of squared magnitude of coefficients of each sub-band.
- Entropy: Entropy of the decomposed coefficients, formula for entropy is same as the one used in Eq. 2

4.3 MATLAB Classification Learner

For sleep onset detection, classification learner is used. It performs supervised learning tasks, to interactively explore the imported data, selecting specific features, training models and allows comparison and evaluation of results. The training data sets are imported to the classification learner separately, and the classification models are trained.

5 Results and Discussions

5.1 Training Using Set One

Using MATLAB classification learner, the extracted features were imported into the application. Tables 2 and 3 show the range of values of the features extracted after computing the power spectral density of wake and sleep signals, respectively. These features are used as inputs for training the classification models. All the classification models were trained and self-validated. Using "select specific feature" option in classification learner, models were trained using different feature combinations. From Table 4, it can be concluded that the fine tree, course tree (decision tree), and quadratic SVM give the most accurate prediction. Quadratic SVM has the fastest prediction speed (total number of observations per second) but has an accuracy of 99.8%. Course tree which is trained using all the twelve features has 99.9% accuracy, even though its prediction speed is slower than SVM, it took the least amount of time for training.

5.2 Training Using Set Two

Finite sum and entropy of each brainwave were computed. Tables 5 and 6 show the range of values of the features extracted after DWT of wake and sleep signals, respectively.

These features are fed into the classification learner as training dataset. From Table 7, it can be inferred that decision tree gives the most accurate prediction. Course tree which is trained using six features has 99.9% accuracy and took the least amount of time for training, but the prediction speed of fine tree is better than the rest.

Table 2 Range of feature values obtained for wake state

Feature	Delta	Theta	Alpha	Beta
Energy	1.434–10.81	1.89–7.78	1.45–5.62	1.45–5.62
Entropy	6.74–2183.82	16.10–612.24	7.82–316.41	6.02–746.93
Mean of PSD	0.24–97.78	0.008-2.67	0.09–30.98	0.64–99.03

Table 3 Range of feature values obtained for sleep stage 1 state

Feature	Delta	Theta	Alpha	Beta
Energy	1.13–15.11	0.95–8.88	0.71–2.89	0.73–1.28
Entropy	2.53–3688.18	2.48–975.56	0.74-59.2	0.80–7.95
Mean of PSD	0.296–98.08	0.046–3.48	0.26–40.14	0.34–96.97

Table 4 Most accurate models obtained after training and validation

Model	Fine tree	Coarse tree	Fine Tree	Medium Tree	Quadratic SVM
Accuracy (%)	99.9	99.9	99.9	99.9	99.8
Misclassification cost	1	1	1	1	2
Prediction speed	24,000	30,000	26,000	28,000	31,000
Training time (s)	2.055	0.3178	1.0748	0.38717	1.3422
Total features used	12	12	4	4	6

Table 5 Range of feature values obtained for wake state

Feature	Delta	Theta	Alpha	Beta
Finite sum	4.59–290.66	1.52–24.4	2.32–36.23	1.08–55.78
Entropy	14.47–23.49	3.41–135.42	6.49–199.60	2.05–329.86

Table 6 Range of feature values obtained for sleep stage 1 state

Feature	Delta	Theta	Alpha	Beta
Finite sum	2.49–431.23	0.85–21.18	0.56–11.28	0.41–0.99
Entropy	6.35–3645.92	1.42–105.91	0.74–51.09	0.38–2.045

Table 7 Most accurate models obtained after training and validation

Model	Fine tree	Medium tree	Coarse tree	Fine tree
Accuracy (%)	99.9	99.9	99.9	99.9
Misclassification cost	1	1	1	1
Prediction speed (obs/s)	28,000	23,000	24,000	34,000
Training time (s)	0.9401	0.4008	0.38798	0.87645
Total features Used	8	8	8	6

Table 8 Performance results

Trainingdata	Testingdata	Correctly detected	Accuracy (%)
Wake-500	50	Setone-47	Setone-94
Sleep-316		Settwo-49	Settwo-98

5.3 *Output of Trained Models*

Coarse decision tree is selected as the training model for both the dataset. Set 1 includes values of energy, entropy, and mean of PSD. Set 2 includes values of finite summation and entropy of decomposed coefficients. Table 8 shows the output of the trained models, the test epochs containing both the wake stage and the sleep stage one was given as input to the classification model. Model trained using set one classified 47 out of 50 epochs into wake or sleep stage correctly, with an accuracy rate of 94%. Model trained using set two classified 49 epochs correctly, therefore yielding an accuracy of 98%.

6 Conclusion

This work focused in designing an online automatic sleep onset detection system. EEG signals were acquired from an online database "PhysioNet.Org". After preprocessing the signal, the sampling frequency is changed 100, 120 Hz using MATLAB. The test data was segregated into delta, theta, alpha, and beta sub-bands using FFT. FFT is applied on each sub-band, and its power spectral density is computed. The wavelet transform is computed using Daubechies filter of order 2 with 4 levels. Certain features were extracted using the power spectral density and the wavelet transform which was later used to train the models. MATLAB classification learner is used to train the models which showed that the features extracted through wavelet transform gives better accuracy than the features extracted after FFT. The models were trained using wake and stage 1 sleep where stage 1 is the drowsy state, thus making the system suitable for drowsiness detection. As a future work, the designed method can be implemented on hardware and an alarm system can be designed to alert the user when he/she is drowsy.

References

1. Ingre, M., Akerstedt, T., Peters, B., Anund, A., Kecklund, G.: Subjective sleepiness, simulated driving performance and blink duration: examining individual differences, **15**, pp. 47–53 (2006). https://doi.org/10.1111/j.1365-2869.2006.00504.x.
2. Otmani, S., Pebayle, T., Rog'e, J., Muzet, A.: Effect of driving duration and partial sleep deprivation on subsequent alertness and performance of car drivers **84**, 715–24 (2005). https://doi.org/10.1016/j.physbeh.2005.02.021
3. Boyraz, P., Hansen, J.: Active accident avoidance case study: Integrating drowsiness monitoring system with lateral control and speed regulation in passenger vehicles **293–298**, (2008). https://doi.org/10.1109/ICVES.2008.4640863
4. Thiffault, P., Bergeron, J.: Monotony of road environment and driver fatigue: a simulator study **35**, 381–91 (2003). https://doi.org/10.1016/S0001-4575(02)00014-3
5. Dinges, D., Grace, R.: Perclos: a valid psychophysiological measure of alertness as assessed by psychomotor vigilance (1998)

6. Wierwille, W.W., Wreggit, S., Kirn, C., Ellsworth, L.A., Fairbanks, R.: Research on vehicle-based driver status/performance monitoring; development, validation, and refinement of algorithms for detection of driver drowsiness, final report (1994)
7. Hanowski, R., Blanco, M., Nakata, A., Hickman, J., Schaudt, W., Fumero, M., Olson, R., Jermeland, J., Greening, M., Holbrook, T., Knipling, R., Madison, P.: The drowsy driver warning system field operational test: data collection methods final report (09 2008). https://doi.org/10.13140/RG.2.2.24144.25608
8. Li, G., Chung, W.-Y.: Estimation of eye closure degree using EEG sensors and its application in driver drowsiness detection. Sensors (Basel, Switzerland) **14**, 17491–17515 (2014). https://doi.org/10.3390/s140917491
9. Li, G., Chung, W.-Y.: Detection of driver drowsiness using wavelet analysis of heart rate variability and a support vector machine classifier. Sensors (Basel, Switzerland) **13**, 16494–16511 (2013). https://doi.org/10.3390/s131216494
10. Jung, S.J., Shin, H.-S., Chung, W.-Y.: Driver fatigue and drowsiness monitoring system with embedded electrocardiogram sensor on steering wheel **8**, 43–50 (2014). https://doi.org/10.1049/iet-its.2012.0032
11. Swarnkar, V., Abeyratne, U., Hukins, C.: Objective measure of sleepiness and sleep latency via bispectrum analysis of eeg **48**, 1203–1213 (2010). https://doi.org/10.1007/s11517-010-0715-x
12. Iber, C., AncoliIsrael, S., Chesson, A., Quan, S.: The aasm manual for the scoring of sleep and associated events: Rules, terminology and technical specifications (2007)
13. Khushaba, R., Kodagoda, S., Lal, S., Dissanayake, G.: Driver drowsiness clas- sification using fuzzy wavelet-packet-based feature-extraction algorithm **58**, 121–31 (2011). https://doi.org/10.1109/TBME.2010.2077291
14. Akin, M., Kurt, M., Sezgin, N., Bayram, M.: Estimating vigilance level by using EEG and EMG signals **17**, 227–236 (2008). https://doi.org/10.1007/s00521-007-0117-7
15. Hal, B., Rhodes, S., Dunne, B., Bossemeyer, R.: Low-cost eeg-based sleep detection **2014**, 4571–4 (2014). https://doi.org/10.1109/EMBC.2014.6944641
16. Rechtschaffen, A., AA, K.: A manual of standardized terminology techniques and scoring system for sleep stages in human subjects, **55** (1968)
17. Hori, T., Sugita, Y., Koga, E., Shirakawa, S., Inoue, K., Uchida, S., Kuwahara, H., Kousaka, M., Kobayashi, T., Tsuji, Y., Terashima, M., Fukuda, K., Fukuda, N.: Proposed supplements and amendments to 'a manual of stan- dardized terminology, techniques and scoring system for sleep stages of human subjects', the rechtschaffen kales (1968) standard, Vol. 55, pp. 305–310 (2001). https://doi.org/10.1046/j.1440-1819.2001.00810.x
18. Zhang, Z., Guan, C., Eu, T., Yu, J., Ng, A., Zhang, H., Kwoh, C.-K.: Automatic sleep onset detection using single eeg sensor **2014**, (2014). https://doi.org/10.1109/EMBC.2014.6944071
19. da Silveira, T., Kozakevicius, A., Rodrigues, C.: Awake/sleep scoring through wavelet analysis associated to decision tree algorithms (2015)
20. Wang, C., Zou, J., Zhang, J., Wang, M., Wang, R.: Feature extraction and recognition of epileptiform activity in eeg by combining pca with apen **4**, 233–40 (2010). https://doi.org/10.1007/s11571-010-9120-2

Transition of Ex-Servicemen from Military to Civilian Society: Challenges and Prospects in Reintegration

Molak R. Sharma, Dhiraj Jain, and Vijayeta Sharma

Abstract *Background*: The armed forces need young soldiers to sustain hardships under challenging terrains and win wars. Therefore, about 55,000–60,000 soldiers retire every year who are called Ex-Servicemen (ESM), they face challenges in getting employment and reintegration with civilian society. *Objectives*: Study the reintegration problems and examine the factors acting as enhancers and barriers in the reintegration process. *Materials and Methods*: We did this study in India and included the retired soldiers of Indian Forces. It was conducted online and had fifty questions. *Findings*: Majority of ESM (77%) retire young and need a second career; employment emerged as the most critical factor in the reintegration process; higher education, job-reservation and networking helped get work. ESM were not trained for civilian jobs, and military skills were not helpful. ESM perceived resettlement agencies powerless. Attitudes, social skills, personal and cultural attributes held by an ESM play a crucial part in the reintegration process.

Keywords Armed forces · Ex-servicemen · Employment · Resettlement · Reintegration

1 Introduction

The Indian Armed Forces comprising the Indian Army, Navy and Air Force have about one and a half million personnel. With this strength, India has the world's fourth-largest military force [1] and the largest volunteer army [2]. The system that

M. R. Sharma (✉)
Symbiosis International University, Pune, India

D. Jain
Symbiosis Centre of Management Studies, Pune, India
e-mail: dhiraj.jain@scmspune.ac.in

V. Sharma
East London NHS Foundation Trust, London, UK
e-mail: vijayeta.sharma2@nhs.net

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_52

governs retirement in armed forces is designed around the principle that the forces are kept young and competitive. Therefore, about 55,000–60,000 personnel of the Army, Navy and Air Force retire every year [3] and are called ex-servicemen (ESM). It is they who make exceptional sacrifices and pay substantial personal costs to keep the forces young [4]. This study was set in India to investigate the challenges faced by the ex-servicemen in reintegration with civil society. The reintegration process involves-

- Organising psychological and social traits to form a harmonious and whole personality.
- Bringing of people of different social groups into unrestricted and equal association in society.
- Act of intermixing people who were previously segregated and restoration of different thought processes of a person to a harmonious unified state.
- Development of emotional attachment to the newfound environment, society and culture.

Employment is a crucial factor for an ESM, for commencing a successful reintegration, and it becomes necessary, not only to meet the rising economic demands in the wake of sharply slumped pensionary income but also to get liberated from the situation arisen due to unemployment [5]. Majority of the ex-servicemen had a smooth transition to civilian life, and even those who faced problems in the beginning reported constructive response from society in due course [6]. However, a significant minority (34%) faced considerable difficulties in reintegrating due to reasons attributable to military service, which forced them to make many compromises and adjustments resulting in anxiety and stress. The government has initiated various programs that include training courses, reservations in central & state government offices, public sector undertakings and self-employment schemes. However, previous studies indicated that measures taken for the resettlement of ESM were inadequate and ineffective [5, 7, 8].

1.1 Reintegration Challenges

A section of ESM felt alienated and unable to socialise or communicate with civilians, including their friends and immediate family members [9, 10] for the reasons attributable to military service and the institutions, which should provide support to them, left them feeling helpless, unappreciated and disappointed. Ex-servicemen may also face difficulties in reconnecting with family and relating to civilian people who do not know or understand what military personnel have experienced [11]. ESM, on their return to civilian life, experience a disastrous sense of loss, and the most profound loss is of their 'self-identity' and distinctiveness [10] which troubles them to no end. Retirement affects most soldiers adversely in the form of loss of income, professional status, associates, peers, autonomy and self-expression [7], and they face

the problem of spending time and expending energy, which results in various societal and psychological issues. Besides, majority of ESM retire and become jobless at such a time when they have major responsibilities upon them [12]. Military trains soldiers to be aggressive, fierce and combative; then, on retirement, suddenly, they are expected to forget all of that and integrate smoothly into the civilian society, which has an entirely different culture [13]; that is not feasible. By the virtues of military culture and training, armed forces' personnel are proud and independent people with high self-esteem; so, they find it difficult to seek help continuously and end up living a stressful life [14]. The service in the armed forces is stressful, demanding, hazardous and unpredictable; but being reintegrated and settled in civilian life was an equally daunting task for most of the ex-servicemen [15]. Military transforms a civilian into a soldier and gives him a unique identity which is complex and highly respected in the armed forces [16], and the most significant and painful loss, experienced by all ex-servicemen, was the loss of their 'self-identity' [10]. ESM who had less profound military identity had a smoother transition, whereas the others with deep-rooted military identity had a difficult change [6]. On their part, the society should acknowledge the sacrifices of soldiers who even sacrifice their lives in the line of duty [17], so in recognition of these sacrifices, the state and society have a responsibility to ensure that ESM do not face hardships after retirement. On their part, the ESM should focus on fostering closer association with military peers [9], which has been found helpful in previous studies.

1.2 Reintegration Process

Reintegration is the process of soldiers transitioning back from the military to civilian society in personal and organisational roles [18]. It includes economic, political and social assimilation into civilian life, primarily through educational and vocational training and employment creation schemes [19]. A study found that the reintegration was not 'one mega-event' but involved forming multiple associations and assuming many different roles, which could be done better with social support [18]. A previous study described four themes, i.e., personal reintegration, family reintegration, work reintegration and cultural reintegration for a successful reintegration [20]. Resettlement requirements of military retirees involve both social and financial aspects, and it is essential for an ESM to have a second career in civilian life [21], although social elements of integration are essential but rarely considered. ESM, who worked in civilian settings while still in service, had a smoother transition to civilian life [6], implying that positive contact between soldiers and civilian groups promoted goodwill and reduced misunderstandings and negativity.

Reintegration process chart conceptualised by the authors is shown in Fig. 1. We named the factors assisting the reintegration process as the 'enablers'. Between the period of retirement and reintegration, the ESM has to find a suitable job, a school for children, house to live and money to support the family. These factors

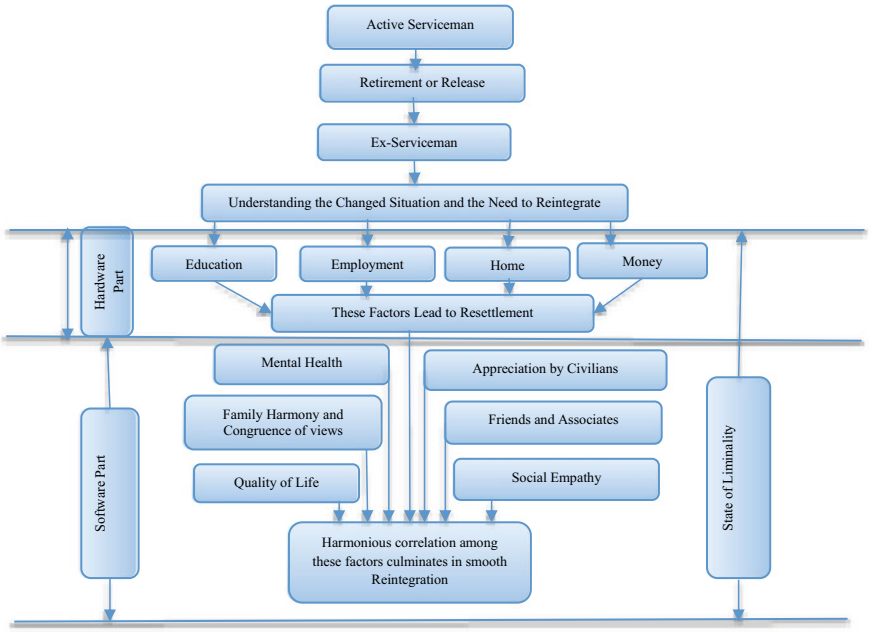


Fig. 1 Reintegration process chart

assist in reaching a significant milestone in the journey towards reintegration, called 'resettlement'.

The elements, which help ESM to get the 'resettlement stage', satisfy their physiological and safety needs [taken from Abraham Maslow's theory of 'Hierarchy of Needs'], and we have termed them as the 'hardware part' of the reintegration process. Final destination in transition is 'reintegration', which mainly pertains to the 'affective domain' or psychological well-being, and includes feelings, emotions, appreciations, harmony in family and friends. These enablers satisfy social and psychological needs and boost the self-esteem of ESM. We termed them as the 'software part' of the reintegration process. For achieving a satisfactory level of reintegration, the ESM should have both the 'hardware' and 'software' parts in equal measures coexisting in harmony. Between the period of retirement and reintegration, an ESM lives in a challenging state of liminality, confusion and uncertainty. The final survey item in this study was to rate the 'Quality of Life' (QOL) an ESM lives after retirement; which, is directly affected by the 'hardware' and 'software' factors. Therefore, the values derived from the responses to this particular survey item (i.e., quality of life) can be taken as a fair indicator of the reintegration status of an ESM. In other words, we considered that an ex-serviceman was reasonably reintegrated if he is having a fair 'Quality of Life'.

1.3 Research Gap

Review of the literature revealed a scarcity of academic research on the reintegration and resettlement problems of Indian ex-servicemen. Singh [7] conducted a study named 'Resettlement Problems of Retired Army Officers' which covered only retired 'army officers' and was limited to Pune district of Maharashtra state. Kishore [5] conducted a study which covered the resettlement problems of retired personnel below officer rank (PBOR) of the Indian Army, born in the Indian states of Haryana and Bihar only. Phadtare [12] conducted a study on resettlement issues of ex-servicemen, confined to district Satara of Maharashtra state. Recent works by Maharajan and Subramani [8] and Maharajan and Krishnaveni [21] covered the resettlement issues of ex-Indian Air Force personnel belonging to two districts (Nilgiris and Coimbatore) of Tamil Nadu state.

Additionally, retired officers of the armed forces also published papers on the military-related topics in military journals such as United Service Institution of India (USI), the Institute for Defence Studies and Analysis (IDSA), the Indian Defence Review (IDR) and the Centre for Land Warfare Studies (CLAWS). We have reviewed the relevant articles and referenced them accordingly. We found that few of these studies were undertaken long back and were confined to particular districts or states and a few select categories of ex-servicemen, and the relevance of their findings may not be valid now. Clearly, there are too many gaps related to comprehensiveness, the period, coverage of the geographical area and limited categories of the respondents included in previous studies.

1.4 Need for the Study

The earlier studies focussed on the 'resettlement' aspect of transition, which may be one of the crucial factors of 'reintegration' process. Researchers of this study could not find literature on the 'reintegration aspects' of ex-servicemen, and from the review of previous studies, it appears that the authors assumed that 'getting resettled' also included 'reintegration', which may not be the correct assumption. This study may be the first work undertaken to study the 'reintegration challenges' faced by the Indian ESM. Further, the strength of ESM and their family members in India is approximately 10 million, and we cannot ignore the existential issues of such a large and important section of the population. The study intended to cover the previous research gaps extensively and discover new information to meet its objectives.

1.5 Objectives of the Study

1. Investigate the reintegration challenges faced by the Indian ESM.

2. Identify and examine the factors helpful in the reintegration process.
3. Identify and analyse the factors hindering the reintegration process.
4. Study and explore the 'Quality of Life' (QOL) of the ex-servicemen.

2 Materials and Methods

This research was confined to the ESM belonging to the Indian Armed Forces as they form the bulk of the ESM population in India. The commissioned officers were excluded because they are comparatively better qualified, well placed and enjoy a higher social status. Presently, there is no universal scale available to measure the degree of 'reintegration', but some research studies have tried to identify and assign measurable parameters [20, 22]. We find that these factors are similar to 'hardware and software' parts described in this study (see Sect. 1.2). Additional parameters included in the current study were (i) employment status, (ii) behaviour of civilian people, (iii) making friends with civilians, feeling comfortable and the level of participation in community functions, (iv) instances of alcohol abuse and drug-related problems, (v) effect on power and prestige after retirement, (vi) occurrence of mental health issues and family problems and (vii) type of QOL after retirement.

2.1 Sampling Design

There are about 27 Lakh ex-servicemen in India [23]. Majority of ESM retire in their mid-thirties and early forties [24] and would have approximately 20 years of productive life post-retirement. So, we have a sampling frame of 12.0 lakh (60,000 $\times$ 20), and the sample size was calculated 385 for a confidence level of 95% ($\alpha = 0.05$, and critical value $Z = 1.96$) with a 5% margin of error. Sample proportion was taken as 50% for a conservative estimate of sample size.

2.2 Procedure

Based on extensive experience in the military service, the authors were aware of the reticent nature of ex-soldiers and their reluctance to participate in any direct 'questioning-answering activity'. We, therefore, decided to conduct an online survey, which increased the likelihood of getting accurate responses as compared to the 'face to face' interviews. To begin with, the researchers' resolve to conduct this study was conveyed on active and popular ex-servicemen groups on the web with an appeal to ex-servicemen to volunteer and participate in this research. In response, over 500 ESM responded and volunteered to participate. The research questionnaire was forwarded to their email addresses; also, a Web link was provided to all respondents.

An online survey software was used to design and share the questionnaire and also to record and tabulate the data. Statistical Package for the Social Sciences (SPSS) was used for data analysis.

3 Results and Analysis

We received inputs from 411 ESM of all the three services from 27 states, making it a pan India survey. The geographical spread of the respondents is shown in Fig. 2 (Table 1).

A vast majority of recruits hailed from villages and towns, and 34% of them could not get jobs. Upward mobility in the educational qualification of the ESM was evident; 20% of the respondents held Army Graduation Certificate, but not all civilian employers recognise this qualification. We found that 34% of the respondents considered higher education as the most critical factor for getting jobs, followed by the 'merit and networking' by 25.6% and reservation by 21%. Further, 42% of the respondents got jobs with self-efforts and only 24.5% with the government help. Majority of the respondents considered the government resettlement agencies as powerless and of little use to ESM.

3.1 Procedure adopted for Determining the Status of Reintegration

We considered the 'reintegration process' comprising hardware and software parts (see Fig. 1), which are further divided into more factors and sub-factors. To determine the correlation, we analysed them by 'Crosstabs' process on SPSS against common factor 'quality of life' (QOL) which has been taken as a fair indicator of the reintegration status.

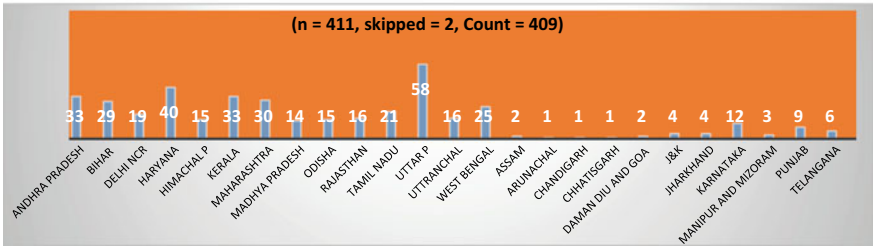


Fig. 2 Geographical spread of respondents in the country

Table 1 Demographics of the respondents

($n = 411$)				
Survey query	Count	%	Total	
			Count	%
From villages or small towns	309	75.6	409 (Skipped = 2)	100
From big cities	95	23.2		
Others	5	1.2		
From the Indian Army	42	10.3	410 (Skipped = 1)	100
From Indian Navy	256	62.4		
From Indian Air Force	112	27.3		
House owning status			410 (Skipped = 1)	100
Did not have their own house	225	54.9		
Had own house	159	38.8		
NA/others	26	6.3		
Children education status			410 (Skipped = 1)	100
Children were studying	310	75.6		
Education completed	23	5.6		
NA/others	77	18.8		
Employment status			410 (Skipped = 1)	100
Did not get employment	140	34.1		
Got employment (0–2 years)	213	51.9		
Got after three or more years	40	9.7		
NA/others	18	4.3		

(n = 411, skipped = 1, Count = 410) for details in following table

Age on recruitment	Count	(%)	Age on retirement	Count	(%)
18–21 years	374	91.2	30–35 years	192	46.8
22–24 years	13	3.2	36–40 years	125	30.5
Above 25 years	23	5.6	41–45 years	49	12.0
			46–50 years	25	6.1
			51 and more	19	4.6
Total	410	100	Total	410	100

Education (recruitment)	Count	(%)	On retirement	Count	(%)
Under matric	5	1.2	Under matric	1	0.2
Matric (Class 10)	122	29.8	Matriculation	30	7.3
HSC (Class 10 + 2)	225	54.9	HSC (10 + 2)	46	11.2
Diploma	17	4.1	Diploma	43	10.5
University graduation	36	8.8	University graduation	106	25.9
Army graduation	00	00	Army graduation	82	20
Postgraduation	5	1.2	Postgraduation	102	24.9

(continued)

Table 1 (continued)

Education (recruitment)	Count	(%)	On retirement	Count	(%)		
Total	410	100	Total	410	100		

3.2 Hardware Part

Factors comprising 'hardware' assist ESM to satisfy their physiological and safety needs and bring to a milestone called 'resettlement'. Pre-requisite to begin a smooth reintegration process is the 'willingness and conscious efforts by the ESM to realise that he is no longer a serving soldier and for all practical purposes he is like any other citizen. But it is difficult to do because of the military identity the soldiers acquire through training and indoctrination of military values and ethos. In our study, we found that 76% of the respondents were trying to reintegrate, and 74% felt that eventually, all ESM would reintegrate.

Employment—We found that 61% of the ESM could get jobs and 80% of them were satisfied. Further, 76% described their QOL as good and 14% as not good. We, therefore, conclude that employment was a crucial factor and a strong enabler in the reintegration process.

House Owning Status—Data revealed that 70% of the ESM who had their own house rated their QOL as good and 11% as not good; in contrast, 60% of those who did not have a house rated their QOL as good and 20% as bad. It showed that the ESM who had their own home had a comparatively better QOL. So, this factor may be classified as an enabler.

Children Education Status—Data showed that over 76% of the ESM had school or college going children at the time of retirement and only 6% had their children education completed. Due to a vast difference between the numbers in two groups, comparison of them would not have yielded fair results. We, therefore, did not do further analysis.

Lack of Funds and Resultant Family Problems—Data revealed that 72% of the respondents felt that the lump sum received at the time of retirement was inadequate to meet their major expenses and the meagre monthly pension insufficient to support the family. To another survey query, 54% of the respondents felt that 'job and pension' together were adequate to support the family, and 83% of them rated their QOL as good. Further, 44% of respondents confirmed that the shortage of funds resulted in family problems; ESM who did not face money problems reported a far better QOL (84%), whereas it was much low (48%) for the ESM who faced a shortage of money. It reiterated the earlier point that the availability of money is crucial to have good QOL and avoid family problems. It supports our previous finding that employment was essential for a satisfactory reintegration.

3.3 Software Part

Factors comprising 'software part' take us from the 'resettlement' point to reintegration stage. We termed them as the 'software part' as they mainly pertain to the 'affective domain' or psychological well-being of a person which includes feelings, emotions, appreciations, harmonious relations, essential for merging back into civil society. These enablers satisfy social and psychological needs and boost the self-esteem of ESM.

Alcohol and drug-related problems—A big majority (77%) of the ESM denied having alcohol or drug problems, and 70% of them felt that they had a good QOL and 15% did not. A comparatively small number (10%) of the ESM confirmed that they developed alcohol or drug-related problems; out of this, 33% reported a good and 32% an impaired QOL. It is evident that the ESM who did not have alcohol or drug-related issues had a superior QOL and better chances of reintegration.

Behavioural Problems (Irritation, Anger, Confusion)—We found that 26% of the ESM confirmed that they became irritated, angry and confused after retirement whereas 59% denied having any such issues. Interestingly, only 48% of the respondents from the first group reported having a good QOL, whereas a massive 84% from the second group reported a good QOL. In response to a related query, one-third of the respondents found it difficult to mix up with civilian people; 50% of them rated their QOL as good and 25% as bad. On the other hand, 47% of the respondents faced no such problem, and 77% of them rated their QOL as good and only 10% as bad. It reflects on how the personal attributes of an individual could influence the QOL and reintegration process.

Behaviour of Family Members and Relatives—Data showed that 40% of the respondents received an acceptable behaviour from their family members and relatives and 71% of these rated their QOL as good and 13% as bad. Further, 25% of the respondents felt that the family behaviour was terrible, and 56% of them rated their QOL as good and 26% as bad. It is evident that the ESM who were welcomed home by the family members and relatives had a better quality of life than the others. This is an essential factor in the reintegration process.

Having Civilians Friends and Attending Community Functions—Data divulged that 62% of the ESM had made civilian friends and 73% of them rated their QOL as good; on the other hand, 22% ESM could not make civilian friends, and 47% of them rated their QOL as good and 26% as bad. To another query, 69% of the respondents said that they regularly attended community functions, and 73% of them rated their QOL as good and 11% as bad; on the other hand, 16% did not attend such events, and only 43% of them rated their QOL as good and 38% as bad. It proved that ESM who had civilian friends and participated in communal events enjoyed a better QOL and better chances of a smooth reintegration.

Feeling Uncomfortable, Awkward or Edgy in Civilian Society—29% of the respondents felt uncomfortable or awkward in the company of civilian people, 51%

of them reported to have a good QOL, and 26% a lousy one. Besides, 48% of the respondents felt comfortable in the civilian company, and 77% of them rated their QOL as good, and only 9% of them rated their QOL as bad. To another query, 54% of the respondents felt that the behaviour of civilians towards ESM was good, 76% of them reported a good QOL, and 15% of them reported a bad one. It indicated that a much higher number of ESM who felt comfortable with civilians and perceived their behaviour as right, reintegrated smoothly.

Effect on Power and Prestige—Having power and prestige in a society is highly valued as it symbolises the social status. In our study, 47% of the respondents felt that their power and prestige increased after retirement, and 66% of them rated their QOL as good and 17% bad; besides 29% thought that they lost power and prestige, and 43% of them reported having a good QOL and 40% as bad. Further analysis revealed that 57% of the job holding ESM felt that their power and prestige have increased; in contrast, only 32% of the unemployed respondents felt the same, whereas a larger 44% felt that they lost it. It showed that—(i) majority of the ESM who held jobs felt that their power and prestige have increased and vice versa, (ii) majority of the ESM, who thought that their power and prestige have increased, rated their QOL as good and vice versa.

4 Conclusion

Majority of ESM reintegrated and merged in civil society smoothly, but there is a significant minority who face problems and need help. Employment emerged as the most crucial factor in the reintegration process; it brings in much needed regular income required to support the family and keeps a person engaged constructively. Retirees lack civilian job skills, and their military skills are not recognised. Higher educational qualification was marked as the most crucial factor for getting jobs, and the ESM falling in the lower age bracket, holding higher educational qualification and good network skills had better chances of gaining employment. Further, the majority of ESM having suitable job felt that their social status and QOL were better than that of the others.

In contrast, the ESM, who could not find jobs, faced a shortage of funds, challenging domestic and personal issues and ended up having a problematic transition. The results support the findings of [9, 22]. The government resettlement agencies are perceived ineffective and of little use; the majority of ESM got jobs with their own efforts. The results are similar to the findings of research studies [3, 7].

State of the mental health, attitudes, social skills, personal and cultural attributes held by an ESM affects the reintegration process. ESM, who worked with civilians colleagues, had civilian friends, felt comfortable in civilian society, enjoyed a better QOL and a smooth reintegration. ESM with family discord, incompatibility and deep-rooted military identity faced a problematic reintegration.

References

1. Global Fire Power: India Military Strength (2018) (Online). Available: https://www.globalfirepower.com/country-military-strength-detail.asp?country_id=india
2. Rana, S.O.: Top 10 Armies in the World (2015) (Online). Available: https://worldnewsreport.in/top-10-armies-in-the-world/. Accessed 24 Oct 2018
3. Taneja, V.: Empirical Research on Second Career Options for Ex-service Personnel in India Incorporated. Centre for Land Warfare Studies, vol. 63, pp. Section 9, 18, 19 (2016)
4. Sethi, S.: Initiatives for exploiting the potential of ex-servicemen. Centre for Land Warfare Studies (CLAWS), 51 p. Section 2, Apr 2015
5. Kishore, S.: Resettlement of Ex-Servicemen in India—Problems Patterns and Prospects, p. 78. Concept Publishing Company, Delhi (1991)
6. Binks, E.: The Transition Experiences of British Military Veterans. Polit. Psychol. **39**(1), 135 (2018)
7. Singh, M.K.: Resettlement Problems of Retired Army Officers. Mittal Publications, Delhi (1985)
8. Maharajan, K., Subramani, B.: A critical study on the resettlement problems of air force ex-servicemen in India: evolving management strategies. Int. J. Res. Manage. Sci. **2**(1), 13–23 (2014)
9. Ahern, J., Worthen, M., Masters, J., Lippman, S., Ozer, E., Moos, R.: The challenges of Afghanistan and Iraq Veterans' transition from military to civilian life and approaches to reconnection. PLOS One 1–13 (2015)
10. Brunger, H., Serrato, J., Ogden, J.: No man's land: the transition to civilian life. J. Aggress. Confl. Peace Res. 86–100 (2013)
11. US Department of Veteran Affairs. Common Challenges During Readjustment to Civilian Life, 2017. (Online). Available: https://www.usveteransmagazine.com/2017/02/common-challenges-during-readjustment-to-civilian-life/. Accessed 02 May 2020
12. Phadtare, R.G.: A study of resettlement of ex-servicemen in Satara district. Kolhapur India (2002)
13. Pemberton, M.: Combat Stress: Ex-Servicemen and Mental Illness. The Telegraph, London (2010)
14. Higate, P.R.: Theorizing Continuity: From Military to Civilian Life, Armed Forces and Society, pp. 443–460 (2001)
15. Morin, R.: The Difficult Transition from Military to Civilian life. Pew Research Centre, pp. 1–6 (2011)
16. Glyndwr University: Leaving the Armed Forces, Double Click Design & Print, Shotton (n.d.)
17. Ashcroft, L.: The Veterans' Transition Review. MoD, London (2014)
18. Currie, S.L., Day, A., Kelloway, K.E.: Bringing the troops back home: modeling the postdeployment reintegration experience. J. Occup. Health Psychol. **16**(1), 38–47 (2011)
19. Chikumbu, A.: Reintegration of Ex-Combatants into Civilian Life, 1st edn. SIA Omni Scriptum Publishing, Riga (2018)
20. Blais, A.-R., Thompson, M.M., McCreary, D.R.: The development and validation of the army post-deployment reintegration scale. Milit. Psychol. 365–386 (2009)
21. Maharajan, K., Krishnaveni, R.: Managing the migration from military to civil society': motivation model for socioeconomic needs in resettlement of veterans in India. Armed Forces Soc. **42**(3), 605–625 (2016)
22. Black, T.G., Papile, C.: Making it on Civvy street: an online survey of Canadian veterans in transition. Can. J. Couns. Psychother. **44**(4), 383–401 (2010)
23. Govt. of India, MoD.: Press Information Bureau, 4-12-2019. (Online). Available: https://pib.gov.in/newsite/PrintRelease.aspx?relid=195438
24. Oberoi, V.: Veterans and the Society. The United Service Institution of India, pp. 1–7 (2010)

Effectual Evaluation on Diabetic Retinopathy

Arun T. Nair, K. Muthuvel, and K. S. Haritha

Abstract DR or diabetic eye disease is an ailment which causes visual impairment in individuals with diabetes. It is discovered to be a procedure reason for preventable visual impairment. The absence of conduction of retinal screening assessment on all diabetic patients has let to numerous undiscovered and subsequently untreated instances of DR. Ideal and precise analyses can decrease the rate vision misfortune if patients with DR are alluded to an ophthalmologist for assessment and treatment. This investigation plans to achieve a strong symptomatic innovation so as to mechanize DR screening. For the mechanized DR identification, an information-driven profound learning calculation was created and assessed as a novel symptomatic instrument. Shading fundus pictures were prepared by this algorithm and ordered them as having DR or sound, recognizing medicinally pertinent cases for referral. Additional clinical survey, all the took in data from the mechanized strategy was promptly imagined through naturally created irregularity heat map, which featured sub-locales inside each info fundus picture. This investigation empowers to recognize cases that ought to be alluded to an ophthalmologist for additional assessment and treatment, with utilizing a completely information-driven man-made consciousness-based reviewing calculation which can screen fundus photographs from diabetic patients. On a worldwide premise, the ramifications of such calculation can definitely help to diminish the pace of vision misfortune brought about by DR.

Keywords Diabetic retinopathy · Proliferative diabetic retinopathy · Pathology · Vitrectomy

A. T. Nair (✉) · K. Muthuvel
Research Scholar - Assistant Professor, Dept. of EEE, Noorul Islam Centre for Higher Education, Kumaracoil, Tamil Nadu, India

K. S. Haritha
Assistant Professor, Govt. Engineering college, Kannur, Kozhikkode, Kerala, India

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_53

1 Introduction

It is assessed that 415 million individuals are influenced worldwide with diabetes that is one in each eleven grown-ups. Diabetic retinopathy (DR) which influences the little veins in the eye is one of the main sources of reparable visual deficiency. The diabetic patients are probably going to have DR, yet the greater part of them may not be even mindful of the malady. Accordingly, it is important that DR ought to be identified early and treated to forestall the loss of vision brought about by this ailment. In any case, the counteraction of DR is not a simple assignment. Typically, ophthalmologist can discover the presence of DR by visual assessment of the fundus and by assessment of shading photographs. In any case, this is a tedious and broad technique considering the quantity of patients influenced by this sicknesses internationally. Exact information on the finding and early treatment of DR are additionally not accessible. Additionally, about 75% of the of the DR patients live at places where there are no offices for location of DR. Disregarding worldwide projects to recognize and treat retinopathy, the infection exists in such a huge scope that compelling individual treatment for everything is preposterous with the outcome that millions actually endure visual perception issue. To counter the issue of lack of offices and prepared people, programmed arrangements dependent on screened shading fundus pictures were proposed. Indeed, even undeveloped professionals can deal with the patients all things considered. Anyway such programmed discovery has a few disadvantages as the calculations are determined dependent on two or three hundred pictures which will be unable to cover enormous scope and various sorts of fundus informational indexes gathered under fluctuating conditions, for example, extraordinary fundus camera, distinctive eye widening technique, and so on. Other than these calculations rely on the DR qualities got through manual hand-tuned highlights to portray anatomic structures in fundus, for example, optic plate, vein, and so on. While these hand-tuned highlight can deal with particular fundus information, this cannot be applied effectively to the fundus pictures got from various socioeconomics. Despite the fact that highlights, for example, speeded-up robust features (SURF) and histogram of oriented gradient (HOG) have been found as the vague technique cannot portray unpretentious contrasts in retinopathy seriousness.

One important factor in the development of diabetic retinopathy is vascular endothelial growth factor signaling. Recently, it is found that the growth factor signaling in endothelial cells is suppressed by leukocyte cell-derived chemotaxin2. The purpose of this study is to find out serum LECT2 levels and the presence of DR. The study covers 230 people with type 2 diabetes mellitus 95 with DR and 135 without DR. It is found by multiple regression analysis that serum LECT2 level is not associated with the presence of DR [1]. In the development of DR, many overlapping and connected molecular pathways are involved. The status of microglia in the process of DR is given in the study which takes into consideration the biochemical mechanisms affecting the connection among neuroretina, vessels, and glial cells [2].

Examine whether DR is a valid predictor of coronary artery disease. The study covers 60 male patients randomly selected from diabetic and ophthalmologic clinics.

Fundus examination for evaluating DR and duplex ultrasound to find out carotid intimal medial thickness was done on them. In all cases, coronary angiography was done [3]. An important cause of blindness including retinopathy of prematurity is neo-vascular eye diseases. Adipokines, like adiponectin which are derived from white adipose tissue, modulate metabolic responses. Evidence shows that retinal neovascularization may occur due to lack of adiponectin [4]. Retinal diseases related to inflammation are normally followed by macrophage/microglial cells activation. Human diabetic donor's assessment of microglia polarization in retinas was evaluated. In retinas of diabetic donors, markers of activated microglia were detected [5].

Comparison of the various components of Prameha/Madhumehajanya Timer with DR and its stages are described in this paper. In Madhumehajanya Timer, all the three doshas together with Rak-thadosha and saptadhatu with four dristi petals of eye are affected. In view of the prolonged and uncontrolled hyperglycemia, Avarana and Share also have an important role in the development of DR. In pathology of DR, Agnimandya-related Ama formation has a role which is similar to the oxidative theory of DR described in modern pathology [6]. The acetone or thelial agents are seen to boost visual acuity and decrease the central macular thickness of the patients. It is also found to be an effective remedy for vitrectomy. Triamcinolone is related to risks of excessive intraocular pressure and cataract [7]. In the ocular side, the fundamental remedy of diabetic macular edema, both severe non-PDR and PDR, is still the laser treatment. Anti-VEGF was found to be effective as an adjunct therapy [8]. Though the major method of diagnosis to detect the microvasculature changes is fundus examination, the visual function test is considered to be an effective alternative having the potential to detect DR at early stages. Visual function components can be characterized by the electrophysiological test of the retina [9]. In cases of proliferative DR after successful vitrectomy and uneventful postoperative course, the preoperative prognostic factors contributing to the extremely poor visual outcome should be investigated. The control group was found to have creatinine level significantly lower (1.23 $\pm$ 0.46 mg/dl; p ¼ 0.003) than the study group (4.07 $\pm$ 4.15 mg/dl), as per multiple logistic regression analysis, chronic muscular detachment and broad fibrovascular proliferation were found to have significant associations with poor visual outcome [10].

The albuminuria and estimated glomerular filtration rate are the supposed biomarkers of DR. In Spain, the prevalence of DR, its relationship with eGFR, and other factors in T2DM was investigated. It resulted in 14.9% prevalence, with more prevalence ($p = 0.0087$) in women and older patients ($p < 0.0001$) [11]. MiR-15 in bone marrow cells and retina was found to be minimized by using a variety of tools and various technologies. The inhibition of miR-15a was found to unregulate a pro-inflammatory molecule acid sphingomyelinase (ASM). It was also found to upregulate an angiogenic molecule expression, vascular endothelial growth factor A in the retinal pigment epithelial and endothelial cells [12]. The existence of serum pro-inflammatory cytokines and acute phase reactants was examined in the patients with and without DR. The case study was conducted using a total of 36 patients divided into three groups. Group 1 consisted of 12 patients with diabetes mellitus and

diabetic retinopathy, Group 2 had 12 diabetic patients without diabetic retinopathy, and 12 healthy patients were grouped as the control group [13]. Three intravitreal injections of ranibizumab were given to a patient at monthly intervals. The new vessels were found to be completely regressed after conducting repeat angiography. The improvement was found in the perfusion of chemical areas in the retina. When compared to the normal pan-retinal photocoagulation, the intravitreal anti-vascular endothelial growth factor injections were found to be valuable for reversing the neovascularization [14]. Modeling approach using m-Medoids combined with a Gaussian mixture modeling to shape a hybrid classifier was proposed in order to improve the classification accuracy. The sensitivity, specificity, accuracy, and operating characteristics of the receiver along with databases of fundus images are used to evaluate the proposed system [15]. The analysis was done on the visually impaired registered adults suffering from diabetic retinopathy aged 18–69 years under the National Council for the Blind of Ireland (2004–2013). The analysis revealed the need for increasing the preventive methods for microvascular complications [16]. From February 1, 2006, to February 1, 2009, a study was conducted on 2435 diabetic patients. 17.90% was found to be prevalent with DR out of which mild-moderate proliferative DR was found in 80.73%, non-proliferative DR was found in 12.16%, proliferative DR was found in 2.29% and diabetic maculopathy in 4.82%. Forty-one patients (1.69) had low-quality retinographies. The improvement of the circuit of communication between primary and specialist care enabling early diagnosis and treatment is the notable benefits of tele-ophthalmology [17].

A combination of the Gaussian mixture model, support vector machine, and an extension of multimodal medoid-based modeling approach was used to present a hybrid classifier, in an ensemble to improve classification accuracy. The publicly available retinal image databases were used to evaluate the proposed system and were found to have higher accuracy in comparison with previously established methods [18]. Studies on various factors which are altered due the influence of endostatin's action are being studied even though the exact mechanism of the action of endostatin is not completely known. This includes the influence of various factors such as down-regulation and activation in, on the progression of angiogenesis. Once the clinical trials show positive feedback, endostatin can be exploited as a durable agent in anti-angiogenic therapy [19].

It appears that the use of some nanoparticles assists in the growth of diabetic retinopathy symptoms, for example, retinal neovascularization. These are also debated considering the thorough management of ocular chronic disease [20].

Diabetic patients used to be directed to an ophthalmologist for frequent screening because the laser photocoagulation treatments can prevent severe vision loss. New inhibitors of vascular endothelial growth factor can provide targeted non-surgical treatment for improving vision in diabetic retinopathy [21]. According to the recent clinical trials data, in addition to hyperglycemia, dyslipidemia plays an important role in the development of DR, which is often overlooked. The main aim of this article is to show the important role of dyslipidemia in DR progression and to underline the novel therapeutic solutions which take advantage of the vital roles played by lipid metabolism in DR progression [22]. The role of connective tissue growth factor in the

pathogenesis of DR relating to extra-cellular matrix remodeling and wound healing actions are discussed and also examined whether CTGF can be an effective novel therapeutic solution in the clinical management of initial as well as subsequent stages of DR [23]. Experiments were done to check whether the diabetic neurodegeneration (DRN) leads to DR. If it is so, the method could be used for the early detection of ocular diabetic damage. Also it could be used to surpass the vision loss by treating DR [24].

The non-dilated retinal fundus images are used to automatically examine lesion exudates. The neighborhood-based segmentation technique produces low contrast images which shows the presence of exudates [25]. Inter- and intra-observer variability can be determined by a computer-aided diagnosis system. The severity of the disease can be evaluated by the proposed methods of various retinal feature extraction and automated analysis. The obtained results are compared with the results obtained from the segmentation technique. The accuracy is determined, and it is found to be 97.89% and 94.76%, respectively, for SVM and PNN [26]. The systematic inflammation of patients with type 2 diabetes mellitus is determined and compared. Similarly, patients having T2DM with and without diabetic retinopathy are also compared. The serum levels of inflammatory cytokines CRP, TNF-alpha, and VEGF were found to be high in patients with diabetic retinopathy than those without retinopathy [27]. The retinal oxygen delivery and metabolic rates are determined to measure retinal hypoxia using visible light OCT. These techniques are used to study more about the vascular pathophysiology of diabetic retinopathy. The reliable and compatible factors to study diabetic retinopathy are to be researched [28].

Figure 1 shows the ordered audit on identification of diabetic retinopathy from the year 2007 to 2018.

Figure 2 speaks to the distinctive AI calculations that are used for grouping the retina pictures into typical and influenced zones from the advanced fundus picture. Here, the AI techniques are ordered as supervised and unaided learning models. The arrangement and relapse go under the managed learning, and the bunching falls under the solo learning. Under the order, models are SVM, fuzzy logic, neural networks, and Gaussian blend models.

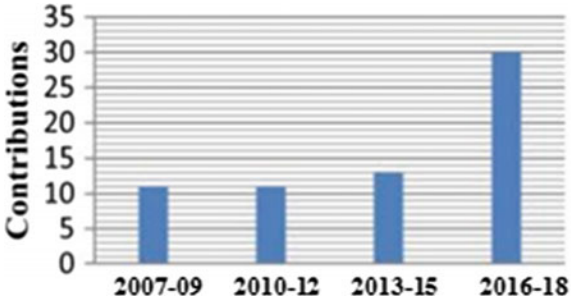


Fig. 1 Bar diagram showing to sequential audit on location of diabetic retinopathy

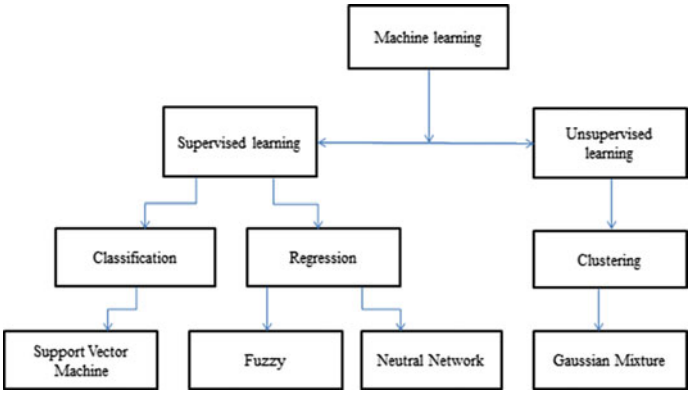


Fig. 2 Scientific classification of AI calculation in diabetic retinopathy discovery

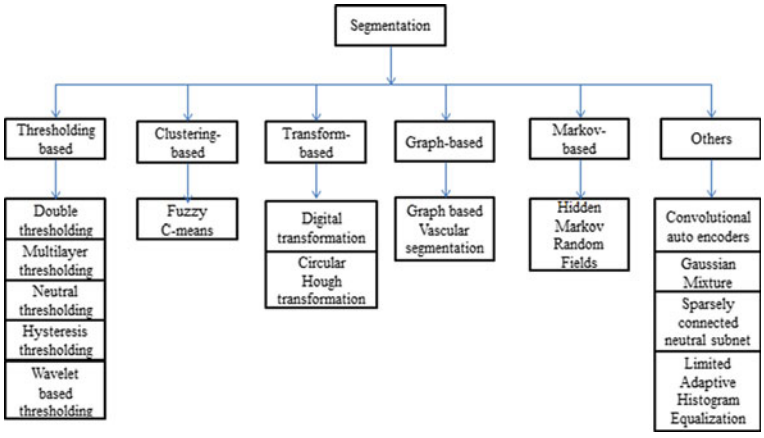


Fig. 3 Segmentation techniques

Figure 3 represents the different segmentation models that are used in the state-of-the-art models.

2 Research Gaps and Challenges

Change from the manual diabetic retinopathy recognition model to the programmed diabetic retinopathy discovery model with the translation of the computer faces an immense number of moves that should be abrogated. The significant test lies with the dependability and exactness in preparing the picture and separation in the typical and anomalous district of the fundus picture. The essential part of the programmed

identification model is to remove the highlights like the optic plate, fovea, and the retinal veins in the fundal picture from even inferior quality shading fundus picture in an exact way. Simply after the discovery of the reference facilitates, the recognizable proof of the obsessive elements, for example, hard exudates, cotton fleece spots, hemorrhages, and MA can be cultivated. Rather than this, the greater part of the examination does not consider on the identification of the reference facilitates, and this makes the recognition model wasteful. Further it is important to have a request for priority independent of inconstancy in shading, brightening levels, and measure of clamor. This will upgrade the amount of the recognition, and quicker the outcome could be acquired at the prior stages. Moreover, the vein by single division approach in vessel division is an unpredictable errand regarding between administrator reputability and reproducibility. This, thusly, decreases the exactness of recognition, and this is not appropriate adversary anatomical locale. So rather than single division approach, the multi-division can be used for valuable and brisk perceptions. Along these lines, rather than single vein division approach, the multi-division can be used for valuable and speedy perceptions. Further, the structure of the veins should be distinguished, with the aim of looking for the conceivable non-vascular injuries that for the most part gets wiped out during the division of retinal vein. The adjustment in the vessels is likewise an explanation for the development of diabetic retinopathy. The greater part of the investigations does not take thought on these alterations, while playing out the division activity. Further to improve the results of the division, an ideal pre-handling and sifting strategies should be determined.

3 Conclusion

Automatic conclusion of diabetic eye disease is possible and attainable through the utilization of all around characterized picture preparing strategies. While numerous triumphs has been recorded in the current advances in mechanization of clinical determination, this examination will in general amplify the huge accessibility of pervasive gadgets and elicitation of past symptomatic strategy that is set toward giving practical, simpler, and quicker finding. Future work will focus on defining and entrancing novel features, set as to improve the accuracy, and also to integrate the module to risk assessment procedure. In addition, more sophisticated pattern recognition methods will be implemented. This study paper had explored a lot of research papers and showed the prevalence just as the advantages of every commitment. At first, the examination had illuminated the different recognition models. Moreover, the investigation additionally centered around different highlights like vein division, received instruments just as information base sets used for determination. The classification of different machine learning algorithms appropriate for the analytic model is done as the most critical examination. The investigation additionally assessed the presentation accomplishments as far as exactness and other significant frameworks too. At last, this paper depicts the diverse existing examination issues which can be helpful for the specialists to do future exploration. Moreover, the future upgrade will

expand on these triumphs by empowering novel variations of meta-heuristic idea in the classifier, with the goal that presentation can be accomplished.

References

1. Okumura, et al.: Reduced serum level of leukocyte cell-derived chemotaxin 2 is associated with the presence of diabetic retinopathy. Clin. Chim. Acta **463**, 145–149 (2016)
2. Sorrentino, F.S., Allkabes, M., Salsini, G., Bonifazzi, C., Perri, P.: The importance of glial cells in the homeostasis of the retinal microenvironment and their pivotal role in the course of diabetic retinopathy. Life Sci. **162**, 54–59 (2016)
3. El Demerdash, F., Refaie, W., Allakany, R., Tantawy, S., Dawood, E.: Diabetic retinopathy: a predictor of coronary artery disease. Egypt. Hear. J. **64**(2), 63–68 (2012)
4. Fu, Z., Gong, Y., Löfqvist, C., Hellström, A., Smith, L.E.H.: Biochimica et Biophysica Acta Review: adiponectin in retinopathy. BBA-Mol. Basis Dis. **1862**(8), 1392–1400 (2016)
5. Arroba, I., et al.: Biochimica et Biophysica Acta Modulation of microglia polarization dynamics during diabetic retinopathy in db/db mice. BBA-Mol. Basis Dis. **1862**(9), 1663–1674 (2016)
6. Sahoo, P.K., Fiaz, S.: Conceptual analysis of diabetic retinopathy in Ayurveda. J. Ayurveda Integr. Med. **8**(2), 122–131 (2017)
7. Abu El-Asrar, M., Al-Mezaine, H.S.: Advances in the treatment of diabetic retinopathy. Saudi J. Ophthalmol. **25**(2), 113–122 (2011)
8. Gupta, N., Gupta, R.: Diabetic retinopathy—an update. J. Int. Med. Sci. Acad. **28**(1), 54–58 (2015)
9. Umashankara, Gunasundari, R.: A review on electrophysiology based detection of diabetic retinopathy. Procedia Comput. Sci. **48**(C), 630–637 (2015)
10. Lin, S.J., Yeh, P.T., Huang, J.Y., Yang, C.M.: Preoperative prognostic factors in vitrectomy for severe proliferative diabetic retinopathy. Taiwan J. Ophthalmol. **4**(4), 174–178 (2014)
11. López, M., Cos, F.X., Álvarez-Guisasola, F., Fuster, E.: Prevalence of diabetic retinopathy and its relationship with glomerular filtration rate and other risk factors in patients with type 2 diabetes mellitus in Spain. DM2 HOPE study. J. Clin. Transl. Endocrinol. **9**, 61–65 (2017)
12. Chakrabarti, S.: MicroRNA15a—a molecule modulating multiple pathologies in diabetic retinopathy. EBioMedicine **11**, 13–14 (2016)
13. Eustolio, S., Mota, H., Soto-bahena, J.J., Viveros-sandoval, M.E., Cardiel-ríos, M.: CIRUGÍA y CIRUJANOS Pro-inflammatory serum cytokines in diabetic. Cirugía y Cir. (English Ed.) **83**(2), 100–106 (2015)
14. Chandra, S., Sheth, J., Anantharaman, G., Gopalakrishnan, M.: Ranibizumab-induced retinal reperfusion and regression of neovascularization in diabetic retinopathy: An angiographic illustration. Am. J. Ophthalmol. Case Rep. **9**(December 2017), 41–44 (2018)
15. Usman Akram, M., Khalid, S., Tariq, A., Khan, S.A., Azam, F.: Detection and classification of retinal lesions for grading of diabetic retinopathy. Comput. Biol. Med. **45**(1), 161–171 (2014)
16. Tracey, M.L., McHugh, S.M., Fitzgerald, A.P., Buckley, C.M., Canavan, R.J., Kearney, P.M.: Trends in blindness due to diabetic retinopathy among adults aged 18–69 years over a decade in Ireland. Diabetes Res. Clin. Pract. **121**, 1–8 (2016)
17. Martínez Rubio, M., Moya Moya, M., Bellot Bernabé, A., Belmonte Martínez, J.: Diabetic retinopathy screening and teleophthalmology. Arch. la Soc. Española Oftalmol. (English Ed.) **87**(12), 392–395 (2012)
18. Akram, M.U., Khalid, S., Khan, S.A.: Identification and classification of microaneurysms for early detection of diabetic retinopathy. Pattern Recognit. **46**(1), 107–116 (2013)
19. Behl, T., Kotwani, A.: Possible role of endostatin in the antiangiogenic therapy of diabetic retinopathy. Life Sci. **135**, 131–137 (2015)
20. Fangueiro, J.F., Silva, A.M., Garcia, M.L., Souto, E.B.: Current nanotechnology approaches for the treatment and management of diabetic retinopathy. Eur. J. Pharm. Biopharm. (2014)

21. Fante, R.J., Durairaj, V.D., Oliver, S.C.N.: Diabetic retinopathy: an update on treatment. AJM **123**(3), 213–216 (2010)
22. Hammer, S.S., Busik, J.V.: The role of dyslipidemia in diabetic retinopathy. Vis. Res. (2017)
23. Klaassen, I., Van Geest, R.J., Kuiper, E.J., Van Noorden, C.J.F., Schlingemann, R.O.: The role of CTGF in diabetic retinopathy. Exp. Eye Res. **133**, 37–48 (2015)
24. Lynch, S.K., Abràmoff, M.D.: Diabetic retinopathy is a neurodegenerative disorder. Vis. Res. (2017)
25. Mahendran, G., Dhanasekaran, R.: Investigation of the severity level of diabetic retinopathy using supervised classifier algorithms q. Comput. Electr. Eng. 1–12 (2015)
26. Rama, M., et al.: Author â€™ s Accepted Manuscript Reference : Keywords (2013)
27. Nalini, M., Raghavulu, B.V., Annapurna, A., Avinash, P., Chandi, V., Swathi, N.: Diabetes & metabolic syndrome: clinical research & reviews correlation of various serum biomarkers with the severity of diabetic retinopathy. Diabetes Metab. Syndr. Clin. Res. Rev. 8–11 (2017)
28. Nesper, P.L., Soetikno, B.T., Zhang, H.F., Fawzi, A.A.: OCT angiography and visible-light OCT in diabetic retinopathy. Vis. Res. (2017)

Image-To-Image Translation Using Deep Convolutional GANs

Nahid Guard and Suneeta V. Budihal

Abstract The paper proposes a framework to build a deep neural network (DNN), established by heaping layers of GANs. The GAN architecture comprises both generator and discriminator layers which are locally trained to generate fake images which are of similar resolution as the original input image. The generated inputs are reconstructed exactly as the original input. The algorithm is a candid variation by stacking the ordinary GAN algorithm. The discriminator network is basically a classification problem of machine learning yielding less classification error. Hence, spanning the performance gap with DNNs and in many of the cases surpassing it results in the reconstruction of the inputs depending upon the training parameters such as the upsurge of the epoch and batch size. It will increase the training period, thus increasing the accuracy of synthesized images. A novel mode of training the GANs is discussed in the proposed model. The key point is to grow generator and discriminator progressively, with a low resolution, adding new layers model increasingly and finding details in the model as training progresses. It speeds up training process and also stabilizes results, allowing the results to be of high quality. Additionally, a non-random noise is added and the noise is fixed for the latent space size which yields stable results. A novel metric for evaluating GANs in terms of image quality and variation is discussed.

Keywords Deep learning (DL) · CELEB image dataset · Unsupervised feature learning · GAN · Deep convolutional neural networks

1 Introduction

Translating input image into a similar output is a major issue in image processing, computer vision and computer graphics. For instance, it can be related to a simple and most common concept of language translations in either French to English. To note as

N. Guard · S. V. Budihal (✉)
School of Electronics and Communication, KLE Technological University, Vidyanagara, Hubballi 580031, India
e-mail: suneeta_vb@kletech.ac.in

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_54

an example, it can also refer to RGB image, edge map or any semantic label map, etc. To be concise about the analogy, wherever there is automatic language translation, there can also be image-to-image translation. As the particular task involves translation of one possible representation into another, given sufficient data to analyze certain parameters such as edge to image detection, computer vision, etc., can be dealt easily.

Traditionally, each of these tasks was tackled separately with different machinery. GAN is a promising candidate to address the issue of scarce dataset for training the neural network model. In order to obtain different applications, different forms of GANs are presented with different levels of synthesis and architecture. Despite the fact, the goal is to develop a framework that is suitable and stable to sustain the variations in the different variety of inputs given to the network and provide optimal outcome. Significant steps are already initiated by the domain experts toward this. Among them is convolutional neural networks (CNNs) that is becoming the most common and basic platform in machine learning (ML) and artificial intelligence (AI) for a broad range of image processing problems.

Although the learning in any CNN is automatic, lots of manual efforts are needed in reducing the losses. Alternatively, to train the CNN model to the required model and expectations, needs lots of dataset. Suppose in a naive approach, if CNN is designed to minimize Euclidean distance between predicted and ground truth pixel, it has a tendency to provide blur results. Averaging all outputs eventually reduces the Euclidean Distance. The loss functions are used in CNN which tend to force the CNN to do what we really want it to do and generally require knowledgeable background. Instead, if it is specified only the high-level goal most of the lower level work will be reduced, which is mostly desirable, like to obtaining the results that are like indistinguishable from reality, and to automatically learn some specific functions such as loss functions which are appropriate to satisfy the goal. Fortunately, this is what is exactly done by the GANs. GANs try to classify output as real or fake, simultaneously training the generator model to reduce the discriminator losses.

Moreover, blur or unrealistic images are not considered since they obviously look fake. As GAN model is capable of learning loss functions and can adapt data, they can be used in various sectors of real-world applications. In this paper, the focus is on generating human indistinguishable fake human images using deep convolutional GAN (DCGAN), which are of high resolution and the model should have good impact in reducing the image generation timings. In most of the generative networks, time consumption is a very big issue faced due to the complex networks and the manipulation of pixels such that every generated image will look different and real. We are trying to solve this problem by coming up with a fully connected network that can be trained in the back-end and enable cloning of the same such that it will reduce the code competency and have better results quickly.

Section 2 discusses about various concepts available for the design of the model which satisfies the specifications as much as possible. It consists of subsections such as the functional block diagram, algorithm and the details about the dataset required. Section 3 discusses the implementation details about synthesis and reconstruction of the human images. There are sub-topics like training strategies and performance

parameters. Section 4 discusses about the obtained results and the detailed discussion about it. This section also discusses the analysis of the model. Section 5 presents the conclusions.

2 Literature Survey

Promising results of some realistic facial images are produced as a result of GANs [1–3]. For example, DA-GAN [4] is an algorithm that has been proposed for generating augmented data automatically. GAN-based models face some issues with the synthesis, i.e., to control the images that are generated due to random distribution used in the input to generator. Some of the modifications in the GAN architecture, such as conditional GAN, have made an attempt in addressing the issue by having some conditions on both the networks, i.e., the generator and the discriminator for some images with conditional synthesis [5–7]. However, the conditional GAN mapping does not encounter any constraint in output to the target; thus, there can be some arbitrary differences between the target and the output.

Face ID-GAN [8] proposed a model that treats as a third player classifier of the face, to the face identity, where a competing generator and discriminator distinguish real and fake out of synthesized images. The GAN-based models for face synthesis have a major shortcoming that may lead to produced images to be inconsistent because of the weak constraints globally. To address this, a SimGAN model is developed which enables the model toward learning of the image synthesis using the inputs instead of the random noise vector [9]. Difficulty in controlling the output of the generator is another issue in vanilla GAN. Recently, there is an addition of conditional information in the conditional GANs in the generator and the discriminator network to have a conditional image synthesis [10, 11]. DR-GAN is proposed that takes a pose code along with the random noise vector as a case input in the discriminator with a goal to synthesize the images of face with same identity to fool discriminator [12].

A couple agent discriminator is introduced, to guide the generator in toward learning; it forms a mask image and provides a controllable and flexible condition during the interface in the CAPG-GAN [13]. The regression of the generator results in some large arbitrary errors, as a consequence of unreliable applications in the real world [6]. The authors in [5], proposed GAN, shows quite appealing results in image transfer [9], image generation [12], super-resolution [12], etc. GAN consists of generator and discriminator model that are trained so that generator tries to synthesis images and the discriminator tries to distinguish real and fake images. They compete with each other, where the generator tries to fool the discriminator by assuming that image generated by the generator is real, and on the other hand, discriminator puts full efforts to identify it as fake. This idea can be used in terms of the discriminator and generator losses, which encourages the result of the synthesized image, close to original dataset. Even though there is a high level of simplicity in training, model-level optimization may lead to collapse, in which the generator produces the same results.

For stable training in GANs, [14] suggested a rendering discriminator omniscient. By the process of learning the loss function which separates generated samples with real samples, LS-GAN [2] focuses on improving the quality of the generated result and thus avoids collapse. Different variations of GAN were proposed for generating images. In [15], the authors applied CNN for GAN to generate images from latent inputs vector. Instead of synthesizing images from latent space vectors, several methodologies used the same idea of adversarial to generate images with much more meaningful input. The following are the contributions of the proposed framework.

- To design a GAN using deep convolutional layers.
- To build a generator for producing new images.
- To build a discriminator network which classifies image in terms of fake or real.
- To train the model such that the new images should not only be machine indistinguishable but also human indistinguishable.

3 Proposed Methodology

It includes a brief description about the block diagram of GAN. To be more specific, it provides the outline for the design and the operation of network. GAN consists of two networks—one is a generator capable of generating new images and other is a discriminator that distinguishes between fake and real images. The block diagram gives a brief description about the generative adversarial network. Initially, a generator network that is capable of synthesizing the images that are as same as that of original input taken from the CelebA dataset. Then, the discriminator network will distinguish between the fake and the real images. In order to obtain this, the losses from generator to discriminator is modeled and analyzed.

3.1 Algorithm for Generative Adversarial Network

GAN consists of several layers of implementation. The layers depend upon the designer, generally as layers of training increase; output will be closer to the original input. Training sample is 190,000 images and testing samples is around 20,000 images, summing up for a total of 200,000 image samples. The training samples are fed to various layers, and as a result, the model is trained. Initially, data is processed and then fed it to the generator G to produce the synthetic image $x\ (y = 0)$ (x, y) input pair. Generally, a fake pair and the real pair $x\ (y = 1)$ are considered and alternatively fed to the discriminator. It calculates the loss for both fake loss and real loss as it is a binary classification neural network and combines them both to be the final loss as D loss. The generator G also does the loss calculation as G loss. Both the losses are fed back to their networks for further learning from the loss. This is where adjusting of parameters happen w. r. t. the loss. The generator G gets stronger in synthesizing

the fake images very close to real, and the discriminator D also gets stronger in distinguishing between the fake and the real images.

3.2 Training Strategies

The output from the GAN is used to form an image, as the network should yield positive output. Hence, ReLU activation function is added at the last output layer, which can also accelerate the training speed. An Adam optimizer is used that can adjust the learning rate on each weight parameter, which guarantees effective training. In order to ensure that the recovered data has less loss, a function $J\,(w, b)$ is used to calculate the differences between them (Fig. 1).

$$J(w, b) = \frac{1}{m}\sum_{i=1}^{m}\left(x' - x\right)^2 \tag{1}$$

Celeb Faces Attributes Dataset (CelebA) is a dataset with face attributes that describe the facial features like hair bangs, contours of the face, position of lips, position of nose, eyes, etc. The dataset can be used as the training dataset and testing dataset for the computer vision tasks such as recognition of face attribute, detection of face, facial part localization and face editing synthesis. Identifying the faces and various applications is a popular element of computer vision and deep learning. This dataset is very effective in training and testing the images of human and to synthesize them, particularly to recognizing attributes such as black hair, or people wearing glasses even people who are smiling can be identified.

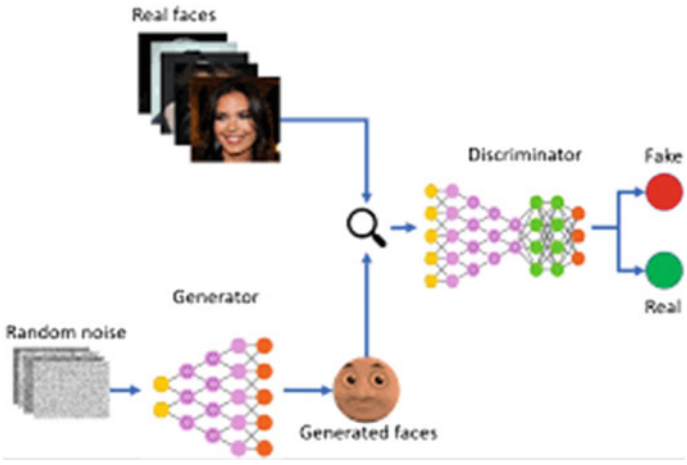


Fig. 1 A functional block diagram for GAN comprising of generator and discriminator

Deep Neural Networks are used as the encoding function and the decoding function. Cross-entropy function is being chosen for the distance function between the input and the output. To be specific, CNN with ReLU activation, leaky ReLU and sigmoid function in generator and discriminator network are used. As DCGAN network can achieve image compression from its hidden neurons' output, because the hidden layer can learn a compressed feature representation for the input data. Decompression is achieved by the output layer because the output value size is the same as that of the input. ReLU function (rectified linear units) is the most common activation function used in neural networks. It returns 0 for negative values, and for any positive value x, it returns the value back, so we can write,

$$f(x) = \max(o, x) \tag{2}$$

The equation of the generator of the generative adversarial network which is used in order to generate the new synthetic data can be expressed as: At generator (G):

$$\text{Gloss} = \log(1 - D(G(Z))) - \log(D(G(Z))) \tag{3}$$

$$\frac{1}{m}\sum_{i=1}^{m} \log\left(1 - D\left(G\left(Z^i\right)\right)\right) \tag{4}$$

$$\frac{1}{m}\sum_{i=1}^{m} -\log\left(D\left(G\left(Z^i\right)\right)\right) \tag{5}$$

The equation of the discriminator of the generative adversarial network which is used in order to detect the data which is real and fake can be expressed as: At discriminator (D):

$$D\text{loss}_{\text{real}} = \log D(X) \tag{6}$$

$$D\text{loss}_{\text{fake}} = \log 1 - D(G(X)) \tag{7}$$

It can allow the model to account for nonlinearities very well. Sigmoid function can take any range of real numbers and returns the output in the range from 0 to 1. The sigmoid function produces a curve which will be in the shape S. The sigmoid function returns a real-valued output. The first derivative of the sigmoid function will be non-negative or non-positive. Another problem with ReLU is the dying ReLU problem where some ReLU neurons die for all inputs and turns out to be inactive irrespective of input. Here, no gradient flows and neural network performance is affected, if large number of dead neurons are there. This can be corrected by using leaky ReLU, where slope is changed left of $x = 0$, thus causing a leak and extending the range of ReLU.

A DCGAN is an unsupervised machine learning algorithm that uses multiple neural networks that pits against each other. In DCGAN, the image is split up into

different levels of resolution, and the process of generation or data synthesis happens to be slow as it has to go through all the stages in of the image. It has been split up into a progressive approach of growing the images takes place which assist in getting better results with much better accuracy when compared to the normal GAN. On the other hand, if we use a simple GAN algorithm to generate human images according to survey and previous work done, the network is much simple to handle complex features of the face, as the computations are operated only on one level of resolution in order to fool the discriminator. Since more the layer activation better the feature extraction and visualization, if it is a classification network, the sigmoid function is used to classify in terms of zero or one, i.e., concluding it to be either fake or real.

Normally, a noise factor or random noise is used in the generator part to make the image synthesis but, by using random noise, we cannot predict the output resolution. In order to achieve a much more stabilized output, we have used the size of the latent space as a factor to decide the noise level which gives much better and stable results. Latent space is a simple representation of the images which consists of some simple features. The latent space of the input images will be same for all the input images as we are using the same sized inputs from the dataset, and such conditions have provided better and stable results (Fig. 2).

Figure 3 demonstrates the generator and discriminator losses incurred during the training the network model for epocs. CelebA is a large-scale face attributes dataset with more than 200,000 celebrity face images, each with 40 attribute annotations. The images in this dataset cover large pose variations and background clutter. It has large diversities, large quantities and rich annotations, including 10,177 number of identities, 202,599 number of face images, five landmark locations and 40 binary attributes annotations per image. Figure 4 shows different results for different noise levels. Initially, the analysis is done by keeping the training and testing the data constant and varying the noise level. From beginning it was checked up till 100% of noise level. As the noise level increases the accuracy increases and the losses decrease.

4 Results and Discussions

See Figs. 3 and 4.

5 Conclusion

We can achieve effective data collection of generated human images with the method of DCGAN with high resolution by having several convolutional layers in the network by just feeding original images to DCGAN and making sure that the output images are similar as the original images. This method guarantees that we do not need to do

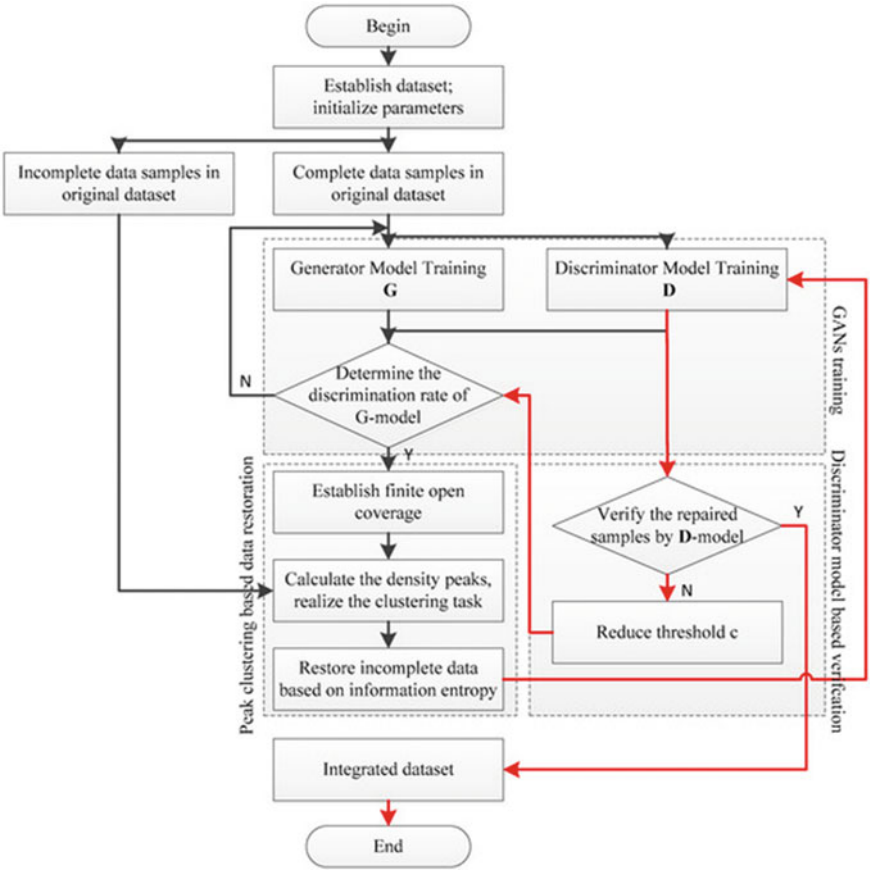


Fig. 2 Action diagram for GAN which gives the step-by-step procedure followed in the proposed model

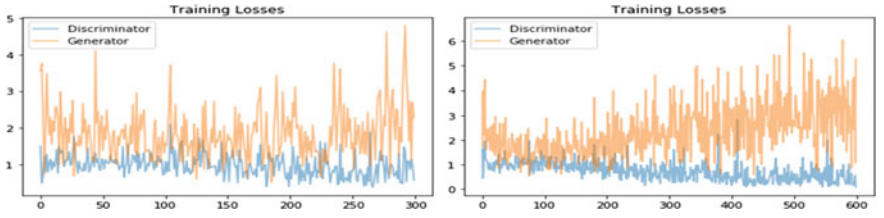


Fig. 3 Generator and discriminator losses as a function of size of epocs

Fig. 4 Rresults for generated fake images for various test images from CELEB dataset

much images preprocessing work and can get the good quality image just from the DCGAN.

References

1. Bao, J.D., et.al.: Towards openset identity preserving face synthesis. In: CVPR (2018)
2. Brock, A., Donahue, J., Simonyan, K.: Large scale GAN training for high fidelity natural image synthesis. In: ICLR (2019)
3. Goodfellow, I., et al.: Generative adversarial nets. In: NIPS (2014)
4. Isola, P., Zhu, J., Zhou, T., Efros, A.: Image-to-image translation with conditional adversarial networks. In: CVPR (2017)
5. Lin, J.Y., et al.: Conditional image-to-image translation. In: CVPR (2018)
6. Radford, A., Metz, L., Chintala, S.: Unsupervised representation learning with deep convolutional generative adversarial networks. arXiv preprint arXiv:1511.06434 (2015)
7. Zhao, J., et al.: Dual-agent GANs for photo realistic and identity preserving profile face synthesis. In: NIPS (2017)
8. Shen, Y., Luo, P., Yan, J., Wang, X., Tang, X.: FaceidGAN: learning a symmetry three-player GAN for identity preserving face synthesis. In: CVPR (2018)
9. Tuzel, O., Susskind, J., Wang, W., Shrivastava, A., Pfister, T., Webb, R.: Learning from simulated and unsupervised images through adversarial training. In: CVPR (2017)
10. Tran, L.Q., et al.: Representation learning by rotating your faces. IEEE PAMI (2018)
11. Hu, Y., Wu, X., et al.: Pose-guided photorealistic face rotation. CVPR (2018)
12. Goodfellow, I., Jean, P.-A., Mehdi, M., Xu, B., David, W.-F., Ozair, S., Courville, A., Bengio, Y.: Generative adversarial nets. In: Advances in Neural Information Processing Systems, pp. 2672–2680 (2014)
13. Radford, A., et al.: Unsupervised representation learning with deep convolutional generative adversarial networks. arXiv preprint arXiv:1511.06434 (2015)
14. Pavaskar, et al.: Real-time vehicle-type categorization and character extraction from the license plates. Cognitive Informatics and Soft Computing, pp. 557–565. Springer (2019)
15. Suneeta, V.B., et al.: Facial expression recognition using supervised learning. In: Computational Vision and Bio-Inspired Computing. ICCVBIC 2019. Advances in Intelligent Systems and Computing, vol. 1108, pp. 275–285 (2020)

Neural Network Based 3-Phase Permanent Magnet Synchronous Motor

Nahid Guard and Suneeta V. Budihal

Abstract An adaptive model for Permanent Magnet Synchronous Motor (PMSM) is highly required in applications such as industrial, domestic, automobile, etc. The downside of PMSM is the pulsation of torque which results in instantaneous torque that pulsates periodically with the position of a rotor. As the PMSM is a synchronous machine, speed and torque are the parameters to be controlled externally. The traditional Proportional Integral (PI) controller is mainly used to control these parameters of the motor in industry because of its robustness, but the performance of PI controller tends to decline as the system opts for dynamic changes. Hence, the conventional PMSM is replaced with Artificial Neural Network (ANN) based PMSM model which will avoid the torque pulsation at the dynamic response. The main purpose of ANN is to achieve accurate prediction of the torque, current of the motor and while the speed and voltage parameters are dynamically changing.

Keywords Permanent magnet synchronous motor (PMSM) · Artificial neural network (ANN) · Forward propagation · Back propagation

1 Introduction

Many electrical vehicles and industrial applications use PMSM because of features like high efficiency, reliability, high torque to inertia ratio, ease of maintenance, a constant power region, high power density, etc. Synchronous motors are used for low to medium-power applications. A magnetic flux in the air gap is produced by the Permanent Magnet Machines that are electromechanical devices with magnets. The AC motors are of two types, asynchronous and synchronous motor. The synchronous motors have a rotor that is free to rotate and fixed stator winding which is in electrical

N. Guard (✉)
KLE Technological University, Hubballi, India
URL: http://www.kletech.ac.in

S. V. Budihal
School of Electronics and Communication, KLE Technological University, Hubballi, India

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_55

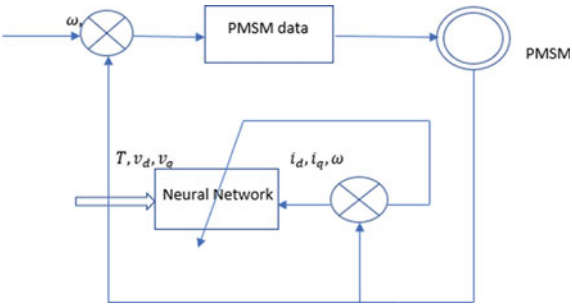


Fig. 1 Overview of neural network model

connection to the AC supply with a source which is exited similar to the DC motor as it has an external connection for the rotor to rotate at the synchronous speed. PMSM has winding [1] on the stator and a permanent magnet rotor.

PMSM generates magnetic flux using permanent magnets in the rotor, driven by the stator through a synchronous rotational field, when the stator is perpendicular to flux. In the rotor, permanent magnet helps in rotation of the rotor due to the magnetic field generated across a stator. Due to the design of substantial air gap magnetic flux, a highly efficient design of the PMSM can be obtained. The conventional [2] PI controller is used to control the speed and torque of PMSM. The PI controllers are very sensitive to disturbances. On the other hand, the intelligent model of the PMSM does not need exact mathematical equations of the motor. The objective of the paper is to develop and analyze an ANN model as shown in Fig. 1 to rectify the traditional problems of dynamic response in the model and improve the time taken to run the model [3]. The idea is to focus on controlling current in the PMSM. The objective is to overcome dichotomize inaccuracy problem present in the traditional PI vector based controller by building a neural-network based controller. The NN is developed by considering the dynamic behavior of PMSM parameters where fully dynamic equations are used, and the training is done to optimize the dynamic programming [4]. The second objective is evaluation of the adaptive and robust performance of the Neural Network (NN) controller against conventional PI vector controller under the variations of motor parameter. In PMSM, usually the inner loop of current control has a hysteresis current controller and outer loop has PI controller for speed control. The paper deals with the development of a NN-based approach to enhance efficiency in control of PMSMs. In industries, the PI controller is mainly used because of its robustness [5]. As there is diversity in the parameters and due to dynamics of the system in some cases, with operating conditions, the performance of the controller gets degraded. One promising solution is to apply Artificial Intelligence to develop a NN-based speed controller to serve the purpose. A feed forward NN is used instead of a PI controller for the vector control scheme of PMSM to control the speed. Analysis of results has ensured performance and robustness against load and parameter variations.

The main aim is to obtain the speed to have an accurate trajectory control, being the load parameters unknown [6]. The ANN captures the unknown nonlinear dynamics

of the motor. The performance is then better than the traditional controllers for a dynamic load. Investigation of ANN-based speed control for the motor in several conditions such as loaded/no-load and open/closed loop is the primary objective of the paper. In industries, the PI controller is mainly used because of its robustness. But, with time as there is diversity in the parameters and due to dynamics of the system the in some cases, with operating conditions, the performance of the controller gets degraded [7]. One promising solution in such case is to apply Artificial Intelligence, to develop a NN-based speed controller. A feed forward NN is replaced for the PI controller of the vector control of the PMSM.

2 Proposed Methodology

It includes the brief description of the block diagram of NN. It provides the outline for the design and operation of network. Electric Motor Temperature Dataset is a large data-set with various attributes that describe the features like Current and Voltages at *d* and *q* axis, torque, motor speed, etc. This dataset has huge diversities, the data-set can be used as the training dataset and testing dataset for the computer vision tasks such as predicting parameters.

2.1 Functional Block Diagram of a Neural Network Model

The block diagram in Fig. 2 provides brief description of the overview of NN to replace the traditional PMSM. Initially, a reference model that has the model for PMSM and required parameters are given as a feedback to the NN block which

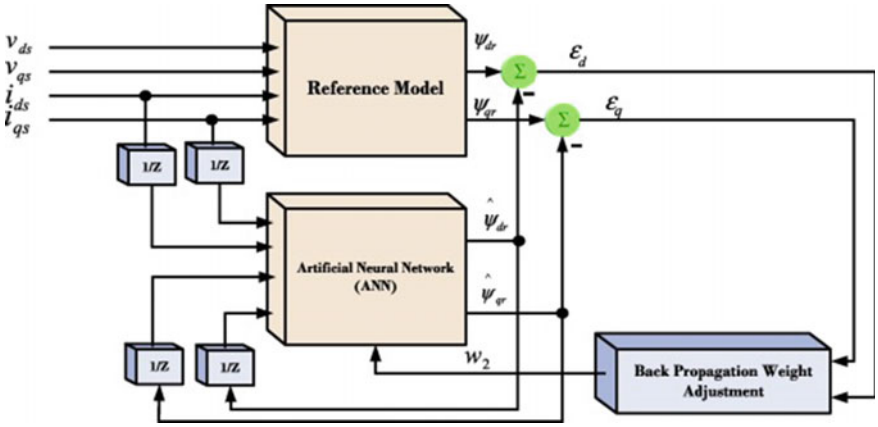


Fig. 2 A functional block diagram of neural network based PMSM

then performs the forward propagation through the layers. Back propagation block minimize the errors and optimize the model. NN for PMSM should yield positive and negative output due to the requirement. Hence, we add LReLU activation at the last output layer, which can account for non-liberates as well. Adam optimizer is chosen that can adjust the learning rate on each weight parameter, which guarantees effective training. In order to ensure the recovered data has less loss, we use a function $J(w, b)$ to calculate the differences between them.

$$J(w, b) = \frac{1}{m} \sum_{i=1}^{m} (x' - x)^2. \tag{1}$$

ReLU, leaky Relu, and sigmoid activation functions are used in the network. Activation function Rectified Linear Units (ReLU) is the most common activation function used in NNs. It can allow the model to account for non-linearities very well. It returns 0 for negative values and for any positive values x it returns the value back. Dying ReLU, where some of the neurons die and remain inactive of all the inputs irrespective of the input supplied, is another issue in ReLU. In such case, no gradient flow is seen and therefore the performance of the NN is affected if a large number of dead neurons are present. In order to solve this issue, make use of another activation function called Leaky ReLU in which the slope is changed in the negative direction thus causing a leakage which extends the range of ReLU.

3 Implementation Details

These neurons do not hold any information rather carries the message from one neuron to another. The NN interprets the sensory data and labels or clusters raw input. All real-world data, be it images, sound, text must be translated into the patterns the NN recognize which are numericals in vectors.

3.1 Mathematical Model for PMSM

This section deals with the mathematics involved in developing and designing the PMSM reference model with permanent magnets. In case of a dynamic model certain assumptions have to be made. A force known as Magneto Motive Force (MMF) is sinusoidally distributed which indirectly balances the windings in stator, the changes in the parameter are neglected, the rotor vs inductance have a sinusoidal position. Expression for each of the 3 phases can be shown as,

$$V_a = R_s i_a + p\psi_a \tag{2}$$

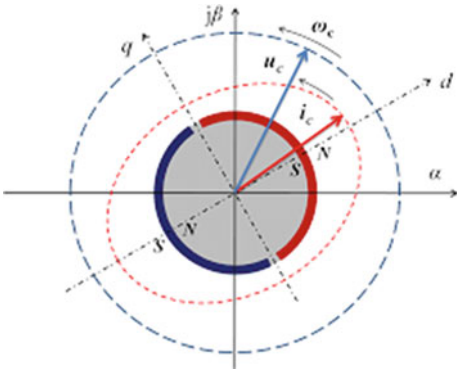


Fig. 3 PMSM reference frame

$$V_b = R_s i_b + p\psi_b \tag{3}$$

$$V_a = R_s i_c + p\psi_c \tag{4}$$

The stator voltages can be symbolized by V_a, V_b and V_c; R_s is the stator resistance; i_a, i_b and i_c are the currents in stator as denoted in the Fig. 3 of reference frame. The stator voltages according to the air gap between the rotor and stator are characterized by the following equations,

$$V_q = R_s i_d + p\psi_d - \omega\psi_d \tag{5}$$

$$V_d = R_s i_q + p\psi_q + \omega\psi_q \tag{6}$$

where i_d and i_q symbolizes the stator currents; The electrical speed of the rotor is symbolized by ω, i.e. omega and ψ_d, ψ_q are the magnetic fluxes.

$$\psi_q = L_q i_q \tag{7}$$

$$\psi_d = L_d i_d + \psi_m \tag{8}$$

Thus, d and q are the new axes of the variables. The Torque (Te) is the electrical torque that is obtained by solving all the above equations.

$$T_e = \frac{3P}{4}\psi_m i_q + (L_d - L_q) i_q i_d \tag{9}$$

In the equation P is a variable that can be 2, 4, 6 or 8, etc. which denotes the number of pole pairs. In exposure to a constant flux where the magnetic conductivity in the air gap remains constant, id current is nearly equal to zero and the electric torque is expressed as,

$$T_e = \frac{3P}{4}\psi_m i_q \tag{10}$$

Relating to the equation of motion torque can be denoted as,

$$\sum T_e = J\partial \tag{11}$$

3.2 Mathematical Model of Neural Network

NNs provide a good classification and clustering accuracy by grouping the unlabelled data according to their historical similarities among the inputs fed to the model. It helps to group data that are unlabelled according to the similarities among the inputs, and thus they also manage to classify sets of data when they have datasets that are labeled to train. There are many different NN architectures that have been tried, some involve more mathematical analysis of the problem, some are completely based on imitating the things seen under the microscope. The most commonly used one is shown in Fig. 4. It has input layer, hidden layer, and output layer. Each layer consists of nodes that are represented by small circles in the diagram. Information flows between the nodes and is indicated by the lines that connect one node to another. We have used a NN with feedback paths termed as back-propagation. A Neural Network built to replace the traditional PMSM has two main parts, one is the Forward Propagation and the other is the Back-Propagation. In Forward Propagation, the input data is fed through the network in forward direction. Suppose we have an input array named x and the weights are w_1, w_2, w_3 then

$$\sum = (w_1 * x_1) + (w_2 * x_2) + (w_3 * x_3) + b \tag{12}$$

$$y_i = \phi(\Sigma) \tag{13}$$

After calculating the output we need to compare the predicted output with the actual,

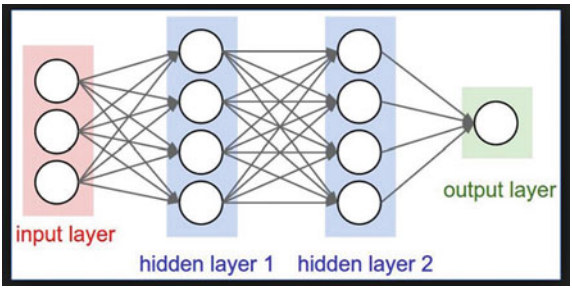


Fig. 4 Layers of neural network

$$e_i = d_i - y_i \tag{14}$$

Weight adjustment has to be done in the next step where, alpha is the learning rate that ranges between 0 and 1 and j is the nth node in case we have more than on hidden layer.

$$w_{ij} = w_{ij} + \alpha * x_i * e_i \tag{15}$$

Weight updation using generalized delta rule demonstrates that we need to adjust the weights to reduce the errors where, the error is now replaced by the delta term,

$$w_{ij} = w_{ij} + \alpha * x_i * \delta \tag{16}$$

$$\delta = \phi'(\Sigma_i)e_i \tag{17}$$

Weight Initialization: We do not want the layer activation outputs to explode or vanish during the course of forward propagation in a deep NN. This is prevented by choosing the weights correctly. Back propagation architecture is one of the standard methods to train a NNs which helps to calculate the loss function gradient w.r.t. the network weights. Using both the FP and BP, a NN is built in MATLAB using two hidden layers and 10 neurons in each hidden layer, the inputs given to the model are motor speed, voltages at d and q axis (direct and quadrature axis), for which three different output currents at d and q axis, torque are obtained.

4 Results and Discussion of Neural Network Model

The NN model built-in the Simulink are shown in Fig. 5, where we have two hidden layers as shown in Fig. 6 and different activation functions used in each layer. The actual model is developed through scripting the NN in MATLAB and then converting the script into a Simulink model by the use of Gensim function which is a built in function to generate a Simulink model through the given script. Figure 7 shows the result of NN model based prediction of torque and current. The inputs are towards the left and the predict vs the actual outputs are towards the right. Inputs given to the model are motor speed, voltage at Q-axis, and voltage at D-axis. Outputs obtained are Torque, current at Q-axis and current at D-axis. It can be clearly observed that the predicted vs actual outputs are quite relevant and the accuracy is huge, the error is very less and thus the model has learned the mapping is a precise one. The learning done by the model in terms of pattern recognition is pretty impressive. Hence we can conclude that the model is perfectly able to understand and predict new data with much higher accuracy if trained with data more that 10,000.

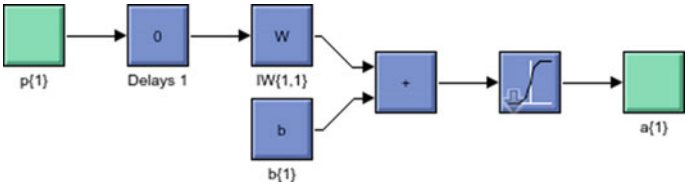


Fig. 5 Simulink model of neural network

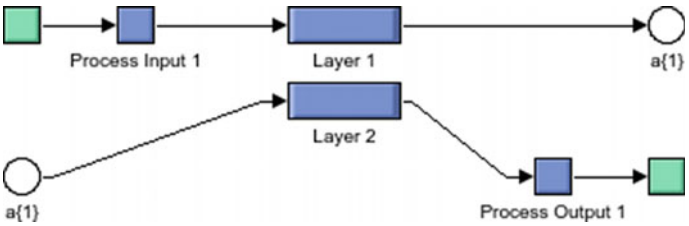


Fig. 6 Layers of neural network model

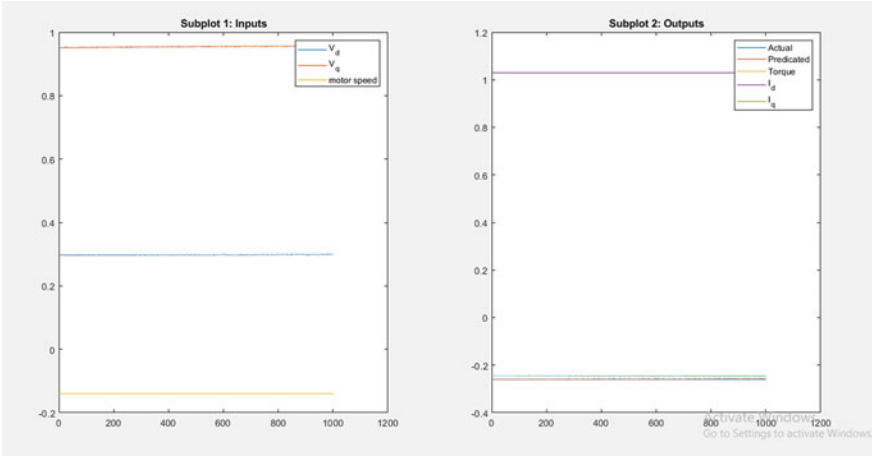


Fig. 7 Results of torque and current from neural network

5 Conclusion

We could effectively predict the torque and currents from given voltage and motor speed with high accuracy by using the Deep learning concept of NN. With CPU access and around 10,000 input data for training the NN is able to provide high amount of accuracy. Although a Simulink model of NN can be obtained which can predict the future response in real-time and can completely replace the PMSM motor model of Open Modelica or MATLAB-Simulink.

References

1. da Silva, C.A., et al.: Analysis of a ThreePhase LSPMM by numerical method. IEEE Trans. Magn. **45**(3), 1792–1795 (2009)
2. Ugale, R.T., et.al.: A novel line start permanent magnet synchronous motor using two part rotor. TENCON -2009 IEEE Region 10 Conference, pp. 1–5 (2009)
3. Ashok, K.R., et al.: Sensor less control of PMSM fed from three phase four switch inverter based on back EMF observer and sliding mode controller with fast reaching law. IJET **7**(2), 725–731 (2018)
4. Ugale, R.T., et al.: Effects of Short Power Interruptions and Voltage Sags on the Performance of Line Start Permanent Magnet Synchronous Motor, pp. 184–188. Machines and Drives, PEMD, in Power Electronics (2008)
5. Pattanashetty, V.B., et.al.: Inclusive Device to Connect People with the Surrounding World. Information and Communication Technology for Competitive Strategies. Lecture Notes in Networks and Systems, vol. 40. Springer
6. Gomez, J.C., et.al.: Induction motor behavior under short interruptions and voltage sags. IEEE Power Eng. Rev. **21**(2), 11–15
7. Pavaskar, et.al.: Real-time vehicle-type categorization and character extraction from the license plates. In: Advances in Intelligent Systems and Computing, vol 768, pp. 557–565. Springer, Verlag
8. Suneeta, V.B., et al.: Facial expression recognition using supervised learning. In: Computational Vision and Bio-inspired Computing. ICCVBIC 2019. Advances in Intelligent Systems and Computing, vol. 1108. Springer

Evaluating the Effects of Hardware Configurations on the Execution of Radar Tracking Simulink Models

Olukayode Oki and Sizolwethu Ndlovu

Abstract Simulation is the evaluation of a large number of alternatives under different realistic scenarios determined by the decision makers. With simulation only, organizations can only simulate how their system technologies will react under given conditions and what outcomes are expected. However, this does not give the optimal simulating solution, and there is no guarantee that the solution will utilize the resources effectively; hence, there is a need for simulation optimization. In exploiting simulation optimization, various approaches have been adopted to develop Simulink models. Simulink model is an environment used for multi-domain simulation, automatic generation of code, verification of embedded systems and uninterrupted tests. However, the optimization of the Simulink execution time still remains an open issue. This paper evaluates the performance of radar tracking Simulink model execution time under different hardware configurations. The results obtained in this study show that the system performance can be optimized by using better hardware, which in turn increases the system resource availability. It was also observed that the execution time can be optimized by reducing the model computational time. Hence, the results obtained in this work show that both the hardware and software resources can be used to optimize the simulation models.

Keywords Radar tracking · Simulink · Execution time · Optimization

1 Introduction

The mixture of simulation and optimizations has been of much expectation for a long period of time; however, significant progress has only been achieved in the past decade [1]. While simulation may be employed in many fields, computer simulation

O. Oki (✉)
Walter Sisulu University, East London, South Africa
e-mail: ooki@wsu.ac.za

S. Ndlovu
University of Zululand, Richards Bay, South Africa

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_56

is the virtual modeling of a real life or imaginary situation on a computing device for studying and evaluation purposes. By varying the model parameters, predictions about the behavior of the system can be made.

Also, a simulation experiment is a test where meaningful changes can be introduced to the model input variables such that one is able to observe changes in the output variables and are able to pinpoint the reasons behind the observed changes. The use of both the large number of input variables and complex simulation model would make the simulation experiment to become computationally expensive and unaffordable. Apart from the high cost of computation, a higher cost is incurred when the selected input variable values are sub-optimal. Hence, the simulation optimization can be defined as the method used to find an optimum value for the input variable from among all possibilities without explicitly evaluating each possibility.

The goal of simulation optimization is to minimize the resources used while maximizing the obtained results from the simulation experiments [2]. The performing of simulation on computing devices requires the use of simulation software. MATLAB (MATrix LABoratory) is one such software that has been used for simulation and is a versatile multi-paradigm numerical computing environment that offers within it the Simulink tool. The Simulink tool was developed by MathWorks, and it is a graphical programming environment for simulating, modeling and analyzing multi-domain dynamic systems. The tool also provides a user-defined block libraries, graphic editors and solvers for dynamic modeling and simulating of systems. The integration of Simulink and MATLAB has enabled the incorporation of MATLAB models into Simulink models, thereby allowing the exportation of simulation experiment results to MATLAB for additional analysis [1].

The major increase in use of Simulink by researchers and other entities (such as military systems) has called for the optimization of Simulink models in order to obtain better result and execution times. The growing complexity of industrial and military systems causes their simulated models to execute more slowly [3]. However, a poor execution time will affect the performance of these systems. Hence, there is a need for the optimization of these models' execution time without reducing the model efficiency.

In recent years, the optimization of simulation has become one of the most significant simulation technologies. However, the increase in the computing power of the latest computer systems is helping in advancing the optimization of simulation. However, in order to achieve optimal execution of models, software implementations should also be considered. The computing requirements for the optimization of simulation models without the software approaches become very difficult. Simulation has been used in a variety of areas where use of the real system is either not feasible or may have serious adverse effects (such as radar systems), or when considering a non-existent system or a system still in its development stage [1, 2].

Radar systems were first developed for military use in the beginning of the twentieth century, and since then, it has grown to have many uses outside the military such as airport traffic control, blind spot detection, automotive applications, fire control and space (ground mapping, earth surveillance) [4].

The radar simulation technology is an important piece for development of radar systems because certain radio wave systems and behaviors are very expensive and dangerous to be tested in real-world environment. Hence, the need to optimize the execution time of radar Simulink models remains an important open issue that needs to be addressed.

This paper seeks to investigate the performance of radar tracking Simulink model when subjected to different hardware configurations using execution time and the CPU usage metrics. The results obtained in this study show that the use of better hardware configurations improves the availability of system resources and hence optimizes the system performance. It was also observed that the radar tracking execution time can be optimized from the software implementations by reducing the model computational complexity time. The remaining sections of this paper are organized as follows: Sect. 2 presents the related works in the Simulink simulation research area. Section 3 details the experimental setup used, while the performance evaluation results were presented in Sect. 4. The paper is concluded in Sect. 5 with the future work.

2 Related Works

Recently, the optimization of simulation concept is one of the major significant challenges in simulation technology. In achieving optimal execution of models, both the hardware and software implementations should be considered. The computing requirements for simulation optimization make the optimization without software approaches very difficult.

The authors in [2] argue that the MATLAB has different performance pattern on OS like Linux, Windows and Mac. The differences in MATLAB performance were attributed to the MathWorks, which is normally used to build models in MATLAB. In MathWorks, products are build using a different compiler on each operating system platforms, with each having their individual performance features. MathWorks also integrates third-party libraries into their products, which makes it perform differently on different platforms. The authors in [2] also argued that the operating systems may cause differing performances, especially when performing intensive disk or graphic operations which may have an effect on the execution of application such as radar tracking Simulink models.

The effect of different hardware components on MATLAB performance was presented in [5]. Their study explained that the performance of MATLAB is impacted by each and every component of a computer system configuration. Computers with a greater number of CPU cores can perform better than the ones with a lower number of cores. The performance outputs could vary, depending on the kind of operating system where the MATLAB application is installed. MATLAB uses multithreading, which helps to make full use of the MATLAB's natural parallelism feature. However, it is not every MATLAB functions that are multithreaded, and the execution speed may vary depending on the algorithm used. In order to improve the capability, the

system toolbox offers parallel programming constructs, which uses the computer multiple cores.

In [6, 7], it was proposed that the general system performance can be improved by using virtual cores. However, it may not have much effect on MATLAB performance. Hyper-threading CPUs suggest that the computer system has double the number of cores than it really takes. While using a tool like task manager, the MATLAB process could seem like using 50% of the available computer CPU cores, while in reality, the remaining "unused" 50% is the virtual cores generated by the hyper-threading.

Computers may suffer a degrading of performance because of the thrashing when MATLAB runs concurrently with other programs and use more than the available physical memory; this could make the computer opt for the use of virtual memory. Hence, a MATLAB application that is using more than 3 GB of memory would require 64-bit version of MATLAB.

The simulation model of real-time applications was presented in [8]. The authors explained that, in order to build up a simulation model of real-time applications, it is beneficial to know the timing behavior of the system. To optimize the execution time of the Simulink model, it is often necessary to know the worst-case execution time of the model source code that each Simulink block in the model produces. This can be achieved by using the source code generator that will produce start and stop markers. Marking the code produced by each block provided the code generation optimization features which lead to overlapping of blocks or block ranges which are not used.

In [3], the authors observed that the Simulink models can perform much faster when using all the CPUs for natural parallelism simulations. This is because of the multithreading and increased processing power present on multiple cooperating CPUs, which can improve the system performance up to 50% than in single CPUs. The time taken to execute a model may also depend on the type of solver used based on the model specifications. The time taken to execute a model may also be reduced by removing the terminator blocks since they allow the transfer of unused data between CPUs consuming resources.

The performance analysis of MATLAB simulation has been widely studied, but none of the existing works has looked at the optimization of the execution time from the context of both hardware and software implementations. The investigation presented in this paper is different from the previous studies in respect of execution time optimization from the context of both the hardware and software implementations using the radar tracking Simulink models.

3 Experimental Setup

This section describes the experimental setup employed for this study. The simulation models used for investigation were the radar tracking models. These models are part of the MATLAB Simulink models. The radar tracking models were selected because of its fairly complex electronic warfare models, which would enable us to develop

a more robust automated approach. Also, radar tracking Simulink models have been used extensively in other research areas [2, 9].

Two Simulink models were selected for the evaluation: aero radmod and radarsys. The aero radmod model was implemented using the Kalman filter algorithm. This algorithm can be invoked by using the "aero_radmod_dsp" command in the MATLAB terminal environment. The radarsys can be invoked by using the "radar_system" command in the MATLAB terminal window.

The aero radmod model consists of three main functions and aims to generate the velocity, acceleration in polar coordinates and position of the aircraft in space. It simulates imprecise readings by the sensor by adding noise measurements with the use of a Kalman filter for estimating position and velocity from noisy measurements. The extensive detail of the aero radmod model can be found in [10].

The radarsys model contains subsystems that model important features of a system. The model is of the type used for target positions and velocity detections. The radarsys model consist of radar pulse generator, simulink representation of a moving object, a radio frequency transmitter subsystem and a radio frequency receiver with a receiving module [11]. The model was executed with its default parameters, which remained unchanged throughout the evaluation process. After the model had been executed with the profiler being attached to the simulation, the model output was displayed which shows the magnitude and phase of the filtered return signal with noise (filtering in time), the "RadarPulse" which displays the time-domain representation of a chip signal with a 10% of duty cycle, and the "Filtering in frequency" chart displaying the imaginary and real response of the filtered return signal with noise through an ideal filter implementation. For extensive detail of the radarsys model functions and operations, see [12].

The two models were simulated on different hardware configurations, using five different computer setups which were configured for the simulation of the aero radmod and radarsys models. Each of the computing devices was using Windows OS, and they were all installed with MATLAB & Simulink 2017a for all the simulation that was conducted. Table 1 contains the additional details of the hardware configuration setup that was used for all the experiments reported in this paper.

Table 1 Hardware configuration setup detail

Device type	Processor type	RAM	CPU cores
A	Intel core i7 (3.40 GHz)	8 GB (7.86 usable)	4 cores, 8 logical
B	Intel core i5 (2.60 GHz)	4 GB (3.84 usable)	2 cores, 4 logical
C	Intel Celeron (1.60 GHz)	4 GB (3.85 usable)	2 cores, 2 logical
D	Intel core i5 (3.40 GHz)	2 GB (1.95 usable)	1 core, 1 logical
E	Intel core i7 (2.56 GHz)	4 GB(3.50 usable)	1 core, 1 logical

On each of the specifications indicated in Table 1, MATLAB and Simulink were installed and properly set up to allow for simulations. While MATLAB environment was opened, the Simulink terminal was opened up and the model script (aero_radmod_dsp OR radar_system) was invoked. Upon invocation of the model, the model was given adequate time to fully load on the MATLAB/Simulink platform.

Thereafter, the models were run on normal accelerator mode, with the Simulink profiler option being enabled. After the execution of the model, the execution time displayed by the profiler report was then captured into the result table for further analysis. This was repeated for each of the specifications mentioned in Table 1, with the average Execution Time Required (ETR) calculated and recorded in each result table.

On each of the executions above, the computer's resource monitor was opened up for the observation of the effect; each model had on the CPU usage of the device as a whole. The CPU usage was then recorded, and the table containing the results of the respective specifications was analyzed and presented in this paper.

4 Performance Results

The performance evaluation results of radar tracking Simulink models are presented in this section. In order to achieve a more statistical accuracy, each of the reported analyses is the average of the results of ten experiments for each setup. The computer specification and their respective execution time are presented in Fig. 1. The execution time is presented in seconds, and it was obtained using the profiler measuring method in MATLAB. The lower the achieved value, the better the performance.

From the results presented in Fig. 1, it can be observed that the most efficient execution times are those with higher hardware configuration. It can also be observed

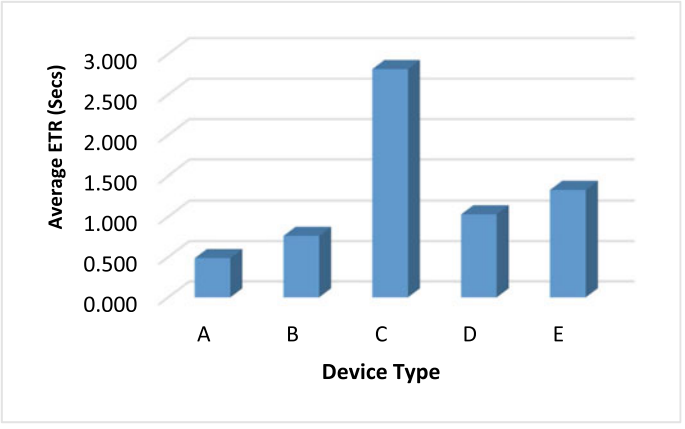


Fig. 1 Aero radmod model execution time

from the achieved results that the clock speed is affected mostly by the processor type and the amount of RAM available which is in-line with the existing literature results.

By reducing the simulation time from 100 to 98 without altering any model parameters, this led to achieving an improvement of 21.36% as presented in Figs. 2 and 3.

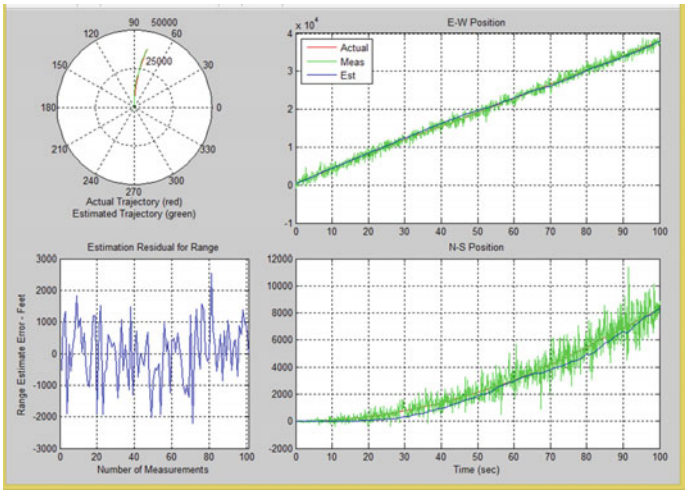


Fig. 2 Results before computational time reduction

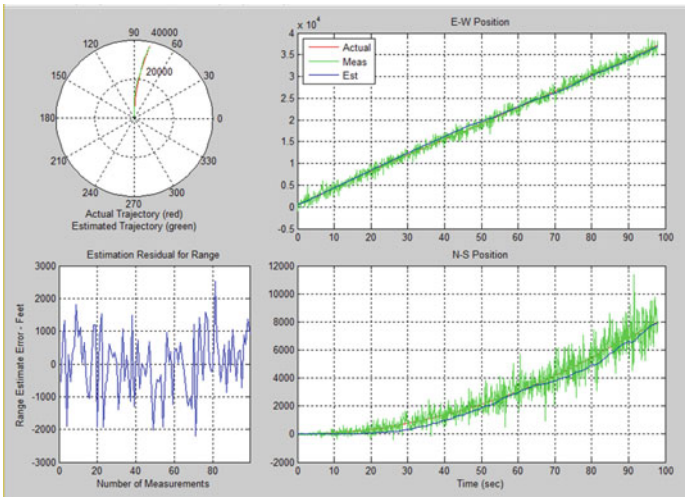


Fig. 3 Results after computational time reduction

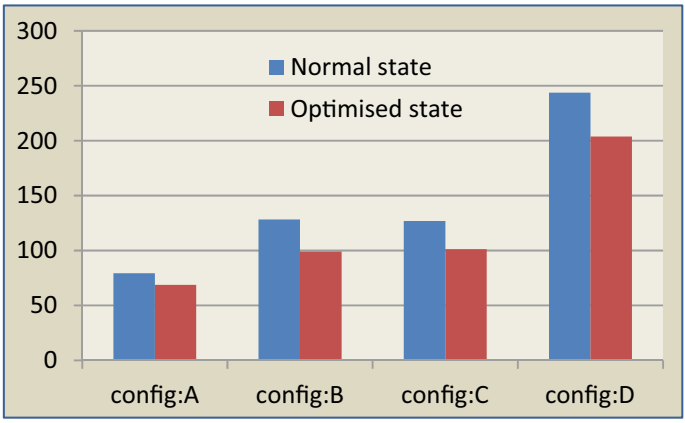


Fig. 4 Radarsys model optimization results

From Figs. 2 and 3, it can be observed that the obtained results are almost the same, but with different detail levels. This performance can be attributed to the simulation stop time being reduced from 100 to 98. The detailed procedures and reason for simulation time reduction were explained in Sect. 3.

The execution time for radarsys model under different configurations are presented in Fig. 4. The optimization of the radarsys model was achieved by reducing the stop time by 0.2 s. The results presented in Fig. 4 show an improvement. The achieved improvement ranges between 13 and 23% approximately.

5 Conclusions

The optimization of simulation execution is one of the major significant issues in simulation technology, especially in recent times. The computing requirements for the optimization of simulation models without the hardware and software approaches become very difficult. Simulation has been used in a variety of areas where use of the real system is either not feasible or may have serious adverse effects (such as radar systems), or when considering a non-existent system or a system still in its development stage.

This paper analyzed the general performance of both aero radmod and radarsys Simulink models, using MATLAB simulation. The authors analyzed how different hardware configurations affect both the execution time and the CPU usage of aero radmod and radarsys Simulink models. The results obtained in this study show that the use of better hardware improves the availability of system resources and optimizes the system performance. It was also observed that the execution time can be optimized by reducing the model computational time.

Based on the results presented in this paper, the authors can reaffirm that the hardware and software resources can be used to optimize the simulation models.

The Simulink resource utilization can still be enhanced. The future work would focus on developing an upgraded version of aero radmod and radarsys with multi-threaded feature. This multi-thread feature would enable both the aero radmod and radarsys to utilize the advantages of modern multi-threaded computers.

Reference

1. Mathworks.: Choosing hardware-Mathlab, 06 June 2016. (Online). Available: https://www.mathworks.com/products/matlab/choosing_hardware.html. Last accessed 2019/10/12
2. Carson, Y., Anu, M., Andradottir, S., Healy, K., Withers, D., Nelson, B.: Simulation optimization: methods and applications. In: Proceedings of the 1997 Winter Simulation Conference (1997)
3. Moore, S.K.: Multicore CPU: Processor Proliferation. IEEE Spectrum report (2011)
4. Mahmod, A., Al-Zubaidy, K.H., Sayidmarie, Saad, S.A.-S.: Radar system simulator using PC and matlab simulink. In: 2006 International Radar Symposium (2006)
5. Wu, H., Zhao, H.-z.: Modeling and simulation of a full coherent LFM pulse radar system based on simulink. In: 2013 2nd International Conference on Measurement, information and Control (ICMIC), Harbin, China (2013)
6. Kirner, R., Lang, R., Freiberger, G., Puschner, P.: Fully automatic worst-case execution time analysis for. In: Proceedings 14th Euromicro Conference on Real-Time Systems. Euromicro RTS (2002)
7. Hrčka, L., Važan, P., Šutová, Z.: Basic overview of simulation optimization. Research Papers Faculty of Materials Science and Technology Slovak University of Technology, Sciendo, vol. 22(341), pp. 1–6 (2014)
8. Ozard, J., Desira, H.: Simulink Model Implementation on Multi-Processors under Windows-NT (2000)
9. Pontus, B., Jerker, B.: Optimisation-based black-box testing of assertions in Simulink models. Technical report, No 711 (2005)
10. Mathworks.: Radar Tracking System—MATLAB & Simulink Example, (Online). Available: https://www.mathworks.com/help/simrf/examples/radar-tracking-system.html. Last accessed 2019/10/13
11. MathWorks Inc.; Simulink Documentation, (Online). Available: https://www.mathworks.com/help/simulink/index.html?s_cid=doc_ftr. Last accessed 2019/10/13
12. MATLAB.: Radar tracking-MathWorks India, (Online). Available: https://in.mathworks.com/examples/matlab-dsp-system/mw/dsp_product-aero_radmod_dsp-radar-tracking. Last accessed 2019/10/23

Straight-Line Detection with the Hough Transform Method Based on a Rectangular Grid

Cheick Amed Diloma Gabriel Traoré and Abdoulaye Séré

Abstract This paper deals with the reduction of straight-line detection time with the standard Hough transform method. The number of processed pixels is reduced and therefore contributes to improve the processing time. In a preprocessing step, the image is divided into rectangular boxes and only boxes with a certain percentage of high intensity pixels are considered in computing Hough transform. The proposed method is evaluated with examples for edge detection in Lena image. The main purpose is to show that how straight-line detection can be done rapidly.

Keywords Hough transform · Pattern recognition · Reconstruction

1 Introduction

The recognition of discrete primitives (straight line, circle) in the image is based on the linear Hough transform [7, 8] allowing first to recognize the straight lines in the image, then with the help of its generalization [2, 5] to detect other geometric shapes (circles, ellipse, parabola).

Jung et al. in [9] have worked on rectangle detection in a digital image. To detect straight lines in a digital image, not all areas of the image are useful. The main purpose of this paper is the utilization of a virtual rectangular grid to identify interesting regions in an image. This grid can be parametrized and leads to the reduction of recognition time for straight lines.

In this paper, we discuss about time reduction for the detection of straight lines with the Hough transform method and we do rectangle selection and their dual computing; all this will give dual space dimension reduction. After this introduction

C. A. D. G. Traoré (✉) · A. Séré
Laboratoire d'Algèbre, de Mathématiques Discrètes et d'Informatique (L.A.M.D.I), Université Nazi BONI (U.N.B), Bobo-Dioulasso, Burkina Faso
e-mail: abdoulaye.sere@recifaso.org

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_57

follows, the preliminaries which recall us a set of definitions and properties to better understand the rest of the paper. The method section presents all the algorithms, in particular those of meshing and recognition of lines with generalized preimage [4, 13] with the accumulator of data. All the algorithms are in pseudocode, strongly inspired by the programming language Python. Section 4 presents the results of the implementation of different algorithms on real images.

2 Preliminaries

This section recalls the definitions of analytical straight line and standard Hough Transform.

Definition 1 (*Analytical discrete straight line* [11]) An analytical discrete straight line with parameters (a, b, μ) and thickness w is defined by the set of integer points (x, y) verifying :

$$\mu \leq ax + by < \mu + w \ , \ (a, b, \mu, w) \in \mathbb{Z}^4, \ gcd(a, b) = 1 \tag{1}$$

The straight line is :

- thin if $w < \max(|a|, |b|)$
- thick if $w > \sum(|a| + |b|)$

Definition 2 (*Standard Hough Transform* [5]) Let $\mathscr{I} \subset \mathbb{R}^2$ be an image space. Let l be the number of columns in an image. Let h be the number of rows in an image. Suppose that the point $(x, y) \in \mathscr{I}$. The dual $S(x, y)$ of (x, y) is the standard Hough transform defined by the set of points :

$$\{(\theta, r) \in [0, \pi] \times [-\sqrt{l^2 + h^2}, \sqrt{l^2 + h^2}] / \ r = x \cos\theta + y \sin\theta\} \tag{2}$$

3 Method

This section presents the proposed method with the virtual meshing and the straight-line recognition algorithms. This method is subdivised in several operations, which compose algorithm 1.

3.1 Building a Rectangular Grid

Here, we create a rectangular grid by subdividing a digital image into several rectangles. Nl or Nc can be prime integers, so they should not be divided by any integer. The

Algorithm 1: Steps of the method

Data: pretreated image
begin
to build a rectangular grid (algorithms 2, 3);
to select the rectangles having $\alpha\%$ of pixels (algorithms 3, 4);
to recognize the straight line in the image with the selected rectangles (algorithms 5, 6, 7, 8);

deterministic Eratosthene's algorithm [10] determined if NL or Nc is a prime integer before their factorization with algorithm 2. Deterministic Eratosthene's algorithm is slower than the probabilistic primality algorithm [1, 6] but we have the best result (the true result) with Eratosthene's algorithm.

Algorithm 2: Integer factorization

Function *factorization(Nl: integer): two integers*
Pre-condition: Nl no prime number
Variables: k, k_1, k_2: integers
tab: table of integers
Output: k_1, k_2: integers
$k \leftarrow 2$;
e$\leftarrow E(\sqrt{Nl})$;
while $k \leq e$ **do**
if $Nl \mod k = 0$ *and erast(k)=True* **then**
tab$\longleftarrow$k ;
$k \leftarrow k + 1$;
%divide by 2 for not considering a prime divider
%too small or too big of Nl
$k \leftarrow E(len(tab)/2)$ %len is the number of %elements in tab;
$k_2 \leftarrow tab[k]$;
$k_1 \leftarrow Nl/k_2$;
return k_1, k_2; % $Nl = k_1 \times k_2$, k_1 *et* k_2 no %necessary a prime number

In the algorithm 2, at the end of "while" loop in order to obtain a mesh having a number of grid k_2 big enough we take k_1 as the element located at the index: integer part of half the number of tab elements. To obtain k_2, we divide Nl by k_1, Nl is then subdivided into k_1 part of length k_2.

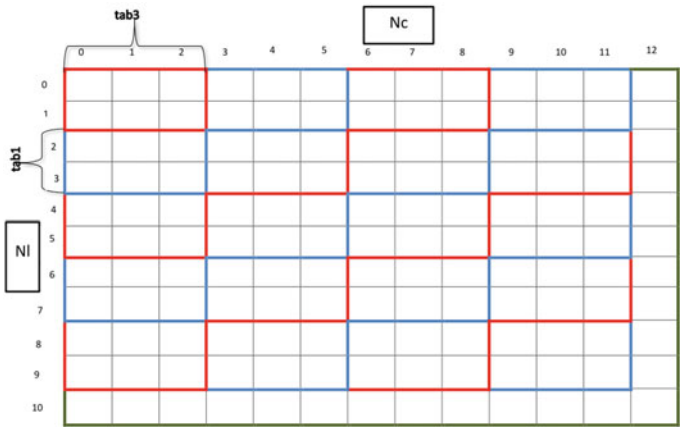
In the algorithm 3, in blocks "if", Nl is prime so Nl−1 no prime and therefore has two factors k_1 and k_2 of Nl. The instruction "tab$\longleftarrow$factorization (Nl)" provides k_2 and k_1, so that we have exactly k_1 rows of length k_2 stored, respectively, in tab[0] and tab[1] and a line like residue that will be considered during the straight-line recognition. The same for Nc but the values are stored this time in tab[2] and tab[3] with a column as the residue that will be considered in straight-line recognition.

Each virtual rectangle has dimension tab[1]×tab[3].

Algorithm 3: Building a rectangular grid

```
Function meshing(Nl, Nc: integers): 2 tables of integers
    Variables: tab: table of 4 integers,
    tabool: table of 2 integers
    tabool[0]←erast(Nl);
    if tabool[0] then
        tab[0], tab[1]←factorization(Nl-1);
    else
        tab[0], tab[1]←factorization(Nl);
    tabool[1]←erast(Nc);
    if tabool[1] then
        tab[2], tab[3]←factorization(Nc-1);
    else
        tab[2], tab[3]←factorization(Nc);
    return tab,tabool;
```

Example 1 if Nl = 11, NC = 13 then tabool=[1,1]; tab[0]=5, tab[1]=2, tab[2]=4, tab[3]=3, the dimension of one rectangle is tab[1]× tab[3]=2 × 3 (see below for a representation).



3.2 *Selection of Rectangles and Straight-Line Recognition*

Having an image provided with the rectangular grid in Sect. 3.1, we will select the rectangles having a rate of $\alpha\%$ pixels turned on with the algorithm 4.

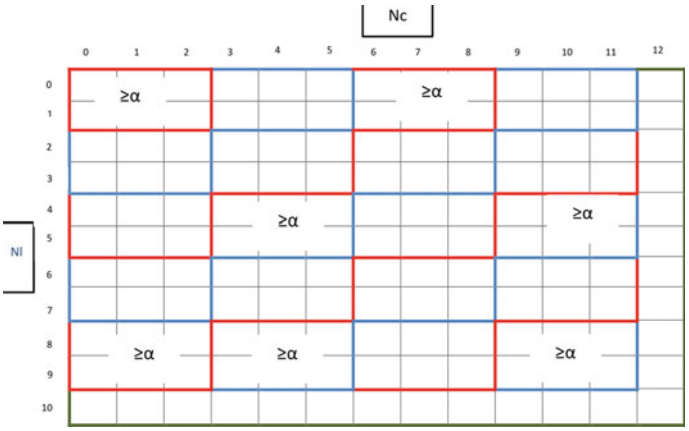
This will allow us in the next step to detect the presence of the straight lines either by using the notion of general preimage or by using a data accumulator [12] for selected rectangle. There is no raison to compute the dual of all rectangles defined by our mesh in the image because few lit pixels in a region characterized by noise or

Algorithm 4: Rate of pixels having the value *n*

```
Function count(A, C: vertices of the rectangle, n: integer, img: image): float
    % A and C is on the same diagonal of rectangle
    Variables: k, l: integer
    k ←0;
    for x between x_A and x_C do
        for y between y_A and y_C do
            if img[x,y]=n then
                k ← k + 1;          % we count the pixels %having the value n
            else
                l ← l + 1;
    return k/(k+l);
```

pixels from a thin straight line, and the robustness to noise of the Hough transform makes it possible not to lose data (for thin straight line we take a low threshold value, see Table 3).

Example 2 The rectangles with a lit pixel rate greater than α are selected. All pixels in the remaining row and column will be processed directly (see below for a representation).



3.3 Straight-Line Recognition with Data Accumulator

The utilization of a data accumulator [7] leads to the detection of straight lines presented in an image and does not depend on the order of addition of the discrete objects. This method has major disadvantages being its large size and the slowness of its execution once programmed. Sere et al. in [12] had optimized the data accumulator. In this paper, we will reduce its size by using the rectangular grid. The discrete line

recognition algorithm with the data accumulator (algorithm 8) will use the algorithm 4 of lit pixels in a grid of our mesh and the algorithm 5.

Let M be a matrix, $n \times m$ means that M has n rows and m columns.

Algorithm 5: Data accumulator definition

```
Function DefAccu(tab, tabool)
    Pre-condition:
    Execute tab, tabool←meshing(Nl,Nc) % algorithm 3
    Variables: acc: integers matrix
    if tabool[0]=0 and tabool[1]=0 then
        return zero matrix tab[0]×tab[2]
    else if tabool[0]=1 and tabool[1]=0 then
        return zero matrix (tab[0]+1)×tab[2]
    else if tabool[0]=0 and tabool[1]=1 then
        return zero matrix tab[0]×(tab[2]+1)
    else
        return zero matrix (tab[0]+1)×(tab[2]+1)
```

The algorithm 5 takes input on the parameters returned by the mesh function and returns a zero matrix that will be a data accumulator. It takes into account all the cases concerning the primality of the number of rows or columns of the image matrix. The number of rows and columns of the data accumulator depends on the rectangular mesh of Sect. 3.1 made on the image.

The straight lines in the data accumulator correspond to the maxima of this accumulation matrix, it is then necessary to determine the maxima of the accumulation matrix, and all the more so that it will allow us to represent the straight lines to be detected if necessary. To do this, the values of the accumulation matrix lower than a given threshold are canceled, the nonzero values representing the maxima of the data accumulator. The algorithm 6 determines the maxima of the accumulation matrix.

Algorithm 6: Maxima in the accumulation matrix

```
Data: seuil: integer, accu: matrix
Pre-condition: 0 < seuil
Variables: i, j, n1, n2: integers
n1 ← number of rows of accu;
n2 ← number of columns of accu;
for i between 0 and n1−1 with a step of 1 do
    for j between 0 and n2−1 with a step of 1 do
        if accu[i, j] < seuil then
            accu[i, j]←0;
```

The algorithm 7 updates the accumulation matrix by using standard Hough transformation.

Algorithm 7: Data accumulator update

Function *UpdateAccu(n, Nl, Nc, z, k: integer, img: image)*
  Pre-condition: Execute meshing function
  Variables: theta_max=π, $r = \sqrt{Nl^2 + Nc^2}$
  acc, r_dim, theta_dim←DefAccu(tab, tabool);
  for τ *between 0 and theta_dim-1 with a step of 1* **do**
    if *img[z,k]* $\neq n$ **then** Continue;
    theta← $\dfrac{\tau \times theta\_max}{theta\_dim}$; r← $z\cos(theta) + k\sin(theta)$;
    ir← $\dfrac{r\_dim \times r}{r\_max}$;
    ir←E(ir) ; τ ← E(τ) ;
    acc[ir, τ]←acc[ir, τ]+1 ;

With our method, the accumulator size is the same than zero matrix return by algorithm 5 instead of image size $Nl \times Nc$. So, we made Hough space size restriction. It follows, then, a reduction in the detection time of the straight lines because there will be less than dual pixels to compute.

Algorithm 8 performs straight-line recognition in calling the algorithms 5, 6, and 7.

4 Simulation and Results on Real Images

In this section, the meshing algorithm (algorithm 3) and the discrete line recognition algorithm with data accumulator (algorithm 8) are illustrated by an implementation of program in python using the OpenCV library applied to images pretreated with the Canny filter [3]. We used a laptop with the following characteristics:

- Processor: AMD A8-6410 APU with AMD Radeon RS Graphics 2.00GHz
- Ram memory utilizable: 4.94 GB
- Operating System: 64 bits, processor $\times 64$.

Results and analysis are illustrated by the following subsections.

4.1 Simulation with Pretreated Lena Image

We will detect the straight lines in Lena pretreated image with our method according to the parameter α and the value of threshold in the accumulator; the results will be presented in the tables. Figure 1 shows Lena pretreated image and some detected straight lines with the presented method.

We illustrated in Tables 1 and 2 the detection times of the straight lines in the Lena image without our method and with our method by varying the rate of lit pixels α in order to explain its role.

Algorithm 8: Digital Straight Line Recognition with data accumulator

```
Pre-condition: 0 < α < 1, 0<seuil
Data: α: float, n, seuil: integer, img: pretreated digital image
Variables: x, y, h, l, d,Nl, Nc: integers,
tab, tabool: table of integers,
acc: matix of integers,
A, C: tuple of 2 integers
begin
    Nl←number of rows in the image; Nc←number of columns in the image ;
    tab, tabool←meshing(Nl, Nc) ;
    h ←tab[0]×tab[1]; l ←tab[2]×tab[3];
    % the grids of the image without taking into account %any residues
    for x between 0 and h-tab[1]-1 with a step of tab[1] do
        for y between 0 and l-tab[3]-1 with a step of tab[3] do
            A=(x, y); C=(x+tab[1], y+tab[3]) ;
            if count(A, C, n, img)≥ α then
                for z between x and x+tab[1] with a step of 1 do
                    for k between y and y+tab[3] with a step of 1 do
                        UpdateAccu(n, Nl, Nc, z, k, img)
    % treatment of any residual pixels
    % x=h-1 and y=l-1 if tabool[0]=1 and tabool[1]=0 then
        % Nl ∈ ℙ there is one more line to process
        x← h;
        for y between 0 and l-1 with a step of 1 do
            UpdateAccu(n, Nl, Nc, x, y, img)
    if tabool[0]=0 and tabool[1]=1 then
        % Nc ∈ ℙ there is one more column to process
        y← l;
        for x between 0 and h-1 with a step of 1 do
            UpdateAccu(n, Nl, Nc, x, y, img)
    if tabool[0]=1 and tabool[1]=1 then
        x← h;
        % Nl ∈ ℙ, Nc ∈ ℙ there is one more line and
        % column to process
        for y between 0 and l with a step of 1 do
            UpdateAccu(n, Nl, Nc, x, y, img)
        y← l;
        for x between 0 and h-1 with a step of 1 do
            UpdateAccu(n, Nl, Nc, x, y, img)
    Algorithm 6 ;
    To trace straight lines (if desired);
```

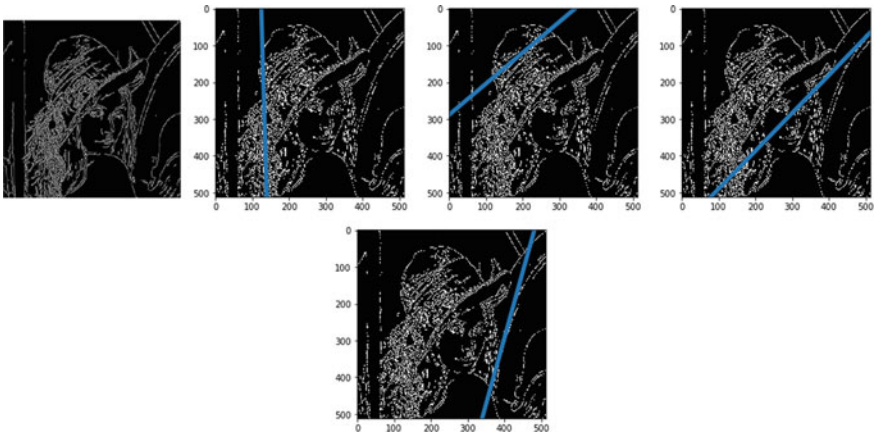


Fig. 1 Examples of detected lines $\alpha = 0.3$ and a threshold value of 150

Table 1 Results without our mesh

Thresh	Time (s)
150	289

Table 2 Results for $0 < \alpha < 1$ and threshold at 150

α	0.1	0.3	0.4	0.5	0.7
Time (s)	261	218	132	66	13.26

Table 3 Results for the variant threshold and $\alpha = 0.6$

Thresh	150	100	60	40
Time (s)	22	21	29	37

To recognize thin discrete lines, a weak value of the threshold is needed. Table 3 illustrates that for $\alpha = 0.6$.

4.2 Simulation Example with a Pentagon Image

In this subsection, the presented method is applied on a pentagon image (Fig. 2). The different results have been presented in Tables 4 and 5.

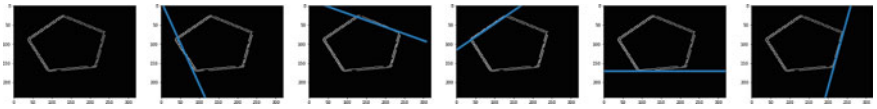


Fig. 2 A pentagon and it detected straight lines

Table 4 Results without our mesh

Thresh	Time (s)
50	12

Table 5 Results for $0 < \alpha < 1$ and threshold at 50

α	0.2	0.3	0.4
Time (s)	11	6	3

4.3 Simulation on Aerial Image

In this subsection, the presented method is applied on a GPS map image (Figs. 3 and 4), and the roads in the map are the straight lines. The different results have been presented in Tables 6 and 7.

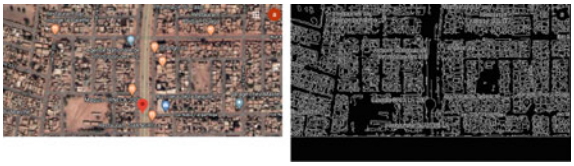


Fig. 3 GPS maps

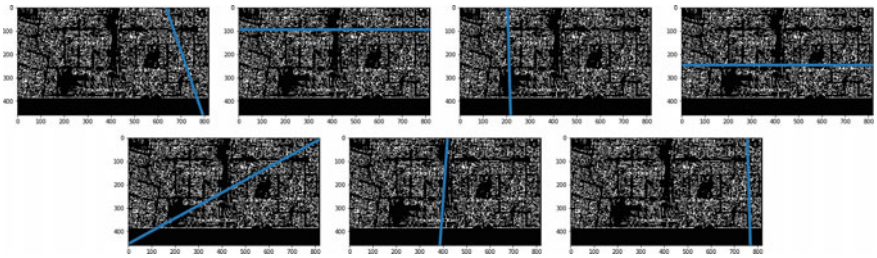


Fig. 4 Some detected straight lines (roads)

Table 6 Results without our mesh

Thresh	Time (s)
80	1403

Table 7 Results for $0 < \alpha < 1$ and threshold at 80

α	0.2	0.3	0.4
Time (s)	225	159	93

4.4 Results Analysis

The discrete straight-line recognition method presented and implemented with algorithm 8 has two great advantages:

- The use of the algorithm 3 in the function of recognition of the algorithm 8 generates additional computation time but to combine with the algorithm 4 makes it possible to compute the dual of a group of pixels and not to calculate for all pixels in the image, thus, to go faster than the usual method of recognition of forms or one calculates the dual of many more pixels. The Hough parameter space generated by this mesh has the smallest dimensions but contains the same lines as the initial image matrix and is therefore optimal in use;
- allows the recognition of both thin and thick discrete lines.

However, the detection of the lines depends on the choice of α and also of the threshold value. Tables 1, 2 and 3 show that with our method the straight lines are detected more quickly, in spite of the redundancy of calculation of the duals of selected connected rectangles. Straight-line detection can even be fastest by using faster factorization and primality algorithms.

5 Conclusion

We have presented rectangular meshing algorithm for straight line detection through Hough transform definition. We have also shown what using it with the rate parameter α allows fastest detection of discrete lines compared to the classical Hough standard transform.

As prospect we will work:

- to combine the proposed method with deep learning techniques to carry out discrete line detection in important quantities of images.
- to extend the proposed method to other geometric shape recognition.

Conflict of interest

The authors declare that, they have no conflict of interest.

References

1. Adleman, L.M., Pomerance, C., Rumely, R.S.: On distinguishing prime numbers from composite numbers. Ann. Math. Sec. Ser. **117**(1), 173–206 (1983)
2. Ballard, D.: Generalizing the Hough transform to detect arbitrary shapes. Patt. Recogn. **13**(2) (1981)
3. Canny, J.: A computational detection to edge detection. IEEE Trans. Patt. Anal. Mach. Intell. **8** (1986)
4. Dexet, M., Andres, E.: A generalized preimage for the digital analytical hyperplane recognition. Discr. Appl. Math. **157**(3), 476–489 (2009)
5. Duda, R., Hart, P.: Use of the Hough transform to detect lines and curves in pictures. Commun. ACM **15**(1), 11–15 (1972)
6. Ezome, T.: Test de primalité et de pseudo-primalité. Publications mathématiques de Besançon, pp. 89–106 (2013)
7. Hassanein, A.S., Mohammad, S., Sameer, M., Ragab, M.E.: A survey on Hough transform, theory, techniques and applications. Int. J. Comput. Sci. Issues **12**(1) 2 (2015)
8. Hough, P.V.C.: Method and means for recognizing complex patterns. United States Patent 3069654, 47–64 (1962)
9. Jung, C.R., Schramm, R.: Rectangle detection based on a windowed Hough transform. Institute of Electrical and Electronics Engineers (IEEE) (2004)
10. Liret, F.: Arithmétique, cours et exercices corrigés. Dunod (2011)
11. Reveilles, J.P.: Structures des droites discrètes. In Journées mathématique et informatique.Marseille-Luminy (1989)
12. Sere, A., Ouedraogo, F., Zerbo, B.: An improvement of the standard Hough transform method based on geometric shapes. In: Future of Information and Communication Conference (FICC), pp. 1–8. Singapore (2018)
13. Sere, A., Sie, O., Andres, E.: Extended standard Hough transform for analytical line recognition. (IJACSA) Int. J. Adv. Comput. Sci. Appl. **4(3)**, 256–266 (2013)

Performance Assessment of IDS Based on CICIDS-2017 Dataset

V. Priyanka and T. Gireesh Kumar

Abstract With the exponential growth of the internet among users worldwide, network engineers pose a great challenge in network security to identify intrusion activities. Intrusion Detection System (IDS) and Intrusion Prevention System (IPS) are the tools used to defend against these intrusion activities. IDS has sometimes been prone to false alarms. Therefore, to improve IDS, machine learning method is used. The largest number of IDS data sets are available till date where some are unreliable to use, whereas some are out of date and some does not cover common updated attacks. CICIDS-2017 data set overcome above major flaws [1]. Consequently, this paper assesses the performance of CICIDS-2017 data set by applying various machine learning algorithms such as Convolution Neural Network (CNN), Naive Bayes (NB) and Random Forest (RF), RF with highly ranked features, RF with feature reduction techniques (PCA and SVD). Then the comparison study is done which shows Random Forest gives good result when compared with other algorithms.

Keywords IDS · CICIDS-2017 · CNN · NB · Random forest · PCA · SVD

1 Introduction

Intrusion Detection System (IDS) used to detect and defend network attacks [2]. Network Intrusion Detection System (NIDS) and Host Intrusion Detection System (HIDS) are two types of IDS. NIDS [3] consist hardware sensor located along the network at many different points. It monitors inbound, outbound and local traffic [4]. HIDS monitors the system once after the system has been installed with IDS and responds by logging the malicious activity and reports the designated authority if

V. Priyanka (✉) · T. Gireesh Kumar
TIFAC-CORE in Cyber Security, Amrita School of Engineering, Amrita Vishwa Vidyapeetham, Coimbatore, India
e-mail: cb.en.p2cys18016@cb.students.amrita.edu

T. Gireesh Kumar
e-mail: t_gireeshkumart@cb.amrita.edu

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_58

any intrusion activity is detected. Signature-based IDS and Anomaly-based IDS are two techniques used to detect intrusions. Signature-based IDS uses pattern matching techniques to detect the attack. Anomaly-based IDS detects malicious activity by identifying deviations from the normal behaviour.

For an IDS to be trained, there should be an effective data set. There are many Intrusion detection data sets that had already been proposed by researchers claiming limited false alarms. But most of the data sets are out of date and unreliable [5]. CICIDS-2017 data set is provided by the Canadian Institute of CyberSecurity set. The data set is extracted using CICFlowmeter which covers all eleven necessary characteristics such as Heterogeneity, Attack Diversity, Feature Set, Complete Capture, Available Protocols, Complete Traffic, Labelling, Complete Interaction, Metadata, Complete Network configuration, Anonymity. It includes most common updated attacks such as Botnet, Portscan, Infiltration, SQL injection, XSS, Brute Force, DDoS, DOS.

The contribution of this paper Naive Bayes, CNN (Convolution Neural Network) Random Forest and PCA (Principal Component Analysis), SVD (Singular Value Decomposition) along with Random Forest are applied on CICIDS2017 data set to evaluate the performance of CICIDS-2017 data set. Comparing all the machine learning algorithm results and finally identifying the best machine learning algorithm that performs well on CICIDS-2017 dataset.

The rest of this paper is organized as follows. Section 2 discusses related works. Section 3 presents problem formulation. Section 4 presents proposed architecture. Section 5 provides performance evaluation. Section 6 discusses results and discussion. Conclusion and future work are discussed in Sect. 7.

2 Related Works

2.1 Towards Generating a New Intrusion Detection Data Set and Intrusion Traffic Characterization

This paper produces a reliable data set, namely CICIDS-2017 [6] which is completely labelled and contains benign and seven common updated attacks. This data set covers all eleven necessary criteria. This paper evaluates the performance of network traffic features and executes seven common machine learning algorithms to indicate the best set of features for detecting specific attack categories. Results show that RF algorithm gives high accuracy and short execution time. Then we compare the proposed data set with other data sets and found that no other data set covers all the eleven criteria.

2.2 *IDSGAN: Generative Adversarial Networks for Attack Generation Against Intrusion Detection*

This paper proposes IDSGAN which generate the adversarial attacks that deceive and evade the detection of defense systems [7]. NSL-KDD data set are used to design training set and testing set. Pitch, deep learning framework is adopted to implement IDSGAN. Multilayer Perceptron performs well. Future work will be applied IDSGAN in many more categories of intrusion attacks and increasing IDS robustness against IDSGAN.

2.3 *Flow-Based Network Traffic Generation Using Generative Adversarial Networks*

This paper proposes a novel method for generating realistic flow-based network traffic using Generative Adversarial Networks (GAN). GAN can only process continuous attributes whereas flow-based data contain both categorical and continuous attributes [8]. Therefore, three different preprocessing methods. First method is numeric transformation. Second method is binary transformation. Third method is embedding transformation. E-WGAN gives good results on evaluation.

2.4 *A Multi-level Intrusion Detection Method for Abnormal Network Behaviours*

This paper proposes multi-level network detection method [9]. Here NSL-KDD data set is used. Supervised learning algorithms such as SVM, Neural Network and Naive Bayes is used here. Future research of this work can be as follows. First, can be identifying possible outliers and to identify the abnormalities. Next compare the proposed work with different network intrusion data sets. Next, compare with unsupervised learning algorithms to determine effectiveness of our current work. Then conducting additional visual analysis. Finally, can apply proposed work on software applications to detect abnormality.

3 Problem Formulation

The number of threats getting increased as the internet is getting advanced. Therefore, there should be an Intrusion Detection System (IDS). Deep learning algorithms developed fast and better accuracy rate than traditional machine learning algorithms in recent years. To add on, generally Convolution Neural Network is most commonly

applied to analyze images and till date there is no work of CNN applied to CICIDS-2017 data set. Then ensemble algorithms such as Random Forest and Random Forest with feature reduction techniques such as PCA and SVD and then with traditional machine algorithm such as Naive Bayes are also applied on CICIDS-2017 data set. Performance are done for both 2 class and multi-class labels.

3.1 Naive Bayes

Naive Bayes introduced by Rev. Thomas Bayes is a supervised machine learning algorithm used for classification purposes [10] based on probability. Naive Bayes classifier assumes that each and every feature is independent. Gaussian Naive Bayes is a type of Naive Bayes algorithm used here.

3.2 Convolution Neural Network (CNN)

Convolution Neural Network [11], deep learning neural network introduced by Alex Waibel et al. CNN, like neural networks consists of neurons. The major role of CNN is feature extraction. By applying a relevant number of filters, convolution network could able to successfully capture spatial and temporal dependencies in the given data set. Then pooling takes place which reduces the dimensionality of the convolved data set. After going through the above process the output of the convolution layers is flattened into one-dimensional array and is fed to a fully connected layer, where the number of classes are mentioned. Classification is done using soft max activation functions.

3.3 Random Forest Classifier

Random Forest [12], proposed by Bremen is a supervised learning algorithm. Random Forest takes thousands of inputs and classifies them based on their importance. Random forest construct decision trees from the data set by selecting randomly. Each and every decision tree predicts results. Voting is performed. The predicted result with the maximum number of votes is chosen.

Dropping 12 features which has zero values when applied feature ranking and then random forest algorithm is applied to remaining 66 features. Principal Component Analysis (PCA) and Singular Value Decomposition (SVD) [13] are dimensionality reduction techniques used. PCA transforms dependent/correlated data into independent/uncorrelated data. PCA reduces the number of features so that the data set is left with more important features. SVD allows us to extract the information. SVD is

a method of breaking a matrix into three other matrices. After reducing features by applying feature reducing techniques such as PCD and SVD in the data set, random forest classifier is applied.

4 Proposed Architecture

Initially, preprocessing is done on CICIDS-2017 data set and then six various kinds of machine learning algorithms applied on both two-class labelled and multi-class labelled file shown in Fig. 1.

4.1 CICIDS-2017 Data Set

CICIDS-2017 data set, published by the Canadian Institute for cyber security contains benign and most updated attacks.

In detail, this data set all-together contains 2,830,743 which is about 8 files of both benign and attacks and each file contains 78 different features with label. In our paper 2 files are taken among 8 files, one file consists of two class labels such as benign and DDos and the other file consists of multi-class labels such as benign, DOS golden eye, DOS hulk, DOS slowhttptest, DOS slowloris, heartbleed for evaluating the performance of CICIDS-2017 data set.

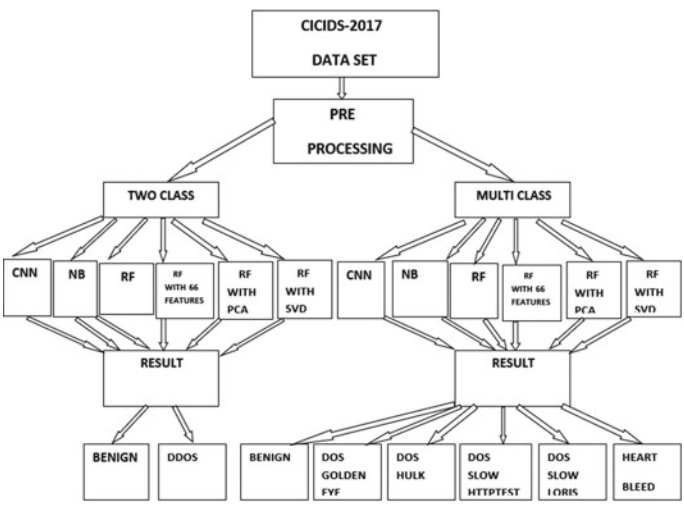


Fig. 1 Proposed architecture

4.2 Data Pre Processing

Data preprocessing involves removing null and infinity values and normalizing data. Here infinity values are replaced by null values and null values are dropped thereby removing infinity values. Labels are factorized and sorted. Then label is encoded using one hot encoding. One hot encoding is a mapping of categorical variables to binary vectors. Then features are normalized in the range between 0 and 1. Normalization equation called as minimum-maximum method is defined as:

$$x_{\text{normalized}} = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \tag{1}$$

$x_{\min}$ is minimum value of feature x
$x_{\max}$ is maximum value of feature x

Machine learning algorithms such as Convolution Neural Network (CNN), Naive Bayes (NB), Random Forest (RF), Random Forest with PCA, Random Forest with SVD are applied on CICIDS-2017 data set for both two-class and multi-class labelled file. Then the results are compared.

5 Results and Discussion

The implementation is done in Jupiter Notebook. Various machine learning algorithms applied on CICIDS-2017 data set and results are shown below (Tables 1, 2, 3, 4, 5, 6, 7 and 8).

Applying machine learning algorithms such as Naive Bayes, Convolution Neural Network, Random Forest classifier, PCA with Random Forest, SVD with Random Forest in two classes labelled file, we observed that Random Forest without feature reduction technique gives higher accuracy than other machine learning algorithms. Even though f1score, Recall and Precision is high in both CNN and Random Forest classifier without feature reduction, accuracy, true positive and true negative values of CNN is slightly less when compared with Random Forest classifier without feature reduction.

Table 1 Two class with feature reduction

Machine learning algorithms	True positive	True negative	False positive	False negative
Random forest with PCA	39,114	51,141	15	15
Random forest with SVD	39,054	51,183	17	31

Table 2 Two class with feature reduction

Machine learning algorithms	Label	Precision	Recall	F1 score	Accuracy
Random forest with PCA	Benign	1.00	1.00	1.00	0.999
	DDOS	1.00	1.00	1.00	
Random forest with SVD	Benign	1.00	1.00	1.00	0.999
	DDOS	1.00	1.00	1.00	

Table 3 Multi-class with feature reduction

ML algorithms	Label	Precision	Recall	F1 score	Accuracy
Random forest with PCA	Benign	1.00	1.00	1.00	0.998
	DOS GoldenEye	0.99	0.99	0.99	
	DOS Hulk	1.00	1.00	1.00	
	DOS slowhttptest	0.98	0.98	0.98	
	DOS slowloris	0.99	0.99	0.99	
	Heartbleed	1.00	1.00	1.00	
Random forest with SVD	Benign	1.00	1.00	1.00	0.998
	DOS GoldenEye	0.99	0.99	0.99	
	DOS Hulk	1.00	1.00	1.00	
	DOS slowhttptest	0.97	0.99	0.98	
	DOS slowloris	0.99	0.98	0.99	
	Heartbleed	1.00	0.80	0.89	

Applying various machine learning algorithms such as Naive Bayes, Convolution Neural Network, Random Forest classifier, Random Forest with PCA, Random Forest with SVD in multi-class labelled file, we observed that Random Forest without feature reduction technique give high Recall, f1 score, precision and accuracy when compared with other machine learning algorithms.

Table 4 Multi-class confusion matrix with feature reduction

ML algorithms	Label	Benign	DOS GoldenEye	DOS Hulk	DOS slow httptest	DOS slowloris	Heart bleed
Random forest with PCA	Benign	175,453	11	143	35	8	0
	DOS GoldenEye	39	3990	10	1	0	0
	DOS Hulk	86	16	92,174	0	1	0
	DOS slowhttptest	16	1	2	2182	18	0
	DOS slowloris	26	1	0	8	2339	0
	Heartbleed	0	0	0	0	0	3
Random forest with SVD	Benign	175,545	21	106	43	11	0
	DOS GoldenEye	37	4015	15	2	0	0
	DOS Hulk	75	5	92,165	0	0	0
	DOS slowhttptest	18	1	0	2136	13	0
	DOS slowloris	23	1	1	11	2314	0
	Heartbleed	1	0	0	0	0	4

Table 5 Two class without feature reduction

Machine learning algorithms	True positive	True negative	False positive	False negative
Naive Bayes	21,647	51,379	17,187	72
CNN	38,832	51,246	100	107
RF	38,963	51,308	3	11
RF with highly ranked features	39,096	51,176	5	8

Table 6 Two class without feature reduction

ML algorithms	Label	Precision	Recall	F1 score	Accuracy
Naive Bayes	Benign	1.00	0.56	0.71	0.81
	DDOS	0.75	1.00	0.86	
CNN	Benign	1.00	1.00	1.00	0.997
	DDOS	1.00	1.00	1.00	
Random forest	Benign	1.00	1.00	1.00	0.999
	DDOS	1.00	1.00	1.00	
RF with highly ranked features	Benign	1.00	1.00	1.00	0.999
	DDOS	1.00	1.00	1.00	

Table 7 Multi-class without feature reduction

ML algorithms	Label	Precision	Recall	F1 Score	Accuracy
Naive Bayes	Benign	0.99	0.26	0.42	
	DOS GoldenEye	0.09	0.95	0.16	
	DOS Hulk	0.52	0.96	0.68	0.51
	DOS slowhttptest	0.41	0.58	0.48	
	DOS slowloris	0.08	0.52	0.14	
	Heartbleed	1.00	1.00	1.00	
Convolution neural network	Benign	1.00	0.99	1.00	
	DOS GoldenEye	0.98	0.97	0.98	
	DOS Hulk	0.99	1.00	0.99	0.993
	DOS slowhttptest	0.94	0.98	0.96	
	DOS slowloris	0.96	0.90	0.93	
	Heartbleed	0.93	0.97	0.99	
Random forest	Benign	1.00	1.00	1.00	
	DOS GoldenEye	1.00	0.99	1.00	
	DOS Hulk	1.00	1.00	1.00	0.999
	DOS slowhttptest	1.00	0.99	0.99	
	DOS slowloris	0.99	1.00	1.00	
	Heartbleed	1.00	1.00	1.00	
RF with highly ranked features	Benign	1.00	1.00	1.00	
	DOS GoldenEye	1.00	1.00	1.00	
	DOS Hulk	1.00	1.00	1.00	0.999
	DOS slowhttptest	0.99	0.99	0.99	
	DOS slowloris	0.99	0.99	0.99	
	Heartbleed	1.00	1.00	1.00	

Table 8 Multi-class confusion matrix without feature reduction

ML algorithms	Label	Benign	DOS GoldenEye	DOS Hulk	DOS slow httptest	DOS slowloris	Heart bleed
Naive Bayes	Benign	45093	12304	3142	84392	30959	0
	DOS GoldenEye	7	1184	24	313	802	0
	DOS Hulk	63	459	1287	126	282	0
	DOS slowhttptest	223	241	14	89046	2464	0
	DOS slowloris	2	131	33	79	3890	0
	Heartbleed	1	0	0	0	0	2
Convolution neural network	Benign	174,639	37	847	78	81	0
	DOS GoldenEye	110	4042	3	2	0	0
	DOS Hulk	253	49	91750	0	5	0
	DOS slowhttptest	28	2	0	2215	9	0
	DOS slowloris	185	2	0	49	2172	0
	Heartbleed	5	0	0	0	0	0
Random forest	Benign	176,105	1	89	3	5	0
	DOS GoldenEye	7	4164	6	3	0	0
	DOS Hulk	41	13	92,181	0	0	0
	DOS slowhttptest	7	0	0	2248	7	0
	DOS slowloris	7	0	1	2	2302	0
	Heartbleed	0	0	0	0	0	2
RF with highly ranked features	Benign	175,812	0	66	4	6	0
	DOS GoldenEye	5	4090	7	5	0	0
	DOS Hulk	36	1	91,984	0	0	0
	DOS slowhttptest	6	1	0	2249	8	0
	DOS slowloris	12	1	0	6	2260	0
	Heartbleed	0	0	0	0	0	4

6 Conclusion

For assessing performance of CICIDS-2017 dataset, various classification machine learning algorithms are applied to CICIDS-2017 data set and comparison study is done in this paper. Naïve Bayes did not perform well as accuracy for two class and multi-class is less. CNN performed well, but Random Forest has a high accuracy rate when compared with CNN. When Random Forest without feature reduction and Random Forest with feature reduction (PCA and SVD) applied on CICIDS-2017 data set, Random Forest without feature reduction shows the best result in terms of accuracy, precision, recall and f1 score as accuracy is higher in Random Forest without feature reduction when compared with feature reduction. In the future work, various other types of machine learning algorithms can be applied to the CICIDS-2017 data set. Same experiment can be performed in real-time environment and extracting features and apply machine learning algorithms. The limitation is that not many numbers of novel attacks included and this work is not performed in real-time environment.

References

1. Gharib, A., Sharafaldin, I., Lashkari, A.H., Ghorbani, A.A.: An evaluation framework for intrusion detection dataset. In: International Conference on Information Science and Security (ICISS), pp. 1–6 (2016)
2. Jose, et al.: Gigabit network intrusion detection system using extended bloom filter in reconfigurable hardware. In: Proceedings of the Second International Conference on Computer and Communication Technologies, 11–19 (2016)
3. Anuraj, et al.: High speed network intrusion detection system using FPGA. In: Proceedings of the Second International Conference on Computer and Communication Technologies, 187–194 (2016)
4. Kumar, G.: Network intrusion detection system based on machine learning algorithms
5. Intrusion Detection Evaluation Dataset (CICIDS2017) (2017) https://www.unb.ca/cic/datasets/ids-2017.html
6. Sharafaldin, I., Lashkari, A.H., Ghorbani, A.A.: Toward Generating a New Intrusion Detection Dataset and Intrusion Traffic Characterization. (ICISSP), pp. 108–116 (2018)
7. Lin, Z., Shi, Y., Xue, Z.: Idsgan: Generative Adversarial Networks for Attack Generation Against Intrusion Detection.(arXiv preprint arXiv:1809.02077)(2018)
8. Ring, M., Schlör, D., Landes, D., Hotho, A.: Flow-Based Network Traffic Generation Using Generative Adversarial Networks, pp. 156–172. Elsevier (2019)
9. Ji, S.-Y., Jeong, B.-K., Choi, S., Jeong, D.H.: A Multi-level Intrusion Detection Method for Abnormal Network Behaviors, pp. 9–17. Elsevier (2019)
10. Frank, E., Bouckaert, R.R.: Naive Bayes for Text Classification with Unbalanced Classes, pp. 503–510. Springer (2006)
11. Kim, J., Shin, Y., Choi, E., et al.: An intrusion detection model based on a convolutional neural network. J. Multimedia Inf. Syst. 165–172 (2019)
12. Anand, et al.: An ensemble approach for algorithmically generated domain name detection using statistical and lexical analysis. Procedia Comput. Sci. **1129–1136** (2020)
13. Tanwar, S., Ramani, T., Tyagi, S.: Dimensionality Reduction Using pca and svd in Big Data: A Comparative Case Study, pp. 116–125. Springer (2017)

Identifying and Mitigating Vulnerabilities of Hardened Windows Operating System

M. Sreerag, M. Sethumadhavan, and P. P. Amritha

Abstract Microsoft Windows 10 is one of the most popular operating systems. So, there should be proper security measures to protect personal data and unauthorized access to the systems. As per HIPAA, patient data within the systems need to be prevented from unauthorized access. System hardening is a way to reduce security risk and is done by expelling and blocking access to all insignificant applications and framework utilities from the system. Antivirus software provides protection from malware, but there will be configuration flaws that could leave systems exposed to attacks. In this paper, vulnerabilities due to insufficient security configurability, security misconfiguration of hardened windows system, recommendations for those vulnerabilities and an approach to reproducing those vulnerabilities are explained. This approach can provide better security to Windows OS.

Keywords Applocker · Unified write filter · Windows hardening

1 Introduction

In this Internet age, most of the computers offer network security to resist the unknown access to the system. Vulnerability analysis has to be done for the security of systems. In [1, 2], the authors have done vulnerability analysis and fixing of it on different networks. Even studies have been done for comparing the vulnerabilities in open- and non-open-source systems [3]. In this paper, we are concentrating on Windows 10 operating systems. Even though antivirus software and other secu-

M. Sreerag (✉) · M. Sethumadhavan · P. P. Amritha
TIFAC-CORE in Cyber Security, Amrita School of Engineering,
Amrita Vishwa Vidyapeetham, Coimbatore, India
e-mail: cb.en.p2cys18023@cb.students.amrita.edu

M. Sethumadhavan
e-mail: m_sethu@cb.amrita.edu

P. P. Amritha
e-mail: pp_amritha@cb.amrita.edu

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_59

rity tools are there which will be preventing from the malicious content to run on the system, still systems are being vulnerable to external attacks [4]. In every field, protection of the systems from attackers is required. One of the main fields which we need proper security measures is the healthcare sector. According to the Health Insurance Portability and Accountability Act (HIPAA), it is mandatory to ensure the confidentiality, integrity and availability of all the patient data. So, HIPAA privacy rule is protecting the privacy of individually identifiable health information. Most of the medical devices in hospitals will be connected to a computer system that will be dedicated to that device. So, all the data management will be done using that dedicated system only. If it is not secured properly an attacker can easily misuse the system as well as the data which are processed. All mainstream modern operating systems will be designed to be secure by default. But, on maximum systems, we are able to add greater safety functions and can modify the default configuration settings to make the machine less at risk of the attacks than it would be with a default setup.

OS hardening is very necessary for things wherever a system faces above-average security risks, like an online server that is exposed to the public Internet or a data server that contains data that is subject to strict regulatory privacy needs [5]. However, given the high rate of cyberattacks nowadays, OS hardening will be the best option even in cases where servers or data face only average security risks. The time it takes to harden an OS is often well worthwhile. Windows is an operating system family containing different proprietary operating systems that are developed by Microsoft [6]. Microsoft itself will be providing many security features and tools to their customers to get rid of the attacks and attackers even though the customers are not utilizing those features. Most of the users think by blocking the direct way to access the data and modify the configuration which will make the system secure. But there are multiple ways of getting access to the same data and configuration settings. Here, we will be discussing some of the hardened system flaws which we identified while doing penetration testing on a hardened Windows operating system.

2 Hardening Windows System

There are multiple ways in which we are able to harden a Windows operating system [7].

2.1 Group Policy

Microsoft provides a feature called Group Policy for its Windows NT family of the operating system. This can be used to control the working environment of user accounts and computer accounts. It also provides the centralized configuration and management of the operating system, user's settings and applications in an Active Directory environment. Mentioned below are some of the policies that can be done

using group policy configuration. There are many more things that we can do using Group Policy here mentioning only a few of them [8].

Hide mechanisms to remove zone information This Policy helps to manage whether users can manually remove the zone information from the saved file attachments. Removing the zone information allows users to open potentially dangerous file attachments which Windows has blocked users from opening.

Registry policy processing Checks whether the registry policies are updated. This policy will affect all the policies in the Administrative Templates folder and any other policies which will be storing values in the registry.

Use Certificate Rules on Windows Executables for Software Restriction Policies This policy setting will check if digital certificates are processed when software restriction policies are enabled and a user or a process attempts to run the software with a .exe file name extension.

2.2 Bitlocker

Bitlocker is a fast and secure way to encrypt the files inside the windows system. Encrypting the entire drive protects against unauthorized changes to our system. We will be having different partitions in our system (C, D, E, etc.), and it is possible to use Bitlocker for encrypting specific drives also. Hence, even if the encrypted information is stolen, it would be unusable [9].

2.3 Firewall

Windows firewall is a built-in network security system and will be preventing unauthorized access to and from our private networks. We can also provide specific rules to whitelist only our own application traffic and blocking the rest of the traffic.

2.4 User Accounts

In a Windows system, there can be multiple accounts that too with multiple types. The three major account types with different levels of permissions are administrator account, standard account and guest account. Administrator account comes under the administrator group and has the right to install software and modify the system configuration. Standard user account users will have limited access to administrator features, and guest account will have a blank password. So, an attacker can easily get in to a guest account without providing any credentials. Thus, considering that

the guest account offers secret access to your computer and suppose if proper care is not taken the account running as guest may allow attacker to gain system access and later on may lead to privilege escalation. Disabling guest account will reduce security risk, and it is a best practice.

2.5 AppLocker

AppLocker helps us by controlling which apps and files users can run. These will include executable files, scripts, Windows Installer files, dynamic-link libraries (DLLs), packaged apps and packaged app installers. AppLocker can also define the rules based on file attributes that persist across app updates, such as the publisher name (derived from the digital signature), product name, file name and file version. We can also create rules based on the file path and hash. Assigning a rule for a security group as well as an individual user is also possible [10].

2.6 Registry Edit

The Windows Registry is a central Windows operating system archive. Settings and details that govern hardware configuration, operating system configuration, software configuration installed and user preferences are all stored in the registry.

3 Vulnerabilities on Hardened Windows System

During the security assessment on some of the hardened window systems, we were able to find several interesting ways by which an attacker could infiltrate into the system. Once the attacker is inside the system, he will be able to execute commands as an "Administrator" user or the privilege user with what the application is running. Then, the complete system can be compromised. Different possible scenarios that we found to infiltrate the system are described below.

3.1 Event Viewer

During the security assessment of a thick client application that was installed in a hardened windows system, we were able to get a reverse shell with limited user privileges on the target system by abusing Windows Event Viewer. We were also able to further escalate our privileges on the system to Administrator and "NT Authority System" via different vectors that are discussed below in "Steps done" section. As

a system user, we were able to execute all operating system commands with full privileges on the target box. For exploiting this vulnerability, the attacker has to be connected to the same network and need physical access for a brief period. The impact on this vulnerability will be the compromising of complete system.

Steps Done:

1. Create a PowerShell reverse shell script and copy it to a pen drive.
2. Plugged in the pen drive to system, navigate through the application to get event viewer. Event Viewer −− Windows Logs −− Application, and then selected one of the events in the log.
3. Clicked on "Save All Events As..." option in the right panel and copy the reverse shell code into "I" drive. ("I" is an FTP Drive. Several entry points were found to access I. It will be detailed after the Steps done section)
4. After copying, clicked over "System" events in the left pane and filtered the logs for event id "20003."
5. Selected one instance in the log and click on "Attach Task To This Event" option in the right pane.
6. Choosed to "Start a Program" and provide "PowerShell" in "Program/script" section.
7. Given the following command in the "Add arguments" box:
   -NoLogo -NonInteractive -nop -ep bypass -File "I:.ps1
8. In attacker machine the netcat listener should be there to listen for incoming connections: *Nc -nlvm 9077*.
9. Complete scheduled task creation on the device by clicking the "Finish" button.
10. Plugin a new pen drive into the device, wait for a while and you should get a PowerShell reverse shell connection on the attacker machine.

We observed several entry points through which you could copy any kinds of files into the FTP drive on the device. This is essential for an attacker to copy his reverse shell code. The methods by which we enumerated are listed below:

- "Save All Events As...", "Save Selected Events..." and "Save All Events in Custom View..." option in Event Viewer.
- Event Properties – Event Log Online Help – Read our privacy statement online – Internet Explorer – Select text and click "Ctrl + U" – File – Save
- Help – Event Log Online Help – Read our privacy statement online – Internet Explorer – Select text and click "Ctrl + U" – File – Save
- Select any event – In the details pane – General – Event Log Online Help – Read our privacy statement online – Internet Explorer – Select text and click "Ctrl + U" – File – Save
- Right click over any of event log files in the left pane – clear log – save & clear
- Right click over any of event log files in the left pane – properties – clear log – save & clear
- Right click over any of event log files in the left pane – help – event log online help – Read our privacy statement online – Internet Explorer – Select text and click "Ctrl + U" – File – Save

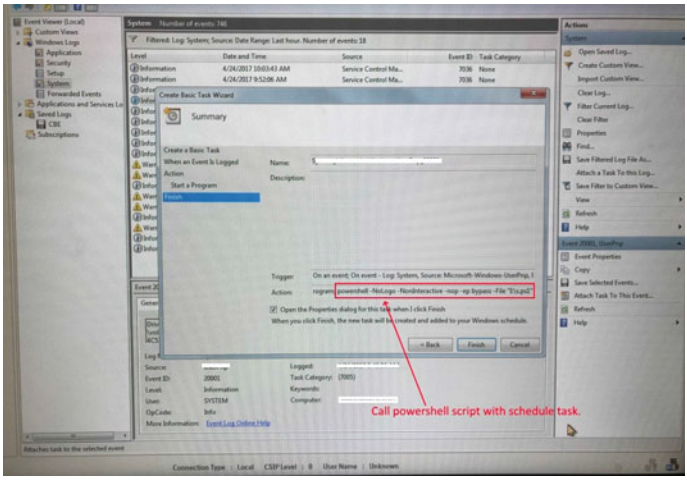


Fig. 1 Complete file system accessible via rescue toolkit network boot

- Action – Help – event log online help – Internet Explorer – Select text and click "Ctrl + U" – File – Save
- Windows Keyboard shortcut: Windows Key + L , Alt + Shift + PrintScreen , Ctrl + Alt + Delete (Fig. 1)

Recommendation:
It is recommended to securely configure Windows Event Viewer to mitigate command execution issue.

3.2 Windows HotKeys

We found that the following Windows hotkeys were not filtered on the system and an attacker could use any of these to infiltrate.

- Alt + Shift +PrintScreen
- Ctrl + Alt + Del
- Filter and Toggle keys shortcut (Press down "Shift" key for 8 s or "Num" key for 5 s).

Hotkey "Alt + Shift +PrintScreen" will bring up the "HighContrast" dialog box, and you will be able to navigate to the control panel by clicking over the "Ease of Access" link within the dialog box.

Hotkey "Ctrl+Alt+Del" will take you to the GINA screen in Windows 7. Although you do not have many options there, "Ease of Access" utility is available. The "Magnifier" and "Narrator" tools could be made use of to infiltrate. Select "Magnifier" or "Narrator" tool and click "Apply" Ensure the tool's window is visible on screen.

Now, click the "Cancel" button in GINA. Even though you are back in the application, you will observe that the "Narrator"/"Magnifier" window is still there. Navigating to "Help" menu in Narrator/Magnifier, you will be able to get to the Windows explorer

Holding down the "Shift" key for 8 s (or "Num" key for 5 s) while on the screen that shows the application version (screen just after operating system boot up and before application launch) would pop up the Filter keys (or Toggle keys) menu. You could navigate to the Windows control panel from the link inside the dialog box. Pressing "F8" while the application is being loaded would bring up the "Field Service Login" box. The aforementioned filter/toggle keys hotkey would work there as well.

In another application, it was having a service window that could be opened with the hotkey combination "Ctrl + Alt + S." We Navigated to "Utility" tab and open "Event Viewer." We got plethora of options within Event Viewer to open up Windows explorer window.

Recommendation:

Disable all unwanted windows Hotkeys or override with your custom implementation for same shorcuts.

3.3 BIOS Password Bypass

Even though the system BIOS is protected with a password, the system can use network boot without the need for a password. The primary Ethernet card within application was configured to obtain an IP automatically via DHCP (secondary card can be configured via the application to use DHCP). We configured a PXE server on our test laptop and connected it with a device having the application via a crossover cable (on primary Ethernet port). On restart, we were able to make system boot from the network.

Recommendation:

It is recommended to disable network-based booting or use Trusted Platform Module (TPM) (Fig. 2).

3.4 Security Misconfigurations

During the security assessment of the product, it was observed that the application accounts (Administrator, Standard) could be modified by any user in the system. Moreover, by logging in as an emergency user (Guest User), an attacker could delete, add/modify other accounts in the system. Sensitive system information like hotfixes installed, services running on the target system, internal and exposed ports, WiFi profiles, system and application usernames and roles, etc., can be exported by an attacker using emergency mode of operation and also in scenarios where there are no application accounts configured on the system. Attacker would require brief physical

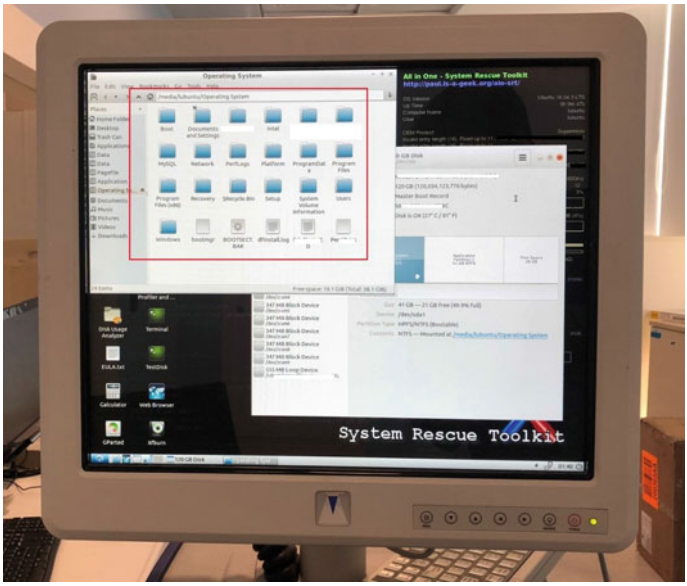


Fig. 2 Complete file system accessible via rescue toolkit network boot

access to the system to gather all sensitive details. Exploitation does not necessarily need physical access. Attacker who successfully exploits this vulnerability could modify or delete patient information and scan details on the target system.

Recommendations:

It is recommended to configure the system in such a way that:

- During installation phase, the end user/field service agent should be forced to create one application "Administration" account on the target.
- The default configuration of the system should enforce authentication-based login.
- In guest mode, user account management features should be disabled.

3.5 Privilege Escalation

Once inside the system, normal user has permission to modify most of the files present on the system (except one's in folders restricted by Windows in system drive). We found a myriad of vectors through which an attacker can escalate his privilege once inside the system. The techniques that we found to be successful on the target are listed below.

1. **Privilege escalation by modifying windows batch files(*.bat)/ PowerShell scripts (*.ps1)**:
   Issue: The target system uses a lot of batch files (and PowerShell scripts) to setup the environment at boot time. These files are modifiable by normal user. At boot time, these scripts are run with higher user privilege, and by just modifying these files, we were able to escalate privilege to administrator/system account.
   Mitigation: Normal user should not be given permissions to modify any of the batch or PowerShell scripts that are run through higher privilege accounts.
2. **Privilege escalation by modifying configuration files**:
   Issue: Several of the applications within the system loads configuration from *.xml files. Many of these files have hardcoded paths to other executables. By replacing executable/dll path in these configuration files, an attacker will be able to execute his own malicious code and gain administrator/system privileges.
   Mitigation: Normal user should not be given permissions to modify any of the configuration files used by other apps.
3. **Privilege escalation via dll preloading**:
   Issue: Most of the dll's used by different application/services in the system are not digitally signed. Since the normal user has permission to access/modify contents in most of the drives/directories where applications reside, we were able to replace several *.dll files in different applications with malicious dll's to escalate our privileges on the target system.
   Mitigation: Digitally sign *.dll files and verify signature before loading them by the application. Also, restrict permissions on these files so that a clinical user will not be able to modify them. Remove all *.dll reference in the code, for which actual dll's does not exist. Specify fully qualified path for dll's.
4. **Privilege escalation via misconfigured services**:
   Issue: Normal user can replace many of the Windows service executables used by different applications in the system. All of these services are run with accounts having high privileges. An attacker could replace these service executables with his own code to achieve privilege escalation.
   Mitigation: Properly assign permission to these executables, so that a low privilege user will not be able to replace or modify them.

4 Conclusion

During our assessment, it was observed that due to weak security configuration, the application was at risk and eventually leading to violation of confidentiality–integrity–availability. From the vulnerabilities that were found we recommend the product team to implement the relevant fixes. Based on the feasibility of available solutions, teams can choose the best possible solution if multiple options are available. The outcome of these assessments are, there are ways where hardening is done; however, there are chances where developers still miss out due to security vs usability trade-off, such has to be addressed in product risk assessment with the presence of physical/external network security.

References

1. Rajesh, M., Sreevidya, B.: Vulnerability analysis of real-time operating systems for wireless sensor networks. In: Advanced Computing and Intelligent Engineering, pp. 449–460. Springer (2020)
2. Kumar, B.K., Raj, N., Dhivvya, J.P., Muralidharan, D.: Fixing network security vulnerabilities in local area network. In: International Conference on Trends in Electronics and Informatics, pp. 1349–1354. IEEE (2019)
3. Achuthan, K., SudhaRavi, S., Kumar, R., Raman, R.: Security vulnerabilities in open source projects: an India perspective. In: International Conference on Information and Communication Technology, pp. 18–23. IEEE (2014)
4. Parimi, M.R.: Critical analysis of software vulnerabilities through data analytics. In: International Conference on Industrial Engineering and Operations Management, IEOM Society, pp. 923–934 (2020)
5. Zamora, P.M., Kwiatek, M., Bippus, V.N., Elejalde, E.C.: Increasing windows security by hardening PC configurations. In: EPJ Web of Conferences, vol. 214, p. 08019. EDP Sciences (2019)
6. Security Strategies in Windows Platforms and Applications, Michael G. Solomon
7. The Windows Server Hardening Checklist, https://www.upguard.com/blog/the-windows-server-hardening-checklist
8. Hardening GPO, https://social.technet.microsoft.com/Forums/en-US/2a7da18f-cfb6-4fbf-98ab-cc41a1c67da6/hardening-gpo. Last accessed 6 Jan. 2020
9. Security Best Practices for Your Windows 10 Computer, https://www.securicy.com/blog/security-best-practices-hardening-windows-10/. Last accessed 7 Jan. 2020
10. Windows AppLocker, https://docs.microsoft.com/en-us/windows/security/threat-protection/windows-defender-application-control/applocker/applocker-overview

Optimized Convolution Neural Network (OCNN) for Voice-Based Sign Language Recognition: Optimization and Regularization

Suman Kumar Swarnkar, Asha Ambhaikar, Virendra Kumar Swarnkar, and Upasana Sinha

Abstract Human–computer interactions (HCIs) are very helpful for current technology world. Later, the trendy data policies restrict the instinctive nature and velocity of HCI use, and the sign language acknowledgment framework has multiplied many significances. Completely one of a kind sign language is frequently accustomed to specific intentions and intonations or for dominant units like domestic robots. In previous work, ICNN was once in contrast to a baseline CNN and in preceding work OCNN gain 99.96% recognition rate. The foremost center of attention of this evaluation is to enhance the loss and execution time of ICNN model by using and optimizing the preference for alternate OCNN. New model is improving configuration by applying optimizing function in bottom layer. Due to the fact it is previously acknowledged, one optimizer will never give higher accuracy for all situation. The desire for the optimizer is to be created through thinking about the variability of facts and, consequently, the nonlinearity degree of the connection designs that happen inside the facts. As a result of the theoretical calculation, it is not always sufficient to work out the easiest optimization function. The study is required to visualize the complete optimization function of model that definitely suits the behaviors of the facts being examined. Gradient descent method is well known for its convenient computing and step of coming together on real-time data. This research paper compared three optimization types of gradient descent method, i.e., stochastic gradient descent algorithm, adaptive gradient algorithm, root mean square propagation algorithm. From the test result, it is recognized that RMSProp is that the first-class optimizer to preserve the loss of model performs of OCNN capacity in managing sign language recognition.

Keywords Convolutional neural network · Optimization · SDG · RMSProp · ADAGrad · Regularization · Deep learning · Sign language recognition

S. K. Swarnkar (✉) · A. Ambhaikar
Kalinga University, Raipur, India

V. K. Swarnkar
Bharti College of Engineering and Technology, Durg, India

U. Sinha
JK Institute of Engineering Bilaspur, Bilaspur, India

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_60

1 Introduction

Gradient descent should be a universal method for finding the nearest minimal cost. In extraordinary CNN-based algorithms, gradient descent algorithm works to decrease error rate for the duration of the training technique and internal configuration [1]. In general, the optimization method is labeled as first and second order of optimization rule. Optimization rule order category is having important role in gradient descent algorithm. The first order spinoff affords records related to the path of the error function at a particular purpose, whether or not it is developing or reducing. Here, the experiment is employed to inform that downward mapping of error method will do until it will not reach to native minimal [2]. It is verified to be ineffective for batch-wise gradient descent algorithm in the ordinary approach. To furnish unique replace in the internal configuration, it computes the whole training facts gradient (that is in huge amount). This approach makes your model very slow. To overcome the weaknesses of such ordinary methodology, many algorithms had been established as follows: stochastic gradient descent (SGD), adaptive gradient (ADAGrad), and root mean square propagation (RMSProp) [3].

2 Previous Improved CNN Architecture: ICNN

In previous work, CNN structure was improved by using tuning technique and adaptation strategies utilized to the CNN structure are discussed [4]. The complete CNN performance used to be increased utilizing parameters tuning to include public initialization, regularization, and information augmentation [5].

2.1 Data-Augmentation-Technique

For extra wonderful performance, learning algorithm is used to the huge data. And as soon as the information is of reduction quality, the algorithms will simply operate higher, as long as beneficial facts will be extracted with the aid of the model from the unique dataset [6]. Model will increase the variation of training data by transforming the base data using data augmentation. The converted snapshots are generally twisted from the unique snapshots with very slight computations and are produced throughout the training data [7].

2.2 Network Initialization

Terrifically, initializing the weights can drastically impact how without problems the community learns from the training. The ICNN structure makes use of the unchanging He_initialization (he_uniform) for all ReLU layers and the uniform_Xavier_initialization (glorot_uniform) for the output softmax layer (for advantageous simplification of the logistic characteristic for more than one input) [8].

2.3 L2-Regularization-Technique

The objective of this technique is to minimize the complication of the model while preserving the identical parameter count. L2-regularization does via refining weights with massive amounts, utilizing and minimizing their L2-regularization. It makes use of a hyperparameter $\lambda = 0.0001$ to require the absolute significance of using the L2-norm to reduce the loss on the training time [9].

3 Proposed Optimized CNN Architecture: OCNN

This optimized model, OCNN, is proven in Fig. 1 The loss (training and Validation) of the previous ICNN was once increased by way of exceptional optimization in convolutional layer which consists of community initialization and regularization. The proposed strategy additionally consists of records augmentation [10] (Figs. 2, 3, 4, 5, 6 and 7).

4 Experimental Methodology

The convolutional-NN is the enhancement supported with the gradient-descent technique. Once classification result was carried out and in contrast by making three one of a kind fashions. Different type of optimizer like The SGD for model-1 (M1), RMSProp for model-2 (M2), and Adagrad for model-3 (M3) are used for improving, where the categorical cross-entropy is used as cost functions. The parameters of every model are defined in Table 1, for enhancement of cost function model discovering network from the last [11].

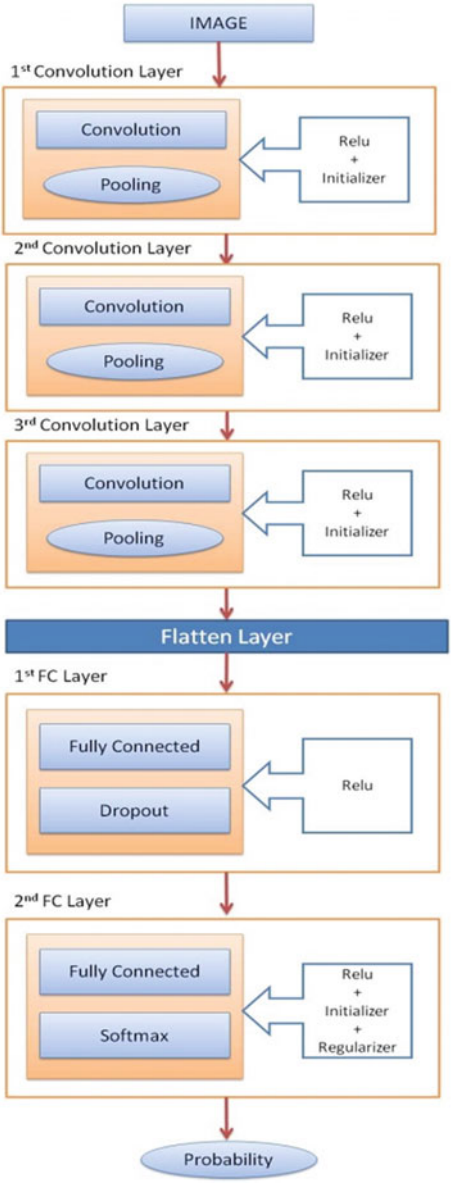


Fig. 1 Architectural diagram of the OCNN

5 Conclusion

It is proven from the scan that all optimizers exhibit one of a kind overall performance throughout distinct OCNN architectures. However, all optimizer operates constant accuracy. Among three optimizers, RMSProp offers faster education on the OCNN

Fig. 2 Training time results analysis of Model 1, Model 2, and Model 3

Fig. 3 Training loss results analysis of Model 1, Model 2, and Model 3

Fig. 4 Training accuracy results analysis of Model 1, Model 2, and Model 3

design. SGD additionally suggests the flexibility to take care of the training time nevertheless excessive inside the subsequent batch as soon as it reaches the minimal on precise batch size. On the contrary hand, the Adagrad had been unsuccessful also as RMSProp conjointly provides higher lead to training loss and validation loss in

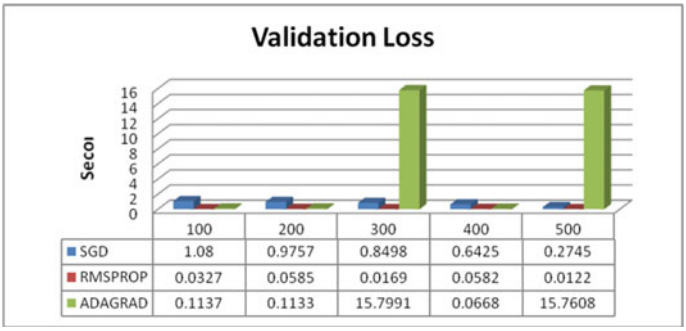


Fig. 5 Validation loss results analysis of Model 1, Model 2, and Model 3

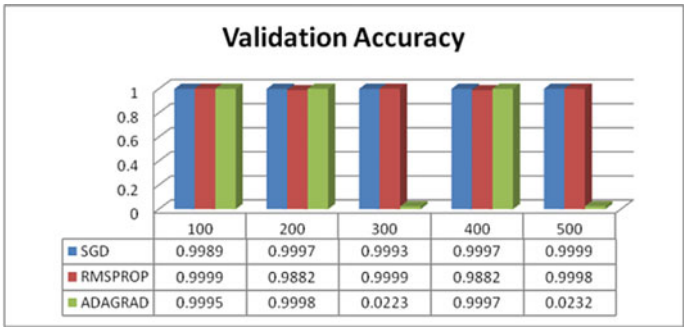


Fig. 6 Validation accuracy results analysis of Model 1, Model 2, and Model 3

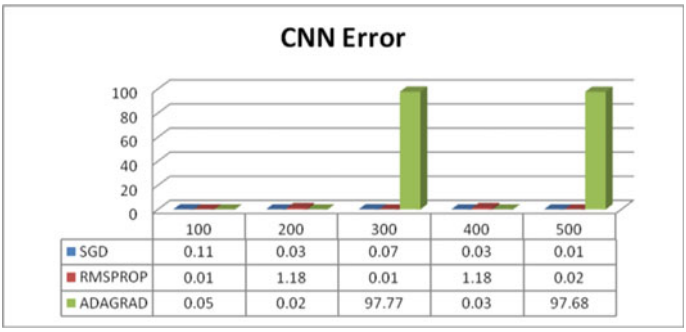


Fig. 7 CNN error results in the analysis of Model 1, Model 2, and Model 3

evaluating SGD. As per CNN error result, SGD and RMSProp provide greater result than Adagrad.

Table 1 Parameters of model with different network

Network	SGD (M1)	RMSProp (M2)	Adagrad (M3)
Number of training samples	88,000	88,000	88,000
Activation function	RELU-softmax	RELU-softmax	RELU-softmax
Learning rate	0.01	0.01	0.01
Epoch	5	5	5
Cost function	Categorial cross-entropy	Categorial cross-entropy	Categorial cross-entropy
Optimization	Adam	Adam	Adam
Data augumentation	Yes	Yes	Yes
Network intialization	ReUL (he_uniform) Softmax (glorot_uniform)	ReUL (he_uniform) Softmax (glorot_uniform)	ReUL (he_uniform) Softmax (glorot_uniform)
Regularization	L2	L2	L2

References

1. Li, G., Tang, H., Sun, Y., et al.: Hand gesture recognition based on convolution neural network. Cluster Comput. **22**(S2), 2719–2729 (2019). https://doi.org/10.1007/s10586-017-1435-x
2. Rao, G.A., et al.: Deep convolutional neural networks for sign language recognition. In: Conference on Signal Processing and Communication Engineering Systems (SPACES), vol. 2018, pp. 194–197 (2018)
3. Huang, J., et al.: Sign language recognition using 3D convolutional neural networks. IEEE, 1–6 (2015)
4. Flores, C.J.L., et al.: Application of convolutional neural networks for static hand gestures recognition under different invariant features. In: International Conference on Electronics, Electrical Engineering and Computing, pp. 1–4 (2017)
5. Masood, S., Srivastava, A., Thuwal, H.C., et al.: Real-time sign language gesture (word) recognition from video sequences using CNN and RNN. In: Advances in Intelligent Systems and Computing, pp. 623–632 (2018)
6. Ruder, S.: An overview of gradient descent optimization algorithms. Arxiv preprint arXiv: 1609.04747 (2016)
7. Toulis, P., et al.: Stable Robbins-Monro approximations through stochastic proximal updates. Arxiv preprint arXiv:1510.00967v3 (2018)
8. Duchi, J., et al.: J Mach Learn Res. **12**, 2121–2159 (2011)
9. Hinton, G., et al.: Overview of mini batch gradient descent, Lecture 6a. ClassLecture, Computer Science Department, University of Toronto (2015)
10. Kingma, D.P., Ba, J.L.: Adam: a method for stochastic optimization. In: International Conference on Learned Representacion, San Diego (2015)
11. Dozat, T.: Incorporating Nesterov momentum into Adam. In: International Conference on Learning, Representations Workshop, 2013-6 (2016)

A Comparative Study of Various Wireless Network Optimization Techniques

Kanika Chauhan, Kanika Gupta, Prashant Vats, and Manju Mandot

Abstract A very important aspect is being played by the wireless optimization in the field of wireless services network (WSN) schemes. The various network optimization techniques play different roles in different implemented real-time scenarios. These techniques are used to provide better performance, reduction of error and much more during wireless transmission of binary data over a computer network. In this research paper, we have presented a comparative literature review of these WSN optimization schemes based on the work done by various researchers in the field of wireless network optimization technique. We have further tried to review the various aspects of these techniques in terms of their area of application and other relevant factors.

Keywords Wireless optimization technique · Wireless sensor network · Ant colony optimization · Genetic algorithm · Particle swarm optimization

1 Introduction

Wireless optimization techniques have played a vital role in the field of networking. Optimization techniques have a different role in different areas like broadband radio access networks (BRANs) such as 3 g/4 g cellular networks, wireless sensor networks, wireless mesh networks, and many more [1]. As we know wireless sensor networks have become more consequential day by day and widely used in the areas of patient monitoring, environmental monitoring, etc. With unprotected ideas of WSNs, sometimes sensors affect the functionality of networks in terms of error, delay, energy, and bandwidth. All these issues are accomplished via different optimization problems. Also, various optimization techniques are mainly used to reduce the error rate, enhance the performance, and also help to improve the resulting actions

K. Chauhan (✉) · K. Gupta · P. Vats
Indira Gandhi Delhi Technical University for Women, New Delhi, India

M. Mandot
Janardan Rai Nagar Rajasthan Vidyapeeth, Udaipur, Rajasthan, India

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_61

[2]. However, the best solution to this problem is to compose a protocol that uses less energy and improve the functionality of the network.

Somehow, wireless network optimization is a key of wireless networks, i.e., the main goal of optimization in the wireless network is to make a network full of quality of coverage and quality of service; therefore, it also includes various task where the first step is collection of data which is achieved by the help of drive test tools like phone, laptop, antenna and then analyzes the data which is collected. This is the main and important step where all the experts analyze the data and further comprises what are the issues, problems that assault the network. All this analyzation is achieved with the aid of network analysis tool that is mainly used by the experts after doing PlayStation. Last step is an optimization which may lead to formulating solution of a correspondence problem. Therefore, different optimization techniques are also used for the same, but differently like improve the longer network lifetime, to choose an optimal path, high success rate, and target to the data coverage of networks. Nowadays, it is important because our whole world is totally interconnected by a single channel and digitally services are available 24/7. Through these services, in network our network optimization becomes more critical. Therefore, optimization core benefits are listed as choosing an optimal path over the network, tracks the performance metrics, also provides the different analytical and simulation to the network.

2 Various Wireless Optimization Techniques

In this section of the paper, we have discussed various types of wireless optimization techniques.

2.1 Ant Colony Optimization

This technique is an optimization technique for the WSN which is based on the probabilistic analysis of the wireless network. It is a self-managed organized system that is based on the ant's behavior and is made by keeping in mind the nature of the ants as they find the ideal solution from the origin to the target. According to the nature of the ants, they first find the origin of their target, and from there, they return to their colony leaving behind pheromones so ants left behind can reach the target following the path. This technique is used for finding the ideal and minimum distance to reach the target. For this technique to work, the origin and target have to be known beforehand and distance between them should not be a straight path. To calculate the value of the pheromone, it is required to know the value of hops between the origin and the target. As soon as the nodes get the information, it makes a routing table which consists of the pheromone value (PH), next hop (NH), destination address (DA) [3]. The disadvantage of this algorithm is that it is difficult to analyze w.r.t

theory and provides more experiment approach than theory and for every iteration the probability distribution changes.

2.2 Particle Swarm Optimization

In this technique, WSNs were optimized by recursively trying to provide an improvised solution by candidate selection w.r.t. given quality measures. This technique has multiple objectives and has strong optimization. In this technique, the group of animals is considered having no head. Similarly, this technique consists of particles which are actually the optimal solution. Here two positions are considered first; the local best position lbest and the second the global best position, i.e., gbest. To calculate particle swarm optimization technique, there are ways:

1. Particle's current position
2. Particle's current velocity

The changes are made to the position and velocity of the particle as it changes the position. The updates of velocity and acceleration are made the lbest and gbest locations [3]. The major advantages of this algorithm are as follows: easy implementation, robustness, lesser parameters to adjust, computation time is low and fast convergence. The disadvantages are problem in initializing the initial parameters and cannot solve problem of scattering.

2.3 Firefly Optimization

This technique was introduced for multimodal optimization of WSN. The fireflies find the location of the particle with flashing behavior. $I(r) = \text{Is}/r2$ here, I is intensity, r is radius, and Is is intensity of source which is according to the inverse square law. As the light intensity changes with changes in distance as R changes. This algorithm not only changes the throughput but also increases the life expectancy of the network. This optimization is based on the following methods:

1. These are unisexual as they can attract each other in flashlight.
2. This attractiveness is based on the fact that light is directly proportional to the brightness.
3. Intensity of light decreases when the distance between them is increased.
4. Brightness of the firefly is the most important part of this optimization [3].

The major advantage of this algorithm is that it can be used to solve nonlinear problems and dimensional problems. The disadvantage of this algorithm is that in a reasonable amount of time, it is difficult to reach the best result.

2.4 Genetic Algorithm

This algorithm is used for genetic algorithm based on WSN optimization. It is based on the genetic procedure of the living. The technique is used to find solution for optimization problem. The main ingredient in this algorithm is the chromosomes which are either binary numbers 0 or 1 or real numbers and are search space for various positions. The usage of this algorithm is to find clusters with their masters. The application of this algorithm is as follows: (1) The fitness value is calculated when in the search space, the population is produced in a random way. (2) The various ways of producing the population are repair, mutation, crossover, and selection which can be described as mechanisms. In the selection mechanism, the parents are selected for crossover and mutation. In the crossover mechanism, the children are produced by mixing of genetic material between the two parents. The control of population is done by this algorithm by making the changes in the chromosomes to make new generation. This goes on until the population produced is very large; the algorithm locates its cluster and cluster heads in locate environment, and there is no mode of reaching out to the global process. This all happens when this algorithm is used in wireless sensor networks [3].

3 A Comparative Review of Various WSN Optimization Tech

This section provides a comparative review about various WSN optimization techniques that have been presented into tabular format in Table 1.

4 Literature Review of Various Wireless Optimization Techniques

This section provides information about various wireless optimization techniques that are used in various aspects. The descriptions of the wireless optimization techniques have been presented into tabular format in Table 2.

5 Conclusion

In this research paper, we have carried out a comparative study of the various wireless network optimization techniques concerning a different area of application where optimization technique is used. We have also performed a comparative analysis of the different author's work executing in different optimization techniques used in the

Table 1 Some wireless network optimization techniques

S. No.	Author	Technique used	Merits	Demerits	Area of application	Tool used
1	Aggarwal et. al. [9]	Ant colony optimization (ACO)	Search population in parallel + guaranteed convergence	Difficult theoretical analysis + every iteration the probability distribution changes	Optimization techniques	Not addressed
2	Aggarwal et. al. [9]	Particle swarm optimization [PSO]	Robust and easy implementation + few parameters to adjust as per fast convergence	Not suitable for scattering problems	Optimization techniques	Not addressed
3	Aggarwal et. al. [9]	Fire fly optimization [FFO]	Can deal non linear problems + high convergence	Deals with complex problems + discrete in nature	Optimization techniques	Not addressed
4	Aggarwal et. al. [9]	Genetic algorithm	Compatible to work with randomly deployable nodes	Problem in using dynamic datasets	Optimization techniques	Not addressed
5	Wang et al. [10]	Whale swarm optimization algorithm	The initial population distribution more reasonable, provides solution to negative movement of nodes	Not addressed	Optimization techniques	Not addressed
6	Sharma et al. [14]	Packet profile based scheme	Increases lifetime of network + reduces power consumption	Dysfunctional nodes causes damage	Power optimization	Not Addressed
7	Hasegawa s et al. [4]	Rigorous optimization algorithm	Useful for networks using centralized server	Signaling delay may affect the performance	Cognitive radio network	Not Addressed

(continued)

Table 1 (continued)

S. No.	Author	Technique used	Merits	Demerits	Area of application	Tool used
8	Hasegawa et al. [4]	Distributed algorithm	Useful for networks with decentralized server	Optimality of solution cannot be made sured	Cognitive radio network	Not Addressed
9	Rai et al. [17]	Key performance indicator optimization	Based on network statistics and easily retrieve the KPI using drive testing tool	Mainly not preferable due to limited sample size	4G wireless technology	Drive Test Tool
10	Liu et al. [5]	Deep learning based link evolution	Improve efficiency of solving large scale optimization problems	Not addressed	Deep learning	Not Addressed
11	Oshima et. al. [31]	Cognitive cycle technique	Improves whole network performance mainly throughput	Problem occurs in scarce radio resources and interference among nearby devices	Machine learning	Not addressed
12	Li et al. [19]	Cross layer optimization	Efficient data transmission using Lagrangian multiplier method	NUM challenge of dynamic-routing WSNs with the transmission rate, routing, link, and battery capacity constraints	Optimization techniques	Not addressed
13	Zhao et al. [21]	Quantized conflict graph	It provides a finer-grain measurement of interference condition and can deal with the accumulative effect of multiple small amount interference	Rough representation of interference condition, which is inaccurate and will cause sub-optimal results for wireless network optimization	Wireless network optimization	Not addressed

Table 2 A comparative review of the literature work onto wireless network optimization technique

S. No.	Authors	Issues addressed	Key points	Merits	Demerits	Tools used
1	Abrao et al. [1]	Discussion about the challenges in resource optimization	Focus on the EE-SE problems in multiuser environment	Mentioned about the EE-SE problem and solution in multiuser environment	Not addressed	Not addressed
2	Wang et al. [2]	Survey conducted on the security issues of the physical layer with review	Discussed about the security and its designs in the physical layer	Discussed about the security designs of the physical layer	Need exploration related to adversary issues, wireless channels, hardware impairments ands commercial wireless networks	Not addressed
3	Li et al. [3]	Proposed an approach for optimizing the problem of IEEE 802.11	Designing algorithms on the model to optimize the network with simulation	The model proposed can handle real life issues and uses small set of constraints	Not addressed	Linux NetGear WAG511 NIC
4	Hasegawa.et al. [4]	Discussion to optimize algorithm for decision making	Optimization algorithms have been proposed for different scenarios	distributed algorithm works using proposed design on IEEE1900.4	Other factors are there which can be optimized	Not addressed
5	Liu et al. [5]	Investigation of approach to reduce the optimization problem	Linking the information flow while exploiting deep learning	The approach improves large scale optimization problem	Not Addressed	Not addressed

(continued)

Table 2 (continued)

S. No.	Authors	Issues addressed	Key points	Merits	Demerits	Tools used
6	Enaya et al. [6]	An extension of frozen-time algorithm is proposed	Three different optimization strategies are proposed	With other changes in the network the methodology can be applied	Not addressed	Not addressed
7	Lin et al [7]	Overview of optimization approaches for resource allocation	Provided how imperfect scheduling is used in cross layer framework	The algorithm obtained is applied to different levels of protocol	Not addressed	Not addressed
8	Chakraborty et al. [8]	Analyzation of factors that take IEEE 802.11b access points	Simulation of test bed scenario with Xlite soft phones	Active queue management is comfortably applied with threshold	Not addressed	NetSim and MATLAB
9	Aggarwal et al. [9]	Discussion about the optimization techniques	Description and comparison of optimization techniques	Discussed about the various optimization techniques	Wireless network has cons like communication and energy	Not addressed
10	Wang et al. [10]	Proposes wireless network optimization model on Whale algorithm	Reverse learning is introduced in the original Whale algorithm	Improvement in the network coverage and reduction in number of iterations	Not addressed	Not addressed
11	Zappone et al. [11]	Discussion about wireless network designs in deep learning	Different case studies are used to show that deep learning is useful	Deep learning is an indispensable tool for design of network	Not addressed	Not addressed
12	Chakravorty et al. [12]	Comparative analysis in of various optimization techniques	Provides an understanding how each technique can benefit us	Proxy based servers are useful for application performance	Not addressed	Not addressed

(continued)

Table 2 (continued)

S. No.	Authors	Issues addressed	Key points	Merits	Demerits	Tools used
13	Ravindranath et al. [13]	Proposes an architecture of wireless protocol architecture	The idea is evaluated using sensor-augmented wireless protocols	Sensor hints improves network performance	Cons of using sensor hints is related to energy, privacy and calibration across devices	Not addressed
14	Sharma et al. [14]	Proposes packet profile based scheme for power optimization	The results are compared with the flooding scheme	The results shows effectiveness of the approaches in PPB	Not addressed	Not addressed
15	Karaki et al. [15]	A survey on routing techniques in wireless network	The design tradeoff is studied for each routing paradigm	The routing techniques have a scope in sensor network	Not addressed	Not addressed
16	Karthikeyan et.al. [16]	Analysis of wearable antenna design for wireless body networks	Discussed about the size of antenna using HFSS software	Wearable antenna is used for body parameters measurements	Antenna has a wide range of operational frequency	Not addressed
17	Rai et al. [17]	Study about LTE on KPI optimization	Discussion about LTE technologies	Results have been shown in terms of graph	Not addressed	Not addressed
18	Sharm et al. [18]	Discussion on optimization techniques in wireless network	Optimization techniques are discussed in detail	All the optimization techniques improves performance	Not addressed	Not addressed
19	Li et al. [19]	The optimization solutions are developed by solving the optimization model	Discussed about cross layer optimization sol. Using Lagrangian multipliers	Result obtained by lagrangian multiplier is perfectly achieved	Not addresses	Not addressed

(continued)

Table 2 (continued)

S. No.	Authors	Issues addressed	Key points	Merits	Demerits	Tools used
20	Shabdanov et al. [20]	Formulate a cross-layer optimization framework for the routing problem	Techniques involved successive interference cancellation, superposition coding etc.	Obtained a feasible throughput in wireless mesh network	Not addressed	Not addressed
21	Zhao et al. [21]	Quantized conflict graphs for wireless network	Experiments are performed using wireless networks	Low estimation error for system settings	Not addressed	Not addressed
22	Dong et al. [22]	Multi objective optimization method for spectrum allocation	Simplification method and intelligent optimization algo. used	Maximum use of user bandwidth	Contradicts the shortage spectrum resources of user traffic	Not addressed

wireless sensors network. While using different optimization techniques and algorithms in the field of wireless network, possibly enhance the quality of coverage and quality of service of the network. Based on our research, we have concluded that optimization techniques provide real-time reporting that helps the network manager to manage the whole network and we can ensure that using the optimization techniques, one can easily enhance the functionality of the network.

References

1. Abrao, T.: Resource optimization in wireless networks. J. Telecommun. Syst. Manag. **1**(1) (2012)
2. Li, Y., et. al.: Predictable performance optimization for wireless networks. J. ACM SIGCOMM Comput. Commun. Rev. **38**(4) (2008)
3. Aggarwal, R., et al.: Various optimization techniques used in wireless sensor networks. J. Int. Res. J. Eng. Technol. **3**(6) (2016)
4. Wang, L., et al.: Wireless sensor network coverage optimization based on whale group algorithm. J. Comput. Sci. Inf. Syst. **15**(3), 569–583 (2018)
5. Sharma, U., et al.: Power optimization techniques in wireless sensor network by using packet profile bases scheme. Int. J. Comput. Sci. **9**(3) (2012)
6. Hasegawa, M., et al.: Optimization for centralized and decentralized cognitive radio networks. J. Proc. IEEE **102**(4) (2014)
7. Rai, D., et al.: LTE theory to practice—KPI optimization (A 4G wireless network technology). Int. J. Innov. Technol. Explor. Eng. **8**() ISSN: 2278-3075 (2018)
8. Liu, L., et al.: Deep learning based optimization in wireless network. In: Proceedings of IEEE International Conference on Communications (ICC) (2017)
9. Li, C., et al.: Data transmission optimization algorithm for network utility maximization in wireless sensor network. Int. J. Distrib. Sens. Netw. **12** (2016)
10. Zhao, Y., et al.: Quantized conflict graphs for wireless network optimization. In: Proceedings of IEEE Conference on Computer Communication (2015)
11. Wang, D. et al.: A survey of optimization approaches for wireless physical layer security. J. IEEE Commun. Surv. Tutor. **21**(2) (2018)
12. Enaya, Y.A., et al.: Network path optimization under dynamic conditions. In: Proceedings of IEEE Congress on Evolutionary Computation (CEC) (2014)
13. Lin, X., et. al.: A tutorial on cross-layer optimization in wireless networks. J. IEEE J. Sel. Areas Commun. **24**(8), 1452–1463 (2006)
14. Chakraborty, T., et al.: An optimization technique for improved VoIP performance over wireless LAN. J. Netw. **7**(3) (2012)
15. Zappone, A., et al.: Wireless networks design in the era of deep learning: model-based, ai-based or both? J. IEEE Trans. Commun. **67**(10), 7331–7376
16. Chakravorty, R., et al.: Performance optimization of wireless wide area networks: comparative study and experimental evaluation. In: Proceedings 10th Annual International Conference on Mobile Computing and Networking (2004)
17. Ravindranath, L., et al.: Improving wireless network performance using sensor hints. In: Proceedings of 8th USENIX Conference on Network System Designs and Implementation (2011)
18. AI-Karaki, J.N., et al.: Routing techniques in wireless sensor network: a survey. IEEE Wirel. Commun. **11**(6), 6–28 (2004)
19. Karthikeyan, S., et al.: Design and analysis of wearable antenna for wireless body area network. In: Proceedings of International Conference on Frontiers in Material and Smart System Technologies (2019)

20. Sharma, S., et al.: Review paper on optimization techniques in wireless sensor network. Int. J. Innov. Res. Electr. Electron. Instrument. Control Eng. **4**(8) (2016)
21. Shabdanov, S., et al.: Cross-layer optimization using advanced physical layer techniques in wireless mesh networks. IEEE Trans. Wirel. Commun. **11**(4) (2012)
22. Dong, X., et al.: Multi-objective optimization method for spectrum allocation in cognitive heterogeneous wireless network. AIP Advances (2019)

Text Summarization Using Natural Language Processing

G. Sreenivasulu, N. Thulasi Chitra, B. Sujatha, and K. Venu Madhav

Abstract It is very difficult for human beings to manually extract the summary of a large document of text. There are plenty of text materials available on the Internet. So, there is a problem of searching for relevant documents from the number of documents available and absorbing relevant information from it. In order to solve the above two problems, the automatic text summarization is very much necessary. Text summarization is the process of identifying the most important meaningful information in a document or set of related documents and compressing them into a shorter version preserving its overall meanings. The different dimensions of text summarization can be generally categorized based on its input type (single or multi-document), purpose (generic, domain specific, or query-based), and output type (extractive). The main purpose is to provide reliable summaries of web pages or uploaded files depend on the users. This paper summarizes how a web application can convert generate keywords from the entered text or file or URL. It also converts the resultant summary into the audio file (.mp3).

Keywords Text · Document summary · DTD

1 Introduction

Impelled by the cutting-edge mechanical advancements, information to this century is what oil was to the past one. Today, our reality is parachuted by the social event and dispersal of tremendous measures of data. In actuality, the International Data

G. Sreenivasulu
Department of CSE, ACE Engineering College, Hyderabad, India

N. Thulasi Chitra (✉)
MLR Institute of Technology, Dundigal, Hyderabad, India

B. Sujatha
Department of CSE, University College of Engineering, Osmania University, Hyderabad, India

K. Venu Madhav
Department of CS&IS, School of Sciences, BIUST, Palapye, Botswana, South Africa

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_62

Corporation (IDC) ventures that the aggregate sum of computerized information coursing every year around the globe would grow from 4.4 zettabytes in 2013 to hit 180 zettabytes in 2025. That is a great deal of information!

With such a major measure of information coursing in the computerized space, there have to create AI calculations that can consequently abbreviate longer messages and convey precise outlines that can smoothly pass the expected messages. Besides, applying content synopsis diminishes understanding time, quickens the way toward exploring for data, and expands the measure of data that can fit in a zone.

We cannot, in any way, shape, form, or make rundowns of the entirety of the content physically; there is an incredible requirement for programmed strategies. In their 2014 book regarding the matter titled "Programmed Text Summarization," the writers give six reasons why we need programmed content synopsis instruments.

- Summaries diminish understanding time.
- When investigating archives, outlines make the determination procedure simpler.
- Automatic rundown improves the viability of ordering.
- Automatic rundown calculations are less one sided than human summarizers.
- Personalized outlines are valuable being referred to noting frameworks as they give customized data.
- Using programmed or self-loader outline frameworks empower business unique administrations to build the quantity of writings they can process.

Here is an example: Joseph and Mary rode on a jackass to go to the yearly occasion in Jerusalem. In the city, Mary brought forth a youngster named Jesus.

Extractive outline: Joseph and Mary go to occasion Jerusalem. Mary birth Jesus.

As should be obvious over, the words in striking have been removed and joined to make an outline—albeit now, and then, the rundown can be linguistically peculiar.

Deliberation-based summarization: The reflection method involves rewording and shortening portions of the source record. At the point when reflection is applied for content synopsis in profound learning issues, it can defeat the language structure irregularities of the extractive method. The abstractive content rundown calculations make new expressions and sentences that hand off the most valuable data from the first content—simply like people do. Therefore, deliberation performs superior to extraction. Nonetheless, the content rundown calculations required to do reflection are progressively hard to build up, that is, the reason the utilization of extraction is as yet famous.

2 Literature Review

Result of the data recovery gets essential for client to discover solid data for the deliberation as a result of the stridently acceleration of information on the web. Web is broadly utilized by individuals to go over data utilizing capable data recovery (IR) instruments, for example, Google, Yahoo, AltaVista, and so on, where discoveries are inexhaustible. In the greater part of the cases, clients feel held on for the exceptionally monotonous and tedious activity to uncover the fundamental essence of the result

of the IR. Scholastics and scientists are especially profited by utilizing programmed content outline framework as an apparatus to diminish the measure of time spent physically extricating the central considerations from huge reports. Notwithstanding the above reason, programmed content outline additionally furnishes its clients with various advantages just as:

1. Increase proficiency of different inquires about to pick archives/data from web indexes' yield, which as a rule contain an abundance measure of recreated data.
2. Solve the confinement of data introduction on little specialized gadgets, for example, PDA, cell phone, and so on, which can show shortened rendition of the full archive.
3. The running time of machine for interpretation is essentially decreased if a short form of content is given.

American look into libraries brought forth the underlying enthusiasm for programmed content shortening during sixties [1]. A significant assemblage of research in the course of the most recent sixty years has investigated various degrees of examination of content, to help figure out what data in the content is notable for a given rundown task [1]. Radev et al. [2] in 2002 characterized a synopsis as a book that is induced from at least one writings, which passes on basic data of the first text(s), and that is no longer than half of the first text(s) and for the most part strikingly not as much as that. Basically, a synopsis content is a subsidiary of a source content consolidated by choice as well as speculation on significant substance [3]. An enormous report is gone into the PC, and a reiterated substance is returned, which is a non-excess extricate from the first entry. Programmed content synopsis can be grouped into single report content outline and numerous archives content rundown [4]. This paper concentrated on single record content outline as single archive was the objective from the beginning of such research on programmed deliberation.

The remainder of the paper is sorted out as pursues. Area 2 quickly portrays about natural language processing. Segment 3 displays a thorough writing survey about various systems of programmed single archive content outline. Initially, correlation among the different procedures is portrayed in area 4. Area 5 turns determination with a brief about this paper.

2.1 Natural Language Processing

Common language processing (NLP) is a field of software engineering, man-made brainpower, and etymology as every one of those predetermined fields brought it into play. For the most part, it manages the collaborations among machines and human dialects that achieve task on investigating, understanding, and creating the language, which human utilizes normally so as to cooperate with PCs in both oral and composed settings utilizing common human dialects rather than programming languages [5]. It is an interdisciplinary region dependent on adaptable field of study including PC building, which gives strategies to display representation, calculation devise, and achievement; semantics, which orders etymological structures and practices; science,

which gives formal models and techniques; brain science, which studies models and hypotheses of human conduct; insights, which offers methods for anticipating estimates dependent on test records; science, which goes around the hidden engineering of phonetic procedures in the human cerebrum [6].

2.2 Audit on Automatic Single Document Text Summarization

In the earliest reference point of the examination in the field of propelling man-made consciousness to create compressed variant of a huge report, it revealed the standards for extricating striking highlights.

Programmed content outline process model can be isolated into three stages [7] in the preprocessing step source content translation to source content portrayal, (2) source portrayal change to rundown content portrayal with a calculation, and (3) in the last advance, synopsis content age from outline portrayal.

2.3 Spearheading Works

Edmundson in 1969 achieved a remarkable advancement following ten years of the start of the examination on content restatement. He depicted three extra techniques with the standard catchphrase strategy, ignoring the exceptionally high recurrence basic words to decide the sentences' weight. Those are as follows:

- Cue Method: The speculation of this strategy is that the nearness or nonattendance of certain prompt words will register the essentialness of a sentence.
- Title Method: The heaviness of a sentence is determined as an aggregate of all the substance words emerging in the title, headings, and sub-headings of a book.
- Location Method: Here, significance is accepted based on area, communicated that sentences occurring in starting situation of sections have a higher likelihood of being appropriate.

The outcome was productive and accepted that by utilizing a blend of these three plans, the best connection between the programmed and human-caused modified works can be accomplished. This paper underscored on demonstrative digests as opposed to on the creation of instructive modified works. Along these lines, if a client does not think about the archive to be abridged before, it cannot get synopsis through this proposed strategy.

2.4 Algebraic Methods

Kupiec et al. in 1995 investigated an inventive logarithmic strategy. Utilizing innocent Bayes classifier, this grouping system deals with each sentence as deserving of being

in theory or not. A prepared classifier is created from the outset, where the creators utilized a corpus of 188 couples of complete archives/rundowns. The distinctive highlights utilized capitalized words, length of sentence, structure of expression, position in passage other than word recurrence. Hovy et al. in 1999 endeavored to make a strong computerized content synopsis framework and named it SUMMARIST. Rather than sporadic term checking, SUMMARIST joins emblematic world information (encapsulated in WordNet, lexicons, and comparative assets) with solid NLP preparing (utilizing IR and measurable strategy) to steadfast idea pertinence. Their methodology dependent on the accompanying condition is as follows:

Rundown = subject recognizable proof + elucidation + outline age

Yong et al. in 2005 presented a programmed book rundown framework that fuses learning capacity by joining a measurable methodology, watchwords extraction, and neural system with solo learning. It was asserted that their proposed system can remove 83.03% of noteworthy substance. The method is comprised of three stages, as pursues:

1. Text pre-handling subsystem: Two pre-preparing techniques are applied, one is stop words expulsion to evacuate words like "the," "a," "can," "will," and so on, and another is stemming to change over each word to its stem by killing additions and prefixes.
2. Keywords extraction subsystem: from a gathering of test heterogeneous content records, fundamental highlights of each word are dictated by processing term recurrence tf(w, s) as the occasions that the word w shows up in the sentences s, and the converse sentence recurrence is f(w) is determined as the quantity of sentences in which the word w happens. From the produced tf-isf(w, s) framework, most incessant terms are recorded as watchword in the content to be condensed.
3. Summary creation subsystem: In the last advance, the framework picks sentences, which contain the catchphrases, as a feature of the synopsis. It is recommended to go through another round of stop words checking methodology before choosing sentences for being guaranteed that there is no stop word is filling in as watchword.

2.5 Methodology Dependent on Content Relationship

Recently talked about techniques utilized likelihood for picking a sentence for creating condensed form of content. Yet, those scenes are completely stupid about the union of sentences with one another. Thus, the relations between ideas in a book cannot be caught utilizing extractive strategies. On the off chance that a sentence is separated which contains connection to any past setting, at that point, the synopsis will be hard to comprehend [10]. In those conditions, different research works accomplished for investigating the strong properties involve relations among articulations of content.

Hasan in 1984 stretched out the examination to incorporate the idea of durable agreement, which adds lexico-linguistic structure to word bunches by first partitioning them into two kinds: (I) personality of reference chains that join reference and lexical attachment, (ii) similitude chains that utilized just old style relations, and connecting these chains with syntactic intra-sentence relations.

Expository structure theory (RST) of content rundown can likewise be incorporated into the gatherings of strategies dependent on content connection as the content rationality is ascribed essentially to the nearness of explanatory connection. RST was created during 1980s by analysts in characteristic language age, which models the talk structure of a book by methods for a various leveled tree graph [10].

3 Methodology

The objective of the programmed content rundown is to lessen a literary archive to a synopsis that holds the urgent purposes of the first record. The examination about content outline is dynamic, and during the most recent years, numerous rundown calculations have been proposed.

3.1 Algorithm

Step 1: Create the word recurrence table: we make a lexicon for the word recurrence table from the content. For this, we should just utilize the words that are not part of the stop words cluster.

Step 2: Tokenize the sentences: Now, we split the text _string in a lot of sentences. For this, we will utilize the inbuilt strategy from the nltk.

Step 3: Score the sentences: score a sentence by its words, including the recurrence of each constant word in a sentence. Potential issue with our score calculation is that long sentences will have a bit of leeway over short sentences. To fathom this, we're partitioning each sentence score by the quantity of words in the sentence.

Step 4: Find the threshold: Here, we are thinking about the normal score of the sentences as a limit. You can utilize different techniques to ascertain the edge.

Step 5: Generate the outline: Select a sentence for a synopsis if the sentence score is more than the normal score.

3.2 System Design

There are four modules:

- Text to Summary

- URL to Summary

- File to Summary

- Summary to Speech

- Keywords from Text

As per the above diagram, it explains various methods dealt in generating summary. They are mainly classified into four categories: (1) text to summary, (2) URL to summary, (3) summary to speech, and (4) keywords from text.

3.2.1 Keywords from Text

Keyword mining (alo known as "*keyword detection*" or "*keyword analysis*") is a text analysis method that consists of mechanically extracting the most significant words and vocabulary in a text. It helps abridge the content of a text and be familiar with the core topics which are being discussed.

Keyword mining can help you mechanize some of our workflows, like tagging incoming review responses or responding to critical client queries, allowing you to save a more time. It also provides you with actionable insights that you can use to construct enhanced business decisions. But the finest thing about keyword extraction models is that they can comfortably install and work.

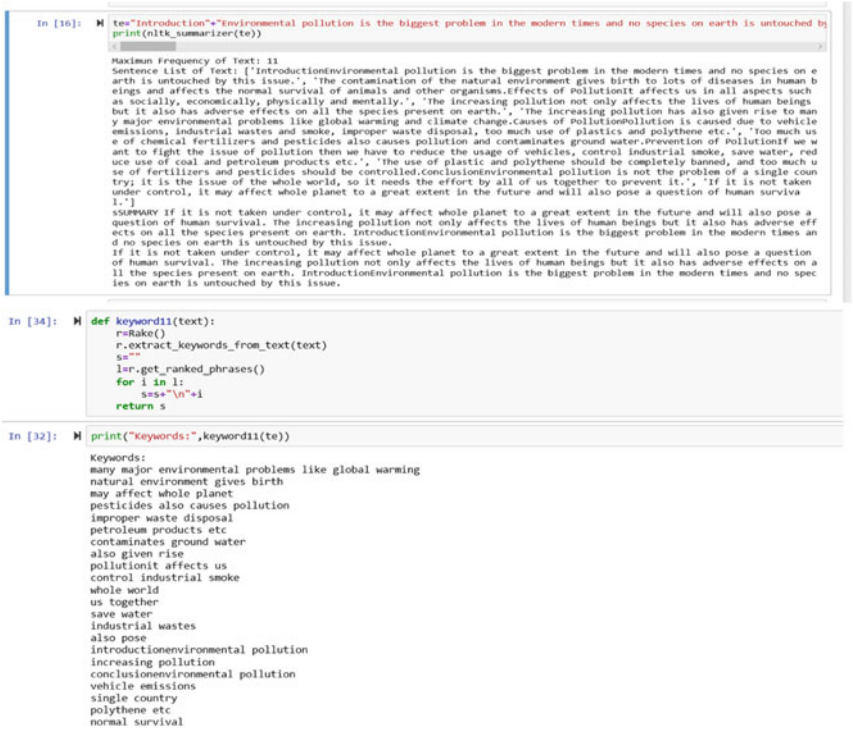
There are various techniques you can use for automated keyword extraction. From a simple statistical approach that discovers keywords by including word frequency, more advanced machine learning approaches allow you to create more compound models that can learn from the earlier examples.

Steps of Keyword Generation:

- Tokenize each document:
- Lemmatize each document
- Remove stop words
- For every term in the current document and every document in the set, compute the term frequency (how many times the term occurs in the document)
- For every term, compute the inverse document frequency
- For every term in the current document and every document in the set, compute the tf-idf
- Choose the top k words by tf-idf [2] score (Fig. 1).

4 Results

4.1 Summary of Input



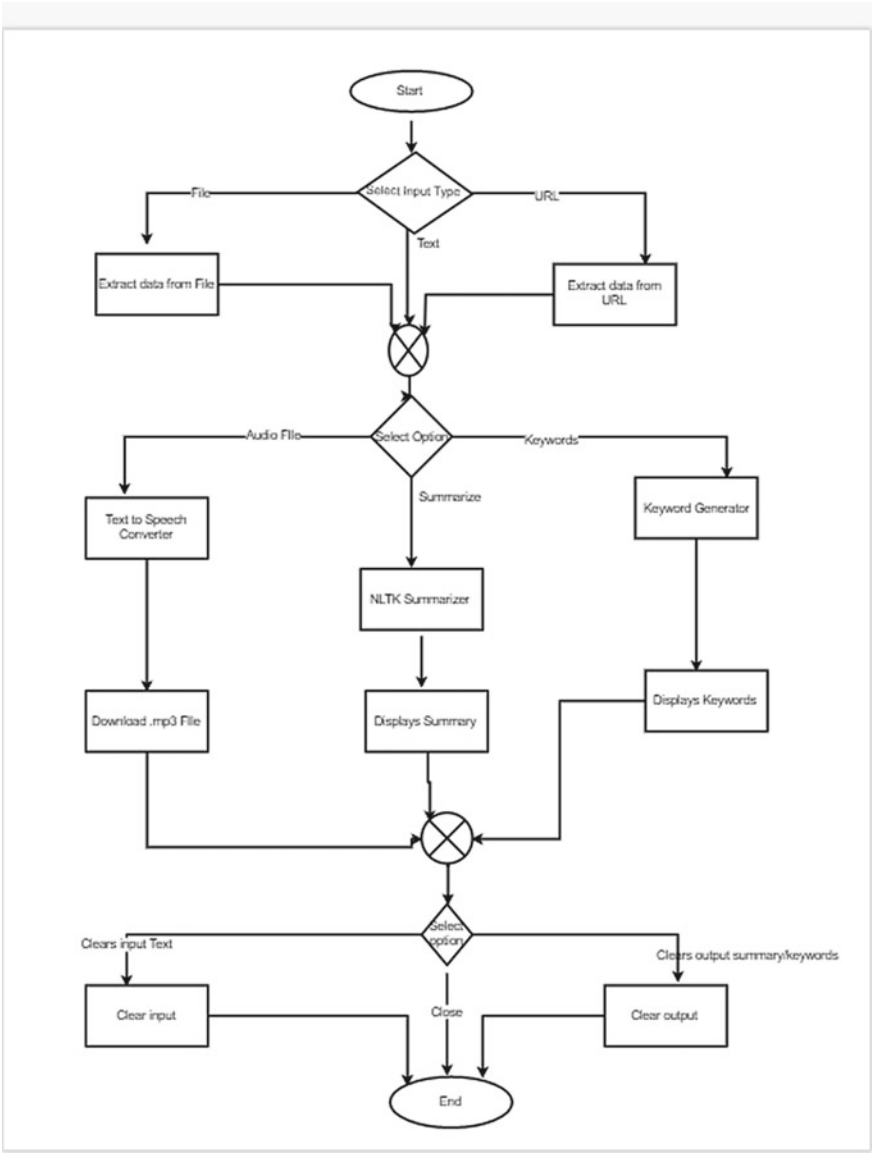


Fig. 1 Architecture of text summarization

5 Conclusion and Future Scope

We propose a completely information driven way to deal with abstractive sentence rundown. The model is basically basic, and it can without much of a stretch be prepared start to finish and scales to a lot of preparing information. This framework

changes over the entered content into abridged content utilizing regular language preparing (RLP). It changes over the information which put away in a document to an abridged book effectively. It changes over the information accessible over overall web (www) to the outlined content by utilizing beautiful soap library.

It proficiently creates watchwords from the information utilizing nltk Rake library.

It proficiently changes over the outlined content into mp3 record utilizing gTTS library.

Favorable circumstances of the proposed framework:

- Some additionally scoring procedures, in view of the arrangement of sentences could have been actualized.
- Reading from word records and HTML documents could have been executed.
- Natural language handling could have been executed

The potential outcomes in this undertaking are unending. With the advancement of regular language handling (NLP), the accompanying does not stay minor considerations…

Filling up structures, given content containing the important information.
Creating a bio-information, from a printed detail if the individual.
Generating the newspaper features from given article.
For programmed paper adjustment framework.

References

1. Gong, Y, Liu, X.: Generic text summarization using relevance measure and latent semantic analysis. In: Proceedings of the 24th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (pp. 19–25), New Orleans, Louisiana, United States (2001)
2. Radev, R., Teufel, S., Saggion, H., Lam, W., Blitzer, J., Qi, H., Celebi, A., Liu, D., Drabek, E.: Evaluation challenges in large-scale document summarization. In: Proceeding of the 41th Annual Meeting of the Association for Computational Linguistics (pp. 375–382), Sapporo, Japan (2003)
3. Ahmad, K., Vrusias, B., Oliveira, P.C.F.: Summary evaluation and text categorization. In: Proceedings of the 26th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (pp. 443–444), Toronto, Canada (2003)
4. Hynek, J., Ježek, K.: Practical approach to automatic text summarization. In: Proceedings of the ELPUB '03 Conference (pp. 378–388), Guimaraes, Portugal (2003)
5. Berry, M.W., Dumais, S.T., O'Brien, G.W.: Using Linear Algebra for Intelligent Information Retrieval. SIAM Review (1995)
6. Edmundson, H.P.: New methods in automatic extracting. J. Assoc. Comput. Mach. **16**(2), 264–228.
7. Kupiec, J., Pedersen, J., Chen, F.: A trainable document summarizer. In: Proceedings of the Eighteenth Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (pp. 68–73), Seattle, Washington, United States (1995)
8. Riyazuddin, Y. Md., Susmithavalli, G., Victor Daniel, G.: An approach for the implementation of double guard method for detecting the Ids in web applications. Int. J. Adv. Res. Manag. Architect. Technol. Eng. **4**(6) (2018)

9. Sai Prasad, K., Pasupathy, S.: Deep learning concepts and libraries used in image analysis and classification. TEST Eng. Manag. **82**, pp. 7907–7913. ISSN: 0193-4120
10. Sai Prasad, K., Miryala, R.: Histopathological image classification using deep learning techniques. Int. J. Emerg. Technol. **10**(2), 467–473 (2019)
11. McKeown, K., Robin, J., Kukich, K.: Generating concise natural language summaries. Inf. Process. Manag. Int. J. **31**(5) (1995)

Floor Plan Designer Application by Predicting Spatial Configuration Using Machine Learning

T. Rajasenbagam, S. Jeyanthi, and N. Uma Maheswari

Abstract In the field of architecture, floor plan design is the way to generate various floor plans for houses. Designing of floor plan is one of the most essentials for all kinds of infrastructures. In spite of this, there are many computerized and manual ways are available today. Nowadays, the burden of manual work has been greatly reduced. Even though the manual works have been greatly reduced, still there are some manual works like dragging and dropping, aligning the rooms, placing in desired locations, etc. We cannot determine the exact lengths of rooms, hall, kitchen, etc. Floor plan designing is one of the highly iterative processes and it demands extensive human labour; however, they are limited to fully automate the creative process. This proposed work faces so many real challenges because of the many implicit and explicit rules in order to create viable floor plans. In this work, a floor plan designer is suggested for an automated floor plan generating application that can generate floor plans automatically when the total area of the land and choice of BHK (B = Bedroom. H = Hall. K = Kitchen) is provided by the customers. In this research work, deep convolutional neural network (DCNN) methods are proposed to handle floor plan parsing and spatial recognition methods. Spatial recognition [1] among objects plays a fundamental role in the human perception and understanding of design. The various spatial recognition methods are used to edit the generated floor plans based on their willingness and acceptance.

Keywords Floor plan · Machine learning · Spatial recognition

T. Rajasenbagam (✉)
Department of Computer Science and Engineering, Government College of Technology, Coimbatore, Tamil Nadu, India
e-mail: trajasenbagam@gct.ac.in

S. Jeyanthi · N. Uma Maheswari
Department of Computer Science and Engineering, PSNA College of Engineering and Technology, Dindigul, Tamil Nadu, India
e-mail: sjeyanthi@psnacet.edu.in

N. Uma Maheswari
e-mail: numamahi@psnacet.edu.in

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_63

1 Introduction

In ancient times, people used to draw the outline for all their construction works which is referred to as blueprints. A blueprint is a reproduction of a technical drawing using a contact print process on light-sensitive sheets, rapid and accurate production of an unlimited number of copies. Blueprints become a tedious process because they are manual work and consume more time. Hence, people made this work into digitalized process which is later referred as floor plans.

In design and building, a floor plan [2] is an attracting to scale, demonstrating a view from above, of the connections between rooms, spaces and other physical highlights at one degree of a structure. Measurements are normally attracted between the dividers to determine room sizes and divider lengths. Floor plans may likewise incorporate various subtleties of fixtures for the convenience of users. Floor plans may also incorporate for development to indicate completes, development strategies or images for electrical things. Time required to create an official floor plan is the most tedious work. Year by year the time required to create an official floor plan [2] has been reduced considerably. In the year 2013, home style application was released which greatly reduced the time required. Since 1982, AutoCAD [3] has been used to create floor plans, but they required more time to progress. We tried to reduce the time required to make floor plans and we achieved this with the help of machine learning. The main objective is to automate [4] the creation of floor plans and allows the user for easy modification of the plans. Our product provides the facility of not wasting the user's time by getting their floor plans from the engineer, who saves a lot of time and provide better suggestion in their plans. The required application can be accessed anytime and at anywhere. People who use our application feel comfortable and technical risk free. This research work is organized into the following sections which are an analysing space probability for land space, proposed system, implementation, model evaluation and conclusion.

2 Analysing Space Probability for Land Space

A floor plan is a structural arrangement, scene design report and a detailed design building attracting of proposed upgrades [5]. Setting up recognized patterns is the most vital part of this project. In this part, the recognized patterns [6] of rooms, kitchen, portico and other parts of house with respect to the outline of the house are placed accordingly without any overlaps. It is also the toughest part because we have to prevent one part from overlapping on other patterns. The outline of the house will be drawn from the details provided by the user. The patterns will also be generated according to the details of the user [7]. Once the patterns are created, they are placed in their respective locations. There will be many random generations of the plan [8] and the user can select a plan according to their wish. Once they selected the plan

they like and if they wish to modify the plan, modification plan will also be provided. From which the user can modify their plan as they wish [9].

Disadvantage of Existing System

i. In existing applications, the designers have to manually design their floor plans and its arrangements which take a lot of time.
ii. There is no automated designing of floor plans in existing applications.
iii. For better suggestion, the designer needs to entirely change the plan and has to think for the new alterations in the plan.
iv. These existing applications are platform dependent.

3 Proposed System

The basic idea of this system is to combine the existing system outcomes to produce an effective floor plan generation technique [10]. The floor plan problem can be broken down into two sub-problems:

i. The model needs to learn how to generate each room type in the context of rooms that have already been generated
ii. The model needs to learn how to assemble the various rooms together.

This proposed floor plan designer application software system as shown in Fig. 1 allows the user to create and adjust the floor plan of their choice within the provided spatial information.

Easy way of creating a floor plan with just a click is purely possible with the design. This application is to optimize the time consumed to create floor plans. This application provides the facility of not wasting the user's time by getting their floor plans from the engineer, who saves a lot of time and provides better suggestion in their plans [2]. The required application can be accessed anytime and at anywhere. People who use our application feel comfortable and technical risk free. The various classes involved in floor plan designer are shown in Fig. 2.

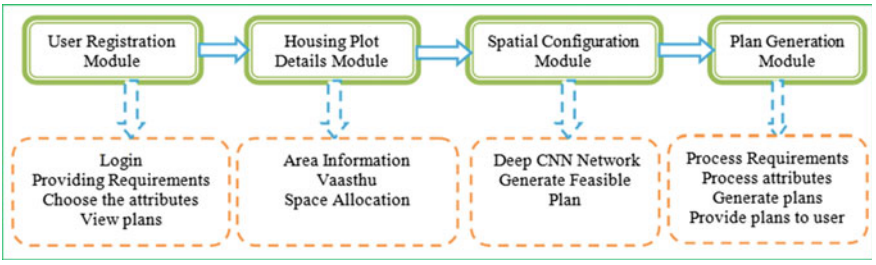


Fig. 1 Floor plan designer

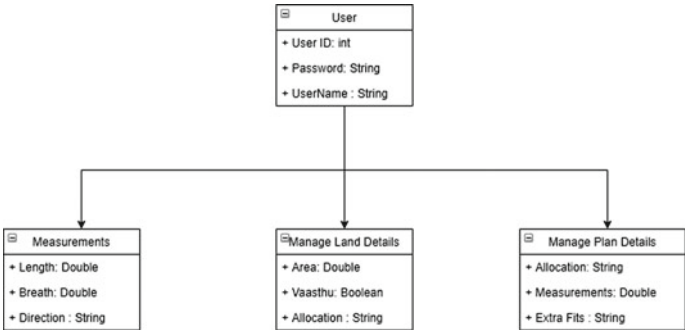


Fig. 2 UML class diagram for floor plan designer

Working Methodology of Floor Plan Designer

The user is allowed to make a registration progress to access our application through the login portal of our software. When the user is completed with his registration process, the user is validated and he/she is allowed to login to the application via the backend controllers. Then, the user is directed to provide the dimensions of the land area and the possible requirements of the floor plan via the HTML forms. When the user completed the required inputs, the values are transferred to the database via the backend controller which is written in PHP. Then, the algorithm in the server is initiated, and the input of the user is fetched from the database and processed, then the output is stored in the database. The algorithm, which uses the Raphael's JS framework, will use the output values and results of the floor plan suggestions in the form of scalable vector graphics which is helpful and understandable to the user. The user can view all the possibilities of the floor plan which is generated by our algorithm. When the user logs in again, the recent works done by the user can be viewed and obtained easily. When the user chooses for his/her satisfied suggestion, he/she can download the plans in the form of printouts, PDFs, images, etc.

Features

i. Time consumption is well reduced.
ii. A platform-independent application which allows all users to access it from anywhere and at any time.
iii. The alteration in their plan is well simplified and it can make an easier approach.
iv. Better suggestions can be made and the user can gain more suggestions beyond normal designs.

4 Implementation

The proposed work consists of various phases are user module, register module, user panel module and plan generation module. The high-level design for floor plan

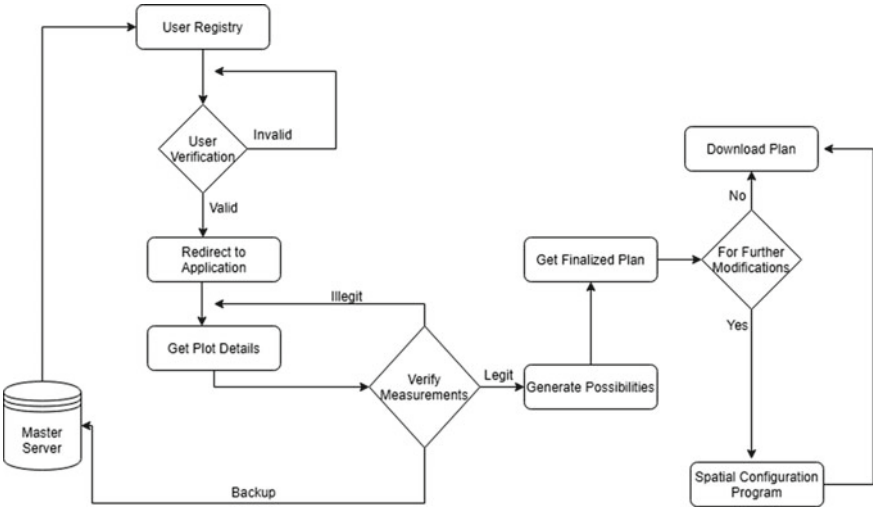


Fig. 3 High-level design for floor plan designer

designer is shown in Fig. 3. System requirements carry front end details like HTML5, CSS3, Bootstrap and Raphael's JS; back end details like PHP, MySQL and software tools are Laragon, Sublime Text 3 to run smoothly and efficiently.

I. User Module:

In this module, a new user is allowed to create an account to manage the user details and the plan works, generated to that user, and on the other hand, existing user will login through this module which is shown in Fig. 4 in order to enter the application in any platform. The user data will be maintained in a confidential manner by providing the server with great layer of firewall protection.

II. Registration Module:

Every user who wants to use the application must register themselves in the application with the required details which is shown in Fig. 5. This module helps in supporting the user for future considerations and enhancement of the application in providing the user with the best visual way to access the application at ease. The user can also log in with the help of social media like Google, Facebook, in order to verify the user.

III. User Panel Module:

This module is the major component of the application where the user will enter the available plot details and required additions to their floor plan, which will be included to the algorithm to generate the plan which is shown in Figs. 6and 7.

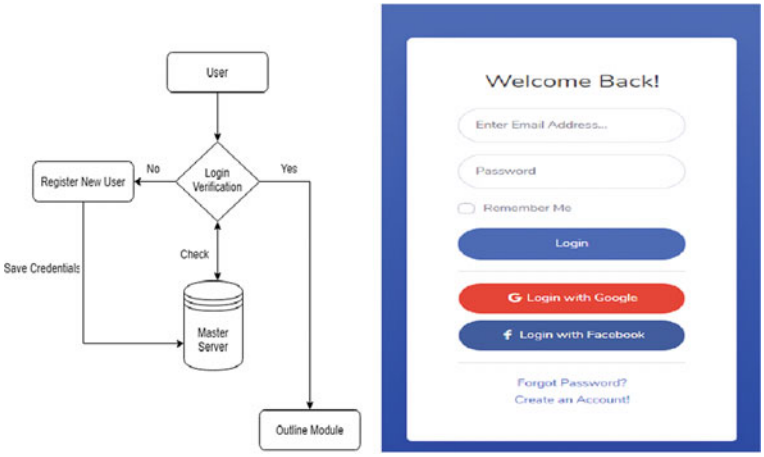


Fig. 4 Login module in mobile platform

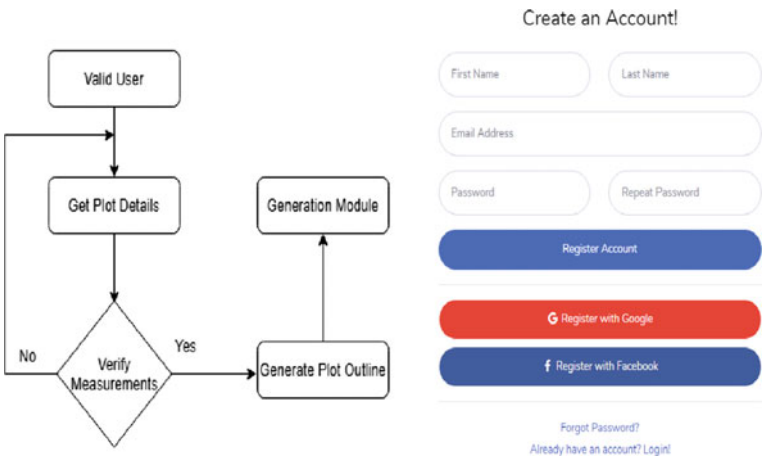


Fig. 5 Registration module in mobile platform

5 Model Evaluation

In this proposed work, mean absolute error (MAE) and the mean relative error (MRE), precision, recall and FI score are considered for the performance metrics. MAE is defined by

$$\mathrm{MAE} = \sqrt{\frac{1}{X}\sum_{x=1}^{X}\left(U_x^{\mathrm{PM}} - U_x^{\mathrm{GTM}}\right)^2}$$

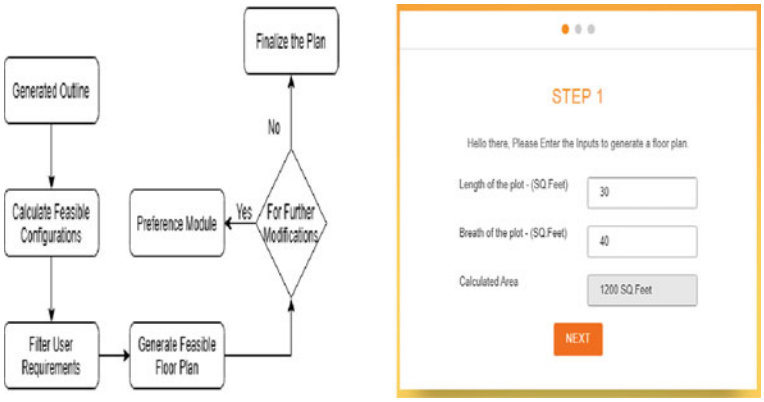


Fig. 6 Basic input form

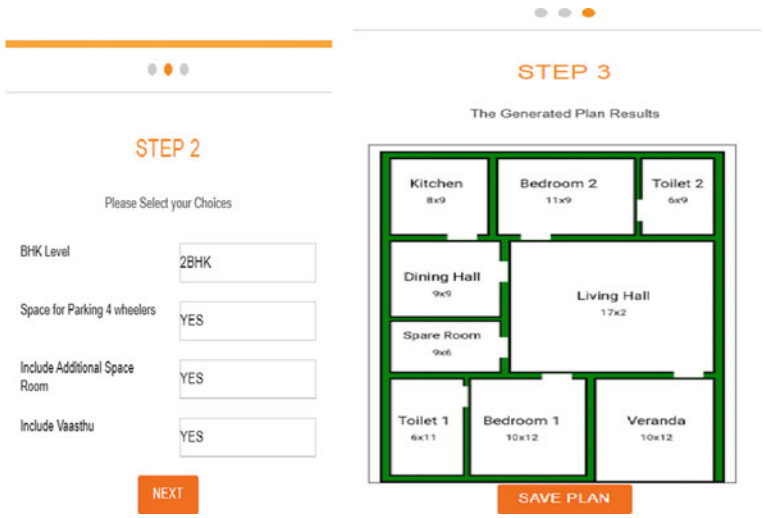


Fig. 7 Generation module

where X represents the total number of entries, x is the single entry and U^{PM}, U^{GTM} denote prediction and ground truth matrix, respectively.

MRE is defined by

$$\text{MRE} = \frac{\text{MAE}}{\max U^{\text{GTM}} - \min U^{\text{GTM}}} \times 100\%$$

MAE is estimated based on the maximum and minimum values of ground truth matrix, respectively.

Let precision is defined as the number of positive class predictions that actually belong to the positive class.

$$\text{Precision} = \frac{\text{True Positives}}{(\text{True Positives} + \text{False Positives})}$$

Recall is defined as positive class predictions made out of all positive examples in the dataset.

$$\text{Recall} = \frac{\text{True Positives}}{(\text{True Positives} + \text{False Negatives})}$$

FI score provides a single score that balances both the concerns of precision and recall in one number.

$$\text{FI score} = \frac{2 \times (\text{Precision} \times \text{Recall})}{(\text{Precision} + \text{Recall})}$$

6 Conclusion

In this paper, we have discussed the machine learning approaches like spatial recognition to set up the recognized patterns, generate a plan and allow user to modify the plan. Here, the dataset is collected and processed, then the respective outline, rooms, kitchen, portico and other features of house are generated and placed within the outline of the plan. When Raphael framework compared to all other frameworks, it gives improved and more performance in terms of drawing shapes and placing the shapes [11]. By using this framework, it will be easier to generate shapes and modification of the shapes when compared to other frameworks.

In the future work, we expand the process to include the rearrangement of patterns and also including the pattern handling and space recognition which will be more effective in identifying the space and to reduce unnecessary spaces in the plan using space recognition. By this, we plan to reduce the wait time of the user until the plan they desire appears.

References

1. Ostwald, M.J.: The mathematics of spatial configuration: Revisiting, revising and critiquing justified plan graph theory. Nexus Netw. J. **13**(2), 445–470 (2011)
2. Liu, C., et al.: Raster-to-vector: revisiting floor plan transformation. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 2195–2203 (2017)
3. Ozler, L.: AutoCAD WS for Android (2011)

4. Wang, H., et al.: Graphgan: graph representation learning with generative adversarial nets. In: Thirty-Second AAAI Conference on Artificial Intelligence (2018)
5. Jia, Y., Shelhamer, E., Donahue, J., Karayev, S., Long, J., Girshick, R., Guadarrama, S., Darrell, T.: Caffe: Convolutional Architecture for Fast Feature Embedding (2014)
6. Michalek, J., Choudhary, R., Papalambros, P.: Architectural layout design optimization. In: Engineering Optimization, pp. 461–484 (2002)
7. Good, I., et al.: Generative adversarial nets. In: Advances in Neural Information Processing Systems, pp. 2672–2680 (2014)
8. Tom Ravenscroft.: Wallgren Arkitekter and BOX Bygg Create Parametric Tool that Generates Adaptive Plans (2019)
9. delas Heras, L.P., Terrades, O., Robles, S., Sanchez, G.: A new database for structural floor plan analysis and its ground truthing tool. Int. J. Doc. Anal. Recogn. (2015)
10. Jaques, N., et al.: Generating music by fine-tuning recurrent neural networks with reinforcement learning. In: Deep Reinforcement Learning Workshop NIPS (2016)
11. Okoi, M.: FreeCAD—A 3D Modelling and Design Software for Linux (2019)

Investigation Study on Secured Data Communication with Blockchain and IOT in Green Cloud Computing

Nithya Rekha Sivakumar

Abstract Green cloud computing is used for utilization of computer resources in well-organized way. Internet of Things (IoT) is a developing model through heterogeneous technologies with smart objects that are seamlessly connected to the Internet. Blockchain technology in green cloud provides the secrecy and preserves the control system in energy-efficient manner for real-time applications. Blockchain extracts the information from IoT devices and accumulate the records to preserve the transparency at different locations. Different blockchain methods were discussed to provide secured communication in green computing. But, the existing techniques failed to improve the security level and reduce the energy consumption during data communication. In order to address these problems, an IoT-based least square orthogonal regressive whirlpool cryptographic hash blockchain (IoT-LSORWCHB) method is introduced in green computing. The key aim of IoT-LSORWCHB method is to collect and analyze the patient data with blockchain for performing the secured data communication in green cloud environment. Simulation is carried out on factors such as data confidentiality level, processing time, and energy consumption with respect to number of patient data.

Keywords Green computing · Heterogeneous · Blockchain · IoT devices · Secured communication · Whirlpool cryptographic hash blockchain

1 Introduction

Internet of Things (IoT) is a heterogeneous technology with smart objects that are faultlessly connected to the Internet. The objects provide the innovative services in different domains like smart cities, smart health, and smart communities. The security is an essential one to guarantee the safety and system effectiveness. Security problems on IoT are addressed through blockchain mechanism. A blockchain

N. R. Sivakumar (✉)
Department of Computer Sciences, College of Computer and Information Sciences, Princess Nourah Bint Abdulrahman University, Riyadh 11671, Kingdom of Saudi Arabia
e-mail: NRRaveendiran@pnu.edu.sa; rekhasiva24@gmail.com

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_64

is decentralized ledger of transaction across the network. Blockchain technology employs the hash encryption to secure the information. Secure transmission is the process of transferring the confidential data over the secure channel.

This paper is structured as follows: Sect. 2 describes the existing secured data communication techniques with blockchain and IoT. Section 3 explains the existing and proposed secured data transmission methods with IoT and blockchain. Section 4 provides the experimental settings with possible comparison between them. Section 5 portrays the limitations of secured data communication techniques. Section 6 concludes the paper.

2 Literature Survey

An efficient lightweight integrated blockchain (ELIB) model was introduced in [1] to address the IoT needs. The designed model was used to validate its applicability in different IoT situations. The resource constrained created the shared keys to transmit the data for processing the incoming and outgoing requests. Blockchain-based trusted networks connect protocol called BTNC was introduced in [2] to guarantee the reliability in IoT. BTNC was used to authenticate the security of terminal device in the IoT networks. A lightweight framework was designed in [3] to use blockchain with anomaly detection model in distributed environment. Lightweight scalable blockchain (LSB) was introduced in [4] for providing the end-to-end security.

A lightweight IoT information sharing security framework was introduced in [5] depending on blockchain technology. Though the security level was improved, the computational cost was not minimized by designed framework. The literature review analysis was carried out in [6] on blockchain with IoT. The challenges were addressed through creating the blockchain architecture for IoT. A novel blockchain-based contractual routing (BCR) protocol was introduced in [7] for untrusted IoT devices. A central authority (CA) was essential one to assist identification and authentication of every device. However, the processing time was not minimized by BCR protocol.

A secure wireless mechanism was introduced in [8] to protect the transparency for smart sensor activity. Blockchain technology extracted the proceedings of every record into number of blocks. Routing protocol over low power and lossy networks (RPL) supported IoT-low power and lossy network (LLN) in [9] with control messages and configuration information. But, the data confidentiality rate was not improved. A blockchain layout was constructed in [10] depending on enhanced security with help of IoT devices for E-voting application. But, the processing time consumption was not minimized by blockchain layout. A decentralized authentication and access control mechanism was introduced in [11] with lightweight IoT devices. The designed mechanism was depending on fog computing and public blockchain ideas. But, the security level was not improved by designed mechanism.

3 Secured Communication in Green Cloud with Blockchain and IoT

Green computing is the developing technology with information resources for preserving or improving the overall performance. Green computing is the durable computing that minimized the electricity usage and environmental waste. Every transaction is fixed to the previous transactions or records.

3.1 n Efficient Lightweight Integrated Blockchain (ELIB) Model for IoT Security and Privacy

An efficient lightweight integrated blockchain (ELIB) model was introduced to address IoT requirements. The designed model was deployed in smart home environment to validate its applicability in different IoT scenarios. ELIB model constructed the overlay network with equipped resources to verify their security and privacy. The consensus algorithm restricted the number of blocks created through cluster heads (CHs) in consensus period. ELIB used CC method for reducing the computational overhead with new blocks. DTM scheme was introduced to change particular system variables in dynamic way for guaranteeing the throughput of public BC in network. The designed ELIB model reduced the processing time, energy consumption, and overhead.

3.2 BTNC: A Blockchain-Based Trusted Network Connection Protocol in IoT

A blockchain-based TNC protocol termed BTNC was introduced to guarantee the reliability of terminals in the distributed IoT environment. Due to the immutability features of blockchain, BTNC authenticated the security of terminal devices in IoT networks. The platform measurements were included in decentralized blockchain mode. The designed scheme joined Diffie–Hellman over bitcoin with TNC framework to perform the key agreement for user authentication in distributed environment.

3.3 Lightweight Collaborative Anomaly Detection for the IoT Using Blockchain

Due to fast development, Internet of Things (IoT) becomes the key characteristics of daily lives. A lightweight framework was introduced to employ the blockchain in

distributed environment. A new framework was introduced for modeling the control-flow application and for identifying the abnormal patterns. The designed agent used blockchain protocol to revise the anomaly detection model through self-attestation and consensus on IoT devices. Through collaborating among agents, false-alarm-rate was minimized because of the shared experience.

3.4 Least Square Orthogonal Regressive Whirlpool Cryptographic Hash Blockchain

Cloud computing is the dynamic field of information and communication technology for environmental security. Cloud infrastructure has become the key environmental problem through considering the energy consumption and carbon emission. Green computing denotes the energy-efficient personal computers and executes the policies to improve the computing resource efficiency and to minimize the energy consumption. The existing methods failed to reduce the energy consumption and improve the data confidentiality level during secured data communication. In order to address the existing problems, IoT-based least square orthogonal regressive whirlpool cryptographic hash blockchain (IoT-LSORWCHB) method is introduced. The main objective of IoT-LSORWCHB method is to collect and analyze the patient data with blockchain for performing the secured communication in green cloud environment. IoT-LSORWCHB method performs three processes, data collection, data classification, and secured communication. The architecture diagram of IoT-LSORWCHB method is described in Fig. 1.

Figure 1 describes the architecture diagram of IoT-LSORWCHB method in green environment. Initially, the patient information is collected with help of IoT devices for performing the secure data communication. The IoT-LSORWCHB method performs

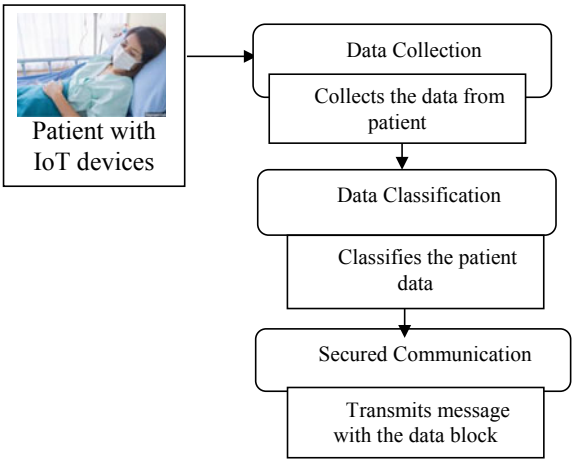


Fig. 1 Architecture diagram of IoT-LSORWCHB method

the data classification using least square orthogonal regressive function. The classified data is transmitted to the data block of blockchain for the secure transmission. IoT-LSORWCHB method uses the whirlpool cryptographic hash blockchain for generating the hash value for every patient data by using one-way compression function. The data with hash value is transmitted into the server through Internet. Therefore, the IoT-LSORWCHB method guarantees the security to preserve the data with better communication.

Data Classification

After collecting patient data, the proposed IoT-LSORWCHB method performs the data classification through the least square orthogonal regressive function. The regression function estimates the relationship between the dependent variable (i.e., output classes) and one or more independent variables (i.e., input patient data). The radial basis kernel function is applied to measure the relationship between the two variables.

Figure 2 illustrates the block diagram of whirlpool cryptographic hash blockchain constructions with one-way compression to create fixed hash for each input patient data. '$B_1, B_2, B_3, \ldots, B_m$' denotes the message block with the fixed size. After partitioning the data, block of message is transmitted to compression function (C_{fm}) with m-bit message block (B_m) and m-bit chaining value (ρ_τ) (i.e., final hash). By whirlpool cryptographic construction, the generated hash is constructed through iterating the compression function to process message of fixed length. Therefore, IoT-LSORWCHB method guarantees the security to preserve the data during communication. Sured communication is with a whirlpool cryptographic hash blockchain employs the one-way and compression function to transmit the length of message into fixed-length hash value. The generated hash value of an input patient data is considered as fingerprint of the message. It causes severe change in the hash value

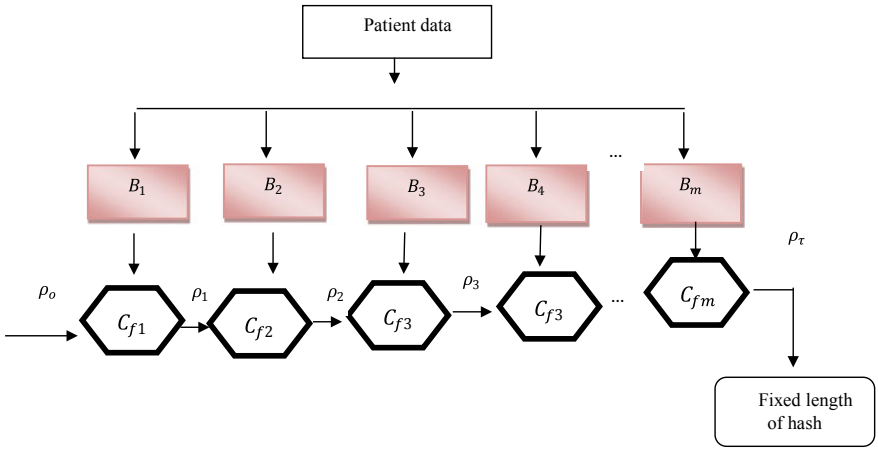


Fig. 2 Whirlpool cryptographic hash blockchain with one-way compression

when there is any small change or modification in input patient data. As a result, whirlpool cryptographic hash blockchain function is used for performing the secured data transmission with better data confidentiality level. The compression function is illustrated in Fig. 2.

4 Experimental Settings and Result Discussion

Experimental results of proposed IoT-LSORWCHB method and the other three related methods, namely ELIB model [1], BTNC protocol [2], and lightweight framework [3] are implemented in Java language. MHEALTH dataset is used to determine the performance of proposed technique for secure data communication. The MHEALTH dataset is taken from the UCI machine learning repository. The mobile HEALTH (MHEALTH) dataset includes the body movement and essential sign recordings for ten volunteers of the various report while performing the different physical activities. The IoT devices are positioned on different positions like subject chest, right wrist, and left ankle to determine the activity experienced by diverse body parts namely, acceleration, rate of turn, and magnetic field orientation. Three parameters are employed to compute the performance, namely

- Energy consumption,
- Processing time, and data confidentiality level.

Impact on Energy Consumption

Energy consumption is defined as the amount of energy consumed by the patient data for performing secured data communication. It is defined as the product of number of patient data and energy consumed by one patient data. It is formulated as,

$$E_C = N * \text{Energy consumed by one patient data} \tag{1}$$

From (1), 'N' denotes the number of patient data. When the energy consumption is higher, the method is said to be more efficient.

Impact on Processing Time

Processing time 'T_P' is defined as amount of time consumed for performing the secured data communication with IoT and blockchain. It is defined as the difference of ending time and starting time of secured data communication. It is formulated as,

$$T_P = \text{Ending time} - \text{Starting Time of Secured Data Communication} \tag{2}$$

From (2), the processing time is determined. It is calculated in terms of milliseconds (ms). When the processing time is lesser, the method is said to be more efficient.

Data Confidentiality Level

Data confidentiality level is defined as number of patient data accessed only by the authorized user. The formula for determining the data confidentiality level while performing the secured data communication is determined as,

$$DC_L = \left(\frac{\text{Number of data accessed by authorized user}}{\text{Number of data collected by IoT devices}} \right) * 100 \quad (3)$$

From (3), 'DC_R' represents the data confidentiality level. It is measured in terms of percentage (%). When the data confidentiality level is higher, the method is said to be more efficient.

The first metric used for performing the secured data communication is energy consumption. In order to determine the energy consumption, the number of patient data considered as an input ranges from 100 to 1000.

Table 1 explains the energy consumption comparison of our proposed IoT-LSORWCHB method with other existing techniques. From the results, energy consumption of IoT-LSORWCHB method is lesser than the three existing approaches ELIB model [1], BTNC protocol [2], and lightweight framework [3]. While conducting the experiments with 100 mobile nodes, IoT-LSORWCHB method consumes 18 J of energy consumption, whereas ELIB model [1], BTNC protocol [2], and lightweight framework [3] consumes 23, 33, and 41 J, respectively. The comparative result analysis of energy consumption is illustrated in Fig. 3.

From Fig. 3, the energy consumption of four methods during secured communication is shown in four different colors. The orange color bar represents the energy consumption of IoT-LSORWCHB method. The blue color, red color, and green color bar represent the energy consumption of existing approaches ELIB model [1], BTNC

Table 1 Tabulation for energy consumption

Number of patient data (number)	Energy consumption (J)			
	ELIB model	BTNC protocol	Lightweight framework	Proposed IoT-LSORWCHB method
100	23	33	41	18
200	25	35	44	22
300	27	38	47	26
400	30	42	51	29
500	35	46	53	31
600	38	49	55	34
700	41	51	58	37
800	44	53	62	39
900	46	57	65	40
1000	49	60	68	42

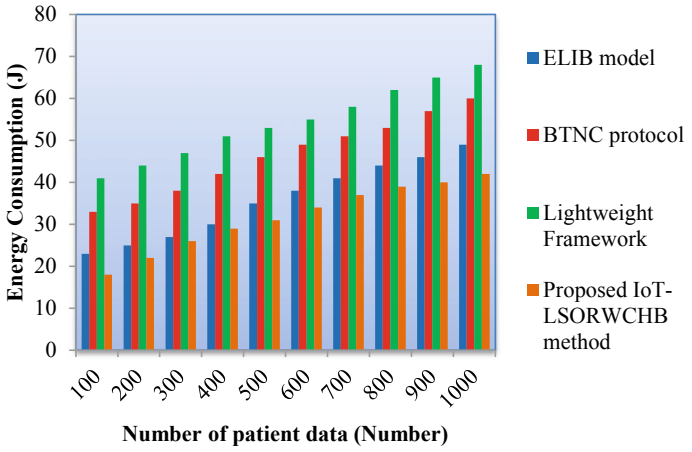


Fig. 3 Measurement of energy consumption

protocol [2], and lightweight framework [3]. The graphical results show that energy consumption of IoT-LSORWCHB method is lesser than the existing techniques. This is due to the application of regressive function and blockchain function for classifying the patient data and for generating the hash value for each patient data through the one-way compression function. The hash value is transmitted to the server for performing the secured communication in green computing. The result proves that the energy consumption of IoT-LSORWCHB method is reduced by 11, 32, and 42% when compared to the existing methods [1–3], respectively, in green computing.

The second metric used for performing the secured data communication is processing time. In order to determine the processing time, the number of patient data is taken as input.

Table 2 describes the processing time comparison of our proposed IoT-LSORWCHB method with other existing techniques. From the attained results, processing time consumption of IoT-LSORWCHB method is lesser than the three existing approaches ELIB model [1], BTNC protocol [2], and lightweight framework [3]. While conducting the experiments with 600 mobile nodes, the processing time consumption of IoT-LSORWCHB method is 36 ms, whereas the processing time consumption of ELIB model [1], BTNC protocol [2], and lightweight framework [3] is 45, 53, and 63 ms, respectively. The comparative result analysis of processing time is explained in Fig. 4.

From Fig. 4, the processing time of four methods is illustrated by four different colors. The orange color bar represents the processing time consumption of IoT-LSORWCHB method. The blue color, red color, and green color bar represent the processing time consumption of existing approaches ELIB model [1], BTNC protocol [2], and lightweight framework [3]. The above graphical results verify that the processing time of IoT-LSORWCHB method is lesser than the existing techniques. This is because of applying the regressive function and blockchain function to

Table 2 Tabulation for processing time

Number of patient data (number)	Processing time (ms)			
	ELIB model	BTNC protocol	Lightweight framework	Proposed IoT-LSORWCHB method
100	34	42	48	25
200	37	45	51	27
300	39	47	53	29
400	41	49	56	31
500	43	51	61	33
600	45	53	63	36
700	47	56	66	38
800	50	59	69	40
900	52	62	71	43
1000	55	64	73	46

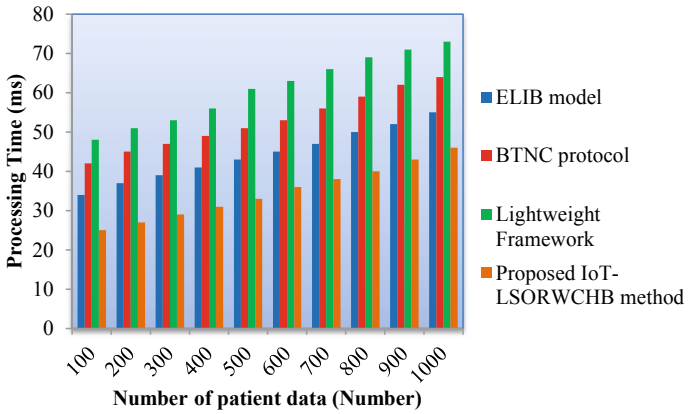


Fig. 4 Measurement of processing time

classify the patient data and to generate the hash value through one-way compression function. By this way, the secured data communication is carried out with minimal time consumption in green computing. The result proves that the processing time of IoT-LSORWCHB method is reduced by 22, 35, and 43% as compared to existing methods [1–3], respectively.

The third metric used for performing the secured data communication is data confidentiality level. In order to determine the data confidentiality level, the number of patient data is taken as input.

Table 3 explains the data confidentiality level comparison of our proposed IoT-LSORWCHB method with three existing techniques. From the results obtained, data

Table 3 Tabulation for data confidentiality level

Number of patient data (number)	Data confidentiality level (%)			
	ELIB model	BTNC protocol	Lightweight framework	Proposed IoT-LSORWCHB method
100	71	78	80	85
200	73	81	82	87
300	75	83	84	89
400	77	85	86	91
500	79	89	90	93
600	81	91	89	95
700	78	88	86	92
800	76	86	88	90
900	74	84	86	94
1000	77	87	89	93

confidentiality level of IoT-LSORWCHB method is higher than the three existing approaches ELIB model [1], BTNC protocol [2], and lightweight framework [3]. While conducting the experiments with 900 mobile nodes, the data confidentiality level of IoT-LSORWCHB method is 94%, whereas the data confidentiality level of ELIB model [1], BTNC protocol [2], and lightweight framework [3] is 74, 84, and 86%, respectively, in green computing. The result analysis of data confidentiality level is illustrated in Fig. 5.

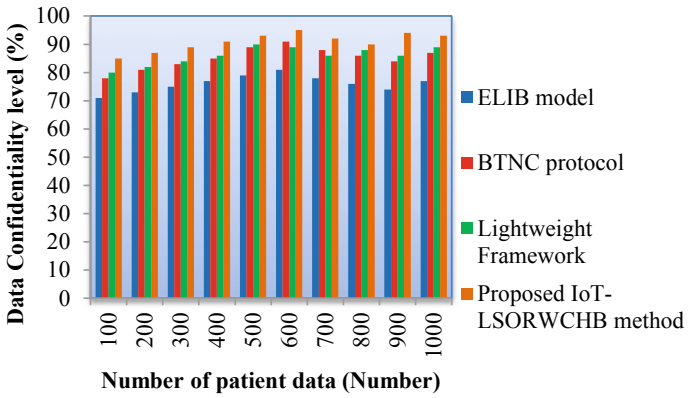


Fig. 5 Measurement of data confidentiality level

From above mentioned Fig. 5, the data confidentiality level of four methods is shown by four different colors. The above graphical results confirm that data confidentiality level is improved using IoT-LSORWCHB method than the existing techniques. This is due to the application of least square orthogonal regressive function and whirlpool cryptographic hash blockchain function. The regression function examines and classifies the patient data into various types of classes. The whirlpool cryptographic hash blockchain generates the hash value for every patient data by one-way compression function. The data with hash value is transmitted into the server in green computing. The comparison results prove that the data confidentiality level of IoT-LSORWCHB method is increased by 19, 7, and 6% when compared to the existing methods [1–3], respectively.

5 Limitations on Secured Communication with Blockchain and IoT

ELIB model verified their applicability in different IoT situations. ELIB model created the overlay network where equipped resources combined with public blockchain to authenticate the security and privacy. But, the data confidentiality was not improved by ELIB model. BTNC guaranteed better reliability for terminals in IoT. An illegal replacement attack described the security goals. The blockchain trusted network recognized mutual user authentication between terminals in IoT. However, the energy consumption was not minimized by BTNC. Lightweight framework used blockchain for anomaly detection in distributed atmosphere. Blockchain framework modernized the anomaly detection among IoT devices. But, the processing time consumption was not minimized by lightweight framework. Therefore, IoT-LSORWCHB method is introduced to analyze the patient data with blockchain for secured data communication.

6 Conclusion

A comparison of existing and proposed secured data transmission methods with blockchain and IoT is carried out. From the discussion, it is observed that the existing methods not reduced the energy consumption by BTNC. The review demonstrates that the processing time consumption was not minimized by lightweight framework. In order to address these problems, IoT-LSORWCHB method is introduced with higher authentication accuracy and lesser time consumption. The wide range of experiments on existing and proposed method determines performance of secured data transmission methods with its limitations. Finally from result, the proposed IoT-SORWCHB method reduced energy and time consumption with higher data confidentiality level during secured data transmission techniques.

References

1. Mohanty, S.N., Ramya, K.C., Sheeba Rani, S., Gupta, D., Shankar, K., Lakshmanaprabu, S.K., Khanna, A.: An efficient Lightweight integrated Blockchain (ELIB) model for IoT security and privacy. Fut. Gener. Comput. Syst. **102**, 1027–1037 (2020)
2. Zhang, J., Wang, Z., Shang, L., Di, Lu., Ma, J.: BTNC: a blockchain based trusted network connection protocol in IoT. J. Parallel Distrib. Comput. **143**, 1–16 (2020)
3. Mirsky, Y., Golomb, T., Elovici, Y.: Lightweight collaborative anomaly detection for the IoT using blockchain. J. Parallel Distrib. Comput. **145**, 75–97 (2020)
4. Dorri, A., Kanhere, S.S., Jurdak, R., Gauravaram, P.: LSB: a lightweight scalable blockchain for IoT security and anonymity. J. Parallel Distrib. Comput. **134**, 180–197 (2019)
5. Si, H., Sun, C., Li, Y., Qiao, H., Shi, L.: IoT information sharing security mechanism based on blockchain technology. Fut. Gener. Comput. Syst. **101**, 1028–1040 (2019)
6. Pavithran, D., Shaalan, K., Al-Karaki, J.N., Gawanmeh, A.: Towards building a blockchain framework for IoT. Cluster Comput. **23**, 1–15 (2020)
7. Ramezan, G., Leung, C.: A Blockchain-Based Contractual Routing Protocol for the Internet of Things Using Smart Contracts. Hindawi Publishing Coorporation, Security, Privacy, and Trust on Internet of Things, vol. 2018, pp. 1–14 (2018)
8. Rathee, G., Balasaraswathi, M., Prabhu Chandran, K., Gupta, S.D., Boopathi, C.S.: A secure IoT sensors communication in industry 4.0 using blockchain technology. J. Ambient Intell. Human Comput. **12**, 1–13 (2020)
9. Sahay, R., Geethakumari, G., Mitra, B.: A novel blockchain based framework to secure IoT-LLNs against routing attacks. Computing **102**, 1–26 (2020)
10. Krishnamurthy, R., Rathee, G., Jaglan, N.: An enhanced security mechanism through blockchain for E-polling/counting process using IoT devices. Wirel. Netw. **26**, 2391–2402 (2020)
11. Khalid, U., Asim, M., Baker, T., Hung, P.C.K., Tariq, M.A., Rafferty, L.: A decentralized lightweight blockchain-based authentication mechanism for IoT systems. Cluster Comput. **23**, 1–21 (2020)

Analysis of PCA Usage to Detect and Correct Skew in Document Images

Rajashekhar Salagar and Pushpa Patil

Abstract In recent years, all the documents are digitized by using different tools like camera and scanner. During the digitization, many mistakes are taken place, such as blur image, noise, and skew formation on entire document. These mistakes are directly influence on the consistency and productivity of the document image segmentation and analysis. In these mistakes, skew detection and correction are important part, and for this many approaches are introduced by different researchers. In this paper, principal component analysis (PCA) procedure is used to find the skew angle on full page. First convert the input document into binary form image, then use Sobel and Gaussian filters to find edges and suppressing the noise, respectively. Then apply the PCA to find the skew angle exists in a document. The PCA procedure first finds the covariance matrix, then generates the eigen values and eigen vectors, and from this calculate the unit vector for principal component. Later using principal component, identify relevant skew angle of the document. To correct the skew, rotate the test sample by preferred angle, and it can remove the skew occurred in document while acquiring the image. The evaluation of system methodology is carried out for different documents; the experimental results are significantly better.

Keywords Skew detection · Correction · Text document · PCA

1 Introduction

The most significance step in standard optical character recognition system is preprocessing, which emphases on improving the captured image to strengthening the easiness of characteristic parts obtaining and to return for the ultimate degradation of the scanned document image. The different document texts are scanned or captured for digitization, and input raw image may bring a certain quantity of distortions at that times. If the capture image holds distortion, it is subjected to the preprocessing

R. Salagar · P. Patil (✉)
Department of Computer Science and Engineering, B.L.D.E.A's, V.P. Dr.P.G.Halakatti, College of Engineering and Technology (Affiliated to Visvesvaraya Technological University, Belagavi 590018), Vijayapur 586103, Karnataka, India

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_65

stage where the removing of distortion of the image takes place. In addition, when a document is fed in to the scanner by the user, a certain quantity of skew deviation is introduced which is unavoidable. After the skew induction as part of the image skew correction is required, it is a procedure aims at identifying the deflection of the text document orientation for skew angle in the direction of the horizontal or vertical. In document image analysis and OCR approach, the detection and correction of skew are essential preprocessing steps. The OCR digital systems usually accept that the text documents are printed or handwritten with an uninterrupted text line direction and that the acquisition process did not add any relevant skew. The flat digital scanners needed to process massive amounts of documents at high accuracy made the above assumption unreliable, and the introduction of the skew estimation and correction phase has become compulsory. In fact, a small deviation of skew in the sheet is regularly introduced during processes such as scanning or copying or capturing from camera. Today, all the scanned documents are never freer styled. Furthermore, text line up along diverse directions is not an uncommon feature. The succeeding steps of OCR systems primarily depend upon the accuracy of preprocessing stage. So, the estimation and correction of skew are the important preprocessing document restoration steps before the actual document analysis is done. Imprecise deskew will considerably decline the succeeding processing stages which may lead to incorrect layout analysis, erroneous character or word or line segmentation, and misrecognition. The complete enhancement of the document processing and analysis system will be rigorously decreased due to the presence of skew.

Several skew estimating methods and procedures have been developed for overall document image processing. These procedures [1] are classified into three categories, based on the interline cross-correlation, Hough transform, and nearest neighbor cluster techniques. The proposed work of this paper is based on principal component analysis (PCA) method.

The left over part of the paper is organized into four sections. Section II deliberates associated hardwork carried out by many researchers. Section III proposes methodology for determining and correcting the skew using block diagram and explanation of each block. Section IV presents the analysis of the end results. Section V accomplishes the work carried out and some future work to be done.

2 Related Work

Many of the methods have been proposed as variations for skew angle detection in document images, but each method has some limits. Many of them have adopted different ways to succeed the detection and correction of the skew such as presented [2], Nandini et al., where initially Prewitt mask is used to find the text edges, then dilation is applied for object thickens. Later, connected component analysis is performed; for every component, coordinates of extreme pixels are calculated, and bounding box is defined which determines average height of all the components. Then finally, Hough transform is used to find skew angle. Shanbehzadeh et al. [3] proposed survey

on procedures for calculation of skew by use of projection profile analysis, Hough transform, cross-correlation, nearest neighbor cluster, and many more. Sablatnig et al. [4] introduced the focused nearest neighbor cluster outperform the results restricted to binarized images. Jain et al. [5] calculated the skew angle on Assamese text documents and compared using horizontal and vertical projection profile approach. Mehta et al. [6] used first and second level wavelet decomposition technique, then converted into square matrix, and then find the centroid moments of fragmented input image to determine the skew in document. Jundale et al. [7] used a methodology for detection of single line skew in handwritten script by using axes parallel rectangle and linear regression. Finally, the rotation transmission technique is used for correction of skew in a document. Ramanan et al. [8] used Wiener filter, radon transform, and smearing technique for skew detection, and to correct the skew rotation, transformation is used in opposite direction. Rezaei et al. [9] used adaptive algorithm and supervised learning technique to improve performance of the current skew detection methods with existing features. The approach is considerably improved with time and accuracy of skew estimation. Harraj et al. [10] applied Gaussian filter to convert gray scale; to extract edges, Canny edge detector is used, and then Hough transform with progressive probabilistic feature is applied to separate the object; at last, nearest neighboring interpolation is taken for correction of skew. Michelucci et al. [11] taken thresholding procedure to get the binary image, and morphological skeletonization is used to get the medial axes representation of the image. The line detection is done by progressive probabilistic Hough transform; then trigonometric concept applied for skew angle detection; lastly, the image is turn around in opposite direction as optimistic alignment. Jundale et al. [12] find skew of word or line or paragraph in handwritten text which introduced an approach by using pixels of the axes parallel rectangle and bisection procedure. First peripheral pixels are calculated for input binary image, and then axes parallel rectangle is formed. Next, by using bisection method, calculate the skew angle with minimum rectangular area. Then, direct method will get corrected image. Vinod et al. [13] considered a methodology of run-length-smoothed image algorithm (RLSA) to make correction of skew in Kannada text documents by using. First Haar wavelet decomposition is used for extracting text without losing any data, and then horizontal-vertical projection profiles are applied to find page border. Finally, RLSA is used to detect skew angle. To get the corrected image, the segmented line is rotated in desired angle. Liu et al. [14] proposed skew detection and correction procedure which uses the baseline features of Tibetan text and Hough transformation to obtain baseline, and then determine the skew of the original input image. The skew correction is done by acquiring the position of the boundary to understand the region of segmentation. To get the boundary, sequence of the activities to be performed by using median filtering, Sobel edge, Gaussian smoothing, and edge smoothing, removing the minor parts, and finding the border position. Raman [15], proposed different existing procedures for detection and correction of skew in printed documents containing dissimilar scripts like Tamil, English, Sinhala, and mixed-script.

3 Methodology

The methodology of proposed work contains two sub-parts. Part-1 is the preprocessing; it describes the transformation of gray scale input image into the binary image with thresholding optimization technique developed by Otsu [16]. Part-2 is the detection and correction of skew. It consists of calculating principal component and skew angle detection by using PCA.

3.1 Preprocessing

The binarization is the important phase as part of document image analysis and segmentation. The document image contains less intensity level with variations, and also brightness in the input image or overlapping of one object with another is imposed, which will affect the loss of important text information, by using threshold optimization method; we can resolve the above problems. In this method, gray scale input image is changed into binary image. In this output binary image, all the pixels of input are replaced by 1 (white) and 0 (black). The substitution is based on the luminance value; if the value is greater, then 1 is substituted and 0 is taken for remaining pixels. The specified range is [0, 1]. The range used is relative to the signal levels possible for the image's class.

3.2 Calculating Principal Component

The output of preprocessing phase binary image is considered as an input. In this image, foreground text pixels are black and background of text pixels are white. First, it is essential to map the obtained pixels of text image to vector space with a two dimensional. Consider an example, if an image contained pixels with black value in the direction of horizontal (x) and vertical (y), it is mapped in to the corresponding vector space (x; y). Then, unit vector of the projection profiles is calculated with maximum deviation.

Now, consider principal component of the given vector taken as unit vector with maxima value, and it is same to the eigen vector that is related with the biggest absolute eigen value of the set. The skew angle of the corresponding image is obtained based on the direction of the principal component.

3.3 Skew Detection and Correction

The principal component analysis is important tool in digital image processing application like image compression, face identification, and to find the skew orientation of automobile license plate. This method is useful for finding the patterns in information of the high dimension images. The method can also define as dimensionality reduction procedure for take-out only the important information from the large records set. The important computation for this method is calculation of eigen value and finding the vectors. The procedure is carried out as follows: first find the covariance matrix of the vector dataset; here size of the matrix is 2×2 dimensions, because the use of vectors having two dimensions. Then find the eigen values and its related eigen vectors. Arrange all the eigen value in maximum to minimum order. The order will show the importance of all the components of image, ignore the components having least value. The required principal component is selected based on the eigen vector with highest eigen value. Consider the principal component to calculate the skew angle; it is vertical to the text line direction in the image. The final skew corrected output is gained by turning the image in anticlockwise direction by that extent of skew angle; output is shown in Fig. 2. The complete algorithm is shown below.

The flow of PCA methodology is shown in Fig. 1 in which first take an input document image as shown in Fig. 2, then convert image into binary image as shown in Fig. 3, by using thresholding technique. Once binary image extracted from input gray image, then apply the Sobel filter to find the edge of the input document. Use the Gaussian filter to remove the noise, which exists in the image, may affect for image processing. Find the covariance matrix, and then calculate eigen value and eigen vector. Based on the eigen vector, find the principal component with respect to unit vector. Find the skew angle by using Principal component direction with perpendicular to text line. At last, skew corrected output image as shown in Fig. 4 obtained by rotating image in anticlockwise direction.

4 Experimental Results

The skew detection and correction of the text document are performed by using PCA approach. The proposed method is executed in MATLAB R2014a. The experiment was carried out on different 60 document images. The document image samples are taken from Tobacco-800 dataset [17]. All the document in dataset contains with diverse skew angle like the full skew on whole document are used. Table 1 displays the results of real skew angle and detected skew angle by using PCA technique. The overall skew detection percentage of the document database is 90.00%.

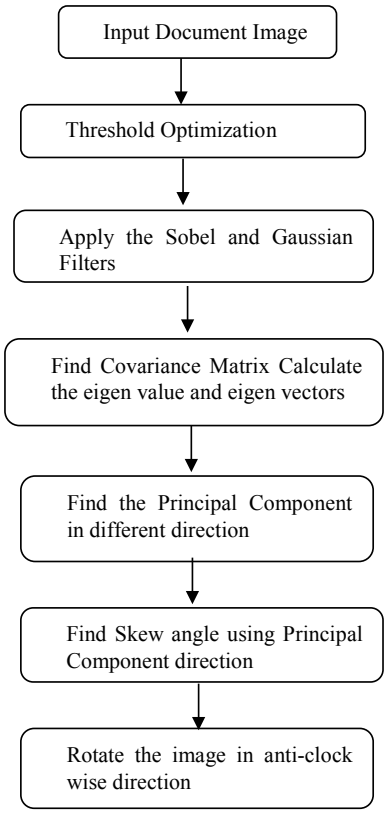


Fig. 1 Methodology

Fig. 2 Input image

Fig. 3 Binary image

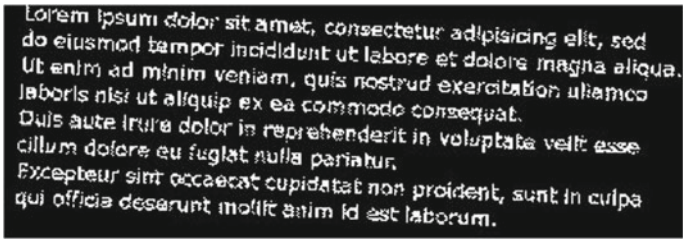


Fig. 4 Skew corrected output image

Table 1 Result analysis: skew angle in original document versus skew angle detected by proposed methodology

Document image name	Actual angle (°)	Detected angle (°)
Image_001	24.26	24.26
Image_002	25.50	25.10
Image_003	24.50	24.50
Image_004	24.20	24.20
Image_005	36	26
Image_006	26.10	26.10
Image_007	−20.30	−20.30
Image_008	23	23
Image_009	−18	−18
Image_010	−13	−13
Image_011	40	28

5 Conclusion

This detection and correction of skew approach are based on the principal component analysis method (PCA) for text documents. The methodology has two sub-sections. The first section consists of preprocessing of input image by using [16] Otsu method for to get binary image. The Sobel and Gaussian filters are used for edge detection and noise suppression. Second part consists of skew detection and skew correction using PCA. The PCA is applied for skew angle detection on entire document. In PCA, we are finding covariance matrix to check the relationships between dimensions in high-dimensional data, eigen values, and eigen vectors correspond to the dimensions that have the strongest correlation in dataset for finding principal component in the document. After detection of skew, we rotate image in anticlockwise direction for skew corrected document image. The presented work is encouraging on full page documents with good performance. The skew detection success rate for the document dataset is 90%. The future work is to support the efficiency of the procedure to get the more skew detection rate with varied skewed angles having complex background.

References

1. O'Gorman, L: The document spectrum for page layout analysis. IEEE Trans. Pattern Anal. Mach. Intell. **15**(11), 1162–1173 (1993)
2. Nandini, N., Srikanta Murthy, K., Hemanth Kumar, G.: Estimation of skew angle in binary document images using Hough transform. J. World Acad. Sci. Eng. Technol. **32**, 50–55 (2008)
3. Rezaei, S.B., Sarrafzadeh, A., Shanbehzadeh, J.: Skew detection of scanned document images. IMECS **1**, 451–456 (2013)
4. Kleber, F., Diem, M., Sablatnig, R.: Robust skew estimation of handwritten and printed documents based on gray value images. In: International Conference on Pattern Recognition, pp. 3020–3025 (2014)
5. Jain, B., Borah, M.: A comparison on skew detection of scanned document images based on horizontal and vertical projection profile analysis. Int. J. Sci. Res. Pap. Publ. **4**(6) (2014)
6. Mehta, S., Walia, E., Dutta, M.: A new fast approach for Skew estimation using moments and wavelet transform. In: International Conference on Image Processing Theory, Tools and Applications, pp. 221–226 (2014)
7. Jundale, T.A., Hegadi, R.S.: Skew detection of Devanagari script using pixels of axes-parallel rectangle and linear regression. ICESA 480–484 (2015)
8. Ramanan, M., Ramanan, A., Charles, E.Y.A.: A preprocessing method for printed Tamil documents: skew correction and textual classification. ICICIS 495–500 (2015)
9. Rezaei, S.B., Shanbehzadeh, J., Sarrafzadeh: Adaptive document image skew estimation. Int. Multi Conf. Eng. Comput. Sci. **1**, 423–433 (2017)
10. EL Harraj, A., Raissouni, N.: An improved algorithm for document image enhancing. CSCI 703–708 (2016)
11. Boudraa, O., Hidouci, W.K., Michelucci, D.: An improved skew angle detection and correction technique for historical scanned documents using morphological skeleton and progressive probabilistic hough transform. In: 2017 5th international conference on electrical engineering—Boumerdes (ICEE-B), Boumerdes, pp. 1–6 (2017)
12. Jundale, T.A., Hegadi, R.S.: Skew detection and correction of Devanagari script using interval halving method. Recent Trends Image Proces. Pattern Recogn. Commun. Comput. Inform. Sci. **709**, 28–38 (2017)

13. Vinod, H.C., Niranjan, S.K.: Multi-level skew correction approach for hand written Kannada documents. ICITS **2018**, 376–386 (2018)
14. Liu, H., Bi, X., Wang, W.: Layout analysis of historical Tibetan documents. In: 2019 2nd International Conference on Artificial Intelligence and Big Data (ICAIBD), Chengdu, China, pp. 74–78 (2019)
15. Ramanan, M.: A hybrid approach for skew detection and correction in the multi-script scanned document. Asian J. Res. Comput. Sci. **4**(2), 1–8 (2019)
16. Otsu, N.: A threshold selection method from gray-level histograms. IEEE Trans. Syst. Man Cybern. **9**(1), 62–66 (1979)
17. https://tc11.cvc.uab.es/datasets/Tobacco800_1

Viterbi Decoder with Configurable Constraint Length with Bit Error Correction for Satellite Communication

B. N. Akash, K. M. Amogh, Sindhu Sridhar, and Priyanka Agarwal

Abstract The Viterbi algorithm is an efficient method to decode convolution encoded data using the concept of maximum likelihood estimation. This paper presents the design and FPGA implementation of a Viterbi decoder for satellite communication. The design is coded in Verilog HDL using the Vivado 2017.4 tool and QuestaSim for behavioral and post-layout simulations. The decoder is designed to decode the output of rate half convolution encoder. It successfully decodes data for any constraint length and corrects error up to 4 bits. The implementation is done on the Zynq-7000 development board. The maximum operating frequency achieved is 221.9 MHz with a power consumption of 37.62 mW.

Keywords Viterbi decoder · Convolution encoder · Block code · Convolution code · FEC—Forward error correction · Reed–Solomon coding · Fano decoding · Bit error correction · Bit error detection · Constraint length · Window size · ARQ—Automatic repeat request · BPSK—Binary phase shift keying · FSM—Finite state machine · FPGA—Field-programmable gate array · Power analysis · Static timing analysis · Resource utilization · LUT—Lookup table

1 Introduction

Communication over a wireless medium is affected by signal attenuation and distortion due to noise at receiver which may result in incorrect data reception. Error correction and detection schemes can be used to ensure reliable communication. Forward error correction (FEC) schemes can be incorporated to achieve maximum throughput. It provides the ability of correcting data at the receiver without requesting for a retransmission. Block codes and convolution codes are two categories of forward error control schemes. Reed–Solomon coding is a scheme, where data is divided into blocks with a fixed length [1]. To improve the efficiency, the block length must be

B. N. Akash (✉) · K. M. Amogh · S. Sridhar · P. Agarwal
PES University, Bengaluru, India

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_66

increased, which may be difficult to accommodate in storage systems [2]. Convolution codes, on the other hand, process data as a stream [3]. Thus, they can be used to decode a continuous stream of incoming data, making them very suitable for digital video, mobile, and satellite communications [4].

Fano decoding and Viterbi decoding are two techniques that are widely implemented for convolution codes. The efficiency of Fano decoding is best seen in convolutional codes with long constraint lengths. But Fano decoding shows variable decoding times [3]. Viterbi decoding, on the other hand, has a high operational speed and gives good bit error correction results. It is also very easy to implement Viterbi decoders and hence used in many communication systems [3]. This paper presents a configurable Viterbi decoder which can operate on convolution codes encoded using any constraint length [5].

The algorithm presented can be implemented for varied constraint lengths and window sizes, and the decoding logic to iterate through all registers has been optimized. This reduces the time required to decode large streams of data as constraint length and window sizes increase. The implementation is successful in decoding a continuous stream of data, irrespective of its size. This makes the decoder extremely suitable in telecommunication and satellite communication applications. Since live telemetry can suffer from idle channel sequences which alters channel synchronization, the encoded data is inverted at every alternate bit to overcome this potential issue.

The paper is organized as follows: Sect. 2 gives an overview of the theory behind convolution encoder. Section 3 briefly describes the logic used in the Viterbi decoding algorithm and states the importance of bit error detection and correction. Section 4 explains the complete implementation of the design, Sect. 4.1 briefs the steps followed in the algorithm implemented, Sect. 4.2 depicts the RTL design of the encoder and decoder, Sect. 4.3 shows the test setup where the design is implemented on FPGA, and Sect. 4.4 gives a detailed explanation of the simulation results obtained while testing the design, organized under behavioral and post-layout simulation results; Sect. 4.5 gives the error correction results obtained when error was induced to the telemetry data; Sects. 4.6 and 4.7 give results for power and static timing analysis and the resource utilization on the FPGA, respectively. Section 5 summarizes the results obtained and compares them with other implementations that are similar to the work presented in this paper. Section 6 concludes on the work carried out in the paper. Finally, the references are used to carry out the implementation.

2 Convolution Encoder

Convolutional codes are used to continuously encode a serial input into an output with more number of bits by means of adding redundancy. The encoded output depends on the current input as well as some of the previous inputs. A convolutional code is characterized by two important parameters, namely the constraint length and the code rate. Constraint length (K) describes the number of memory stages required to

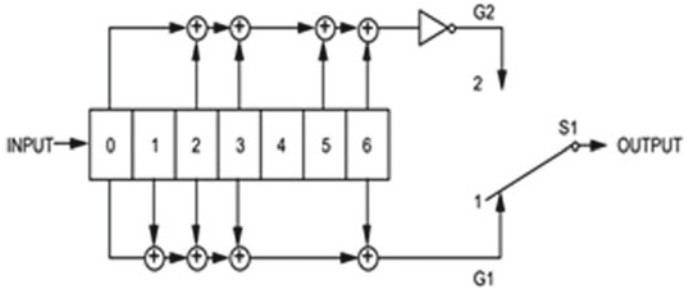


Fig. 1 Convolution encoder block for $K = 7$

output n bits of data from a k-bit input and is analogous to a sliding window. The constraint length directly affects the code's resistance to bit errors, the time taken for decoding, and the hardware complexity of the decoder.

Code rate is the ratio of the number of input bits (k) to the number of output bits (n), where ($n > k$). A half-rate code means that for every input bit, two bits of output are produced. Some practical applications also require code rates up to $(n - 1)/n$.

A convolution encoder can be viewed as a finite state machine (FSM) of shift registers. The output is a modulo-2 operation of the bits from a selection of shift registers. A rate half convolution encoder is implemented for varying constraint lengths to validate the Viterbi decoder. The connection vectors chosen for constraint length when $K = 3$ are $(111)_2$ and $(101)_2$, when $K = 7$ are $(1{,}011{,}011)_2$ and $(1{,}111{,}001)_2$, and so on.

The two channels of output of a half-rate are represented as Y1 and Y2 in this paper. The outputs of Y1 and Y2 are interleaved to produce a single stream of output with double the input data rate. Y1 and Y2 are dependent on the current and previous inputs through a predefined XOR logic represented as a connection vector. The symbol generation polynomial operation for constraint length 7 is shown in Fig. 1 [6].

In practical applications, there is a possibility of having long strings of 0's or 1's as the incoming data sequence. This can create a loss of clock synchronization at the receiving end due to which the exact number of 0's or 1's cannot be determined. The convolution encoder implemented in this paper and designed for satellite application is CCSDS standard compliant [6]. It inverts one of the output channels before transmission, thereby ensuring that the decoder never receives a long, idle data sequence.

3 Viterbi Decoder

Viterbi decoder works on the maximum likelihood logic using Trellis diagram to decode convolutional codes [7]. The most likely path is backtraced along a Markov

graph. When implemented practically, the output of the decoder may differ from the actual input due to channel noise. The paper successfully implements data regeneration up to 4-bit errors for every 35 incoming bits [8].

3.1 Bit Error Detection and Correction

Bit error detection and bit error correction are two important operations required for end-to-end reliable communication. Bit error detection involves the addition of redundant bits (or check bits) to the original data. Bit error correction ensures reliable data transmission of digital data by reconstructing the original data. Automatic repeat request (ARQ) schemes use acknowledgments for error detection and retransmissions for reliable delivery.

4 Implementation

The design consists of two major blocks: Viterbi node and Viterbi decoder.

Viterbi node instantiates the nodes of the Trellis diagram based on the encoding function used in the convolution encoder. It adds the branch metric to the path metric for two cases of input bits, i.e., 1 and 0, compares the two path metrics at each stage, and finally selects the bit corresponding to the lower path metric.

After receiving the decoded bits for the data size equal to the window size times the constraint length, the Viterbi decoder selects the set with the lowest path metric.

The consolidated operation of the Viterbi node and the decoder is shown as a block diagram in Fig. 2.

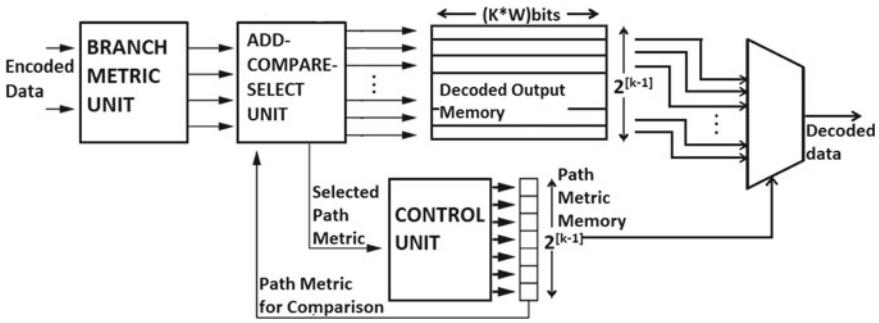


Fig. 2 Block diagram of the Viterbi decoder with constraint length '*K*' and window size '*W*'

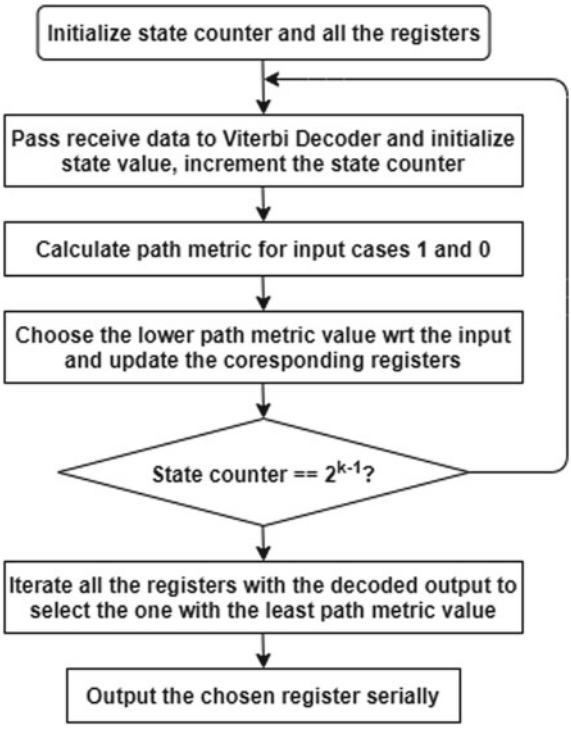


Fig. 3 Flowchart of the decoding algorithm

4.1 Flowchart

Figure 3 depicts the stepwise implementation of the algorithm. This flow is replicated for each window traversed along the encoded stream of data.

4.2 RTL Design

Figure 4 shows the RTL model implemented using Vivado 2017. The design consists of a 1 bitstream of data into the convolution_encoder, which is encoded to generate 2 bitstreams. These encoded data streams are input to the next block, i.e., the viterbi_decoder. The Viterbi decoder block has $2^{(K-1)}$ number of node instantiation blocks called viterbi_node. The final bitstream of output from the decoder is labeled as viterbi_out.

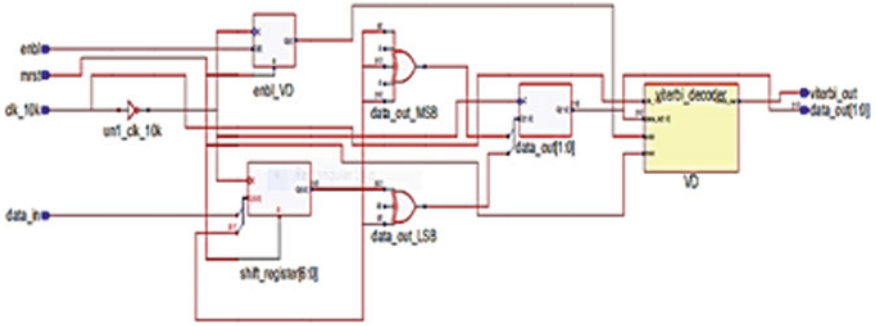


Fig. 4 RTL schematic of convolution encoder and Viterbi decoder

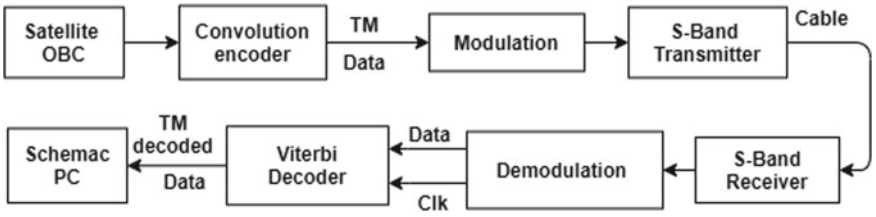


Fig. 5 Test setup for Viterbi decoder

4.3 FPGA Implementation of the Design

The design was implemented and tested on Xilinx Zynq-7000 development board (Zedboard). The data encoder for the satellite telemetry data (TM) was implemented on the Microsemi Proasic3e board. The coded data at 20 KBPS was transmitted by BPSK modulating the locally generated S-band carrier 2240 MHz. The modulated data was down converted to 70 MHz on the receiver side, demodulated, and then fed to Viterbi decoder for data retrieval. The retrieved satellite data was displayed on the Schemac PC for verification. Figure 5 shows the test setup for the designed scheme, which was part of the actual implementation.

4.4 Simulation Results

Libero SoC SmartTime tool was used to perform post-layout timing and power analysis. The maximum operating frequency was found to be 221.9 MHz.

Behavioral Simulation. The behavioral simulation was performed on the testbench designed for the convolution encoder and Viterbi decoder. It was observed that the

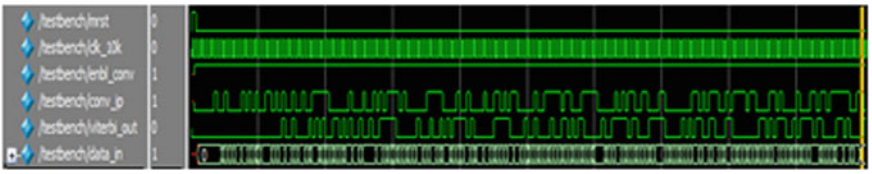


Fig. 6 Post-layout simulation of the design using Libero SoC design suite

decoded output (viterbi_out) was just a delayed version of the data input into the convolution encoder (conv_ip).

Post-Layout Simulation. The post-layout simulation was performed after performing the place and route of the design using Libero SoC (for Proasic3e). The results for the same are shown in Fig. 6. The signal 'conv_ip' is the randomly generated incoming serial data to the convolution encoder. The 2 bit output of the convolution encoder is shown as the signal 'data_in', which is passed into the Viterbi decoder. The output of the Viterbi decoder is another bitstream of data shown as the signal 'viterbi_out', which matches with the 1 bit input stream 'conv_ip'. This proves that the output data is consistent with the input data, or in other words, the encoding and decoding operations are performed coherently.

4.5 Error Correction Results

The bit error correcting capabilities of the decoder was verified through simulation. The design was able to correct up to 4 bits of continuously injected errors. The simulation results of the same are shown in Fig. 7. It is observed that the signal 'viterbi_out', i.e., the output of the Viterbi decoder, is the same, but a delayed version of the signal 'conv_ip', i.e., the input to the convolution encoder. The signals 'onebit_out' and 'error_out' differ by 4 bits per every ($K * W$) bits, i.e., $(7 * 5) = 35$ bits.

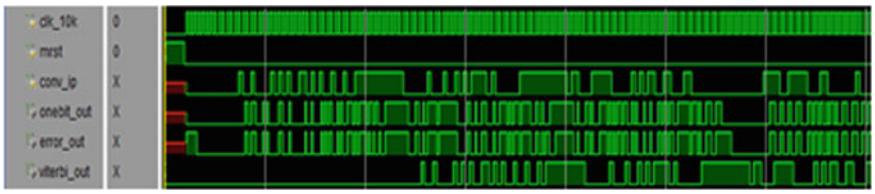


Fig. 7 Error correction simulation for the design using Vivado for $K = 7$ and $W = 5$

4.6 Power and Static Timing Analysis

The static power analysis for the Viterbi decoder was carried out using the Libero SoC SmartTime tool. The total power usage of the design was found to be 37.62 mW. The power analysis results are shown in Fig. 8.

The static timing analysis for the design was also carried out for both min delay and max delay constraints, considering both the worst case and the best case scenarios. A non-negative slack was observed in both max and min timing analyses, inferring that the design can handle internal delays without delaying the final output. The min timing slacks are shown in Fig. 9, and the max timing slacks are shown in Fig. 10.

The worst negative slack (WNS) was found to be 5.494 ns, worst hold slack (WHS) was found to be 0.048 ns, and the worst pulse width slack was observed to be 3.00 ns when the time period was 10 ns. The positive values indicate that all setup and hold timing requirements are met, and the design passes through the entire path.

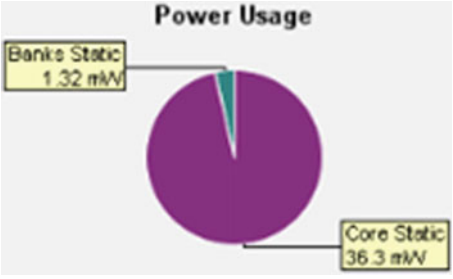


Fig. 8 Power used by the system when $K = 7$

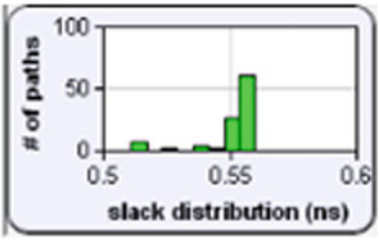


Fig. 9 Min timing report when $K = 7$

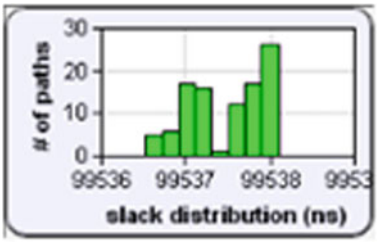


Fig. 10 Max timing report when $K = 7$

Table 1 Resources used when $K = 7$ for Proasic3e

Resource	Used instances
Core: Combinational + Sequential	14,399 = (13,099 + 1300)
IO (with clock)	7
Global (chip)	6

Table 2 Resources used when $K = 7$ for Zedboard. Resource utilization (Vivado)

Resource	Used instances
Slice LUTs	7127
Slice registers	1120
MUXs: F7 + F8	2728 = (1989 + 739)

4.7 Resource Utilization Results (Libero)

Libero SoC Design Suite was used to synthesize the design and implement the same for Microsemi Actel Proasic3e FPGA. The results for the total resource utilization after place and route of the design with constraint length 7 are given in Table 1. It is also observed that the percentage utilization of all resources is below 20.

Vivado 2017.4 tool was used to synthesize and implement the design for Xilinx Zynq development board (Zedboard). The results for the same are given in Table 2. The percentage utilization is negligible when compared to the number of available resources.

For a constraint length of 3, the number of lookup tables (LUT) utilized is found to be 95.

5 Results and Discussions

The efficiency of the decoder can be measured with respect to its frequency of operation, i.e., 221.9 MHz, resource utilization, power usage, i.e., 37.62 mW, and bit error correction, i.e., effective data retrieval for up to 4 continuous bit errors in the data stream.

Tang et al. [8] present a paper with similar parameters of constraint length (k) equal to 7 and window size equal to 5 to decode convolutionally encoded data at the rate of half. The bit error correction is successful if less than 5 bits out of a 41 bit data stream, which is achieved using 7 * 5 message bits per window tailed with 6-0 bits, are corrupted. The resource utilization in terms of combinational ALUTs is 13832 for Stratix III FPGA. The proposed architecture can correct up to 4 continuous erroneous bits in a 35 bit stream of data. The combinational resource utilization, however, is 13099 for Proasic3e FPGA.

Tessier et al. [9] discuss a reconfigurable Viterbi decoder for constraint lengths 4–14. The average power consumption is 98.8 mW. The decoder designed in this paper has an average optimized power consumption of 37.62 mW.

Middya et al. [11] present a paper which implements a decoding algorithm for constraint length 3 and operating frequencies ranging from 246 to 426 MHz. The resource utilization in terms of LUTs is 112. The design explained in this paper is optimized to 95 LUTs for constraint length 3.

Zhao et al. [12] discuss a forward traceback methodology to reduce decoding latency. The average power consumed is about 46 mW at a frequency of operation of 600 MHz.

6 Conclusion

This paper presents a flexible design of the Viterbi decoder which can be configured for any constraint lengths. By inverting one of the 2 bits of the encoded data, synchronization problems are avoided when there is a long string of 0 s or 1 s. There is efficient usage of power, results showing 37.62 mW for constraint length 7 [12]. The bit error correction performance is better than previous design implementations, as the proposed design showed correct decoding results for up to 4 continuous bit errors in the encoded data [8]. The design architecture uses 95 LUTs [11]. The algorithm implemented can be easily scaled across multiple constraint lengths. It is very suitable for satellite communication and related applications, as it has the ability to decode continuous streams of data.

References

1. Liu, Y., Guan, Y., Zhang, J., Wang, G., Zhang, Y.: Reed-Solomon codes for satellite communications. In: IITA International Conference on Control, Automation and Systems Engineering (2009)
2. Sklar, B.: Reed-Solomon codes
3. Gupta, K., Ghosh, P.K., Piplia, R.N., Dey, A.: A comparative study of Viterbi and Fano decoding algorithm for convolution codes. In: International Conference on Methods and Models in Science and Technology (ICM2ST-10) (2010)
4. Viterbi, A.J.: A personal history of the Viterbi Algorithm
5. Dhaliwal, S., Singh, N., Kaur, G.: Performance analysis of convolutional code over different code rates and constraint length in wireless communication. In: International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC 2017) (2017)
6. Consultative Committee for Space Data Systems (CCSDS) Historical Document-Telemetry Channel Coding (1999)
7. Yamamoto, H., Itoh, K.: Viterbi decoding algorithm for convolutional codes with repeat request. In: IEEE Transactions on Information Theory, vol. IT-26, No. 5 (1980)
8. Tang, J.: Design and FPGA implementation of a viterbi decoder: a case study using system Verilog and co-simulation. In: IEEE International Symposium on Signal Processing and Information Technology (ISSPIT) (2009)

9. Tessier, R., Swaminathan, S., Goeckel, D., Burleson, W.: A reconfigurable, power-efficient adaptive Viterbi decoder. In: IEEE Transactions on Very Large Scale Integration (VLSI) Systems, vol. 13, No. 4 (2005)
10. Mamarde, R., Khoje, S.: Viterbi decoder using Zynq-7000 AP-SoC. In: Proceedings of the Second International Conference on Intelligent Computing and Control Systems (ICICCS 2018) (2018)
11. Middya, A., Dhar, A.S.: Real-time area efficient and high speed architecture design of Viterbi decoder. In: International Conference on Advances in Electrical, Electronics, Information, Communication and Bio-Informatics (AEEICB16)
12. Zhao, X., Li, H., Wang, X.: A high performance multi-standard Viterbi decoder. In: 7th IEEE Conference on Electronics Information and Emergency Communication (ICEIEC) (2017)

Dengue Fever Prediction

Md. Habibur Rahman, Md. Omar Faroque, and Farhana Sarmin Tithi

Abstract Dengue is a very common disease typically found in a widespread hot region. It is a deadly disease caused by female Aedes mosquitoes. Patients who have diagnosed with different types of dengue needed different types of treatment. Many experts are experimenting to recognize and exploring some of the new features on dengue disease from last few years. In this research paper, we have predicted dengue fever by implementing various machine learning algorithms such as two boosted decision tree, two-class Bayes point machine, multiclass decision forest, and boosted decision tree regression in a dataset. This dataset is collected from our generated survey of different people who are currently affected by dengue fever or already suffered from dengue fever. We also used tenfold cross-validation to estimate the performance of our machine learning model. We have also used Azure machine learning studio to predict and evaluate data, and we also compare the performance of all techniques that we have used. We also showed the accuracy of all the classification and regression technique. This research paper is the first unique paper based on Bangladesh region which will be used to detecting the dengue fever with 95% accuracy. People will be aware of this dangerous disease and can take necessary actions.

Keywords Dengue fever classification · Dengue fever regression · Tenfold cross-validation · Machine learning · Prediction analysis

Md. Habibur Rahman (✉)
Department of Computer Science and Engineering, Daffodil International University, Dhaka, Bangladesh

Md. Omar Faroque
Department of Computer Science and Engineering, Move Fitness, Frisco, USA
e-mail: omar@letmove.fit

F. S. Tithi
Department of Computer Science and Engineering, United International University, Dhaka, Bangladesh

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_67

1 Introduction

Dengue is one of the most dangerous malady caused by dengue fever germ that injects in the human body by a female mosquito called Aedes. Headache, blooding, joint pain, muscular pain and rush proof, high fever, pain behind the eyes, swollen glands, and vomiting are the indication of dengue [1]. It is also known as a bone-shaking malady [2, 3]. Around all over the world, about 2.5 billion people have infected by dengue fever. About fifty to hundred million people all around the world are suffering from dengue fever every year [4]. Bangladesh is also affected by dengue viruses, and it is increasing rapidly [5]. The main biting insect or tick that transmits the virus that causes the disease called dengue is the Aedes female mosquito. Human is infected by the virus while a female Aedes mosquito bite a person and inject the virus into the body through feeding on the blood. Once the human is infected, they became a main bearer and multipliers of the virus. The virus circulates in the infected human blood for two to seven days. Approximately at the same time, human are suffering from dengue fever [4].

A variety of techniques can be used to classify dengue fever such as two boosted decision tree, two-class Bayes point machine, multiclass decision forest, and boosted decision tree regression. These techniques are evaluated based on four of the mentioned machine learning algorithm to determine or compute accuracy, precision, sensitivity, specificity, and negative rate, and tenfold cross-validation is used to estimate and verify the accuracy and performance of our machine learning model.

2 Literature Review

There has been many extensive research on the detection of dengue fever with the use of ML algorithms and surveys. Most of those works worn out the context of developed countries. Although the dengue information is not systematic and also there is some scarcity of information in these research work, a number of researchers have researched on dengue fever prediction and classification such as Tener et al. and Tarig et al. Tenner's group classified approximately 1200 patients and found some notable features using the decision tree approach. They have got about 84% accuracy in their experiment [6]. Tarig's group have used clustering and made patients into two groups, in this two sets, they used ML feedforward neural networks (MFNN) and self-organizing MAP (SOM), but they have got only 70% accuracy [7]. Fatimah Ibrahim et.al have got 90% accuracy by using machine learning perceptrons (MLPs) [8]. Darani et al. recommended to categorize dengue patients from two datasets using the decision tree method [9]. They have got very good accuracy 97.6 and 96.6% from the first and second research, respectively. The accuracy of both research was greater than 90% in the unseen test set, but the 0 accuracy was too low in a day research, and the decision tree was found to be over fitted. Therefore, the tentative results showed that the decision tree approach did not counterpart this task too much.

3 Objective

The main goal of this paper is to use some classification and regression algorithm methodology to detect dengue fever by analyzing a few data. We have collected 1746 dengue patient's data by survey and generate training and testing data based on the survey. We compare different algorithm score and evaluate with the help of classification techniques. We also use cross-validation for estimating the skill of our machine learning model. We have used Azure machine learning studio to implement machine learning model.

4 Methodology

We have used Microsoft tools called Azure machine learning studio for building, testing, and executing our model. Microsoft Azure machine learning studio is a GUI based joint development environment, where you can use drag and drop tools to build, test, and deploy learning models as Web models that are easily used by custom tools. We are using some machine learning algorithm such as two boosted decision tree, two-class Bayes point machine, multiclass decision forest, and boosted decision tree regression for predicting dengue fever. Our main goal is to check whether or not a patient is stricken by dengue fever or not by making use of some characteristics. We

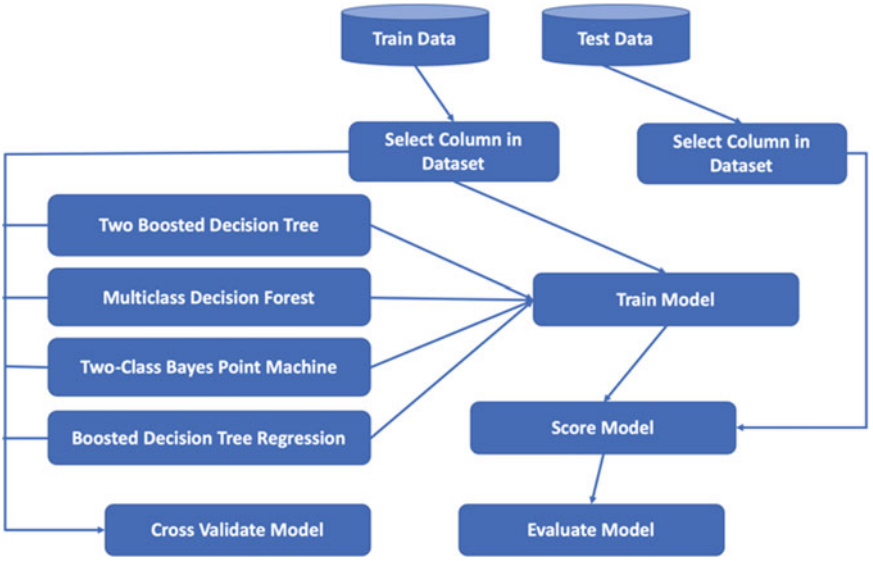


Fig. 1 Working process of dengue prediction model

will show the accuracy of all the classification and regression technique. In "Fig. 1," the process is shown in a diagram.

4.1 Dataset

The dataset which is used in dengue prediction model is a collection of related sets of information. In our dataset, Table 1 has 14 columns and 1746 rows in total. Dataset is collected from our own created online Google form. We have created an online form and shared this form on Facebook, LinkedIn, and many other social site where people put their data in our online form. In each, column 0 represent "No" and 1 represent "Yes" except gender, age, number of days, and UID. UID has been used for finding the unique number of rows in our dataset. In our research paper, we have used "Table 1" for attribute description.

Table 1 Attribute description

Attribute name	Description
Dengue	1 for Yes 0 for No
Age	0 represent 1–2 years old person, 1 represent 3–4 years old person, 2 represent 5–9 years old person, 3 represent 10–14 years old person, 4 represent 14–19 years old person, 5 represent 20–24 years old person, 6 represent 25–29 years old person, 7 represent 30–34 years old person, 8 represent 35–39 years old person, and 9 represent 40 or more than 40 years old person
Gender	1 represent "Male" and 0 represent "Female"
Days	0 represent a person suffering from fever for 1–2 days, 0 represent a person suffering from fever for 1–2 days, 1 represent a person suffering from fever for 3–4 days, 2 represent a person suffering from fever for 5–9 days and 3 represent a person suffering from fever for 10–20 days
High fever	Fever temperatures in °F as an Example (103°F)
Rash	1 for Yes 0 for No
Muscle pains	1 for Yes 0 for No
Joint pains	1 for Yes 0 for No
Blooding	1 for Yes 0 for No
Vomiting	1 for Yes 0 for No
Severe headache	1 for Yes 0 for No
Pain behind the eyes	1 for Yes 0 for No
Swollen glands	1 for Yes 0 for No
UID	Unique ID

4.2 Data Preprocessing

We have collected data from 1746 people survey using our generated online form. The data we collected from dengue affected a person and who are suffering from fever for evaluating and training the model we proposed. The sample of our dataset description is given in "Table 1". In our dataset, we have 14 columns and 1746 rows in total. Dataset is collected from different online sector. There are some data we have collected are not related to our proposed dataset. These kinds of data are noisy and less important or less related data. Firstly, we find out all the unique column. Unique features indicate how similar information is entered into datasets with similar features. After that, we transformed categorical values to numerical values, and we removed UID for getting an accurate result.

4.3 Categorical Features Factoring

Before data visualization, feature analysis is a vital part, and without feature analysis or lacking in feature factoring, prediction accuracy will not be ideal and overfitting situation may occur. The dataset must be cleaned and factored prior to model training and testing. In the sample data, we have faced a minor problem that is some column with missing information. We have solved this problem using Microsoft Excel for finding those missing field, and we have deleted those field to make our dataset clean.

4.4 Training the Dataset

In the previous section, we have cleaned our dataset for training and evaluation of our model. We have used two boosted decision tree, two-class Bayes point machine, multiclass decision forest and boosted decision tree regression for training our model. Short description of each algorithm is given in "Table 2".

4.5 Score Model

To generate predictions, we used score model. Score model is Azure machine learning module which is used for predictions by the trained classification and regression model. To generate a group of metrics used for evaluating the model's accuracy (performance), we tend to connect the scored dataset to judge model. Score model outputs a predicted value for the class and moreover because of the probability of the predicted value. We have used four machine learning algorithms; the description

Table 2 Used algorithm description

Attribute name	Description
Two boosted decision tree	A two boosted decision tree is the aggregate learning system in which the second tree corrects errors of the first tree and third tree correct errors of first and second tree and so forth. The entire aggregate of trees together makes the prediction. Output are given in Table 3
Two-class Bayes point machine	This algorithm uses a Bayesian approach to a linear classification called the "Bayes point machine". The algorithm efficiently approximate the Bayes point, and it is not prone to overfitting the model. Output are given in Table 4
Multiclass decision forest	A multiclass decision forest is an aggregate learning system model that very swiftly makes several decision trees while learning from tagged data. Output are given in Table 5
Boosted decision tree regression	This is a supervised learning method where each tree is dependent on prior trees. Trees are assembled to improve accuracy. Output are given in Table 6
Cross-validation	Cross-validation is for estimating the skill of our machine learning model. Output is given in Table 6

Table 3 Two boosted decision tree evaluation result table

Attribute name	Measure
Accuracy	0.955
Recall	0.944
Precision	0.985
F1 score	0.964
Threshold	0.5
AUC	0.968

Table 4 Two-class Bayes point machine evaluation result table

Attribute name	Measure
Accuracy	0.855
Recall	0.926
Precision	0.857
F1 score	0.890
Threshold	0.5
AUC	0.924

Table 5 Multiclass decision forest evaluation result table

Attribute name	Measure
Average accuracy	0.938326
Averaged precision	0.930799
Averaged recall	0.940079

Table 6 Boosted decision tree regression evaluation result table

Attribute name	Measure
Mean absolute error	0.204113
Root mean squared error	0.857143
Relative absolute error	0.147135
Relative squared error	0.17953
Coefficient of determination	0.82047

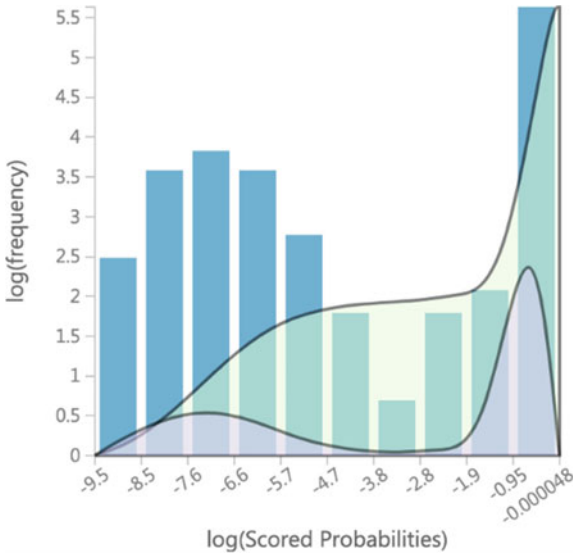


Fig. 2 Probabilities histogram of two boosted decision tree

is given in "Table 3". For each algorithm, scored probabilities histogram and cross-validation are given in "Figs. 2, 3, 4, and 5".

4.6 Evaluation Model

We have used evaluation model to determine or compute the accuracy of trained data. Evaluation model is a module created by Azure machine learning which is used to measure the accuracy from the score model. It also gives us accuracy, precision, recall, F score, AUC, average log loss, and training log loss. Receiver operating characteristic of two boosted decision tree and receiver operating characteristic of two-class Bayes point machine graph are given in "Fig. 5". For each algorithm, measurement is given in "Tables 4, 5, and 6," respectively.

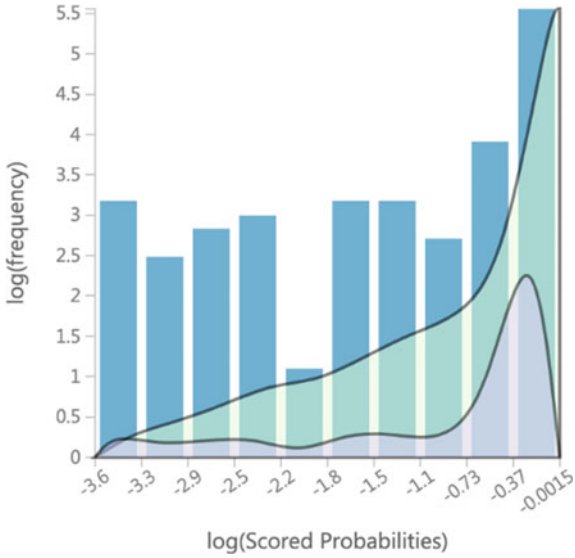


Fig. 3 Probabilities histogram of two-class Bayes point machine

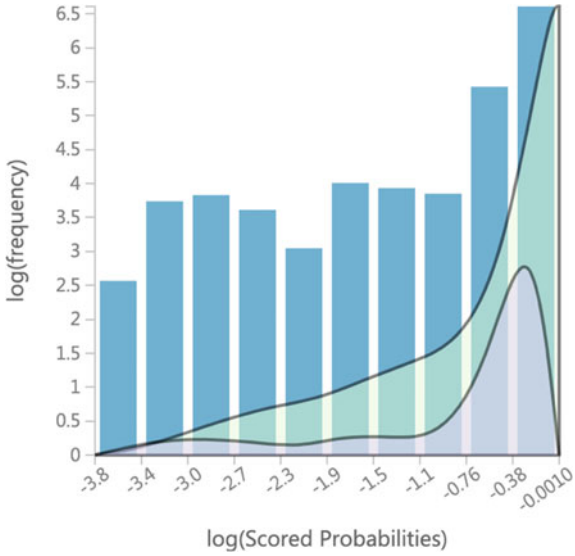


Fig. 4 Probabilities histogram of two boosted decision tree cross-validation

5 Comparison

With four algorithms of machine learning, we have completed training and got the accuracy of two boosted decision tree, two-class Bayes point machine, multiclass decision forest, and boosted decision tree regression. Two boosted decision tree gives

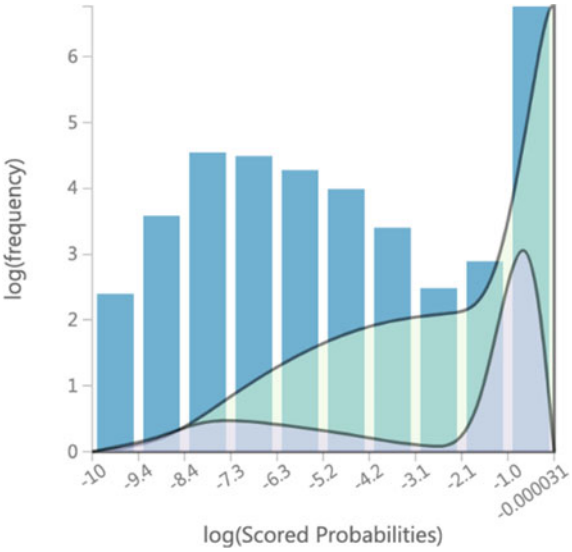


Fig. 5 Probabilities histogram of two-class Bayes point machine cross-validation

Table 7 Accuracy comparison table

Algorithm name	Accuracy
Two boosted decision tree	0.96
Two-class Bayes point machine	0.855
Multiclass decision forest	0.938326
Boosted decision tree regression	0.9333

us the best accuracy 0.96. The below table gives the evaluation of all the techniques and algorithms accuracy on "Table 7".

6 Conclusion

The main objective of this paper is to predict dengue fever using machine learning. For prediction, we have used four algorithms and some module from Azure machine learning. Two boosted decision tree gives us best accuracy 96%. By using our tool, people can know when they need a diagnostic test for dengue fever or not. We will hope that many lives will be saved by using this tool.

References

1. Farooqı, W., Ali, S., Wahab, A.: Classification of Dengue fever using decision tree. VAWKUM Trans. Comput. Sci. **3**(2) (2014)
2. Rigau-Pérez, J.G., Clark, G.G., Gubler, D.J., Reiter, P., Sanders, E.J., Vorndam, A.V.: Dengue and dengue haemorrhagic fever. The Lancet **352**(9132), 971–977 (1998)
3. Mayo Clinic: Retrieved from Retrieved from https://www.mayoclinic.org/diseases-conditions/dengue-fever/symptoms-causes/syc-20353078 (1998)
4. Farooqi, W., Ali, S.: A critical study of selected classification algorithms for dengue fever and dengue hemorrhagic fever. In: 2013 11th International Conference on Frontiers of Information Technology, pp. 140–145. IEEE. 1955 (2013)
5. World Health Organization Bangladesh: Retrieved from https://www.searo.who.int/bangladesh/dengue/en/ (2009)
6. Vong, S., Khieu, V., Glass, O., Ly, S., Duong, V., Huy, R., et al.: Dengue incidence in urban and rural Cambodia: results from population-based active fever surveillance. PLoS Neglect. Trop. Dis. **4**(11), e903 (2006–2008) (2010)
7. Faisal, T., Ibrahim, F., Taib, M.N.: A noninvasive intelligent approach for predicting the risk in dengue patients. Exp. Syst. Appl. **37**(3), 2175–2181 (2010)
8. Ibrahim, F., Taib, M.N., Abas, W.A.B.W., Guan, C.C., Sulaiman, S.: A novel dengue fever (DF) and dengue haemorrhagic fever (DHF) analysis using artificial neural network (ANN). Comput. Methods Progr. Biomed. **79**(3), 273–281 (2005)
9. Thitiprayoonwongse, D, Suriyaphol, P., Soonthornphisaj, N.: Data mining of dengue infection using decision tree. Entropy **2** (2012)

IoT Assisted Predictive Maintenance and Worker Safety: An Initiative

B. C. Kavitha and R. Vallikannu

Abstract The Internet of Things (IoT) is a seemingly and fast developing ecosystem which can transform our lives in a better way in which the 'things' act in a way what we really want them to. IoT can build a world of automation by connecting devices, machines, and people. Industrial IoT (IIoT) is a field where IoT can contribute much to improve the productivity, reliability, efficiency thereby contributing more toward the economic growth of our country. Industrial safety and environmental responsibility goes in par with the above advancements. Advances in IoT technology have given rise to concept of connected worker which can be realized using wearable and embedded sensors by which the workers can be monitored to prevent accidents due to fluids or gas leakage, explosion, overexertion, and falls. Predictive maintenance is a most sought-after area in IIoT. This paper deals with the predictive maintenance and development of wearables for enhancement of worker safety in the industrial floor.

Keywords Internet of Things · IIoT · Sensors · Predictive maintenance · Wearable technology · Big data analytics

1 Introduction

The Internet of Things (IoT) is an ecosystem of connected physical objects that are accessible through the Internet. IoT permits billions of devices containing embedded technology to seamlessly connect, be managed, and firmly move over the network [1, 2]. Internet of Things will improve the standard of human life by making the objects act in the means what individual wish them 'to act'. At present, there is a great demand for improving industrial production and efficiency. Development of a cost-effective system which can provide high efficiency with no compromise on worker safety is the need of the hour. This can be made possible by deploying various types of sensors and wireless devices in the industry floor. This in turn paved the

B. C. Kavitha (✉) · R. Vallikannu
Department of ECE, Hindustan Institute of Technology and Science, Chennai, India

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_68

way for developing IoT systems for industrial applications, which was named as Industrial IoT (IIoT). By the effective deployment of sensors at different locations of the factory floor, big amount of data can be collected. The big data from different sources and time periods can be analyzed to derive cost-effective solutions, which can enhance the productivity and safety of workers providing ecofriendly solutions. The 'smart objects' collect great deal of information, communicate among themselves, and might connect via net which can provide suggestions to regulate them from any part of the world [3]. The introduction of IIoT can enhance lives providing suggestions for automation raising the potency, productivity at the same time taking into thought safety, and environmental responsibility. The wearable technology join hands with the Internet of Things for the gathering, integrating, and analyzing big data gathered from sensors. Wearable technology when coupled together with innovative cognitive capabilities and external sources like environment/ weather can pave the way to a better and safer work place.

Nowadays, IIOT is mainly focused on predictive maintenance and worker safety. This paper is based on improving the safety of workers, by continuously monitoring the conditions inside the Industry and providing security measures to workers to avoid accidents. Predictive maintenance in industries can be achieved only if a sensor interface is introduced to connect all the active sectors of an Industry in an IoT based environment, and the concept can be realized by integrating wireless sensor networks (WSN) and IoT. The idea of enhancing worker safety can be implemented by utilizing and implanting sensing elements like different sensors for temperature, gas, flow, humidity, MEMS, heart beat sensor, and GPS sensors that aid worker safety providing environmental responsibility [4]. The paper is organized as follows: methodology, results and discussion followed by conclusion and references.

2 Methodology

The twenty-first century is an era of intelligence and information. Human beings expect a safe and secure lifestyle constructed by an intelligent social infrastructure of IoT. This expectation motivates the IIoT to emerge as the next big revolution in manufacturing and process optimization, providing a platform for business innovation at reduced costs. IIoT is about connecting embedded system technology, bringing together two different technologies resulting in full automation, i.e., with no human intervention [5]. The objective of this work is predictive maintenance and to enhance safety of workers in the industrial floor. The main features of IIoT include (i) improving safety of workers, (ii) improve productivity with environment friendly solutions, and (iii) environmental monitoring.

The system in Fig. 1 contains four units: Sensor unit consists of different sensors like temperature sensor, humidity sensor, gas sensor, flow sensor, optical sensor, GPS sensor, MEMS sensor, and heart beat sensor. Server unit consists of MCP 3008, Raspberry Pi. Power management unit consists of battery. Measured parameters are transmitted to cloud. Here, analog-to-digital convertor (ADC) MCP 3008 can be

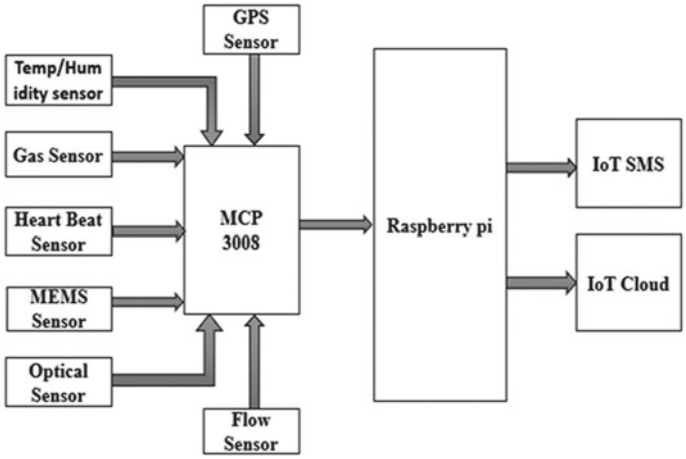


Fig. 1 Block diagram of IIoT

incorporated with Raspberry Pi to help conversion of input signals. Data from all sensors are sent to the MCP3008 which can be sent to Raspberry Pi which acts like a credit card size computer that are used by sensors to collect information. The entire system is networked via a wireless protocol such as Bluetooth, Wi-Fi, or a custom message system. The networking protocols are chosen to support the distribution of sensors nodes and the quantity of information to be collected [6]. This information is distributed over the network via the computer or laptop. The large amount of information from the different sensors can be collected, stored in memory, and analyzed [7]. The results of the analysis can be used to initiate actions. Security is an additional requirement, so that the information propagates over this network assuring that it can be viewed solely by the systems licensed to receive the information. Finally, the measured parameters are transmitted to cloud. For example, the 'smart objects' collect large amount of data, communicate among themselves, and can be connected via Internet providing means to control them from any part of the globe.

The steps involved can be represented in the form of algorithm as:

1. Collect data using various sensors.
2. Feed the data collected by the sensors to the ADC Converter MCP 3008.
3. Feed the output of MCP 3008 to the Raspberry Pi processor.
4. Send the processed information from the Raspberry Pi to the IoT Cloud.

This work aims at controlling and monitoring the different sensors which are connected to the microcontroller, and finally, the collected data will be transmitted to the cloud.

3 Results and Discussion

3.1 Predictive Maintenance

Predictive maintenance in industries can be achieved only if a sensor interface is introduced to connect all the active sectors of an industry in an IoT based environment [4]. The concept can be realized by integrating WSN and IoT. In this paper, a prototype is developed which represents an industrial scenario where three industrial floors are constantly monitored for predictive maintenance [8]. Here, the real-time data are collected from different sensors placed across different boilers, which are monitored. The data can be transmitted to authorized systems and officials via IoT [9].

Each boiler in the industrial floor is equipped with various sensors such as the vibration sensor, the gas sensor, and the temperature sensor as in Fig. 2. The alert signal in case of any failure or emergency can be sent to the officials via SMS. Machine queuing telemetry transport (MQTT) protocol aids in communication. In case, the readings from the sensors connected to a single boiler crosses the pre-set threshold value, and an alert SMS is sent to the corresponding floor manager which can be used

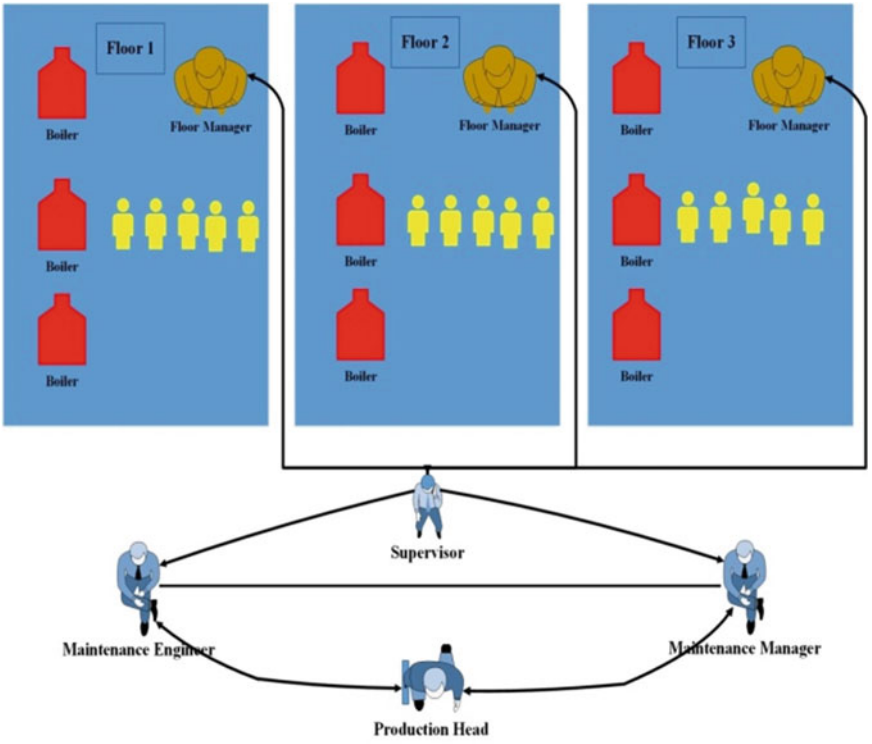


Fig. 2 Design flow

to rectify the fault before system failure. When two boilers show malfunction, then two separate messages are sent to the floor manager and the supervisor as well. The supervisor then directs the other two floor managers to take up the work from this floor or just directs the previous floor manager to do the required amendments. The entire data regarding the defects and faults are sent to the maintenance manager and the production head via MQTT protocol. This lightweight protocol can be used to transmit data at high speed with minimum data loss. The design flow gives a view of how an entire industrial floor is monitored and the hierarchy in which the messages would be sent when an alert situation occurs.

An additional feature can be included by introduction of a standby boiler which gets activated following an unexpected breakdown of a boiler and replaces the faulty boiler. This can provide an effective solution for a breakdown condition wherein anyone of the boilers stops functioning and the efficiency of an entire floor gets affected. The four sets of sensors attached to all the above boilers would be active but only the readings from the three working boilers would be recorded in the cloud network. The above process continues till any one of the working boilers experience a breakdown. Once a breakdown occurs, an interrupt is generated in the output which raises a flag to stop recording the values from the affected boiler and to switch over to the standby boiler. The entire data is then collected and stored in ThingSpeak, an IoT platform that is used to store the data from the boilers.

The final prototype developed for this phase is as shown in Fig. 3 which shows a floor consisting of three boilers, each having piezoelectric vibration sensors and gas sensors. These sensors are set to a certain threshold value which if crossed; an alert message will be given to the concerned official. The entire reading and monitoring of data from the floor can be displayed in the command window. Any abnormality in any sensors triggers the system to send an alert message to the official with the readings of each sensor. The alert messages are sent to the user when the sensor readings have crossed the threshold.

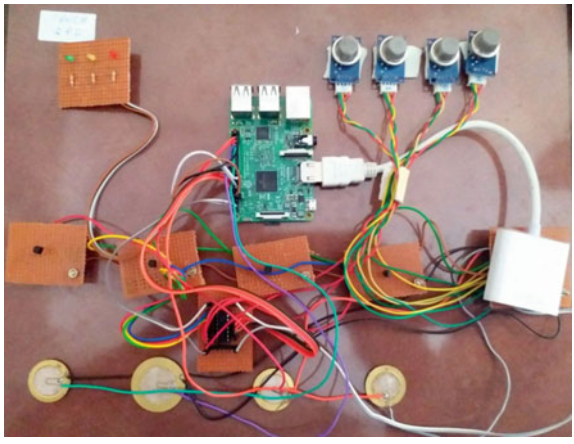


Fig. 3 Prototype developed

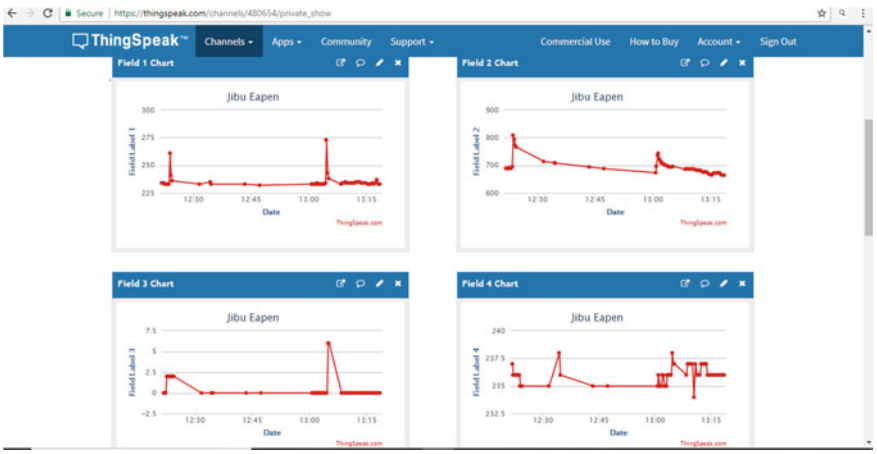


Fig. 4 Cloud data for different sensors

The alert messages generated in case of failure if any will be sent to the concerned officials. The entire data can be stored in the cloud network; ThingSpeak stores the breakdowns and glitches that occur in the entire plant. The concerned officials can check the flaws and initiate the necessary action. The data from the cloud network is as shown in Fig. 4.

The red dots in the above graph (Fig. 4) indicates the breakdowns or the extreme values recorded by the sensor. The sensor values assigned in this prototype are actually set in accordance with the industrial standards.

3.2 Enhancement of Worker Safety

Leakage of toxic gases is an important issue to be addressed in petrochemical plants which can cause serious threat to environment and human kind. Development of an intelligent system to detect the leakage of toxic gases is the need of the hour. Static wireless nodes can be deployed in various parts of the operating area which can detect the presence of toxic gases. In case of the presence of toxic gases above the predetermined level, an alarm can be initiated to prevent fatal accidents. This concept can be realized using four components, namely sensor-embedded wireless devices, static wireless sensor nodes, wireless base stations (Wi-Fi enabled), and remote monitoring center [10, 11]. The wearable devices in the form of helmet and wrist watch can be worn by workers [12, 13], as shown in Fig. 5. This in collaboration with static wireless nodes senses the environment and collects data. These data are sent to remote monitoring center by means of Wi-Fi-enabled base stations. Collaborative networking is made possible among the various wireless devices. This results in

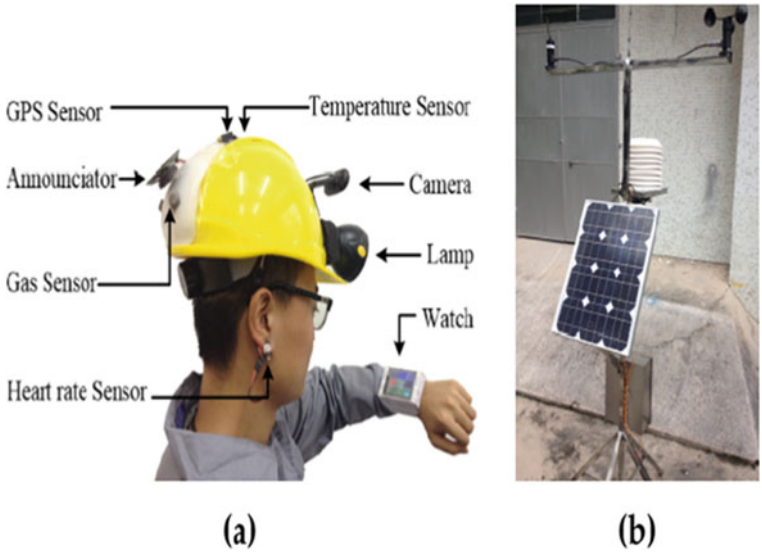


Fig. 5 Wearable sensors embedded in smart watch and helmet

construction of a problem solving network which can detect the leakage of toxic gases.

Constant monitoring of workers is made possible by the deployment of smart helmets and smart wrist watches as shown in Fig. 5. Here, large spatiotemporal information are gathered by the smart helmets which are equipped with smart sensors. The gathered information by the smart helmets and sensors nodes are submitted to the observation center. Finally, the large amount of information are analyzed to support the collaborative sensing intelligent (CSI) framework [9, 14]. The collaborative networking enables the different wireless devices to construct a problem solving network, and the results of the analysis can be sent to the wrist watch by which the worker can himself monitor the status of the parameters in his working area. The information can be transmitted to the floor manager who can monitor the entire industry floor and can initiate rescue actions in case of any malfunctions or leakage of gas. The discharge of poisonous gases can be detected, and in case, the situation goes above the predetermined level, alarm can be initiated to prevent accidents.

A prototype of the above-explained concept was developed. As an initial experiment, four sensors namely heart beat sensor, gas sensor, temperature sensor, and MEMS sensor were included. Heart beat sensor can detect fluctuations in the usual pattern which may also be due to the impact of the changes in the industrial floor. Gas sensor can detect the presence and level of toxic gases above the predetermined limit. The temperature sensor can detect the condition of increase in furnace/equipment temperature above the threshold level. MEMS sensor can detect the position of worker which can detect the case of worker collapsing, which may be due to the any abnormalities in working area or body reflex [15]. The information from different sensors

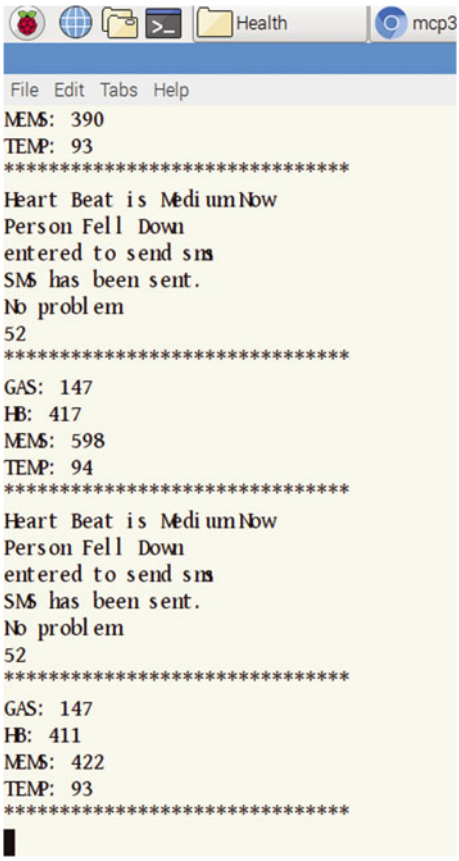


Fig. 6 Output of different sensors

are collected at different location and time points. The output of the all sensors can be collected and displayed in personal laptop and can be sent as SMS to mobile phone of the floor manager as shown in Fig. 6. The threshold value of the parameters can be set, and in case, the current situation goes above the predetermined values, alerts can be given to workers as well as floor mangers to avoid accidents.

4 Conclusions

The paper focuses on the predictive maintenance and industrial safety using IoT which can pave way to improved operational efficiency and productivity with no compromise on safety. Detecting the exact location of the source of leak is a challenge without tracking the change of concentration of toxic gas since it constantly changes with shift of location and time. In such a dynamic environment, collaboration of different sensor nodes is a necessity to track the changes in concentration. Decision

making with regard to the type of gas sensors is a matter of interest since there is a possibility of chemical reaction among different gases which may not be detected by the deployed sensors. Embedding all types of sensors to a detecting system is not possible from the view point of cost effectiveness. This system can be innovatively enhanced by introducing machine learning techniques in the present system. Another future expansion that could be thought of is controlling the movement of machines inside an industry globally. Since the present scenario only detects the flaw in systems, accidents in industries occur partly because of improper placement of machines or partly because of negligence of coworkers. The above-discussed conditions can be kept in mind while developing the new system which can open up an eye for the researchers.

References

1. Lin, H., Bergmann, N.W.: IoT privacy and security challenges for smart home environments. J. Inform. (2016)
2. Patil, S., Lonhari, T., Patil, S.: Internet of Things: current research, trends and applications. Int. J. Innov. Res. Comput. Commun. Eng. **3**(12) (2015)
3. Datta, S., Bonnet, C.: Connect and control things: integrating lightweight IoT framework into a mobile application. In: 9th IEEE International Conference on Next Generation Mobile Application, Services and Technologies (2015)
4. Chi, Q., Yan, H., Zhang, C., Pang, Z., Xu, L.D.: A reconfigurable smart sensor interface for industrial WSN in IoT environment. IEEE Trans. Industr. Inform. **10**(2), 1417–1425 (2014)
5. Vermesan, O., Friess, P., Patrick, G., Sergio, G., Harald, S., Alessandro, B., Ignasio, S., Margaretha, M., Harrison, M., Markus, E., Doody, P.: Internet of things strategic research roadmap. Internet of Things: Glob. Technol. Soc. Trends. **1**, 9–52 (2011)
6. Scott Geller, E.: Behavior-based safety in industry: realizing the largescale potential of psychology to promote human welfare. Appl. Prevent. Psychol. (2014)
7. Zubaĭ, M.: IoT gateway and industrial safety with computer vision. In: IEEE 14th International Symposium on Applied Machine Intelligence and Informatics (2016)
8. Kallappa, B.B.T.: Industrial safety parameters monitoring in IOT environment. IJAREEIE (2016)
9. Chen, Y., Lee, G.M., Shu, L., Crespi, N: Industrial internet of things-based collaborative sensing intelligence: framework and research challenges. J. Inform. **16**, 215 (2016)
10. Datta, S.K., Bonnet, C., Nikaein, N.: An IoT gateway centric architecture to provide novel M2M services. In: Internet of Things (WFIoT), IEEE World Forum (2014)
11. Zhang, H., Zhu, L.: Internet of Things: key technology, architecture and challenging problems. In: Computer Science and Automation Engineering (CSAE), IEEE International Conference (2011)
12. Sharma, G., Kalra, S.: A light weight user authentication scheme for cloud IoT based healthcare services. Iranian J. Sci. Technol. Trans. Electric. Eng. **43**, 619–636 (2018)
13. Haghi, M., Thurow, K., Stoll, R.: Wearable devices in medical internet of things: scientific research and commercially available devices. Health Care Inform. Res. **23**, 4–15 (2017)
14. Lamas, P.F., Fernandez, T.M., Suarez, M., Castedo, L., Gonzalez, M.: A review on internet of things for defence and public safety. J. Sens. **16**, 1644 (2016)
15. Tahir, H., McDonald, K.: On the security of consumer wearable devices in the internet of things. PLoS ONE **13**(4), e0195487 (2018). https://doi.org/10.1371/journal.pone.0195487

SISA: Securing Images by Selective Alteration

Prutha Gaherwar, Shraddha Joshi, Raviraj Joshi, and Rahul Khengare

Abstract With an increase in mobile and camera devices' popularity, digital content in the form of images has increased drastically. As personal life is being continuously documented in pictures, the risk of losing it to eavesdroppers is a matter of grave concern. Secondary storage is the most preferred medium for the storage of personal and other images. Our work is concerned with the security of such images. While encryption is the best way to ensure image security, full encryption and decryption is a computationally intensive process. Moreover, as cameras are getting better every day, image quality, and thus, the pixel density has increased considerably. The increased pixel density makes encryption and decryption more expensive. We, therefore, delve into selective encryption and selective blurring based on the region of interest. Instead of encrypting or blurring the entire photograph, we only encode selected regions of the image. We present a comparative analysis of the partial and full encryption of the photographs. This kind of encoding will help us lower the encryption overhead without compromising security. The applications utilizing this technique will become more usable due to the reduction in the decryption time. Additionally, blurred images being more readable than encrypted ones, allowed us to define the level of security. We leverage the machine learning algorithms like Mask-region-based convolutional neural network (RCNN) and you only look once (YOLO) to select the region of interest. These algorithms have set new benchmarks for object recognition. We develop an end-to-end system to demonstrate our idea of selective encryption.

Keywords Image processing · Image security · Encryption · Decryption · Deep learning · Blurring

P. Gaherwar (✉) · S. Joshi · R. Khengare
Department of Computer Technology, Pune Institute of Computer Technology, Pune, Maharashtra, India

R. Joshi
Department of Computer Science and Engineering, Indian Institute of Technology Madras, Chennai, Tamil Nadu, India

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_69

1 Introduction

Data privacy is a crucial part of the digital era. Businesses and individuals must give a significant amount of thought to how data is gathered, retained, applied, and disclosed. When there is unauthorized access to private or personal information, a privacy breach occurs. It often leads to legal action if the violation is severe. Every business unit, government, and private sector frequently uses the digital image to transfer critical data. These images accessible over the Internet may not be secure; therefore, image security has always been an important subject. A widely used solution to the data breach is encryption, which encodes data with encryption algorithms so that even unauthorized access cannot derive meaningful information [1].

While encryption is the best technique to safeguard our images, full encryption may not always be required [2]. For example, a family photograph may contain background scenery, which may not be necessary. It would be sufficient to encrypt only the people in the images. Therefore, encryption hides the unnecessary data along with the necessary one, which increases the processing time. We propose SISA, which explores the idea of partial encryption [3] to reduce this overhead (Figs. 1 and 2). Our work is concerned with the security of images which are mostly stored in the secondary storage. Although we mainly discuss pictures stored on our hard drives, it is equally applicable to cloud storage.

1.1 Importance of Security

Every year, even the common man generates exabytes of data just in the form of images. These images comprise personal data like copies of passport, social security documents, bank details, or biodata. So the prospect of image theft resulting in the loss of important personal data is an extremely serious problem at hand.

The security measures to safeguard critical information are implemented at multiple levels. On one side, we have machine-level security employing firewalls, whereas on another, we have storage-level security that uses authentication and authorization

Fig. 1 SISA results on an image. The left-most image is the original image; the next image is annotated with a bounding box. The image at the center is the output of the face recognition algorithm used (you can see the person's name on the bottom-left of the bounding box). Next image is an output of the blurring technique (we have elaborated the selective blurring using instance segmentation in the later sections), and the right-most image is the selectively encrypted image

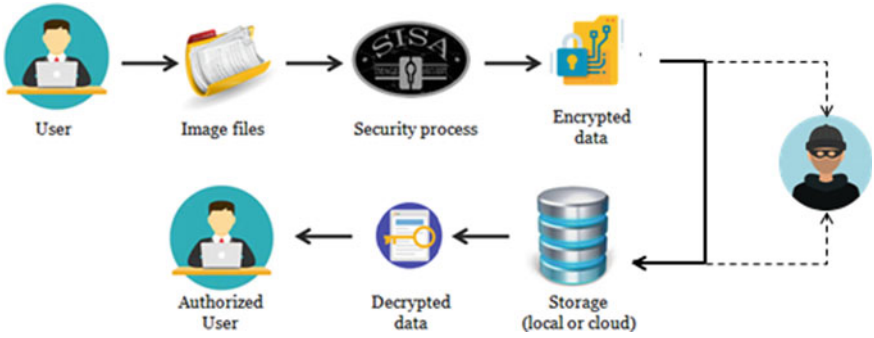


Fig. 2 High-level architecture

as safety measures. However, it is very much possible to break into machine-level security and single-step storage security. In a survey by Varonis [4], security breaches have increased by 11% since 2018 and 67% since 2014. It is too big of a percentage to be ignored or side-lined.

In notable cases, information leakage completely invades personal privacy. For example, the malicious spread of photographs in private online albums or patients' medical diagnosis images may cause numerous losses to the individual or society. According to IBM [5], health care spent the most time in the data breach lifecycle in 2019, 329 days. As the essential data transfers are carried out extensively using digital photographs, safety is the primary concern for a system to maintain integrity, confidentiality, and authenticity.

1.2 Proposal

The widely used technique of complete image encryption while storing the data has some significant disadvantages. These techniques encrypt each pixel one by one, so when the pixel count rises, the encryption time increases as well [1]. Besides, the image quality has evolved. If we consider an HD image, which is the most preferred quality these days, it has 20L pixels (1920×1080), whereas a 4k image has whooping 80L (3840×2160) pixels. Encrypting and decrypting 20–80L pixels will be a too time-consuming process. Moreover, the encryption and decryption algorithms are quite resource-intensive, making it difficult to run on low-end devices. So an application encrypting and decrypting bulk images on a local device will be far from usable.

This paper presents an approach based on partial blurring and encryption. The aim is to ensure image security while reducing the overhead of processing time. SISA detects the objects in the image and assigns them a priority value based on user preference to choose the most critical area. The default priority value is set as

per the affinity toward the center coordinates of the image, i.e., in general cases, the highest priority object will be located at the center of the image, and the one with the least priority will be at the edges. Next, we apply selective encryption to the file (e.g., jpg, png) which encrypts only parts of data instead of encrypting the entire data. In essence, with our approach we try to combine the benefits of prioritization and encryption.
The main contributions in our paper are as follows:

- Our scheme identifies the significant parts of the user data based on the prioritization technique, which we later use for selective alterations of the image.
- We focus on selective alteration such that in event of a successful attack, the meaningful information is not visible to the attacker. Also, we compare full and partial encryption and highlight the challenges from a deployment perspective.

Architecture The image to be secured is passed to the SISA layer before storing it on disk. The layer appropriately modifies the image based on preferences selected by the user. One of the strong user preferences is the level of security required. The image is encrypted for higher security levels, whereas, for lower levels, the image is blurred. Although blurring is used with low levels of security, it is imperative from a usability perspective. A blurred image is more visually accessible as compared to an encrypted image and hence more usable.

When these images are transferred through an insecure channel, even if the intruder eavesdrops on the data on the wire, no usable information will be revealed. These images will also be present in the storage in the modified format, thus hacking into the storage will also render the data unusable. Only the authorized user will be able to access the valid data using a security key.

The level of security also controls the percentage of the image to be encrypted. For the lowest level, 30% of the image is encrypted. This percentage is linearly increased for higher levels of security. The benefit of our approach is the intelligent selection of 30% region (refer Fig. 3). Rather than randomly selecting pixels, we use region selection and prioritization algorithms to select relevant pixels.

Fig. 3 Percentage of blurring in the left image is 30, and in the right image, it is 70. Here, we can see that the object of highest priority (user's image) is blurred in the left image, while in the right, other people are also blurred

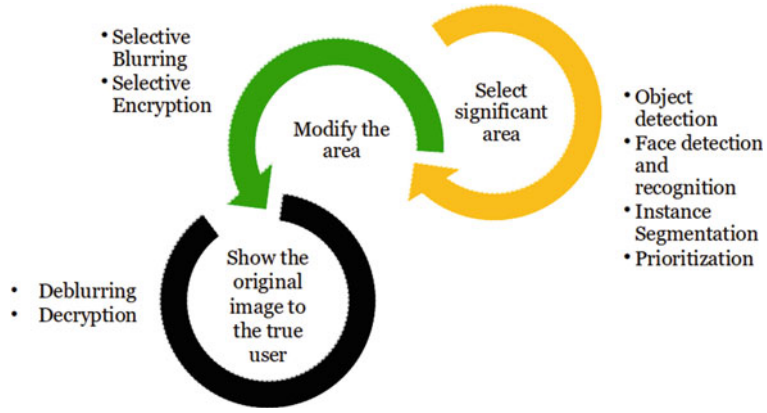


Fig. 4 Process flow

Process Flow As shown in Fig. 4, the process flow will be as follows:

1. Select a significant area
   (a) As the first step, the algorithm of object detection is used to produce bounding boxes of objects in the picture.
   (b) The Mask-RCNN algorithm outlines the object's edges with a bounding box and marks the object instance.
   (c) The prioritization scheme then assigns a priority to these objects. The priority is based on user preferences. Simple heuristics like importance to the foreground, background, or face are part of the preferences.
   (d) In addition to this, we also carry out text detection, face detection, and face identification to select a significant area. These extra selections are part of user preferences and appropriately triggered.
2. Modify the area under consideration
   (a) For a low level of security, the image is blurred, and for a higher level, it is encrypted. The level of security also controls the percentage of the image regions to be altered.
   (b) As per the selections, all these modifications are carried out by altering the image's required part.
3. Display the original image
   (a) The original image can be retrieved using a deblurring or decryption algorithm.

2 Methodology

2.1 Processing on the Image

Object Detection It is a classification and localization problem. Generally, we cannot fix on the number of objects in the image. The model expects data normally in the form (I, A), where I is an image, and A is an array containing (class_label, p, q, l, b). For A = (class_label, p, q, l, b);
p = x-coordinate of top-left corner of the bounding box
q = y-coordinate of top-left corner of the bounding box
l = bounding box length in pixel
b = bounding box breadth in pixel
The necessary steps followed are:

1. Take an input image
2. Extract region maps
3. Compute attributes of each submission using an extensive convolutional neural network
4. Classify each region using class-specific linear support vector machine (SVM)

Face Detection and Recognition The face recognition consists of a two-step process:
(1) face detection (bounded face) in image
(2) face identification (person identification) on the detected fixed face
The face detection algorithm detects all the faces in the image and then puts a bounding box around it. Face identification generates the detected bounded face's embedding and matches it against the embedding of the training faces in the user-specific database.

Instance Segmentation Instance segmentation aims to find specific objects in an image and develop a mask around the object. Instance segmentation can be considered as object detection where the output is defined as a object mask. This is in contrast to normal object detection where the output is a bounding box. Unlike semantic segmentation, which tries to classify each pixel in the image, instance segmentation does not label every pixel in the picture [6].

Several enhancements over RCNN for object detection have resulted in Mask-RCNN [7]. RCNN applies two-step process to generate an object-label and the bounding box. Firstly, it adopts selective search method to create region maps and then each obtained region is transformed. Fast RCNN [8] is an improved RCNN with a further improved version labeled as Faster RCNN [9]. The Mask-RCNN algorithm is an enhancement over Faster RCNN by parallelizing a mask predicting branch with the class-label and bounding box prediction branch. Despite of increasing a tiny expense over the Faster RCNN network, it still provides optimum speed and is the most efficient of all.

Prioritization The prioritization step involves assigning priorities to the objects detected by the object detection and instance segmentation algorithms. In day-to-day life, while taking a picture, we usually keep objects of most importance in the center of the frame. Using this generic idea, we developed an algorithm that selects objects closest to the image's center. The interactive learning of the critical regions can improve this default scheme according to the user specifications.

2.2 Modifying the Image

Blurring and Deblurring Gaussian blur is carried out using a convolution kernel. We prepare a 2D matrix of normalized precalculated Gaussian function values. This matrix is used as weights to calculate weighted sum of all adjacent pixels to find new value of each pixel [10].

Partial blurring is achieved by applying Gaussian blur on the area covered by prioritized objects.

$$G(x, y) = \frac{1}{2\pi\sigma^2}(\mathrm{e}^{-}\frac{x^2 + y^2}{2\sigma^2})$$

Encryption and Decryption Encryption is the process of encoding a simple image into a ciphered image using encryption algorithms so that unauthorized users cannot access the original image. Furthermore, the process of recovering an image from a ciphered image using a set of instructions is called decryption. It is carried out using a secret key.

We use the standard AES algorithm with CFB mode for encryption. The selected group of pixels is modified to form an encrypted image. The standard process is used for decryption as well.

Metadata The information required to reconstruct the image is stored in the image's metadata using ExifTool. Trivially this includes the encrypted portion's coordinates. Any mismatch in encrypted contents length is also added to the metadata.

In Fig. 5, we have shown the implementation of all the techniques used.

3 Literature Survey

1. DeepPrivacy: A Generative Adversarial Network for face Anonymization: The paper introduces an innovative model that automatically anonymizes faces present in the pictures while preserving the primary data distribution. It ensures the complete anonymization of each face seen in the picture by creating the images solely based on the privacy-safe data. Their model is established on a conditional gen-

Fig. 5 First image is the input on which we have generated the following outputs. The second image shows the bounding boxes and segmented instances. The faces of the people that can be seen in the picture are detected and recognized in the next image. The bottom-left image is the output of the blurring algorithm, and the last image is the output of the encryption algorithm

erative adversarial network, developing photographs by taking the existent pose and image environment into consideration [11].

2. You Only Look Once (YOLO): YOLO presents the unified model for the detection of objects. It spatially separates objects' bounding boxes and related class probabilities, viewing object detection as a regression problem. As the detection pipeline comprises only a single network, end-to-end optimization based on detection performance is possible. The convolutional layers are pretrained using the ImageNet 1000-class competition dataset. An accuracy of 88% was achieved [12].
3. Fast Face-swap using convolutional neural networks: This paper presents a feed-forward neural network that achieves high realism levels in generated face-swapped images. This method uses a multi-image style loss, thus approximating a manifold describing a style rather than using a single reference point. The trained networks allow the model to perform face swapping in real time [13].
4. Techniques for selective encryption of uncompressed and compressed images: This paper proposes two approaches for selective encryption of the image. The first approach focuses on raster images. The second approach adapts the JPEG compression scheme, which puts forth, one, a constant bit rate and another, format compliance. It involves encryption of the sign and magnitude of nonzero DCT coefficients. This approach also elaborates on its usage for several selective encryptions [14] (Table 1).

Table 1 Findings from the literature survey

Title	Pros	Cons
DeepPrivacy: a generative adversarial network for face anonymization [11]	• Automatic anonymization of all faces •Loss of mask allows class prediction without competition	• Can generate unrealistic images
You Only Look Once (YOLO) [12]	• A single network completes the whole detection pipeline	• Can miss small objects which appear together • Can fail for unidentified objects, aspect ratios, or configurations
Fast face-swap using convolutional neural networks [13]	• Feed-forward NN achieves photorealism • Style transfer	• Quality depends on the collection of styled images
Techniques for selective encryption of uncompressed and compressed images [14]	• Constant bit rate and format compliance	• PEG specific • Applicable only to monochrome images

4 Mathematical Model

$$S = \{s, e, X, Y, DD, ND, ff, fme\}$$

$I = \{i | i$ is the image uploaded by user$\}$
$P = \{p | p$ is tuple of object properties in an image$\}$
Tuple = {label, coordinates of the top-left corner, width, height, mask}
$Z = \{z | z$ is the tuple of label and priority$\}$
$O = \{o | o$ is the altered image$\}$
s : Initialize
$I = \Phi, P = \Phi, Z = \Phi, O = \Phi$
e : Update prioritization model
X : Input
$I, l, k \in X$
l : level of security
k : security key
Y : Output
DD : Deterministic data
ND : Non-deterministic data
$ND = \Phi$
$O \in Y$
ff : Friend function
$f1(x)$: Object detection function
$I \rightarrow P$

$f2(x)$: Alteration function
$Z, P, k, l \rightarrow O$
As per the value of l and the sets present, blurring/encryption/morphing takes place
fme : Generated Function
$f3(x)$: Prioritization function
$P \rightarrow Z$

5 Results

The COCO dataset is used to train the Mask-RCNN model, and custom made dataset to train the face recognition algorithm. For example, to generate the results of Fig. 1, we trained the face recognition model on the subject's photographs taken from various angles. The performance estimation of existing encryption and SISA algorithms is done on a machine with a 64-bit OS and an 8GB RAM.

From a deployment perspective, decryption is more critical than encryption. The reconstruction time of the image should be less as it will determine the usability of the application. Conversely, encryption and backup can be done in the background without exposing this latency to the user. To show the effectiveness of selective encryption in terms of usability, we compare our approach with full encryption and decryption. Figure 6 shows the comparison of encryption or blurring times for different approaches. Figure 7 shows the corresponding timing required during decryption. The time needed for blurring is significantly less as compared to encryption and decryption. So if we can compromise on the level of security, then blurring is a good option. However, blurring is a lossy operation, and hence the blurred object needs to be stored separately to reconstruct the original image precisely. The objects stored separately are also encrypted, so deblurring time is equivalent to decryption time. We show that as the size of the image increases, the processing time for altering the complete image increases substantially; comparatively, the time required for SISA is almost 46% less.

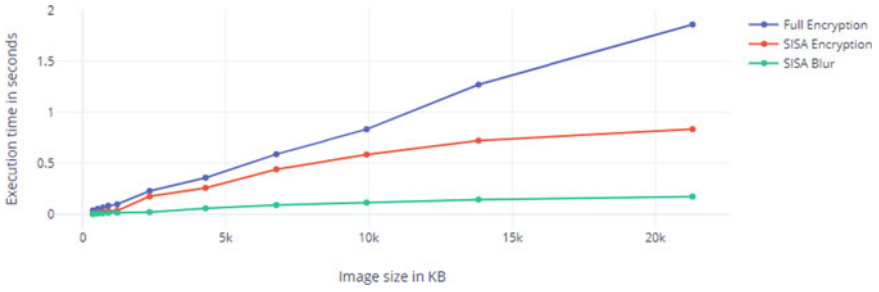


Fig. 6 Full encryption versus SISA encryption versus SISA Blur

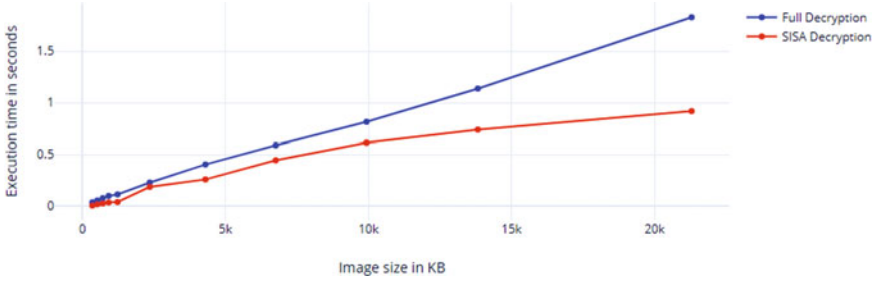


Fig. 7 Full decryption versus SISA decryption

6 Conclusion

In this paper, we present an approach for securing images by selective alteration. Here, we encode an image by selective encryption or blurring, which cuts down the time overhead without compromising security. As the image is encrypted with various constraints and modification types for each image, it becomes difficult for the intruder to access the original image. Although blurring is a less secure alternative, it is comparatively more user-friendly. When integrated with security applications, SISA can provide an efficient interface for secure data storage. Due to partial encryption, the data's actual image recovery time will be substantially less, and these applications will be more serviceable.

7 Application and Future Scope

SISA can be used to secure banking documents, digital signatures, personal confidential snapshots, and medical records. It can also be used with personal identification documents like PAN card and Aadhaar card. Moreover, it can be incorporated in both cloud-based and local image sharing applications. With the popularity of mobile cameras, photographs are a very convenient way to store capture information these days. So, any critical image with confidential content can be passed through the SISA layer before storing it on disk or cloud. SISA can be further enhanced by including morphing of images and text encryption inside the image. In the context of medical records like X-rays and MRI scans, we will need a more specific object detector as compared to our current general object detection approach. Currently, we rely on user preferences to rank the objects for encryption or blurring. It can be fully automated to have an end-to-end machine learning-based pipeline.

Acknowledgements We would like to take this opportunity to show our gratitude toward the L3Cube group, which sponsored the project and guided us throughout, and credit one of our teammates, Garvit Gulati, who was also a part of this project's implementation.

References

1. Mohammad, O.F., Rahim, M.S.M., Zeebaree, S.R.M., Ahmed, F.Y.: A survey and analysis of the image encryption methods. International Journal of Applied Engineering Research **12**(23), 13265–13280 (2017)
2. Khatod, V., Ingale, S., Gund, K., Gorde, S., Joshi, R., Khengare, R.: Enigma: A hybrid approach to file security in cloud. In: Proceedings of ICETIT 2019, pp. 1005–1015. Springer (2020)
3. Yu, J., Zhang, B., Kuang, Z., Lin, D., Fan, J.: iprivacy: image privacy protection by identifying sensitive objects via deep multi-task learning. IEEE Transactions on Information Forensics and Security **12**(5), 1005–1016 (2016)
4. 110 must-know cybersecurity statistics for 2020 | varonis. https://www.varonis.com/blog/cybersecurity-statistics/
5. Ibm study shows data breach costs on the rise; financial impact felt for years, 23 Jul. 2019. https://newsroom.ibm.com/2019-07-23-IBM-Study-Shows-Data-Breach-Costs-on-the-Rise-Financial-Impact-Felt-for-Years
6. Jiao, L., Zhang, F., Liu, F., Yang, S., Li, L., Feng, Z., Qu, R.: A survey of deep learning-based object detection. IEEE Access 7, 128837–128868 (2019)
7. He, K., Gkioxari, G., Dollár, P., Girshick, R.: Mask r-cnn. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 2961–2969 (2017)
8. Girshick, R.: Fast r-cnn. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 1440–1448 (2015)
9. Ren, S., He, K., Girshick, R., Sun, J.: Faster r-cnn: towards real-time object detection with region proposal networks. In: Advances in Neural Information Processing Systems, pp. 91–99 (2015)
10. Filip, S.: An investigation of fast real-time GPU-based image blur algorithms (2014). https://software.intel.com/enus/blogs/2014/07/15/an-investigation-of-fast-real-time-gpubased-image-blur-algorithms
11. Hukkelås, H., Mester, R., Lindseth, F.: Deep privacy: a generative adversarial network for face anonymization. In: International Symposium on Visual Computing, pp. 565–578. Springer (2019)
12. Redmon, J., Divvala, S., Girshick, R., Farhadi, A.: You only look once: unified, real-time object detection. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2016)
13. Korshunova, I., Shi, W., Dambre, J., Theis, L.: Fast face-swap using convolutional neural networks. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 3677–3685 (2017)
14. Van Droogenbroeck, M., Benedett, R.: Techniques for a selective encryption of uncompressed and compressed images. In: ACIVS Advanced Concepts for Intelligent Vision Systems, Proceedings, pp. 90–97 (2002)

Voice-Based Gender Recognition Using Neural Network

Kavita Chachadi and S. R. Nirmala

Abstract The human speech contains paralinguistic information used in many speech recognition applications like automatic speech recognition, speaker recognition, and verification. Gender from voice is considered as one of the essential tasks to be detected for such applications. To build a model from a training set, a set of relevant speech features is extracted in order to distinguish gender (i.e., female or male) from a speech signal. This paper focuses on comparison of the proposed neural network (NN) model with the different features like MFCC and mel spectrogram extracted from the speech signal to recognize the gender. Experiments are carried on Mozilla voice dataset and evaluated performance of the network. Experiments show that the combination of MFCC and mel feature sets shows the better accuracy with 94.32%.

Keywords Neural network · Voice · Speech · Gender

1 Introduction

The human speech consists of unique information such as age, gender, emotional state, language accent, tone which conveys linguistic, and non-linguistic features. It also conveys paralinguistic information such as speaker's intention and attention. Each speaker's voice is different in terms of frequencies which in turn depends on the vocal cord model of a speaker. The voice of human speech is an effective communication method consisting of all these features. Identifying gender through voice has been a challenging task in the research areas of speech processing. Because of the importance of recognition of gender based on speaker's voice recognition, the important features which are unique in terms of pitch and frequency are extracted from the speech signal and then classification models are constructed. The deep

K. Chachadi (✉) · S. R. Nirmala
KLE Technological University, Hubballi, Karnataka, India

S. R. Nirmala
e-mail: nirmala.s@kletech.ac.in

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_70

learning techniques which depend largely on training set should include quality of features that determines effectiveness and robustness of classifier.

The efficiency of the classifiers can be improved by extracting proper voice features. Therefore, in the literature, numerous methods have been developed to extract feature sets from the speech signal and have been evaluated on different algorithms.

The spectral characteristics of utterance of speech are represented by mel-frequency cepstral coefficients (MFCC) which is one of the prominent features. Most of the speech processing applications like speaker recognition and verification, speech recognition and classification of noise, MFCC features are widely used. MFCCs represent the spectrum that models the vocal tube of humans, and length of the vocal tube is different from male to female. Deep learning techniques are used to build a model for recognizing the gender from voice by using a training set with the gender as label combined with the features extracted from speech. In this paper, we propose neural network (NN) which recognizes gender from speech signal by combining various feature sets from speech signal. Contributions in this work are:

1. Preprocessing of dataset
2. Proposed NN model
3. Experimental setup
4. Results and discussion.

2 Related Work

Deep learning techniques have shown success in many of speech processing applications for longer period of time. In literature, many such machine learning techniques as well as deep learning techniques have been used for gender recognition from speech signal and shown improving performance on the different database. Pitch and frequency are the main parameters from which gender of the speaker can be identified.

According to Bishop et al. [1], the speech produced by speaker which is acoustic in nature consists of linguistic, paralinguistic, and non-linguistic information corresponds to the perception of the pitch. Fundamental frequency f0 decides gender (male or female) of speaker, and it also depends on tone and high pitch of the speaker. Authors discussed about the variation in the pitch of different speakers.

Gaikwad et al. [2] showed recognition of gender by extracting pitch, energy, and MFCC features from speech signal and classified using SVM. Their work is based on intuition for using the average fundamental frequency for men is in the range of 100–146 Hz and for female it is in the range of 188–221 Hz. Database consists of 20 speakers in which 8 were male and 12 were female. Each word in the vocabulary was recorded 5 times so that it will be good for training and achieved 95% of recognition rate.

Zeng et al. [3] proposed a robust GMM classifier for gender recognition and has obtained up to 98% of accuracy which is proved to be one of the effective methods for detection of speech and gender analysis. Experiments carried on TIMIT database.

Vergin et al. [4] showed improvement in word error rate by using acoustic models dependent on gender. They showed improvement in gender recognition by 14% compared to baseline speaker-independent systems on ATIS corpora by extracting first two formant frequency from the speech signal of the speaker.

Harb et al. [5] built a system for identification of gender in multimedia applications. The system used a set of neural networks with acoustic and pitch-related features. The combination of features and classifiers showed better performance than individual classifier.

Zeng et al. [6] proposed gender classification system based on Gaussian mixture model (GMM) by extracting parameters of pitch and predictive linear coefficients to model characteristics of female and male speech. They achieved classification accuracy as high as 98%.

Metze et al. [7] have compared four approaches for age and gender recognition using telephone speech. They developed Gaussian mixture models based on MFCCs for recognition of gender and age, discussed about the system using Bayesian networks. They also developed a system based on linear prediction analysis and a parallel recognition of phone which is derived from automatic language identification system.

Ververidis and Lin et al. [8, 9] showed that emotional state can be recognized in a better way by considering gender of speaker into account as compared to the gender independent classification.

Xiao et al. [10] showed an accuracy of 94.65% for gender classification on the Berlin dataset where emotional speech recognition system was built based on backpropagation neural networks with sequential forward selection.

Raahul et al. [11] compared different machine learning algorithms like linear discriminant analysis (LDA), K-nearest neighbor (KNN), classification and regression trees(CART), random forest (RF), and support vector machines (SVM) based on eight different metrics in gender classification based on acoustic data of speech signal.

Zakariya et al. proposed new model for extracting new features from Deep neural network (DNN). Transformed mel-frequency cepstral coefficients (T-MFCCs) are generated by training DNN first in an unsupervised manner and then it is trained in supervised manner. The new T-MFCCs feature set is used as the input to the DNN model along with speaker model which improves gender classification accuracy [12].

Kabil et al. [13] proposed recognition of gender in two stages. In the first stage, raw speech signal is fed to CNN to extract features and these features are fed to the next stage where MLP classifies the gender. Authors demonstrated their results on the AVspoof dataset and achieved very good gender recognition rate.

David et al. [14] proposed 2D CNN model with MFCC features as input to the model. They have demonstrated their results on French database with 2284 recordings and achieved accuracy around 96.5%.

Mucahit et al. [15] proposed MLP model to recognize the gender by voice. The dataset used consists of 3168 recordings of both male and female speakers. Acoustic analysis has been done on these voice signals and extracted 22 acoustic parameters and given to the model. They have achieved 96.74% test accuracy.

Maxim et al. [16] proposed CNN architecture with three layers. Authors conducted experiments on two different set of features MFCC and mel extracted from raw speech signal. They achieved around 88.8% accuracy on German Speech corpus.

3 Description About the Dataset

The common voice dataset (Mozilla) [17] consists of speakers with different age group from less than 19 to greater than 89 with gender as male, female, and other. The dataset includes speakers with 16 different accents speaking in English like United States English, Australian English, Indian English, and so on. All the files are stored in .MP3 format. The dataset is split into several sections. If majority of the listeners say the speech matches with text, then those audio clips are considered as valid. If at least two listeners have said those audio clips does not match with corresponding text, then those audio clips are considered as invalid. The audio clips with less than two votes or those having equal valid and invalid votes are considered as other in their name of audio clips. There are separate folders for development and experimentation as dev folder, for training as train folder and for testing as test folder. This dataset is mainly used for training, experimentation, and testing of automatic speech recognition systems and other speech processing applications.

4 Proposed Approach

To extract relevant features, preprocessing stage is required where speech signal is given as input to the different feature extraction techniques. These feature vectors combined with the gender as label form a training set to build a model which recognizes the gender of speaker.

4.1 Preprocessing Stage

MFCC features extraction: MFCC are the common features used in automatic speech recognition and speaker identification. They have also been successfully used in gender recognition problems. These features model the vocal tube closely and present in the envelope of a short-time power spectrum. Following steps show the method for computing MFCC:

1. Windowing—The audio signal is split into frames of duration 20–40 ms because speech signal is quasi-periodic signal where there is change in the samples for a shorter duration of time. The power spectral density of the spectrum for each frame is calculated which identifies frequencies present in the frame.
2. Mel Filter bank—The mel filter bank is applied to estimate the energy in various frequency regions. This filtering is done as per how humans perceives the sound based on variations in frequencies.
3. Log of Filter Bank Energy—The output of the mel filter bank is power spectrum. To take into account, the variations in the energy logarithm of power spectrum are calculated.
4. DCT coefficients—Since there is correlation in the energies calculated by filter bank, the DCT is used to decorrelate the energies. First 2–13 higher DCT coefficients are considered to reduce fast changes in the filter bank energies and discard the rest.

4.2 Mel Spectrogram

An object of mel-spectrogram type represents an acoustic time–frequency representation of sound. The power spectral density is sampled into a number of points around equally spaced times and frequencies (on a mel-frequency scale). The mel-frequency scale is defined as

$$\text{Mel} = 2595 * \log_{10}(1 + f/700) \tag{1}$$

5 Proposed Model

The different features like MFCC and mel are extracted from the speech signal which represents high-energy coefficients from the signal. In this work, the architecture of CNN is shown in Fig. 1. The network receives MFCC and mel features as input which are extracted from speech signal. These features are fed to neural network which consists of five dense layers with 512, 256, 128, 128, and 64 neurons using ReLU as a nonlinear activation function. Dropout with 30% is used after every dense layer. Final dense layer consists of two output neurons used for gender recognition using sigmoid function. The parameter setting for the CNN is 100 epochs with batch size 128 and Adam optimizer is used.

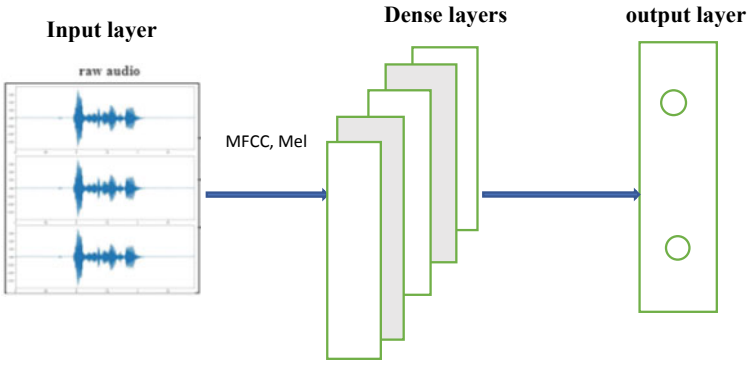


Fig. 1 Proposed neural network

6 Experimental Setup

The standard Mozilla dataset is used for training and testing purpose. This dataset consists of different speakers with different accent, age and speaking in English. The dataset is first filtered based only on gender which consists of 60,000 samples from different age group and variation in accent out of which 30,000 male and 30,000 female speakers are considered. 10% of the dataset is used for validation and testing the model. Then different features (MFCC, mel) are extracted from these speech signals using Librosa [18] before feeding to CNN layers. 250 epochs are used with batch size = 128 and Adam optimizer is used with learning rate 0.01 for training with early stopping. The model is trained on 480,00 speakers consisting of both male and female speakers. Trained model is validated on 6000 speakers and tested on 6000 speakers from the same dataset. Our network architecture based on Keras is trained using NVIDIA programming model on GPU.

7 Results and Discussions

The proposed model is trained on Mozilla dataset. The trained model is validated on 6000 speakers and tested on 6000 speakers from the same dataset. Different feature vectors like only MFCC, only mel spectrogram, combination of MFCC and mel are extracted from the speech and passed through neural network model. Accuracy and loss are given in Table 1.

The model with mel feature achieved the accuracy of 91.38% accuracy on 29 epochs in training whereas the model with MFCC feature achieved the accuracy of 93.13% better than mel-spectrogram feature on 29 epochs in training. The combination of MFCC and mel has achieved accuracy of 94.32% on 35 epochs in training. The results show that combination of MFCC and mel performs better as compared to other feature vectors.

Table 1 Accuracy and loss

Features	Validation accuracy (%)	Testing accuracy (%)	Test losses
Mel	90.09	91.38	0.2334
MFCC	92.65	93.13	0.1951
MFCC + mel	94.13	94.32	0.1554

Table 2 Correctly classified response

Feature	Correctly classified response (%)	
Mel	Male	Female
Male	90.1	9.9
Female	7.4	92.6
MFCC	Male	Female
Male	92.3	7.7
Female	6.1	93.9
MFCC + mel	Male	Female
Male	94.7	5.3
Female	6	94

Table 2 shows correctly classified response for recognizing male and female speakers obtained from confusion matrix. From this, we can observe classification of female speaker is good around 94% and male speakers is with 94.7% with MFCC and mel features as compared to the other features. The model with mel feature set gives less classification response around 90.1% as compared to any other feature sets in recognition of male speaker and 92.6% in recognition of female speaker.

Figure 2 shows graphical representation of comparison of different features with proposed neural network model in recognition of gender on Mozilla Voice Dataset.

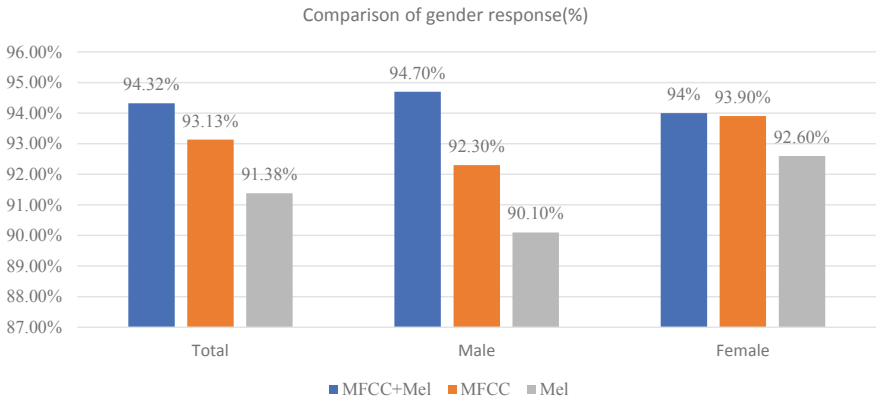


Fig. 2 Comparison of gender response

8 Conclusion

This paper proposed a speaker gender recognition model based on neural network with different features like MFCC, mel, and combination of these features. The model is evaluated on Mozilla database and showed good performance in gender recognition. Our experimental results show that combination of MFCC and mel feature gives very good results in recognizing female speaker as well as male speaker. The results demonstrate that neural networks can capture efficient gender-based features from speech signal and also proved that proposed method is efficient on recognizing gender from speech signal which is useful in speech recognition applications.

References

1. Bishop, J., Keating, P.: Perception of pitch location within a speaker's range: fundamental frequency, voice quality and speaker sex. J. Acoust. Soc. Am. **32–2**, 1100–1112 (2012)
2. Gaikwad, S., Gawali, B., Mehrotra, S.C.: Gender identification using SVM with combination of MFCC. Adv. Comput. Res. **4**, 69–73 (2012)
3. Zeng, Y.M., Wu, Z.Y., Falk, T., Chan, W.Y.: Robust GMM based gender classification using pitch and RASTA-PLP parameters of speech. In: Proceedings of the International Conference on Machine Learning and Cybernetics, pp. 3376–3379 (2006)
4. Vergin, R., Farhat, A., O'Shaughnessy, D.: Robust gender-dependent acoustic phonetic modelling in continuous speech recognition based on a new automatic male/female classification. In: Proceedings of International IEEE Conference Acoustics, Speech, and Signal Processing (ICASSP-96), vol. 2, pp 1081–1084. Atlanta, May 7–10 1996
5. Harb, H., Chen, L.: Voice-based gender identification in multimedia applications. J. Intell. Inform. Syst. **24**(2), 179–198 (2005)
6. Zeng, Y., Wu, Z., Falk, T., Chan, W.Y.: Robust GMM based gender classification using pitch and RASTA-PLP parameters of speech. In: Proceedings of 5th IEEE international conference machine learning and cybernetics, pp 3376–3379. China (2006)
7. Metze, F., Ajmera, J., Englert, R., Bub, U., Burkhardt, F., Stegmann, J., Muller, C., Huber, R., Andrassy, B., Bauer, J.G., Littel, B.: Comparison of four approaches to age and gender recognition for telephone applications. In: Proceedings of 2007 IEEE International Conference on Acoustics, Speech and Signal Processing, vol. 4, pp. 1089–1092. Honolulu, April 15–20 2007
8. Ververidis, D., Kotropoulos, C.: Automatic speech classification to five emotional states based on gender information. In: Proceedings of European Signal Processing Conference (EUSIPCO 04), vol. 1, pp. 341–344, Vienna, Austria, Sep. 6–10 2004
9. Lin, Y.L., Wei, G.: Speech Emotion Recognition Based on HMM and SVM. In: Proceedings of IEEE International Conference Machine Learning and Cybernetics, vol. 8, pp. 4898–4901. Guangzhou, China (2005)
10. Xiao, Z., Dellandréa, E., Dou, W., Chen, L.: Hierarchical classification of emotional speech. Technical Report RR-LIRIS-2007-06, LIRIS UMR 5205 CNRS (2007)
11. Raahul, A., Sapthagiri, R., Pankaj, K., Vijayarajan, V.: Voice based gender classification using machine learning. Published under licence by IOP Publishing Ltd., IOP Conference Series: Materials Science and Engineering, vol 263, Issue 4
12. Qawaqneh, Z., Mallouh, A.A., Barkana, B.D.: Deep neural network framework and transformed MFCCs for speaker's age and gender classification. Knowl.-Based Syst. **115**, 5–14 (2017)
13. Kabil, S.H., Muckenhirn, H., Magimai-Doss, M.: On learning to identify genders from raw speech signal using CNNs. In: Proceedings of Interspeech, pp. 287–291 (2018)

14. Doukhan, D., Carrive, J., Vallet, F., Larcher, A., Meignier, S.: An open-source speaker gender detection framework for monitoring gender equality. In: IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (2018)
15. Buyukyilmaz, M., Cibikdiken, A.O.: Voice gender recognition using deep learning. In: Conference on Modeling, Simulation and Optimization Technologies and Applications (2016)
16. Markitantov, M., Verkholyak, O.: Automatic recognition of speaker age and gender based on deep neural networks. In: International Conference on Speech and Computer (SPECOM) (July 2019)
17. Mozilla: Common voice. Retrieved from https://voice.mozilla.org/ and https://www.kaggle.com/mozillaorg/common-voice
18. McFee, B., et al.: librosa: audio and music signal analysis in Python. In: Proceedings of the 14th Python in Science Conference, pp. 18–24 (2015)

A Survey on Collaboration Technologies and Systems of ICT Application in the Field of Education

Shivam Ribadiya, Dweepna Garg, and Janardan Bharvad

Abstract Information and Communication Technology (ICT), such as the global use of mobile phones and the Internet, has made an unparalleled contribution to the growing world economy. In education, the significance of information and communications technology (ICT) is recognized around the world. Numerous information and communication technology platforms in education are making the teaching and learning process productive, reliable, easier, pleasant, wider, and more comprehensive. Efforts are being made to provide digital information technology resources in academic institutions globally. Institutions of higher education, including government and private-funded (state and central government), have made tremendous investments in enabling ICT institutions. The academic effectiveness of ICTs depends on different factors, like usage, intention, and structure of socioeconomic support. Both modern technologies, equipment and facilities, require routine repairs and updates that are not carried out due to a lack of skilled and committed staff, resulting in their obsolescence. In order to clarify the above observations and some of their consequences in a larger socioeconomic sense, an institutional framework is used. This paper focuses on various ICT platforms and its application for teaching–learning approach.

Keywords ICT · WWW · RFID · AR · VR · MR · Social media networks

S. Ribadiya (✉) · D. Garg · J. Bharvad
Devang Patel Institute of Advance Technology and Research (DEPSTAR), Charotar University of Science and Technology (CHARUSAT), 139, CHARUSAT Campus, Highway, Off, Nadiad - Petlad Rd, Changa 388421, Gujarat, India
e-mail: shivamribadiya@outlook.com

D. Garg
e-mail: dweeps1989@gmail.com

J. Bharvad
e-mail: janardanbharvad.dce@charusat.ac.in

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_71

1 Introduction to ICT Application

Information and Communication Technology (ICT) is a wide term for information technology industry and society in global world. ICT enables all communication technologies including computer, information sharing, mobile phones, software, Internet, wireless communication technology, video meeting, social media network applications, stores of information and user services like access, store and retrieval of digital information from one place to another place with help of wired or wireless medium. Information and Communication Technology (ICT) is a type of education that uses modern information and communication technologies to facilitate, develop, and maximize the provision of information. Global research has also shown that ICT may lead to improved learning for the students and effective teaching strategies. As our development is closely associated with innovation and the process of education is no exception, ICT is certainly the direction to take for organizations, particularly in countries and ensuring that its students pursue high quality of education will decide whether the organization can step on or die. In this paper, a detailed survey regarding Information and Communication Technology (ICT) enables applications and platforms for the institution. Section 2 describes survey on ICT tools and technology that are applied in the field of academic institution like satellite communication, video conferencing, World Wide Web (Web 3.0), ICT application using RFID instrument, Word Wide Web (Web 2.0), blogs and wikies as ICT application platform, LCD projector, computer-aided instruction (CAI), computer-assisted learning (CAL), smart and interactive board, social media network, and future scope of ICT application.

2 ICT Tools and Technology Used in Classroom Teaching–Learning

2.1 Satellite Communication for ICT Application

With arrival of the Early Bird, the first communication satellite, the era of satellite communication dawned in 1962. In 1965 and 1971, the two main multinational Intelsat and Intersputnik satellite networks started operating. India deployed a communications satellite named INSAT, and EDUSAT was deployed in 2004 for education purposes. In the year of 2007, INSAT-4CR satellite was deployed by GSLV_F04 lunch vehicle. After that, INSAT-4CR has been missing and collapse as a backup satellite for INSAT-4C [1].

2.2 Video Conferencing Approach for ICT Application

Video conferencing is a two-way system of communication. Teleconferencing is often referred to as the use of broadcast video and audio systems together with devices to enable digital vision to people, auditory perception and talking to each other in various places. It can enable individuals gathering in different meeting rooms or specially fitted television cabin [2–4].

2.3 World Wide Web (Web 3.0) for ICT Application

One of the online tools built to support, post, coordinate, and have access to information on the Internet is the World Wide Web, which is referred to as WWW, W3, or simply the Web. In the year of 1989, Tim Berners Lee first created the Web while working at CERN which is in Switzerland's European Particle Physics Laboratory [5]. Word Wide Web provides the stage to teaching and learning for school students and higher education students, Word Wide Web acts as a platform of various ICT application for instructors and students in the institution. With the help of World Wide Web, students can grasp the knowledge using ICT enable platform like Google Meet, Microsoft Team, Edmodo, Zoom Meet, Google Jamboard, etc. Worde Wide Web not only the platforms for various Web applications but also a part of ICT-enabled applications.

2.4 ICT Application Using RFID Instruments

Radio frequency identification is the wireless application of electromagnetic fields to transfer information. The purpose of RFID is to identify automatically and trace the tags which is attached to a thing. The electronical information is stored in tags. Tags are operated by induced electromagnetic induction from the magnetic field near the reader. The tag does not actually need to be beyond the reader's line of sight, unlike a barcode, and can be inserted in the tracked object. People can also use library circulation service and crime prevention services for a day's RFID [6, 7]. RFID enables systems go in depth surveillance to be a monitoring method that integrate surveillance with more effective library-based monitoring of materials, including smoother and quicker discharge and processing, inventorying, and handling of materials. This technology enables librarians, while charging and discharging items, to reduce valuable staff time spent scanning barcodes.

2.5 World Wide Web (Web 2.0) for ICT Application

Tim O' Reilly at O' Reilly Media coined the Web 2.0 phrase. Web 2.0 defines sites on the World Wide Web that utilize technologies above existing Web sites' static pages [8]. While Web 2.0 introduces a new model of World Wide Web, it does not apply to any technological standards being updated, but also to systematic improvements in the manner in which Web pages are created and used. It enables people to engage, communicate, and chat synchronously and asynchronously with each other. Social media applications, blogs post, wikis, and video sharing platforms are all based on Web 2.0 framework. Users will connect globally at a negligible cost through Web 2.0 applications. Instead of collecting input from a common source, it helps individuals to communicate and disseminate thoughts with each other.

2.6 Blog and Wikis as ICT Application Platform

Blogs and wikis are essentially Web 2.0, and databases and teaching–learning processes have immense consequences for their global expansion [8]. In the history of media, blogs can also be an even bigger achievement compared to the Web sites. They ensure the accelerated production and consumption of Web-based publications. Blogs are HTML or HTML5 with alterative GUI for better illustration. A blogging Web site is often referred to as an online blog or personal blog, a Web-based document containing mostly of regular documents. Blogs include articles by an individual or a group that are often identical to journal entries. In alphabetical order, the article is dated and listed. Wiki is a shared writing platform available on the Internet. The wiki is a shared online space where material can be created by anybody and material can be changed by anyone. Wikis are built to help communities interact, connect, and create online information sharing portal and are particularly beneficial for time-and-place-divided learners. Wikis offer a more effective approach to community writing and editing than submitting e-mail addons with recorded revisions, the strategy which serves one checker at a time and can cause problems with students having different and inconsistent copies of the same text. The term "wiki" invented from the Hawaiian term "wiki wiki" which means easy or simple. Wikis can be established easily and are easy to understand and update. They provide useful functionality such as graphic editors, content creation, updating and omission, Web site version, information sharing, feeds for subscriptions, search facility, disorganized classification, and statistics for stakeholders.

2.7 Computer-Aided Instruction for Academic ICT

Computer, as an aid to the training, provides a series of programming instructions that are used to develop those skills among the learners in the teaching–learning process. Here, the machine is used to introduce the pupils, exercises, lesson, workout and tutorial sequences, and some time to involve the pupils in a conversation about the content of the teaching. It seems like a teacher's machine is used as a learning tool. So, computer-aided instruction is a category of coaching that is used to accomplish the instruction's purposes [9]. Computer-aided instruction acts as a coaching aid and promotes learners' neurological-based learning, self-pacing, and personalized instruction. In computer-aided instruction, the role of the instructor has shifted from a boss or a guide to the conventional way of presenting lectures. No robot may substitute an instructor. A "teacher" position is quite necessary in the teaching–learning process [10]. The instructor must play too many roles in computer-aided instruction, such as a computer engineer, a lesson author, and a device manager, as computer-aided instruction needs the skill of the specialists listed above.

2.8 Computer-Assisted Learning for Academic ICT

Computer-assisted learning is the utilization of technology to facilitate or encourage people's schooling or preparation. It is more than one word that is used to describe this computer task. Computer-assisted learning, computer-driven, and computer-controlled instruction are other words. The word computer-assisted learning encompasses a selection of computer-based services that typically tend to offer immersive instruction in a particular subject field, and numerous dates back the cyberspace. This can vary from specialized and expensive business packages to project-developed innovations in other research institutions or national programmer to simple ideas for solving a very local problem developed by citizens without financing or support. Computer-assisted learning reflects the next step in the usage of computer-based educational mechanism in which the harmful element of computer-assisted learning has been attempted to address. In a quite short amount of time, computer-assisted learning is to express a huge amount of data [11].

2.9 Interactive ICT Method Using LCD Projector Screen

The liquid-crystal display (LCD) projector is a form of cinematic viewer used on a screen to view video, pictures, or digital information for teaching learning in academic institution, business, government organization, etc. That is a revamped version of a film reel or an over-head projector [12]. Liquid-crystal display (LCD) projectors usually relay lighting through a mirror or series of diagnostic filter from

a metal halide lamp to show images separating light across three polysilicon panels, each for the red, green, and blue component of the video signal. As polarized light passes through the screens, it is possible to design pixel values to allow light to move, or closed pixels will produce a wide range of colors and shades in the reflective surface [13]. It is used to view pictures, diagrams, power point presentation (PPT), digital information, or many multimedia information as a teaching aid in classroom activities.

2.9.1 ICT Application Using Smart and Synergistic Board

The interactive board is an integrated whiteboard that uses user interface human touch enable (e.g., scrolling down and up, right mouse-click using touch, cursor movement, etc.). It is comparable as a standard personal computer or laptop control devices [16]. From a finger, pen, or other solid piece, the white board recognizes touch information. In the school, smart boards can be used instead of blackboards or whiteboards. A "Flick and Scrolling" attribute was incorporated by the interactive board 800 series digital whiteboard. Along with a framework that includes the digital white board, the computer, the projector, and white boarding software, the interactive board digital white board functions as part of either the Smart Notebook cooperative learning platform for instructors or the Smart Meeting Professional Business Applications. The connection medium of components is wireless or USB or serial cable [17]. The screen pictures are seen on the whiteboard by a projector attached to the monitor.

2.9.2 Virtual Reality, Augmented Reality and Mixed Reality (VR/AR/MR) in ICT Application

The virtual reality (VR) technology provides the interactive learning trip experience of a fully immersive book chapter, and millions of teachers and students around the world are going to wish the VR base ICT application for teaching and learning [18]. From history to human anatomy, VR can be used to teach tons of things. The effect by using technologies to superimpose knowledge—sounds, pictures, and text—on the world seen by augmented reality (AR). When using the camera on a mobile phone, AR also applies multimedia elements to a live view. Smart phones are one of the most common ways that AR has invaded daily life [19]. Mixed reality (MR) is the integration between natural and virtual realities, also known as hybrid reality, to create new experiences and visualizations where actual and artificial objects coexist and communicate in real time [20]. Civil engineers having to identify strain points and flaws in their bridge projects, or somebody researching the structure of our mind to see what occurs to the mind whenever, it is put under extreme stress, are perfect examples. In India, experiential learning in the form of virtual laboratories, social media platforms, augmented reality and virtual reality technologies, and game-based education is being introduced. AR/VR/MR game-based learning is an efficient pedagogy that, by incorporating game features into learning contexts, enhances students' interest

Table 1 Future scope of ICT

S. No.	ICT characteristics for future scope
1	By illustrative approach of knowledge
2	Collaborative creativity
3	Cognitive, social, and technological skills
4	Collaborative exploration and building of knowledge
5	Cultivation of critical thinking
6	High observability in education sector
7	Trialability of ICT application for easy use
8	Virtual education and virtual learning environment
9	Collaboration of digital education
10	Technology enhanced learning

and interaction [21]. Unity Technologies, the market-leading forum to produce a wide variety of real-time 3D and AR / VR applications, completed its second version of Unite India in Hyderabad on December 6, 2018.

2.9.3 Social Networks for ICT Application

Social networks are digital platforms based on Web 3.0 or Web 2.0 applications. Facebook is the most popular online source of all social media. People can exchange details, upload comments on photographs/videos, information regarding education, notification of academic notification, etc., on the social media network [22]. That is another way of communicating which is synchronous. People can interact with each other instantaneously.

In Table 1, characteristics of ICT as future scope which are enlarging the area of Information and Communication Technology in education. ICT platforms and application for academic organization intensify the teaching and learning skill for tutors and students.

3 Conclusion

The use of information and communication technology is to promote and expand teaching and learning at all stages, in all areas, and for people from all backgrounds. The core pieces required best achieve the transitions made possible by technology in education are in place, from the automation of the e-learning rate to the development and implementation of publicly approved education services. These tools and mechanisms can also be integrated into their activities by students. These organizations should work together with communities, academics, educational centers, and all other stakeholders to reduce inefficiencies, expand beyond the walls of conventional

schools, and form broad alliances to encourage learning everywhere, all the time. All learners can access opportunities, interactions, preparation tools, and knowledge, regardless of their perceived skills or geographical locations, that can set them on a path to study including the research design unthinkable a generation ago using ICT. Instantaneous ICT enable comfort, learning dashboards, and teamwork, and networking tools can better connect teachers and families. Information and Communication Technology enables greater collaboration, sharing of resources, information, and enhanced experience for these roles, too, so that the mission is shared by everyone and devoted to helping every participant in the organization maximize student learning. So, this paper elaborates the use of information and communication technologies to facilitate learning, and it is a moment of tremendous opportunity and development students and society.

References

1. Rajashekhar, S.L., Ayyangar, G.V., Sharma, R.: Satellite-based distance education in digital paradigm: ISRO perspective. In: 6th International Conference on Digital Content, Multimedia Technology and Its Applications, Seoul, pp. 366–371 (2010)
2. Gharbi, S., Zangar, N., Aitsaadi, N.: Overview: high altitude platform network for disaster and crises application. In: 2019 International Conference on Information and Communication Technologies for Disaster Management (ICT-DM), Paris, France, 2019, pp. 1–2. https://doi.org/10.1109/ICT-DM47966.2019.9032991
3. Yaqoob, A., Bi, T., Muntean,G.-M.: A Priority-aware DASH-based multi-view video streaming scheme over multiple channels. In: 2020 International Wireless Communications and Mobile Computing (IWCMC), Limassol, Cyprus, pp. 297–303. https://doi.org/10.1109/IWCMC48107.2020.9148582
4. Lin, C.R., Zhang, H., Liu, J., Chen, S.: Implementation of secure web conferencing. In: 2020 2nd International Conference on Computer Communication and the Internet (ICCCI), Nagoya, Japan, pp. 12–17 (2020). https://doi.org/10.1109/ICCCI49374.2020.9145972.
5. Daoust, F.: Update from the World Wide Web Consortium (WC3). SMPTE Motion Imaging J **129**(8), 80–83 (2020). https://doi.org/10.5594/JMI.2020.3001776
6. Nyakonda, T., Tsietsi, M., Terzoli, A., Dlodlo,N.: An RFID flock management system for rural areas. In: 2019 Open Innovations (OI), Cape Town, South Africa, pp. 78–82 (2019). https://doi.org/10.1109/OI.2019.8908190
7. Gautam, A.K., Farhan, M., Agrawal, N., Rambabu, K.: Design and packaging of a compact circularly polarised planar antenna for 2.45-GHz RFID mobile readers. IET Microw. Antennas Propag. **13**(13), 2310–2314 (2019). https://doi.org/10.1049/iet-map.2019.0261
8. Vivas, G.M.V.: Web tools 2.0 and its influence on academic performance in collaborative learning environments: an empirical study. In: 2020 15th Iberian Conference on Information Systems and Technologies (CISTI), Sevilla, Spain, pp. 1–4 (2020). https://doi.org/10.23919/CISTI49556.2020.9141113
9. Yang, Y., Lang, H.: Project based instruction for computer aided design and mechatronics courses. In: 2020 15th International Conference on Computer Science & Education (ICCSE), Delft, Netherlands, pp. 406–410 (2020). https://doi.org/10.1109/ICCSE49874.2020.9201888
10. Yu-bao, L., Xiao-jing, W., Qian-li, M., Shao-tang, L.: How to improve the quality and effect of computer aided instruction's application in classroom teaching in institutes of higher learning. In: 2010 Second International Workshop on Education Technology and Computer Science, Wuhan, pp. 658–661 (2010). https://doi.org/10.1109/ETCS.2010.325

11. Ye, Y., Nie, S.: Problems and countermeasures in the use of computer aid instruction in mandarin teaching. In: 2011 International Conference on Multimedia Technology, Hangzhou, pp. 5457–5459 (2011). https://doi.org/10.1109/ICMT.2011.6002288
12. Takafuji, Y.: TRIZ-based design of an LCD for projectors. In: 2019 Pan Pacific Microelectronics Symposium (Pan Pacific), Kauai, HI, USA, pp. 1–6 (2019). https://doi.org/10.23919/PanPacific.2019.8696808
13. Matsueda, Y., Shimobayashi, T., Okamoto, N., Yudasaka, I., Ohshima, H.: 4.55-in. HDTV poly-Si TFT light valve for LCD projectors. In: Conference Record of the 1991 International Display Research Conference, San Diego, CA, USA, pp. 8–11 (1991). https://doi.org/10.1109/DISPL.1991.167419
14. Chilcutt, A.S., Brooks, A.J.: Organizing and outlining your presentation. In: Engineered to Speak: Helping You Create and Deliver Engaging Technical Presentations, IEEE, pp. 49–62 (2019). https://doi.org/10.1002/9781119474913.ch5
15. Chen, K., Deng, L.Y.: The study on design speech-control powerpoint presentation tool. In: 2010 International Conference on Machine Learning and Cybernetics, Qingdao, pp. 3009–3014 (2010). https://doi.org/10.1109/ICMLC.2010.5580738
16. Bhatt, D.P., Tiwari, M.: Smart traffic sign boards (STSB) for smart cities. In: 2nd Smart Cities Symposium (SCS 2019), Bahrain, Bahrain, pp. 1–4 (2019). https://doi.org/10.1049/cp.2019.0194
17. Sunitha, D., Patil, V.C., Manjula, H.N., Jebakani, S.: Digital notice board using Smart Phones-speech recognition voice command. In: 2018 International Conference on Current Trends towards Converging Technologies (ICCTCT), Coimbatore, pp. 1–4 (2018). https://doi.org/10.1109/ICCTCT.2018.8551159
18. Paul, S., Hamad, S., Khalid, S.: The role of AR/ VR in an IoT connected digital enterprise for smart education. In: 2019 Sixth HCT Information Technology Trends (ITT), Ras Al Khaimah, United Arab Emirates, pp. 305–308 (2019). https://doi.org/10.1109/ITT48889.2019.9075102
19. Bryan, S.J., Campbell, A., Mangina, E.: Scenic spheres—an AR/VR educational game. In: 2018 IEEE Games, Entertainment, Media Conference (GEM), Galway, pp. 1–9 (2018). https://doi.org/10.1109/GEM.2018.8516456
20. Koumaditis, K., Venckute, S., Jensen, F.S., Chinello, F.: Immersive training: outcomes from small scale AR/VR pilot-studies. In: 2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), Osaka, Japan, pp. 1–5 (2019). https://doi.org/10.1109/VR44988.2019.9044162
21. Sitharan, R., Kian, N.T., Mai, N., Yeen-Ju, H.T., Syahmi Abd Aziz, M., Bin Dollmat, K.S.: Work-in-progress—assisting AR, VR and Hologram learning experience through MOOC. In: 2020 6th International Conference of the Immersive Learning Research Network (iLRN), San Luis Obispo, CA, USA, pp. 366–368 (2020) https://doi.org/10.23919/iLRN47897.2020.9155133
22. Sekhaolelo, L., Kalema, B.M., Chiyangwa, B.: Leveraging social networks to enhance learning at higher institutions. In: 2018 International Conference on Computer, Information and Telecommunication Systems (CITS), Colmar, pp. 1–5 (2018). https://doi.org/10.1109/CITS.2018.8440184

Real-Time Data Monitoring System for User Conveyance

Venkatesh Mane, Shweta Kore, Preeti S. Pillai, C. I. Nalini, and Abhishek Puri

Abstract Advances in real-time data monitoring devices bring the present-day system vision to the near future. Such devices can be automatically discovered, configured, and invoked for a given task. For instance, weather monitoring is a device that comprises various measuring instruments for the measurement of temperature, humidity, pressure, altitude, and many more. Nowadays, the weather information is mostly gathered with the help of devices like smartphones and PCs which are connected to the Internet. The measurements on various factors of weather can be made a few times manually in a single day. But, in an automatic system, the measurements are taken at least once an hour. Therefore, a device can replace as a solution to such difficulties in monitoring weather conditions, smaller in size, and is of less weight, i.e., called a data monitoring device. Also, a device can be developed such that one can get a live update on a particular issue. Here, the plug-and-play device consists of sensors that are used to sense the surroundings to provide the accurate information based on the sensor chosen like DHT11 humidity and temperature sensor, BMP180 barometric sensor, HCSR04 ultrasonic sensor, and gathering some live updates. STM32F407VGT6 microcontroller for interfacing sensors and Nextion display as an output device is used. Using the ESP8266 Wi-Fi module, the variables from all the above-mentioned sensors can be sent on to the mobile or PC that are connected to the Internet.

Keywords STM32F407VGT6 · DHT11 · BMP180 · HCSR04 · ESP8266 · Nextion

V. Mane (✉) · P. S. Pillai · C. I. Nalini · A. Puri
ECE Department, KLE Tech University, Hubli, India
e-mail: mane@kletech.ac.in

P. S. Pillai
e-mail: preeti_pillai@kletech.ac.in

C. I. Nalini
e-mail: nalinic@kletech.ac.in

S. Kore
JSS College, Dharwad, India

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_72

1 Introduction

As modern society is developing too fast, the oldfangled forms of devices that are used to gather any information from surroundings fail to fulfill a human need. Any monitoring of surrounding areas like humidity, pressure, altitude, temperature, or observation of distance for a particular object, getting some live information, searching on Personal Computer or mobile would be a tedious process. Nowadays, the Internet is the most widely used as high speed and easy communication medium across the world. One can have access to the data of any sensor that is interfaced with a microcontroller. This data monitoring device acts as secured and is very much feasible as it makes use of the Internet to send data over smartphones and PCs. This system is a more reliable, light-weight, and compact tool for any industry or home where there is a necessity to gather various changing variables that are surrounding us.

Real-time data monitoring device allows the monitoring of five variables, i.e., temperature, humidity, pressure, altitude, distance, and getting the COVID-19 live updates. The intention behind this system is to act as a helping tool or a device in any industry or home so that the decisions can be made at a faster rate and more accurate. Here, the DHT11 is used for the measurement of both humidity and temperature, and BMP 180 is used for measuring the pressure and altitude of the place. HCSR04 to get the distance of a particular object kept under observation. With the help of the ESP8266 Wi-Fi module, one can monitor the data of a sensor that is interfaced with a microcontroller, sitting anywhere in the world with a smartphone or PC. The data from these sensors also updates on the Nextion display based on the commands given from Nextion display.

2 Literature Survey

The work briefly introduces the need for monitoring weather conditions in storerooms of various industries. A system is developed that comprises Arduino, wireless sensor networks, and cloud. Then to develop a method by which one can transmit the data for monitoring both humidity and temperature. The remote sensor systems are associated with the Web with the help of IoT. Such a device or a system can help in industrial automation, with which one can make quick and intelligent decisions [1]. The ideas in this work suggest that the climate can be monitored, and also, data is transmitted without the help of the Internet. In some of the exceptionally made sure about regions, it is beyond the realm of imagination to expect to utilize mobiles and other Web-associated gadgets, particularly in military and exploration research facilities. To get rid of these problems in monitoring weather with versatile size and less weight gadget is planned. This comprises various sensors that are interfaced with Arduino mega to sense the surrounding environment and provide accurate information or data from sensed from these sensors. Here, the weather parameters are transferred to mobile or PC via the ESP8266 Wi-Fi module [2].

Here, the paper briefly explains about monitoring weather in real time, making use of an IoT and Raspberry Pi. Raspberry Pi board based on ARM and raspbian operating system to choose linux kernel for Raspberry Pi is used. Some of the digital sensors like BMP180, DHT11, and analog sensors like LDR are made use for the measurement of various parameters related to the environment. The HTTP protocol is used to provide networking of servers with the client. The system will be able to monitor pressure, altitude, light intensity, humidity, and temperature [3].

The work briefly introduces a precise method for the estimation of altitude by making use of data fusion algorithms. It provides a precise method to estimate the altitude of a body which is rigid by blending the data of four low-cost sensors namely gyroscope, altimeter, accelerometer, and magnetometer [4].

The paper mainly focuses on monitoring data in real time for the hospitals, industries, schools, homes, etc., where data collection regarding various parameters of weather like humidity, altitude, temperature, and pressure is more necessary. Also, these values of different parameters need to be updated regularly like once in an hour. Using these sensors, Wi-Fi module, and MQTT protocol, i.e., important for transmitting and monitoring the data from sensors at remote locations regularly [5].

3 Methodology

With the help of the literature survey, a different approach is made to get the best possible system. The methodology consists of a functional block diagram of the proposed design. The block diagram mainly consists of the STM32F407VGT6 microcontroller, DHT11 humidity and temperature sensor, BMP180 barometric sensor, HCSR-04 ultrasonic distance sensor, ESP8266 module, and Nextion display (Fig. 1).

As soon as the STM32 board is powered, the Nextion display that is interfaced with the STM32 board will also get powered on. And the first page will appear as soon as the Nextion is powered. The data is received from Nextion display to STM32 using UART communication. The user can move to the next page by clicking on the login

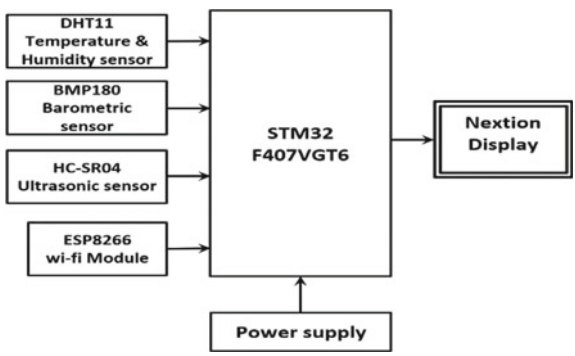


Fig. 1 Functional block diagram

button on the Nextion display.On clicking the first option in the menu, i.e., the DHT11 sensor, the respective page will open up, and thus, the STM32 will initiate the DHT11 sensor. Once the initialization part of the sensors gets completed, data transmission comes into action. Here, the sensor will send a data length of 40 bits. Data transmitted by the sensor comprises integral relative humidity data of 8 bit + decimal relative humidity data of 8 bit + integral temperature data of 8 bit + decimal temperature data of 8 bit + checksum of 8 bit and thus, provides precise data on the page. On clicking the second option in the menu, i.e., the BMP180 sensor, the respective page will open up, and thus, the STM32 will initiate the BMP180 sensor. It makes use of I2C for the communication and sends the values of pressure and temperature which are uncompensated. The calibration values from the EEPROM with the starting address from 0xAA are read. Once the calibration is read, the calculations are done as per the datasheet and and provide the data of pressure and altitude on to the page. On clicking the third option in the menu, i.e.,, the HCSR04 sensor, the respective page will open up, and thus, the STM32 will initiate the HCSR04 sensor. The trigger pin is pulled high for about 10 μs. Now, the sensor transmits an 8 cycle burst of ultrasound at 40 kHz via a transmitter present on the sensor and waits for the detection of any pulse signal coming back to the sensor receiver. The sensor will automatically return a high signal whose width is proportional to the distance of an object kept away from it. The following formula helps to find out the distance between the object and the sensor. Range = time of echo $\times$ speed of sound (340 m/s)/2. On clicking the fourth option in the menu, i.e., the COVID-19 updates, the respective page will open up. The STM32 will initiate the ESP8266-01module and thus starts collecting the data regarding the COVID-19 updates from the worldometer and provides the data of total, deaths, recovered, and active cases on to the page.

3.1 STM32F407VGT6 Microcontroller

The STM32F407xx family is based on the more efficient ARM Cortex-M4. The core, i.e., 32-bit RISC, operates at 168 MHz frequency. The Cortex-M4 core highlights a floating-point unit (FPU), which helps in all the data processing instructions and data types which are single precision. Indeed, even it additionally actualizes a full set of DSP instructions and a memory protection unit which builds up the application security. The family of STM32F407xx includes embedded memories of high speed like the 1 Mb of flash memory, 192 kB of SRAM, 4 kB of backup SRAM, and a wide scope of I/Os and the two APB buses are connected to the peripherals, three AHB buses, and 32-bit multi-AHB bus matrix. Each of these devices provides RTC, twelve general-purpose 16-bit timers as well as two PWM timers for the motor control, two general-purpose 32-bit timers, three 12-bit analog-to-digital converters, two digital-to-analog converters, and true random number generator. They also feature a higher standard and more advanced communication interfaces.

3.2 DHT11 Humidity and Temperature Sensor

DHT11 is a digital humidity and temperature sensor from AOSONG. The main feature or advantage of this sensor is that it provides both humidity and temperature to one decimal place. The sensor provides both humidity and temperature values serially with the help of one data line present on it. The values of relative humidity are being presented in percentages ranging from 20 to 80% with ±5% accuracy and values of temperature are presented in degree celsius ranging from 0 to 50 °C with ±2 °C accuracy.

3.3 BMP180 Barometric Sensor

BMP180 sensor belongs to the family of the BMPXXX series. All the members of this family are designed to measure pressure, altitude, and temperature. This provides high-precision data which is specially designed for consumer applications. Barometric pressure is described as the weight of air applied on any given object or any point. The sensor is considered to be very much efficient as it consumes power less than 1 mA during the time of measurements and about 5 μA during idle. Any microcontroller can have an easy interface to this sensor, as it works on I2C. The sensor can measure the pressure values ranging from 300 to 1100 hPa having 0.02 hPa an accuracy in higher resolution mode. And the temperature measurements ranging from −40 to 85 °C have ±1.0 °C accuracy.

3.4 HCSR-04 Ultrasonic Sensor

HC-SR04 ultrasonic distance sensor can sense the distance of any objects kept up to 13 ft away from it. The ultrasonic transducers, namely the transmitter and receiver, act as the main component of an HC-SR04 ultrasonic distance sensor. The conversion of electrical signals into ultrasonic pulses of 40 kHz is done via a transmitter. The receiver part of the sensor waits for transmitted pulses to return. The width of the pulse that is returned is used for the calculation of the distance of an object which is kept away from the sensor. Its compactness and simple use makes it very much feasible to choose in projects related to robotics.

3.5 ESP8266 Wi-Fi Module

The ESP8266 is a serial Wi-Fi wireless transceiver. It is an on-chip self-contained chip comprising of an integrated TCP protocol stack. Due to this, any microcontroller

that is interfaced with it will have access to the Wi-Fi network. Therefore, the module can host any application or help in uploading Wi-Fi networking operations from other application processors. Each of these modules is installed with an AT command set firmware. By making use of a set of AT commands, the module will be able to have a communication with that of any microcontroller having UART with the required baud rate. The core of this module is based on Tensilica Xtensa consisting of RISC CPU of 32 bits operating at a frequency of 80 MHz and memories such as the instruction RAM of 64 kB, boot ROM of 64 kB, and data RAM of 96 kB. With the help of a serial peripheral interface, it makes it simple to have access to the external flash memory. It is a standalone module of low cost that will be used for endpoint IoT developments.

3.6 Nextion HMI Display

A Nextion is one of the human–machine interface solutions blending an incorporated processor along with a touch display having memory. The Nextion editor helps in the development of human–machine interface projects. The displays are made of resistive touchscreens that make it feasible to create a GUI. It is a highly recommend device in case of situations where monitoring and controlling of processes are of greater interest. The display can be interfaced with MCU peripheral via the UART pins present on both the devices, namely the 5v and ground for power supply, transmitter, and receiver pins to have a communication from both the sides. This type of UART communication helps in getting the event notifications from the Nextion to MCU peripheral. Then, the MCU will in turn update various progresses on to the Nextion display with the instructions that are based on ASCII format [6]. The size Nextion ranges from 2.4″ to 7″. Depending on the user requirement, one can select his or her interest in display size. The 2.4″ sized Nextion display is made use in this project. The display is incorporated with a microcontroller that is based on ARM. Thus, it can control what to display on to it. Also, it can show things like the creation of texts, buttons, and images in the background. This display can communicate with the board that has serial capabilities like Arduino, STM32, Raspberry Pi, ESP32, ESP8266, and so on with a baud rate of 9600. The Nextion editor helps in easy and quick development in the creation of drag-and-drop components such as the sliders, texts, graphics, timers, buttons, texts, dual state buttons, etc.

4 Results and Discussion

The following results show the monitoring of sensors data such as temperature, humidity, pressure, altitude, and the distance of an object kept away from the sensor on Nextion display which will be used as both an input and output device, to get COVID-19 updates from worldometer.On clicking the first option in the menu, i.e.,

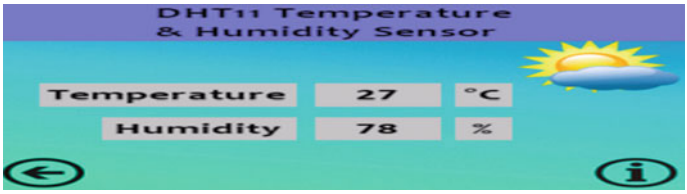


Fig. 2 Temparature and humidity

the DHT11 sensor, the respective page will open up, and thus, the STM32 will initiate the DHT11 sensor and thus provides the data of temperature and humidity on to the page as shown in Fig. 2.

On clicking the second option in the menu, i.e., the BMP180 sensor, the respective page will open up, and thus, the STM32 will initiate the BMP180 sensor and thus provides the data of pressure and altitude on to the page as shown in Fig. 3.

On clicking the third option in the menu, i.e., the HCSR04 sensor, the respective page will open up, and thus, the STM32 will initiate the HCSR04 sensor and thus provides the distance of any object that is being kept away from the sensor. On clicking the fourth option in the menu, it displays COVID-19 data, and the respective page will open up. The STM32 will initiate the ESP8266-01 module and thus starts collecting the data regarding the covid19 updates from the worldometer and provides the data of total, deaths, recovered, and active cases on to the page as shown in Figs. 4 and 5.

Now, the following results show the monitoring of sensor data such as temperature, humidity, pressure, altitude, and distance of object form the sensor on the ThingSpeak server via ESP8266. The sensor's data will be updated on the ThingSpeak server.

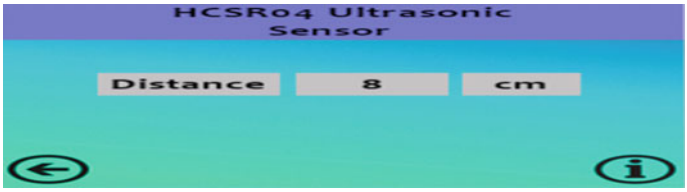


Fig. 3 Pressure, altitude, and distance

Fig.4 COVID-19updates in India

Fig. 5 Pressure and distance data on ThingSpeak

Updating multiple fields at once becomes a requirement because ThingSpeak allows us to communicate once every 15 s.

5 Conclusion and Future Scope

The data monitoring device is designed and implemented which can display data on Nextion and transmit data on ThingSpeak. It assists in better monitoring of various parameters, reliable data transfer, elimination of regular maintenance, and cost-effective and simple to install. The device proposes wireless connectivity providing a reliable, portable, and low-cost tool for an industry where it is necessary to have various data monitoring controls to carry out critical processes. The future work is to increase the number of sensors interfaced with a microcontroller and to have data entry from multiple locations and store the data in the cloud to access the data globally.

References

1. Roy, A., Das, P., Das, R.: Temperature and humidity monitoring system for storage rooms of industries, pp. 99–103 (2017)
2. Pattanashetty, V.B., Iyer, N.C., Ibal, P., Bagewadi, A.: Smart driving assistance. In: 2015 IEEE International Transportation Electrification Conference (ITEC), pp. 1–4. Chennai (2015). https://doi.org/10.1109/ITEC-India.2015.7386870
3. Pattanashetty, V.B., Iyer, N.C., Viswanath, H.L.: Intellectual conveyance structure for travellers. In: Saeed, K., Chaki, N., Pati, B., Bakshi, S., Mohapatra, D. (eds.) Progress in Advanced Computing and Intelligent Engineering. Advances in Intelligent Systems and Computing, vol. 564. Springer, Singapore (2018). https://doi.org/10.1007/978-981-10-6875-1_37
4. Pattanashetty, V.B., Iyer, N.C., Dinkar, A., Gudi, S.: Self-coordinating bus route system to avoid bus bunching. In: Satapathy, S., Bhateja, V., Joshi, A. (eds.) Proceedings of the International Conference on Data Engineering and Communication Technology. Advances in Intelligent Systems and Computing, vol. 469. Springer, Singapore (2017). https://doi.org/10.1007/978-981-10-1678-3_50

5. Pattanashetty, V.B., Iyer, N.C., Viswanath, H.L., Kore, S.: Inclusive device to connect people with the surrounding world. In: Fong, S., Akashe, S., Mahalle, P. (eds.) Information and Communication Technology for Competitive Strategies. Lecture Notes in Networks and Systems, vol. 40. Springer, Singapore (2019). https://doi.org/10.1007/978-981-13-0586-3_49
6. Pattanashetty, V.B., Shamshuddin, K., Iyer, N.C.: A top down approach from job to course. Proc. Comput. Sci. **172**, 204–206 (2020). ISSN 1877-0509. https://doi.org/10.1016/j.procs.2020.05.032

FTRAT: Fault-Tolerant Routing Based on Aggregation Tree to Improve the QoS in Wireless Sensor Networks

P. Manasa, K. Shaila, and K. R. Venugopal

Abstract Wireless sensor network is one of the effective communication fields due to its capability of performing functionalities such as sensing, aggregation and other computational activities. Cluster formation among the deployed sensor nodes has shown to be an effective method in saving life of the network. Cluster head selected supervises other nodes in the network and plays a vital role in data transmission. Since cluster head is responsible of monitoring the whole cluster, so cluster heads are more crucial in the network. If defective cluster head is present in the network, transmission in the network will be worst hit. In the proposed work, the network failure due to defective cluster head can be overcome by considering backup cluster head in the network. Flawless cluster member node energy is coordinated to form the backup cluster head. Aggregation tree formation among the cluster head and backup cluster head in case of cluster head failure in the network helps to improve QoS by adopting beneficiary path from cluster head to the sink.

Keywords Aggregation tree · Backup cluster head · Cluster head · Quality of service (QoS) · Routing and wireless sensor networks

1 Introduction

Wireless sensor networks (WSNs) frequently update to sink node by monitoring the deployed sensor devices. Routing of data from sensor to sink is carried out productively by adopting effective routing methods. Effective routing methods like

P. Manasa (✉)
Research Scholar, Department of Electronics and Communication Engineering, VTU-RC, Vivekananda Institute of Technology, Bengaluru, India

K. Shaila
Professor, Department of Electronics and Communication Engineering, Vivekananda Institute of Technology, Bengaluru, India

K. R. Venugopal
Vice-Chancellor, Bangalore University, Bengaluru, India

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_73

clustering of nodes enhance the quality of network communication [1–4]. Clustering technique groups sensor nodes and identifies cluster head (CH). Aggregation in CH includes compilation of data obtained from all its sensor nodes in order to achieve effective energy consumption. Redundant data transfer from cluster members to the sink can be avoided by using routing techniques via CH. CH extends battery life of a sensor node by accumulating the data obtained from all its cluster members [5–7]. Role of CH in monitoring all its associated members of the clusters becomes a vital function. There is a possibility of failure in CH during routing that completely disrupts the network. Finding faulty CH and avoiding it becomes a major challenge in fault-tolerant system [8–10]. The proposed method chooses a backup CH during failure of existing CH of the network. Choosing of backup CH solely depends on maximum retained energy of the nodes [11, 12]. The proposed algorithm constructs aggregation tree structure among all heads of the cluster and obtains high packet delivery in the network.

2 Related Work

Aranzazu-Suescun and Cardei et al. [13] have proposed four algorithms to monitor the events generated in mobile-sink WSNs which communicates with IoT. Delay involved is more when a routing is established using anchor-based routing algorithm. New tree-based routing algorithm may have a lesser delay compared with anchor-based routing algorithm but fall short in terms of energy consumption. Huang et al. [14] have proposed a method to improve the clustering accuracy by using *K*-subspace and auto-encoder functionality. Method here is to use soft assignments on behalf of hard assignments and optimizes the cluster formation to present in a more favourable way. Compared to other subspace clustering methods, this method outperforms in achieving high accuracy. Lyu et al. [15] have introduced a novel method for abnormality detection using fog computing and clustering algorithms. Fog computing is one of the effective ways to address large data set. The amalgamation of fog computing and clustering comprises clustering of nodes done at the cloud and the fog layer. It incorporates clustering of nodes using hyper-ellipsoidal technique and helps to detect the abnormalities accurately. This method does not address the privacy and security issues of fog computation data exchange. Puschmann et al. [16] have proposed an idea for processing large data set produced due to advancement in IoT system. Number of clusters formed is completely based on the incoming data streams, and hence this method results in adaptive cluster formation. The above scheme uses k-means technique for cluster formation and finds difficult to analyse the data set in case of drift and movable target system.

3 System Model

Sensor nodes $S_1, S_2, S_3 \ldots S_n$ are randomly deployed and n be the number of nodes deployed. Following parameters are considered during network deployment:

- Static sensor nodes and sink node are considered.
- Retained energy and positional information are updated to the sink node by sensor nodes.
- Average energy is calculated using the equation

$$E_{\text{AVG}} = \frac{(E_{n1} + E_{n2} + E_{n3} \ldots + E_{nn})}{n} \tag{1}$$

  Here, $E_{n1}, E_{n2}, E_{n3} \ldots E_{nn}$ are the retained energies of nodes $n_1, n_2, n_3 \ldots n_n$.
  For any of the node to be selected as CH, $E_{nx} \geq E_{\text{AVG}}$.
  Here, $x =$ node that is deployed.
- Distance from one node to other node or cluster head is calculated using the equation

$$d_{n\_\text{CH}} = \sqrt{(x_{n1} - x_{n2})^2 + (y_{n1} - y_{n2})^2} \tag{2}$$

  Here, (x_{n1}, x_{n2}) and (y_{n1}, y_{n2}) are positions of nodes 1 and 2 in the network.
  To transmit b-bits at a distance d_{n_CH}, energy consumed by the ratio is given by

$$E_{Tx}\left(b, d_{n\_\text{CH}}\right) = E_{Tx} - \text{rad}(b) + E_{Tx} - \text{amp}\left(b, d_{n\_\text{CH}}\right) \tag{3}$$

$$E_{Tx}\left(b, d_{n\_\text{CH}}\right) = E_{\text{rad}} \times \text{b} + E_{\text{amp}} \times b \times d_{n\_\text{CH}}^2 \tag{4}$$

  To receive b-bits, energy consumed by the ratio is given by

$$E_{Rx}\left(b, d_{n\_\text{CH}}\right) = E_{Tx} - \text{rad}(m) \tag{5}$$

$$E_{Rx}\left(b, d_{n\_\text{CH}}\right) = E_{\text{rad}} \times b \tag{6}$$

  where E_{rad} is the electronics energy and E_{amp} represents the energy required by an amplifier. Figure 1 shows the deployment of sensor nodes and its cluster head. Detecting defective CH is the essential task carried out during data transmission. This is basically done in the following two ways. Firstly, defective CH is detected, if acknowledgement is not received by the transmitted sensor node and all its cluster members will be aware of its defect. Secondly, sink node is also capable of detecting defective CH, if it does not receive data communicated by the CH.

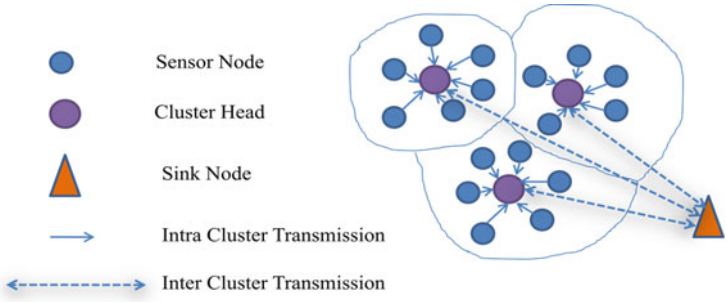


Fig. 1 Communication among nodes and CH

4 Proposed Defect-Tolerant Approach

Defect-tolerant method is proposed to overcome the routing path via defective CHs. This approach enhances data delivery rate by reducing data loss in the network.

4.1 Basic Concept

To achieve defect-tolerant system, the proposed scheme adopts backup CH formation and tree construction scheme. Whenever a defect in CH is detected, it executes the following steps. Sink will have all the details of CH formed. When one of the CH fails, sink node traces the entire fault free CH. Sink node finds fault-tolerant functionality by considering the retained energy of CHs. A tree network is formed among all fault free CHs.

4.2 Tree Formation and Backup CH (BCH) Formation

Tree construction task is driven by the sink node. Control packets consisting of fields such as node id, retained energy, parent node id and position are communicated from sink to all the CHs. All CHs respond to the control message by sending the required field information. During aggregation of data, CH aggregates and transfers information to the logical tree structure.

ALGORITHM 1: Tree Construction Among CHs

Input
Parent Node : Immediate Higher Level Node for CH
Energy : Retained energy of the CH nodes
Node Id : Cluster node identification number
Position : Distance from Sink node
1. Bradcasting tree_formation_msg from sink (CH ID, Parent, Energy, Position)
2. Tree_formation_msg for sink node (Sink ID, nil, Infinite Energy, Level_0)
3. TDMA Schedule considered for CH node
4. **if** (CH doesn't transmit data to the sink) then
5. Switch off the radio
6. **else**
7. Communication of aggregated information to higher level node in the sink direction
8. **end if**

Definition 1 Fault detection of existing CH node is determined by members of the respective clusters.

Each of the clusters consists of a CH whose responsibility lies in aggregation of information communicated from all its cluster members. CHs respond to the information received by sending an acknowledgement to member nodes. During fault occurrence at the CH, information fails to reach CH and acknowledgement will not be received by member nodes. So, member nodes will be aware of the fault and detection can be processed by member nodes.

Definition 2 Fault detection of CH node can be determined by sink node based on the failure of communication path from CH to the sink.

Sink node can determine CH failure by identifying frequent exchange of data packets from CH to sink. CH aggregates information of all its cluster nodes and transfers aggregated data to the sink. This constant exchange of information comes to halt during the failure in CH. Sink node has the capability to determine this failure.

Definition 3 BCH replaces functionality of faulty CH during the fault occurrence in CHs of a network.

After fault detection in the system, BCH is chosen based on maximum retained energy of other member nodes in the cluster. This BCH aggregates information of all its cluster members and takes same responsibility of the CH in a network.

Definition 4 Updated logical aggregator tree is formed using BCH during the CH failure.

The tree structure formed connecting CHs helps the network in achieving data transmission with reduced latency. After BCHs are formed, logical tree structure is updated using CHs and BCHs. Similarly, each BCH sends its status information of the field to sink node. Logical aggregation tree formed by backup CH nodes exists

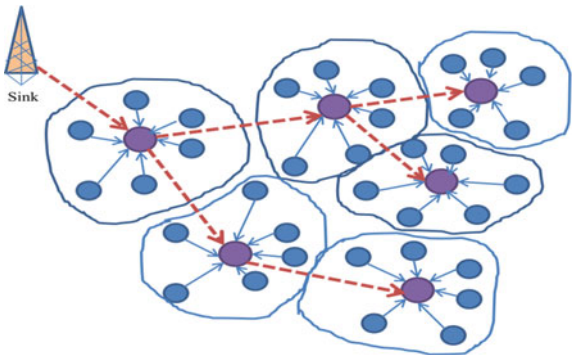


Fig. 2 Architecture of FTRAT algorithm

in the network with sink node as root node. The ratio of CH node goes to off status when it does not have any data for transmission.

Figure 2 shows the architecture of FTRAT algorithm. Cluster formation is done, and selected CH aggregates and routes the information to sink. During the fault occurrence in CH, BCH is selected which plays the prime role in establishing a communication path between sensor nodes and sink. Aggregator tree formation assists the cluster member nodes in communicating information to sink without enquiring sink position in the network.

5 Network Simulation

The simulation of FTRAT is performed on NS-2, and obtained results are compared with the QADA protocol. Various simulation parameters of FTRAT protocol are shown in Table 1.

Table 1 Various simulation parameters of FTRAT

Parameters	Range
Sensor node deployed area	300×300 m^2
Number of static nodes	300
Sink node	Single and fixed
Sink node position	180/90 m
Routing protocol considered	AODV
Packet size	500 bytes
Delay	30 μs
Antenna type	Antenna/Omni antenna

6 Results

Figure 3 shows improved throughput obtained by execution of fault-tolerant system in the network compared with QADA. Tree aggregation schemes enable faster delivery of information and improve throughput of proposed FTRAT method. Thus, improved throughput performance of FTRAT algorithm in kbps with respect to time is plotted and compared with QADA technique.

Figure 4 shows the PDR in FTRAT representing successful packet delivery of cluster member nodes to sink. Since FTRAT algorithm executes fault detection and fault recovery methods during routing of data packets, higher will be the packet delivery rate compared to QADA techniques. Aggregator tree communicates to sink using constructed path connecting CHs and BCH during defect occurrence in CH node. BCH executes the mechanism of a faulty CH node, replaces CH node and reinforces packet delivery in the network.

Average energy consumed for the information processing in FTRAT is less when compared with QADA techniques. Retransmission of data packets due to faulty CH is avoided by BCH formation. Since the tree constructed has a dedicated path to sink, each of the nodes in a cluster will not drain out their energy in finding the route to the sink node. So, FTRAT algorithm consumes lesser energy when compared with QADA techniques and is shown in Fig. 5.

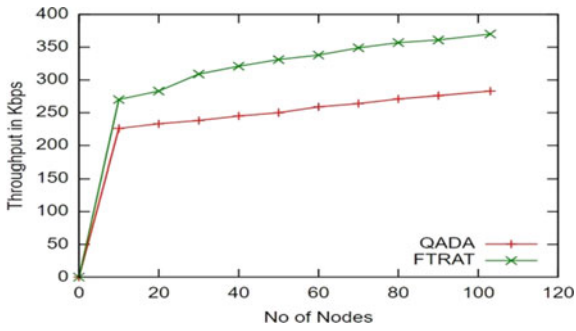


Fig. 3 Throughput

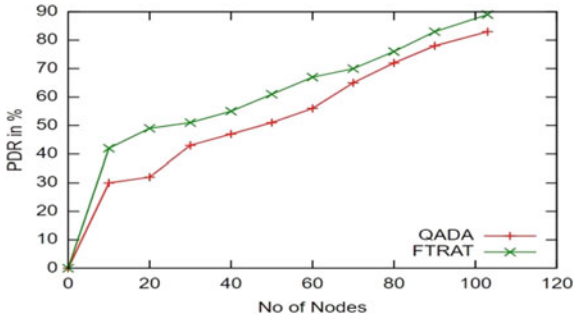


Fig. 4 Packet delivery ratio

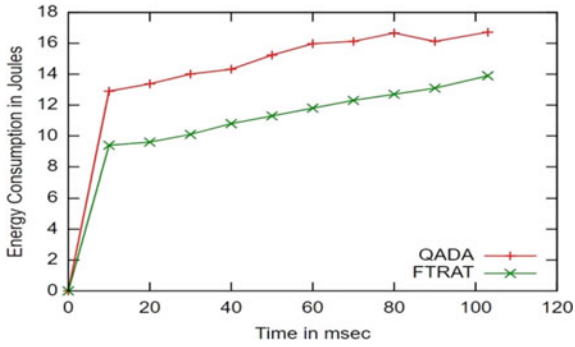


Fig. 5 Average energy consumption

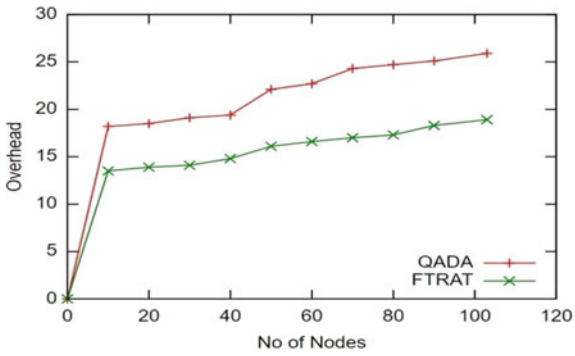


Fig. 6 Routing overhead

Figure 6 shows the graph of routing overhead with respect to number of nodes. Tree structure involved in routing adapts uncomplicated path establishment to the sink from cluster member nodes. This assists the network in obtaining lesser overhead when compared with QADA techniques.

7 Conclusions

In this paper, we have proposed FTRAT, fault-tolerant clustering techniques to communicate the sensed information by formation of CHs. Considered backup CH is successful in presenting the aggregated information on behalf of the defective CH. Tree construction among all the CHs selected helps to obtain lesser latency. Frequent path establishment among the nodes is avoided, and the proposed technique helps to achieve better QoS when compared with QADA in terms of PDR, throughput, average energy and delay. The proposed FTRAT techniques avoid failure of data packet transmission and show improvement of 14% in terms of packet delivery ratio from sender to receiver in the network.

References

1. Bahi, J.M., Makhoul, A., Medlej, M.: A two tiers data aggregation scheme for periodic sensor networks. J. Ad-Hoc Sens. Wirel. Netw. **21**(1–2), 77–100 (2014)
2. Geetha, V., Kallapur, P.V., Tellajeera, S.: Clustering in wireless sensor networks: performance comparison of LEACH and LEACHC protocols using NS2. J. Proc. Technol. **4**, 163–170 (2012)
3. Al-Karaki, J.N., Kamal, A.E.: Routing techniques in wireless sensor networks: a survey. J. IEEE Wirel. Commun. **11**(6), 6–28 (2004)
4. Xu, Z., Chen, L., Chen, C., Guan, X.: Joint clustering and routing design for reliable and efficient data collection in large-scale wireless sensor networks. J. IEEE Internet Things **3**(4), 520–532 (2016)
5. Rajeswari, K., Neduncheliyan, S.: Genetic algorithm based fault tolerant clustering in wireless sensor network. J. IET Commun. **11**(12), 1927–1932 (2017)
6. Kaur, S., Gangwar, R.: A study of tree based data aggregation techniques for WSNs. Int. J. Database Theory Appl. **30**(14–15), 2826–2841 (2016)
7. Zhou, Y., Wang, X., Wang, T., Liu, B., Sun, W.: Fault-tolerant multi-path routing protocol for WSN based on HEED. Int. J. Sens. Netw. **20**(1), 37–45 (2016)
8. Saini, S., Singh, R.S., Gupta, V.: Analysis of energy efficient routing protocols in wireless sensor networks. Int. J. Comput. Sci. Commun. **1**(1), 113–118 (2010)
9. Khanouche, M.E., Amirat, Y., Chibani, A., Kerkar, M., Yachir, A.: Energy-centered and QoS-aware services selection for internet of things. IEEE Trans. Autom. Sci. Eng. **13**(3), 1256–1269 (2016)
10. Morabito, R., Farris, I., Iera, A., Taleb, T.: Evaluating performance of containerized IoT services for clustered devices at the network edge. IEEE Internet Things J. **4**(4), 1019–1030 (2017)
11. Wu, F., Zhang, B., Fan, W., Tian, X., Huang, S., Yu, C., Liu, Y.: An enhanced random access algorithm based on the clustering-reuse preamble allocation in NB-IoT system. IEEE Access **7**, 183847–183859 (2019)
12. Gravalos, I., Makris, P., Christodoulopoulos, K., Varvarigos, E.A.: Efficient network planning for internet of things with QoS constraints. IEEE Internet Things J. **5**(5), 3823–3836 (2018)
13. Aranzazu-Suescun, C., Cardei, M.: Distributed algorithms for event reporting in mobile-sink WSNs for internet of things. J. Tsinghua Sci. Technol. **22**(4), 413–426 (2016)
14. Huang, W., Yin, M., Li, J., Xie, S.: Deep clustering via Weighted *k*-subspace network. J. IEEE Sig. Process. Lett. **26**(11), 1628–1632 (2019)
15. Lyu, L., Jin, J., Rajasegarar, S., He, X., Palaniswami, M.: Fog-empowered anomaly detection in IoT using hyperellipsoidal clustering. J. IEEE Internet Things **4**(5), 1174–1184 (2017)
16. Puschmann, D., Barnaghi, P., Tafazolli, R.: Adaptive clustering for dynamic IoT data streams. J. IEEE Internet Things **4**(1), 64–74 (2016)

SIDA—Secure and Intelligence Data Aggregation in Wireless Sensor Networks

S. Reshma, K. Shaila, and K. R. Venugopal

Abstract Wireless sensor networks are authorized to collect huge data and transmit them in an environment monitoring specific application. Therefore, mitigating redundant data and protecting sensitive information are challenging tasks. The existing algorithms in wireless sensor networks utilized cosine similarity scheme to train nodes to form cluster and redundant data elimination and aggregates data. The proposed protocol, secure and intelligence data aggregation, focuses on providing training to sensor nodes, to dynamically form the clusters as per the network requirements. The protocol concentrates on eliminating outliers and redundant data but protects the sensitive data. This protocol mitigates false data injection in the network. Thus, secure and intelligence data aggregation protocol is simulated with NS2 network simulator, and results show achievement in maximizing network lifetime, maximizing throughput, data accuracy and cluster accuracy.

Keywords Data collection · Environmental monitoring · Network lifetime · Throughput · Transmission · Wireless sensor networks

1 Introduction

A large number of tiny sensor nodes are capable of monitoring environmental condition and forward to their respective station for future environmental prediction. In addition to environmental monitoring, there are numerous numbers of applications in WSNs, viz. military surveillance, forest disaster, etc. The sensor nodes are battery

S. Reshma (✉)
Department of Computer Science and Engineering, Visvesvaraya Technological University—Research Resource Center, Belagavi, Karnataka, India

K. Shaila
Department of Electronics and Communication Engineering, Vivekananda Institute of Technology, Bengaluru, Karnataka, India

K. R. Venugopal
Bengaluru University, Bengaluru, Karnataka, India

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_74

powered with limited resource parameter. It is deployed to monitor the environmental condition where human intervention is not possible. So, it is a challenging task to train sensor nodes. Therefore, in proposed protocol, we train the sensor nodes using unsupervised machine learning techniques to provide efficient and secured data transmission in the network.

There are different secure data aggregation techniques to provide efficient and secure data transmission [1, 2] which consolidates data by elimination of redundant data in the network. Several data aggregation techniques utilized clustering strategy [3] to collect all the sensed data within the cluster and forwards to cluster head (CH). The CH in turn eliminates redundant data and forwards to BS. Therefore, the SIDA protocol uses machine learning technique to guide intelligence for sensors to form the cluster with an intelligence. Thus, the uniform node distribution helps to achieve energy efficiency.

2 Related Work

The existing protocols for secure data aggregation are discussed in this section. Oliveira et al. documented a review of the environmental monitoring applications in wireless sensor networks. Yick et al. [5] reviewed literature on various applications in wireless sensor networks.

Yin et al. [6] proposed data compression model for data aggregation. This model used similarity metric for data aggregation which is strong temporal correlation. This strategy compressed sensor nodes into one cluster and forwards to cluster head (CH). The CH collects the data sensed by sensor nodes and aggregates. To conserve less energy for data aggregation, the author adopted an effective adoptive strategy. Thus, protocol achieves maximum energy efficiency. But this protocol is not suitable for mulithop network. Lin et al., [7] discussed on distributive algorithm for maximizing data collection rate which maximizes optimization problem for WSNs. Therefore, in addition to distributive algorithm, Lagrange multiplier is used to balance an energy effectively in the network.

Suryakant et al. presented a similarity algorithm based on mean measure of divergence. This algorithm addressed a sparsity issue and obtained solution using behavior habit of data between the users. The mean measure of divergence achieves better prediction accuracy compared with cosine similarity and Jaccard similarity algorithms.

The existing approaches for clustering and data aggregation are based on CH selection, cluster formation and similarity of node's behavior and data. In our work, we used neural networks and machine learning strategy for cluster formation and redundant data elimination.

3 Background

SOMDA protocol [9] involves cosine similarity function to find the similarity of neighbor nodes and forms the cluster. This protocol applies weight function to obtain variation of the winner neuron to compute similarity of data during data aggregation. It utilizes interquartile concept to remove an outlier. Thus, SOMDA achieves better performance in forms of network and energy efficiency. In our work, we focused on interquartile and other machine learning techniques to achieve better network lifetime, throughput, data accuracy and cluster accuracy.

3.1 Problem Definition

Wireless sensor networks (WSNs) are composed of n number of tiny sensor nodes and are grouped together to form k clusters. The cluster formation is a challenging task to prolong network connectivity and guaranteed packet delivery. Therefore, it is required to train intelligent nodes to form cluster. In addition to nodes connectivity, it is also necessary to reject false data injected by malicious node and redundant data.

The main objectives of SIDA protocol are as follows:

(i) Prolong node connectivity in the network
(ii) Mitigate data redundancy and
(iii) Mitigate false data injection by malicious node.

4 Implementation and Mathematical Model

The SIDA protocol applies neuro-intelligence paradigm to provide trusted data reception and energy-efficient data transmission to the base station. As per the network requirement, sensor nodes are trained by SIDA protocol to behave spontaneously. Figure 1 shows the proposed protocol training modes which involves three phases.

(i) Neural cluster formation (NCF)
(ii) Intelligence and secured data aggregation (ISDA)
(iii) Data transmission.

4.1 Neural Cluster Formation (NCF) Phase

Consider n number of sensors to monitor a temperature in the network. The randomly elected node is considered as CH. The clusters are formed based on the specified

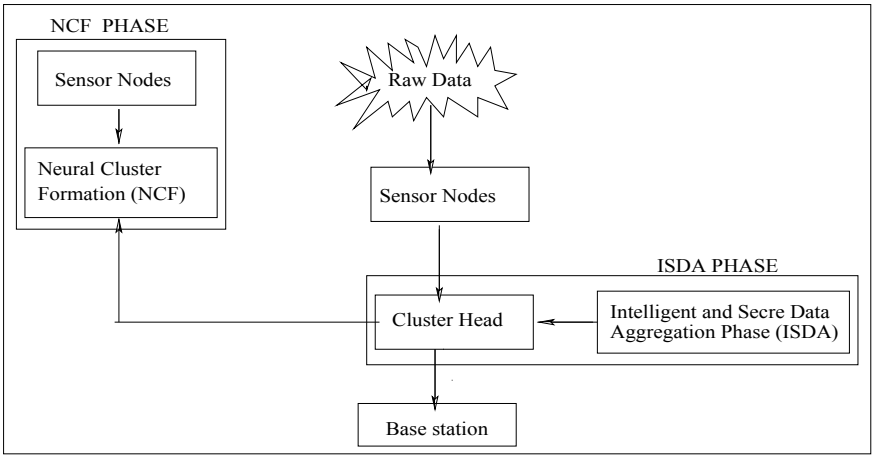


Fig. 1 SIDA model

radius range r from cluster head. The network architecture shown in Fig. 2 illustrates that a small number of nodes are not connected to any cluster group.

The proposed protocol trains these nodes based on distance, node density and rank. The steps involved are as follows:

Step 1: Distance of each sensor to nearest CHs is computed using

$$D = \sum_{i=1}^{m} \sqrt[2]{(x_i - x_{CH})^2 + (y_i - y_{CH})^2} \tag{1}$$

Step 2: Track the node density of nearest clusters.
Step 3: Rank the distance by considering minimum distance as high priority.

$$First\_min\_D = R_{D_1}$$
$$Sec\_min\_D = R_{D_2}$$

Step 4: Rank the node density by considering maximum node density as high priority.

$$First\_max_N D = R_{ND_1}$$
$$Sec\_max\_ND = R_{ND_2}$$

Step 5: Select the clusters that are close to the specified range, which has $min\ R_D$ and $max\ R_{ND}$ among closest cluster.

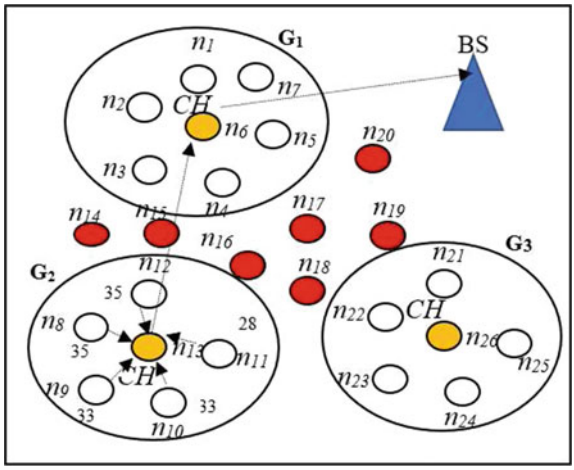


Fig. 2 Example of SIDA network architecture

4.2 Intelligence and Secured Data Aggregation (ISDA) Phase

Sensor node senses the environmental temperature and forwards to CH. Simultaneously, all nodes in the network establish session key for every 2 min. This path key establishes forward secrecy of the data. The CH maintains past session key and also presents session key. This avoids intruder to send packet to CH. CH focuses on eliminating redundant data to save battery life of node and protecting sensitive data. Therefore, CH uses interquartile function to remove outliers of environmental temperature and similarity function to remove redundancy.

The steps followed by the CH during data aggregation are as follow:

Step 1: Initially, it computes variation as shown,

$$Y = X_i - W_C \tag{2}$$

Step 2: Computes interquartile using

$$IQ = Max - Min \tag{3}$$

Step 3: Obtains upper temperature range

$$UpperT_r = Max + 1.5 * IQ \tag{4}$$

Step 4: Computes lower temperature range

$$LowerT_r = Max - 1.5 * IQ \tag{5}$$

Step 5: After obtaining upper and lower temperature range, the outliers range values are removed.
Step 6: The intelligence similarity function is used to minimize redundant data as shown

$$IS = \frac{\sum_{x=1}^{d} D_{x-1} D_{x-2}}{\sum_{x=1}^{d} D_{x-1} + \sum_{x=1}^{d} D_{x-2} - \sum_{x=1}^{d} D_{x-1} D_{x-2}} \tag{6}$$

4.3 Data Transmission Phase

The aggregated data from previous phase is then forwarded to the base station in this phase. The base station uses aggregated data as per the application requirement.

5 Simulation and Performance Evaluation

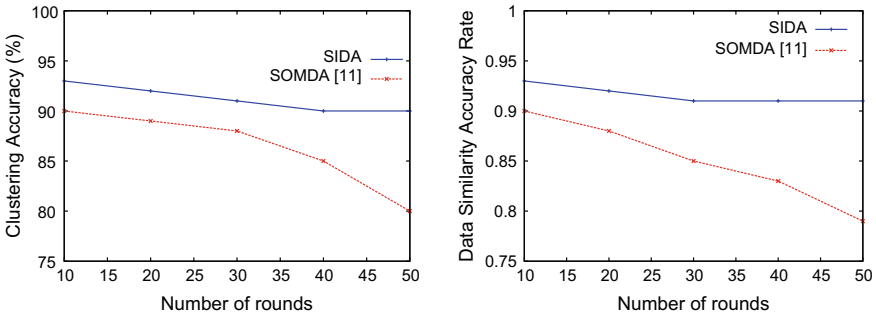
The proposed protocol is simulated by a computer simulator. It is accomplished on NS2 simulator which is developed to simulate WSNs environment. The protocol is tested over 1000 m * 1000 m area with the comparison of SOMDA protocol. The SIDA protocol execution environment is set up for different set of nodes, viz. 50, 10, 200, 250 and 300, respectively, and sink is situated at 500 m * 500 m.

5.1 Performance Metrics

1. ***Custer accuracy***: It is the accurate percentage of cluster selection.
2. ***Data accuracy***: It is the rate of data redundant elimination.
3. ***Network Lifetime***: It is the percentage of number of nodes alive for various rounds.
4. ***Throughput***: It is defined as the rate of number of packets transmitted over a period of time.

5.2 Performance Evaluation

Figure 3a shows the clustering accuracy of SIDA and SOMDA protocol. The training involved in self-learning training for the sensor nodes accurately accomplished such that it belongs to a closest cluster efficiently with respect to node density (ND) and distance (D). In SOMDA, the clustering is done based on cosine similarity function which incorporates SOM (self-organized map). This approach is not suitable for



(a) Clustering Accuracy (b) Data Accuracy

Fig. 3 Clustering accuracy and data accuracy of SIDA and SOMDA protocol

Table 1 Clustering accuracy, data accuracy and network lifetime

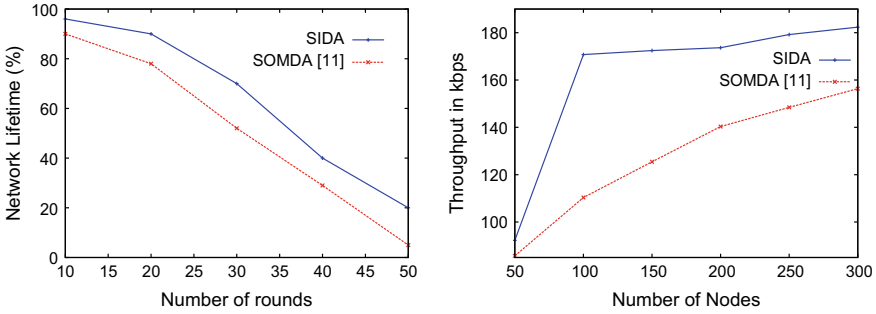
Number of rounds	Clustering accuracy in percentage		Data similarity accuracy		Network lifetime in percentage	
	SOMDA	SIDA	SOMDA	SIDA	SOMDA	SIDA
10	90	93	0.90	0.93	90	96
20	89	91	0.88	0.92	72	90
30	88	90	0.85	0.91	58	70
40	85	90	0.83	0.91	29	40
50	80	90	0.79	0.91	5	20

clustering among large set of clusters. Therefore, SIDA protocol uses ranking strategy to predict exact accuracy of the nodes and their respective cluster. The SIDA protocol exhibits 10% improvement over the SOMDA protocol.

Data accuracy of SIDA protocol is shown in Fig. 3b. The intelligence similarity (IS) equation provides an accurate result compared with cosine similarity. The cosine similarity failed to achieve exact matching when we have a large difference among data. Hence, SIDA protocol shows 12% better efficiency than SOMDA protocol.

The network lifetime of SIDA protocol is shown in Fig. 4a. Table 1 shows the comparison result analysis of clustering accuracy, data accuracy and network lifetime of SIDA and SOMDA protocols. SIDA protocol balances nodes energy that are formed based on node density. The clusters are balanced with uniform node distribution which exhibits longer network connectivity. Thus, SIDA prolongs network connectivity 15% over SOMDA protocol.

The SIDA protocol guarantees packet delivery using key establishment strategy. For interval, it provides promised data delivery to the base station (BS). The SIDA protocol throughput rate is shown in Fig. 4b. When compared with SOMDA, SIDA increases its throughput rate by 26% with SOMDA. Because SOMDA protocol is not focused on security aspect. Table 2 depicts the comparison throughput of SIDA and SOMDA protocol.



(a) Network Lifetime (b) Throughput

Fig. 4 Network lifetime and throughput of SIDA and SOMDA protocol

Table 2 Throughput

Number of nodes	Throughput in kbps	
	SOMDA	SIDA
50	85.678	92.234
100	110.354	170.756
150	125.456	172.453
200	140.352	173.654
250	148.456	179.222
300	156.352	182.352

6 Conclusions

The secure and intelligence data aggregation (SIDA) protocol achieves balanced energy-efficient and secured data aggregation for wireless sensor networks. The SIDA protocol utilized distance, node density and rank between the nodes and closest cluster in the network. This protocol computes interquartile to remove outlier's environment temperature range. It helps to eliminate redundant data with the utilization of intelligence similarity algorithm. The SIDA protocol also uses key establishment strategy to mitigate intruder's false data injection, whereas sensor nodes in the network are trained using an unsupervised machine learning technique. The SIDA protocol effectively achieves improvement of 15% in network lifetime, 26% in throughput, 12% in data accuracy and 10% in cluster accuracy.

References

1. Abid, B., Nguyen, T.T., Seba, H.: New data aggregation approach for time constrained wireless sensor networks. Int. J. Supercomput. **71**(5), 1678–1693 (2015)
2. Huang, C.-F., Lin, W.-C.: Data collection for multiple mobile users in wireless sensor networks. Int. J. Supercomput. **72**(7), 2651–2669 (2015)
3. Jadhav, N.H., Kashid, D.N., Kulkarni, S.R.: Subset selection in multiple linear regression in the presence of outlier and multicollinearity. Stat. Methodol. **19**, 44–59 (2014)
4. Oliveira, L.M., Rodrigues, J.J.: Wireless sensor networks: a survey on environmental monitoring. J. Commun. **6**(2), 143–151 (2011)
5. Yick, J., Mukherjee, B., Ghosal, D.: Wireless sensor network survey. Comput. Netw. **52**(12), 2292–2330 (2008)
6. Yin, Y., Liu, F., Zhou, X., Li, Q.: An efficient data compression model based on spatial clustering and principal component analysis in wireless sensor networks. IEEE Sens. **15**(8), 19443–19465 (2015)
7. Lin, H., Bai, D., Gao, D., Liu, Y.: Maximum data collection rate routing protocol based on topology control for rechargeable wireless sensor networks. IEEE Sens. **16**(8), 1201 (2016)
8. Suryakant, Mahara, T.: A new similarity measure based on mean measure of divergence for collaborative filtering in sparse environment. In: Twelth International Multi-Conference on Information Processing, vol. 89, pp. 450–456 (2016)
9. Ullah, I., Youn, H.Y.: A novel data aggregation scheme based on self-organized map for WSN. Int. J. Supercomput. **75**, 3975–3996 (2019)

Human Anomaly Detection in Surveillance Videos: A Review

K. Chidananda and A. P. Siva Kumar

Abstract This paper presents an exhaustive survey on different approaches for human anomaly detection which has most significance in the field of computer vision due to its many applications in real life. Human anomaly detection is a pressing problem in surveillance video system. We have also presented various feature extraction techniques, classification, clustering, deep learning, and transfer learning models, frameworks used in human anomaly detection area, various kinds of standard datasets which can be used for benchmarking the algorithms, challenges, and applications.

Keywords Human anomaly detection · Computer vision · Deep learning · Transfer learning · Surveillance videos

1 Introduction

Human anomaly detection in videos is an active field in computer vision which is attracting more research attention in recent years. Anomaly detection related to anomalous human activity recognition is a crucial problem in surveillance system, security-based system like CCTV in environments like public crowded places. Efficient algorithms are required for accurate human activity recognition in real-time environments like airports, other transportation systems hospitals, banks, and shopping malls where tight security and surveillance are required.

A smart video surveillance environment is required to alert the people if they identify a potential threat caused intentionally in these surrounding environments. Most crowded places like colleges or universities, bus stations, shopping malls, which require continuous monitoring of surveillance can ensure safety of the people and infrastructure. Hospitals are another environment which requires suspicious human

K. Chidananda (✉) · A. P. Siva Kumar
Department of CSE, JNTUA, Anantapur, Andhra Pradesh 515002, India

A. P. Siva Kumar
e-mail: sivakumar.cse@jntua.ac.in

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_75

anomaly detection to ensure the safety of the patients, and a smart surveillance system is required to detect road accidents. Home surveillance system is also another important environment which avoids theft and burglary.

Video surveillance along with computer vision techniques makes sure security and safety of public in most crowded environments. The following are the different stages in computer vision, gathering information from multiple cameras, preprocessing for feature extraction among video sequences, modeling of environments, detection of motion among objects, tracking, and behavior understanding of objects. The detection of human anomaly is a crucial research problem in video surveillance systems.

Figure 1 shows the general processing and modeling of environments for detection of human anomalies in video sequences, and Fig. 2 shows the steps involved in the classification of human anomaly detection process.

For the detection of human anomaly in videos, most of the models used supervise, semi-supervised, and unsupervised algorithms for classification of either normal or abnormal activity among people. In this paper, Sect. 2 discusses related work, Sect. 3 discusses various techniques used, standard datasets, and challenges.

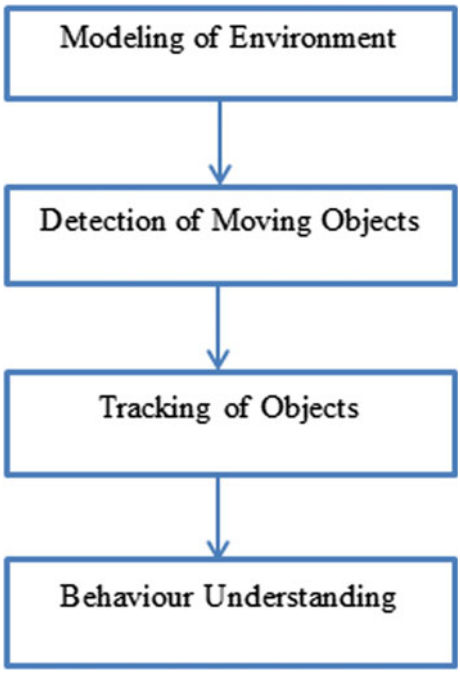


Fig. 1 General steps of human anomaly detection

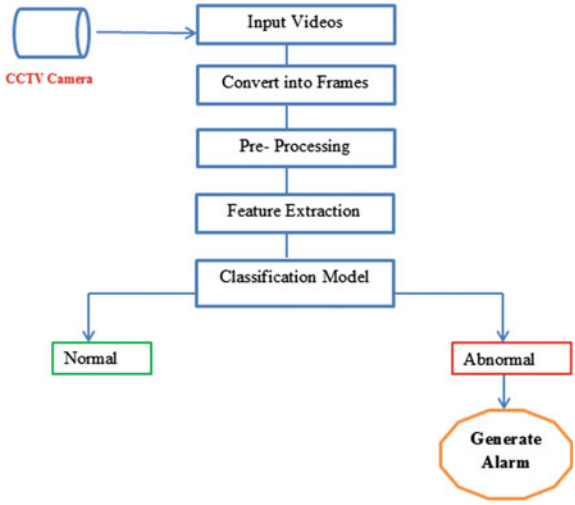


Fig. 2 Human anomaly detection classification

2 Related Work

This section discusses recent related surveys, various techniques used, datasets, and limitations. This paper extended the survey done by the authors of [1] (Table 1).

Ahmad et al. [9] proposed a novel anomaly detection, which is based on an online sequence memory algorithm, called hierarchical temporal memory on Numenta Anomaly Benchmark (NAB) dataset. Ramchandran et al. [10] proposed a hybrid deep learning framework for video anomaly detection using convolutional autoencoder using LSTM model and experimented on three different benchmark datasets such as Avenue, UCSD ped1, and UCSD ped2. Gnouma et al. [11] proposed a motion extraction model based on deep learning which can detect anomaly activities using reasons of interest and evaluated on UCSD and Avenue datasets. Jothi Shri and Jothilakshmi [12] proposed a deep learning technique for anomaly detection in web cameras with an alarm alert with less cost.

Dey et al. [13] proposed an LSTM-based model which can investigate behavior patterns and help to identify anomaly detection in crowded scenes. Kavikuil et al. [14] proposed an optimized model for predicting anomaly video surveillance data by tuning the hyperparameters. Pawar and Attar [15] proposed a deep learning technique for video-based anomalous activity detection using graphical taxonomy. Pradeepa et al. [16]. This proposed an LDA model by using streak flow technique. Qasim et al. [17] proposed an anomaly detection model in video sequences using three different extracted features and implemented on UMN dataset.

Sabokrou et al. [18] proposed an unsupervised model using pre-training and fully convolutional neural networks for detection of anomalies in scenes with the help of cascaded outlier detection. Singh et al. [19] proposed an ensembler model for detecting an anomaly in video data which has crowded scenes by using a pre-trained ConvNets and a pool of SVM classifiers and experimented on benchmark datasets.

Table 1 Recent related surveys

Main focus/topic	References
Online anomaly detection	Doshi et al. [2]
Parallel spatial–temporal convolution neural networks	Hua et al. [3]
Anomaly detection in crowd videos at a global level	Gong et al. [4]
LDA-net-to detect and locate human anomalies in a supervised manner	Qasim et al. [5]
Detect anomaly events in crowded scenes	Huang et al. [6]
Snatch theft detection in unconstrained surveillance videos	Roy et al. [7]
Automatically detects multiple anomalous activities in videos	Chaudharya et al. [8]
Three phases of human activity recognition	Bux et al. (2017)
Human action detection and labeling	Dhulekar et al. (2017)
Anomaly detection based on smart home sensors	Bakar et al. (2016)
Gesture recognition systems	Siddharth and Anupam (2015)
Actions recognition based on spatio-temporal points	Dawn et al. (2015)
Human motion detection	Mishra and Bhagat (2015)
Crowded scene analysis	Teng et al. (2015)
Actions recognition in crowded surveillance scene	de Campos (2014)
Human action recognition	Hassan et al. (2014)
Intelligent video surveillance systems in public spaces	Zablocki et al. (2014)
Human activity recognition and understanding	Sarvesh and Anupam (2013)
Human activity recognition	Shian-Ru et al. (2013)
Abnormal human behavior recognition	Oluwatoyin and Kejun (2012)
Behavior analysis for homeland security applications	Teddy (2008)
Automated video surveillance systems	Valera and Velastin (2005)
Human motion analysis	Aggarwal and Cai (1999, 1997)

Zhang et al. [20] proposed a *K*-NN model identifying homogeneous intra-class similarities and experimented on real-world surveillance videos. Li et al. [21] proposed a spatial–temporal cascade autoencoder model which can detect both spatial and temporal cues from video data.

Hua et al. [3] proposed a localized model for detecting the abnormal behavior in video surveillance using spatial–temporal convolution neural networks.

3 Techniques Used

1. Supervised: Supervised learning algorithms use training data for anomaly detection in video images. These algorithms required labeled dataset for classifications of

video sequences into anomalous or normal event. Based on the guidance of labeled dataset, algorithms learn with the help of training dataset and predict the class labels for new video dataset.

These supervised methods for anomaly detection can be categorized as below:

(1) Video dataset with normal events as training data
(2) Video dataset with anomalous events as training data
(3) Video dataset with both normal and anomalous events as labeled data
(4) Video dataset with multiple classes as labeled data.

Methods 1 and 2 are formulated as single-class classification problem, method 3 is a two-class classification problem, and method 4 is multiclass problem which has specific labeled datasets.

2. Semi-supervised Learning: Semi-supervised learning use unlabeled dataset as training data and almost works like unsupervised technique. As the task of labeling the dataset is time consuming and tedious task, it will be better to have algorithms which process and detects anomalous events in video sequences from unlabeled dataset directly with the help of small portion of labeled dataset.

When a fully supervised method of classification needs a lot of labeled data for learning, there must be a skilled agent to label which is normal and abnormal activity. The cost associated with such a process is expensive, compared with the acquisition of unlabeled data in case of unsupervised learning. Most of the time, it suffers from either overfitting or under fitting. But unsupervised learning effectively works on entirely new patterns of a dataset, but results may be unpredictable. Hence, the authors of [22–24] attempted to combine the advantages of supervised and unsupervised learning in a probabilistic setting.

The author [22] proposed a system for unusual event detection. The methodology was persuaded by the perception that it is ridiculous to obtain large training dataset for unusual events and is unpredictable, and it is conversely possible to do so for usual events, allowing the creation of a well estimated model of usual events?

This scarcity of preparing datasets for irregular functions made the authors of [23] propose the utilization of Bayesian adaptation techniques, which adjust a typical function model to deliver some surprising function models in an unsupervised way. From the point of anomaly detection, not much work has been done using a semi-supervised learning approach.

3. Unsupervised: Unsupervised learning algorithms do not require training data for anomaly detection in video images. These methods process the video images which have both normal and anomalous events without labeled datasets. These techniques work based on clustering techniques in which the anomalies are detected by using the distance of new dataset from the nearest distance or cluster and also detected by using outlier detection.

Several studies [25, 26] detect anomalies with the help of distance between unseen data points which can fit into a nearest cluster. The most common method to detect used is Bhattacharya distance, as this is a measure of divergence or dissimilarity between two probability distributions of data points [27–29]. Unsupervised performs

better than supervised as there more chances of mislabeling, as manual labeling is used in supervised techniques.

Goshorn et al. [30] identifies anomalous points by fitting into training cluster with the help of a change which may cause unseen data point which involves cost, and if all out expense a threshold, the point is considered as anomalous. Other studies [31–36] use various clustering techniques with the help of probabilistic methods log likelihood [32] and log likelihood ratio test [27] for capturing of unseen points. If the probabilistic value falls below threshold value, the event is considered as anomalous event.

3.1. Deep Learning: Deep learning techniques detect anomalies in video images with help of deep layers. [10] Use deep learning framework for video anomaly detection with the help of convolutional autoencoder.

Fan et al. [37] used deep learning for detecting anomalies and Sabokrou et al. [38] used sparse autoencoder which are cascaded. Ravanbakhsh et al. [39] used a pre-trained model with fully CNN to capture temporal patters of local anomalies in video sequences. Sabokrou et al. [40] also used fully CNNs from video regions to extract most discriminative features by using Gaussian distribution model for detecting anomalies. Hasan et al. [41] used fully autoencoder with CNN to extract handpicked features from short video clip by computing regularity score to learn temporal regularity.

3.2. Transfer Learning: Transfer learning techniques use already pre-trained models without starting the training from scratch.

Transfer learning has been utilized in many studies for image classification by using ImageNet dataset Fei-Fei et al. [42]. For classifying real-world images over 14 million labeled datasets. Few models also used to classify medical image data [43, 44]. Transfer learning is also commonly used for the task of semantic segmentation. The main reason for this is the expense of generating annotated data for real urban scenes which typically requires a significant amount of human resources. In order to reduce the cost associated with this task, synthetic annotated data is often used as a training set instead of real-world data. As the usage of synthetic data reduces the time and cost in training the model, as the supervised models requires annotated dataset for detecting anomalies in the dataset which involves manual intervention. In other words, CNN models are trained on a source domain which consists of synthetic annotated images (often from video games or simulations) and then evaluated on the target domain which consists of real-world urban images. However, this is a difficult transfer learning task. Wang et al. [45] suggested using adversarial learning in order to help classifiers learn domain invariant features. In their experiments, they showed that adversarial classifiers outperformed regular CNNs on various datasets—AlexNet, GoogleNet, and VGG Net, etc. (Table 2; Graph 1)

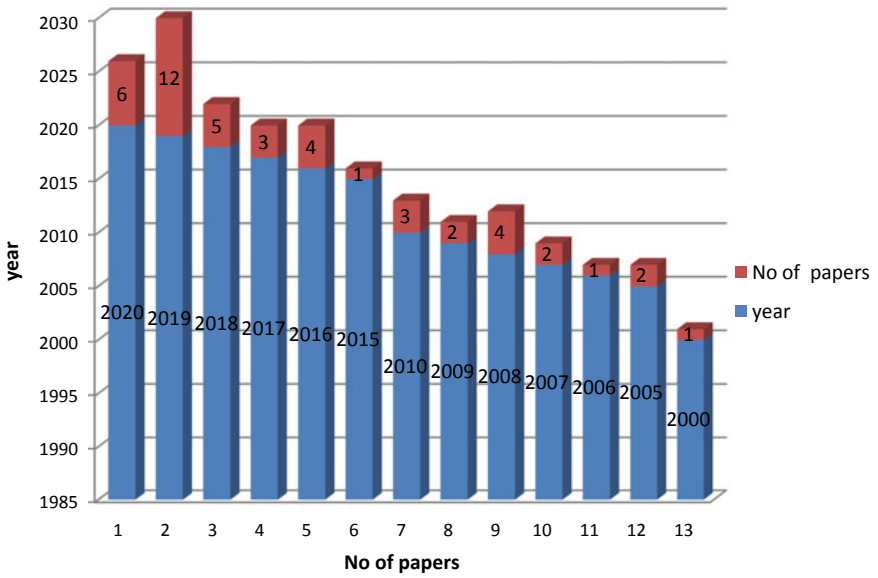
Table 2 Comparison of classification methods categories

Method	Description	Advantages	Limitations
Supervised	Building a model of normal and abnormal behaviors via labeled data	– Good results in detecting known abnormal behaviors – Easy to run and to understand	– Designed to detect specific behaviors, e.g., fighting, loitering and falling – Strongly depends on the training data – Unknown anomalies cannot be detected – Learning all abnormal behaviors is not practical in real world
Semi supervised: rule-based method	– Rule construction using normal patterns – Any new sample that does not fit the rule is an Anomaly	– Easy to perform and to interpret – Very expressive – Close to human reasoning	High memory and computational complexities
Semi supervised: model-based method	– Build a model representing the normal behaviors – Any new sample that does not respect the model is an anomaly	– Model is easy to generate and to understand – A new instance is rapidly classified	– Sensitive to multiple parameters – Unknown normal data can be identified as abnormal (false alarm)
Unsupervised	Learning using statistical proprieties extracted from unlabeled data	– Fast and easy to perform – No prior knowledge required	– Time consuming for result interpretation – Based on the assumption that abnormal behaviors are very rare compared to normal ones

4 Standard Dataset Used

UCSD: Ped1 and Ped2 are two subset datasets in which video sequences are captured from fixed cameras Ped1 consists of unusual videos with dimension of 238 × 158 pixels with 200 frames with 34 normal and 36 unusual videos. Ped2 consists of unusual videos with dimension of 328 × 240 pixels with 200 frames with 16 normal and 14 unusual videos. In case of Ped1 and Ped2 datasets, normal events are the people walking on pedestrians, and abnormal events are the bikers, small cars walking on the pedestrians [37] (Figs. 1 and 2).

Avenue: The selected dataset is an Avenue dataset, and it has 16 training recordings and 20 test recordings. The simulated anomalies are delays and bizarre activities, for



Graph 1 This graph denotes the result of papers searched with the keywords "Human Anomaly Detection" year wise

example, tossing paper or a sack. Casings are extricated independently and remarked as ordinary and strange. The model is prepared on the named informational collection. The casings have a goal of 640 × 300. As a major aspect of the video preprocessing, the frames are extracted and commented. The frames are 128 × 128 in measure and have been overhauled to fit the design. All this preprocessing is done with the help of OpenCV [14]. Tables representing anomalous events are given in Fig. 3.

UMN: The anomaly in the UMN dataset has a place with a global anomaly, for example, frenzy and group flight. The complete season of the UMN informational index is 257 s, and the frame rate is 30 frames for every second. The size of each frame is 320,240. The UMN dataset with 11 recordings is recorded for three distinctive turf,

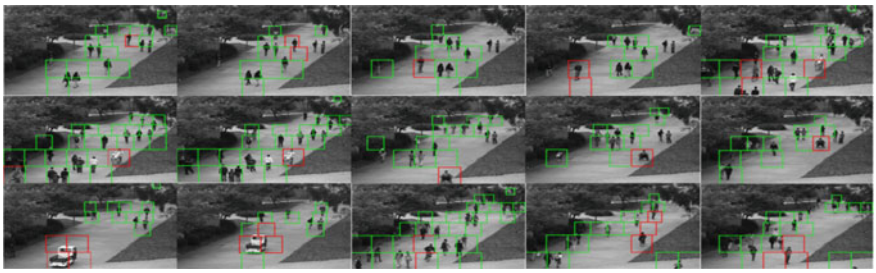


Fig. 3 UCSD Ped1

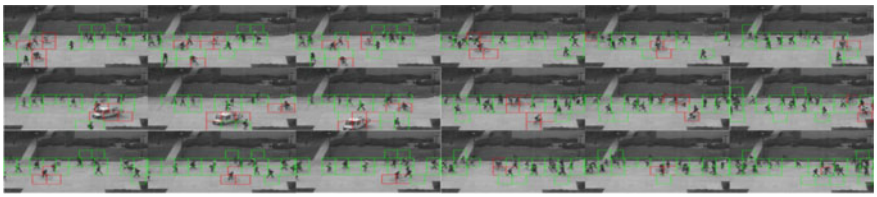


Fig. 4 UCSD Ped2

Fig. 5 Avenue dataset anomalous frames. **a** Strange action, **b** strange action, **c** loitering

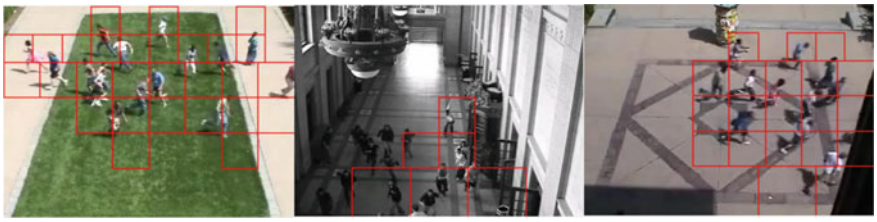


Fig. 6 UMN dataset

indoor, and site scenarios. In each clasp, the individuals in the video grouping scene first walk or wander and afterward begin running toward every path from a specific frame [3]. The example of an experiment to detect anomalies in a UMN dataset is shown in Figs. 4, 5 and 6.

5 Challenges in Human Anomaly Detection

1. Human being performs multiple tasks, and it is difficult to capture the sudden transition among the tasks.
2. The accuracy of anomaly detection fades because of various motion patterns among humans.

3. Availability of training and testing datasets for anomaly detections models are minimal.
4. In low-dimensional video frames, camera jitter leads to degradation of the quality of the video.
5. Climatic conditions like wind and rain may cause dynamic variations in contrast and brightness of the videos.

6 Observation and Gaps

After reviewing literatures, below are few observations:

1. Much work has not been done on detection of multi-crowd scenes in video surveillance systems.
2. Also much work has not been done on night environments which have less clarity video frames; there is scope to develop algorithms for detection of anomalies.

7 Conclusion

In this study, different kinds of machine learning techniques are addressed in human anomaly detection, also reviewed their applicability, implementation details, and limitations, discussed trends, identified gaps, and compared standard datasets, characteristics, and the scope of usage. As most of the supervised learning techniques, use for human anomaly detection process in surveillance videos requires training data which is a time consume step and involves more cost. To overcome these limitations of supervised techniques, deep learning techniques can be used in an unsupervised manner for high-level feature extraction and anomaly detection. Our study also explored that transfer learning techniques can be used for anomaly detection in surveillance videos by using pre-training models which avoids training step prior the detection process. The main goal is to set a few future directions by exploring into that gap in the current computer vision area.

References

1. Mabrouk, A.B., Zagrouba, E.: Abnormal behavior recognition for intelligent video surveillance systems: a review. © 2017 Elsevier Ltd
2. Doshi, K., Yilmaz, Y.: Continual learning for anomaly detection in surveillance videos. IEEE Explore (2020)
3. Hua, Z., Zhang, L., Li, S., Sun, D.: Parallel spatial-temporal convolutional neural networks for anomaly detection and location in crowded scenes, 1047-3203/© 2020 Elsevier Inc
4. Gong, M., Zeng, H., Xie, Y., Li, H., Tang, Z.: Local distinguishability aggrandizing network for human anomaly detection. © 2019 Elsevier Ltd

5. Qasim, T., Bhatti, N.: A low dimensional descriptor for detection of anomalies in crowd videos. © 2019 International Association for Mathematics and Computers in Simulation (IMACS). Elsevier
6. Huang, S., Huang, D., Zhou, X.: Learning multimodal deep representations for crowd anomaly event detection. Hindawi Math. Probl. Eng., 13 (2018). Article ID 6323942. https://doi.org/10.1155/2018/6323942
7. Roy, D., Krishna Mohana, C.: Snatch theft detection in unconstrained surveillance videos using action attribute modelling. Pattern Recogn. Lett. (2018)
8. Chaudharya, S., Khana, M.A., Bhatnagara, C.: Multiple anomalous activity detection in videos. In: 6th International Conference on Smart Computing and Communications, ICSCC 2017. Procedia Computer Science, vol. 125, pp. 336–345. Elsevier (2018)
9. Ahmad, S., Lavin, A., Purdy, S., Agha, Z.: Unsupervised real-time anomaly detection for streaming data, 0925-2312/© 2017 Elsevier
10. Ramchandran, A., Sangaiah, A.K.: Unsupervised deep learning system for local anomaly event detection in crowded scenes. Multimed. Tools Appl. (2019)
11. Gnouma, M., Ejabali, R., Zaied, M.: Video anomaly detection and localization in crowded scenes. Springer Nature Switzerland (2020)
12. Jothi Shri, S., Jothilakshmi, S.: Anomaly detection in video events using deep learning. IJITEE (2019)
13. Dey, A., Mohammad, F., Ahmed, S., Sharif, R., Saifuddin Saif, A.F.M.: Anomaly detection in crowded scene by pedestrians behaviour extraction using long short term method: a comprehensive study. I. J. Education and Management Engineering (2019)
14. Kavikuil, K., Amudha, J.: Leveraging deep learning for anomaly detection in video surveillance. Springer Nature Singapore Pte Ltd. (2019)
15. Pawar, K., Attar, V.: Deep learning approaches for video-based anomalous activity detection. Springer Science + Business Media, LLC, part of Springer Nature (2018)
16. Pradeepa, B., Viji, A., Joshan Athanesious, J., Vaidehi, V.: Anomaly detection in crowd scenes using streak flow analysis. 978-1-5386-9279-0/19/$31.00 © 2019 IEEE
17. Qasim, T., Bhatti, N.: A low dimensional descriptor for detection of anomalies in crowd videos, 0378-4754. © 2019 International Association for Mathematics and Computers in Simulation (IMACS)
18. Sabokrou, M., Fayyaz, M., Fathya, M., Moayedc, Z., Klettec, R.: Deep-anomaly: Fully convolutional neural network for fast anomaly detection in crowded scenes, 1077-3142/© 2018 Elsevier Inc
19. Singh, K., Rajora, S., Vishwakarma, D.K., Tripathi, G., Kumar, S., Walia, G.S.: Crowd anomaly detection using aggregation of ensembles of fine-tuned ConvNets, 925-2312/© 2019 Elsevier
20. Zhang, X., Yang, S., Zhang, J., Zhang, W.: Video anomaly detection and localization using motion-field shape description and homogeneity testing, 0031-3203/© 2020 Elsevier Ltd
21. Li, N., Chang, F., Liu, C.: Spatial-temporal cascade autoencoder for video anomaly detection in crowded scenes. IEEE Trans. Multimed. (2020)
22. Zhang, D., Gatica-Perez, D., Bengio, S., McCowan, I.: Semi-supervised adapted HMMs for unusual event detection. In: Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), pp. 611–618. IEEE (2005)
23. Tu, N.A., Huynh-The, T., Khan, K.U., Lee, Y.: ML-HDP: a hierarchical bayesian nonparametric model for recognizing human actions in video. IEEE Trans Circ. Syst. Video Technol. **29**(3), 800–814 (2019)
24. Li, H., Hu, Z., Wu, Y., Wu, F.: Behavior modeling and abnormality detection based on semi-supervised learning method. J. Softw. **18**, 527–537 (2007)
25. Feng, J., Zhang, C., Hao, P.: Online learning with self-organizing maps for anomaly detection in crowd scenes. In: Proceedings of the 20th International Conference on Pattern Recognition, 23–26 Aug 2010, pp. 3599–3602
26. Sharif, M.H., Uyaver, S., Djeraba, C.: Crowd behavior surveillance using Bhattacharyya distance metric. In: Proceedings of the CompIMAGE, 5–7 May 2010, pp. 311–323

27. Xiang, T., Gong, S.: Incremental and adaptive abnormal behaviour detection. Comput. Vision Image Understand. **111**(1), 59–73 (2008)
28. Calderara, S., Cucchiara, R., Prati, A.: Detection of abnormal behaviors using a mixture of von Mises distributions. In: Proceedings of the IEEE Conference on Advanced Video and Signal Based Surveillance, 5–7 Sept 2007, pp. 141–146
29. Varadarajan, J., Odobez, J.: Topic models for scene analysis and abnormality detection. In: Proceedings of the IEEE 12th International Conference on Computer Vision Workshops, 27 Sept–4 Oct 2009, pp. 1338–1345
30. Goshorn, R., Goshorn, D., Goshorn, J., Goshorn, L.: Abnormal behavior-detection using sequential syntactical classification in a network of clustered cameras. In: Proceedings of the 2nd ACM/IEEE International Conference on Distributed Smart Cameras, 7–11 Sept 2008, pp. 1–10
31. Morris, R.J., Hogg, D.C.: Statistical models of object interaction. Int. J. Comput. Vision **37**(2), 209–215 (2000)
32. Xiang, T., Gong, S.: Video behaviour profiling and abnormality detection without manual labelling. In: Proceedings of the 10th IEEE International Conference on Computing Vision, vol. 2, 17–20 Oct 2005, pp. 1238–1245
33. Tang, Y., Wang, X., Lu, H.: Intelligent video analysis technology for elevator cage abnormality detection in computer vision. In: Proceedings of the 4th International Conference on Computer Sciences and Convergence Information Technology, 24–26 Nov 2009, pp. 1252–1258
34. Andrade, E.L., Blunsden, S., Fisher, R.B.: Modelling crowd scenes for event detection. In: Proceedings of the 18th International Conference on Pattern Recognition, vol. 1, 20–24 Aug 2006, pp. 175–178
35. Bouttefroy, P.L.M., Bouzerdoum, A., Phung, S.L., Beghdadi, A.: Abnormal behavior detection using a multi-modal stochastic learning approach. In: Proceedings of the International Conference on Intelligence Sensors, Sensor Networks and Information Processing, 15–18 Dec 2008, pp. 121–126
36. Ermis, E.B., Saligrama, V., Jodoin, P., Konrad, J.: Motion segmentation and abnormal behavior detection via behavior clustering. In: Proceedings of the 15th IEEE International Conference on Image Processing, 12–15 Oct 2008, pp. 769–772
37. Fan, Y., Wen, G., Li, D., Qiu, S., Levine, M.D., Xiao, F.: Video anomaly detection and localization via Gaussian mixture fully convolutional variational autoencoder. 1077-3142/© 2020 Elsevier Inc.
38. Sabokrou, M., Fathy, M., Hoseini, M.: Video anomaly detection and localization based on the sparsity and reconstruction error of auto-encoder. Electron. Lett. **52**(13), 1122–1124 (2016)
39. Ravanbakhsh, M., Nabi, M., Mousavi, H., Sangineto, E., Sebe, N. (2017) Plug-and-play CNN for crowd motion analysis: an application in anomalous event detection. In: WACV
40. Sabokrou, M., Mohsen, F., Fathy, M., Klette, R.: Deep-anomaly: fully convolutional neural network for fast anomaly detection in crowded scenes. arXiv preprint arXiv:1609.00866 (2016)
41. Hasan, M., Choi, J., Neumanny, J., Roy-Chowdhury, A.K., Davis, L.S.: Learning temporal regularity in video sequences. In: CVPR (2016)
42. Fei-Fei, L., Deng, J., Li, K.: Imagenet: constructing a large-scale image database. J. Vis. **9**(8), 1037 (2010). Retrieved from http://jov.arvojournals.org/Article.aspx?doi=10.1167/9.8.1037
43. Bar, Y., Diamant, I., Wolf, L., Lieberman, S., Konen, E., Greespan, H.: Chest pathology identification using deep feature selection with non-medical training. In: Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization, May 2016
44. van Ginneken, B., Setio, A.A.A., Jacobs, C., Ciompi, F.: Off-the-shelf convolutional neural network features for pulmonary nodule detection in computed tomography scans. IEEE Xplore, 23 July 2015
45. Wang, Q., Gao, J., Li, X.: Weakly supervised adversarial domain adaptation for semantic segmentation in urban scenes. IEEE Trans. Image Process. **28**(9) (2019)

Apple Leaf Disease Detection and Analysis Using Deep Learning Technique

Priyadarshini D. Kalwad, Suvarna G. Kanakaraddi, T. Preeti, Sindhoori Ichalakaranji, Shreyas Salimath, and Shubha Nayak

Abstract Agriculture is a key source of livelihood. Nearly, 70% of the Indian population is dependent on agriculture, and due to lack of technical knowledge, most of Indian farmers adopt manual cultivation which leads to reduction in quality and amount of agricultural production. Various types of diseases on fruit, stem and leaf reduce the crop productivity. Healthy leaves are very important for the fast growth of the plants and hence to increase production of crops and source of revenue. At present, farmers spray pesticides to the plants which has an adversarial effect on human health. Detecting diseases in plants leaves is important for agriculturalists. To identify leaf diseases, many advanced techniques need to be adopted. In this research, authors have developed plant disease recognition model to identify the disease in apple leaf and predicting the percentage of disease affected in leaf, based on leaf image classification and applying convolutional neural network algorithms. This model obtains an accuracy of 97.5%.

Keywords ROC · XGBoost · ALexNet · VGG16 · Gaussian blur · Canny edge detection · CNN · HIS

1 Introduction

Economy of India is majorly dependent on agricultural productivity. Agriculture is a backbone of Indian economy system [1]. Agriculture contributes to more than 60% of employment opportunities and 17% toward total GDP; hence, disease detection in plant plays a major role in agriculture domain. Mango, banana, papaya, orange,

P. D. Kalwad (✉) · S. G. Kanakaraddi · T. Preeti · S. Ichalakaranji · S. Salimath · S. Nayak
KLE Technological University, Hubli, Karnataka, India
e-mail: priyadarshini@kletech.ac.in

S. G. Kanakaraddi
e-mail: suvarna_gk@kletech.ac.in

T. Preeti
e-mail: preeti@kletech.ac.in

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_76

guava, grape, apple, pineapple, sapota, pomegranate, strawberry, litchi, etc. are the major fruits grown in India. Diseases in different parts of plant like stem, leaves and root are inevitable. Identification of the different types of diseases in plant is important because diseases cause major setback in productivity and economic losses [2], as well as the quality and the quantity of the fruit decreases. Hence to make analysis more precise, there is a need of automated tool or software in area of agriculture [3]. "An apple a day keeps doctor away" is a common saying. Apple is commonly consumed fruit and has many benefits on human health. Apple tree is affected by different types of diseases. Disease detection in apple leaf has received increasing attention for the monitoring of large apple or chards. Recently, with the improvement of image processing technique, machine learning [4] is used in industrial food machines which enable the classification of different types of leaves in terms of their size, color, shape and defects.

This paper is organized as follows; Sect. 2 deliberates the survey conducted on disease detection on leaves. In Sect. 3, proposed methodology is discussed. The development process of the research is described in Sect. 4. In Sect. 5, result analysis is discussed. Author concludes the research in Sect. 6.

2 Related Work

Saraansh et al. have used CNN (GoogleNet) model with categorical cross entropy as a loss function and Adam is used as optimizer in back propagation [5]. System is able to achieve best accuracy in case of apple cedar and black rust which is 88%. Size of the dataset has a huge impact on results.

Chaitali G. Dhaware and et al. have discussed that agriculture plays a major role in the Indian economy and the nation's economy hugely depends on its agricultural production. Unhealthy leaves lead to reduction of crops and its productivity. Identifying the unhealthy leaves manually requires colossal amount of work, extreme amount of time and expertise in the leaf diseases. Thus, various image processing techniques are used to discover plant leaf is healthy or unhealthy. Recognition of leaf disease at the initial stage along with some automatic methods reduces the huge efforts of noticing large farms. Image acquisition, image preprocessing, image segmentation, feature extraction and classification are the different stages of disease detection and identification. Different techniques of image preprocessing and image segmentation algorithm used for automatic recognition are discussed [6]. Various plant leaf disease classification algorithms are also discussed.

K. R. N. V. V. D. Aravind and et al. have used gradient boosting and support vector for classification. Here, SVC has performed better than gradient boosting [7]. Accuracy in case of gradient boosting is 87% and for support vector classifier is 91%.

Zhang Chuanlei and et al. have used RGA, GA-CFS and SVM to recognize diseases in apple by considering the extracted features. Powdery mildew, mosaic and rust are the spotted diseases. Result shows that GA-CFS can be used to select useful

features and image processing and pattern techniques for classification of the apple leaf diseases. The proposed method gives 94.22% accuracy [8].

Melike Sardogan and et al. have discussed that the crop quality of tomatoes will be affected by the leaf diseases like, yellow curved, septoria leaf spot, bacterial spot and late blight. To detect the symptoms of leaf diseases, automated approaches are used for plant disease classification. Tomato leaf diseases are classified by convolutional neural network (CNN) with learning vector quantization (LVQ) model [9].

Puspha Annabel et al. [10] have presented that leaf disease detection at the early stage is important for healthy growth of plants. To detect diseases and analyze the data from different aspects, machine learning is applied. Various types of plant diseases are implemented by applying different classification techniques.

Bhavini J. and et al. have extracted color (GCH, CCV) and texture features (LBP, CLBP) using random forest classifier to detect the fruit as healthy or defected. Further if fruit is defected, k-means clustering is used to identify the type of disease. Better results are obtained by Gabor with CLBP and Gabor with LTP compared to other features [11].

Tulshan and Raul [12] presented classification of plant disease using morphological operations and different features such as intensity, color and dimensions. Authors have applied different machine learning classification techniques for various types of plant diseases.

3 Proposed System

This research focuses on disease detection in plant and classify whether leaf is infected or not. Further, it provides the data on what type of disease and percentage of defected disease. Proposed system provides identification of the infected leaf by image processing technique, which is shown in Fig. 1. The existing system can only classify if the leaf is defected or not.

An uploaded image is processed using data preprocessing and data augmentation which involves horizontal flip, rotation, resizing and rescale techniques. For further analysis, useful features are extracted and image processing techniques are applied to the acquired images. Once the dataset is processed, it is fed to train the model as it classifies the apple leaf as healthy or defected. To obtain better accuracy of the model, an XGBoost classifier uses fit method to classify the leaves. The dataset is split into test and train set for validation. Here, feature vectors of image act as predictors and labels of image act as target variables. Augmented data is fed to classification model to classify the leaves as healthy or diseased. Images which are classified as diseased are used to train another classifier; VGG16 is used to recognize three diverse types of diseases in leaves such as black rot, cedar rust and scab. After this disease classification, calculation of the percentage of the infected region is done using gaussian blur, segmentation and contour techniques. To identify the severity of a disease, the affected area is also displayed.

Steps to find the percentage of leaf infected:

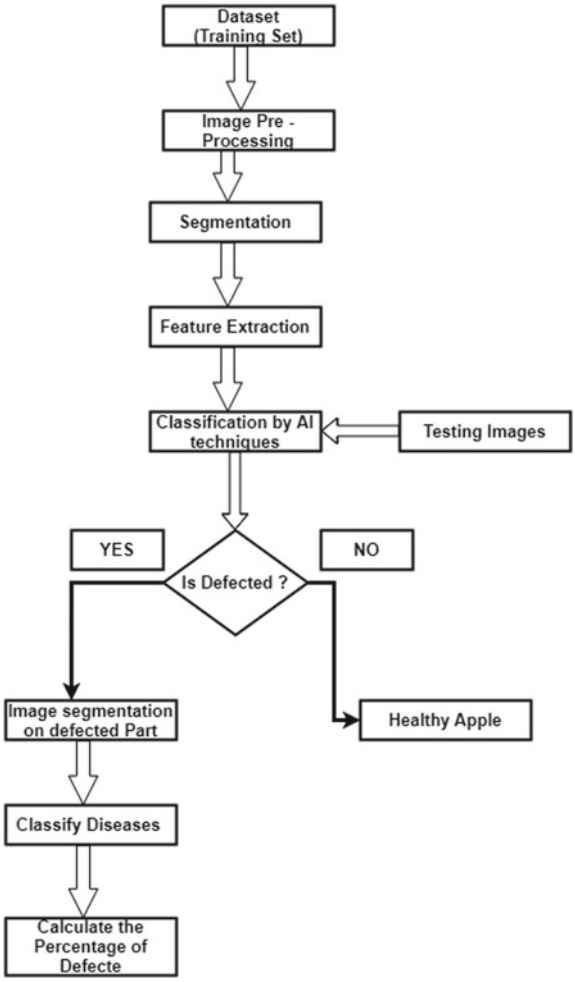


Fig. 1 Proposed system flow diagram

- Read the image
- Compute number of pixels with shade of white (p) to check if exclusion of these pixels is required or not (if more than a fixed %) in order to differentiate the white background or white patches in an image.

$$\text{Totalpixels} = \text{img.shape}[0] * \text{img.shape}[1]$$
$$\text{Per\_white} = 100 * p/\text{Totalpixels}$$

- Exclude all the pixels with color close to white, if they are more than 10% in the image.
- To blur the image and reduce the noise, gaussian blur function is applied.

- After noise reduction, Canny edge detection technique is used to find the boundaries of leaf within image.
- To obtain the outer border of the leaf, contouring method is used to join all the continuous points along with the boundary, having same color or intensity. Using the obtained border, perimeter and total area of leaf are calculated.

$$\text{Perimeter} = \text{cv2.arcLength (contours[maxid], True)}$$
$$\text{Total area} = \text{cv2.contourArea (contours[maxid])}$$

- After this, masking of green pixels is carried to set intensity values of green pixels to zero which will appear as a black background.
- Mask the infected region and apply contouring technique once again to get borders for all infected regions, using which infected area is computed.

$$\text{Infected area} += \text{cv2.contourArea (contours}[x])$$

- Finally using infected area and total area, the percentage of infected region is calculated.

$$\text{Percentage of infected region} = 100 * (\text{Infected area}/\text{Total area})$$

4 Development of Disease Identification and Classification Process

This section discusses about the leaf disease classification. The model involves the following steps: Image acquisition, image pre-processing, feature extraction and neural network-based classification.

4.1 Image Acquisition

The research dataset is collected from Kaggle source. Apple dataset consists of 7300 images. For training, 4610 images are considered, for testing, 1540 images are chosen, and to validate the model, 1150 images are considered.

4.2 Image Preprocessing

To build machine learning algorithms especially as CNN, resizing of the images in the dataset to a defined and standard dimension is needed which implies that dataset

must be preprocessed and scaled to defined identical widths and heights before the dataset is fed to train the algorithm. According to the algorithms, the images in dataset are resized to (224, 224) which are shown in Figs. 2 and 3.

Resize

RGB visualization: To represent colored images, RGB color space is used in most of the training models. All the other colors can be achieved by mixing these primary colors in the appropriate proportions shown in Fig. 4.

Horizontal flip: Shifting all pixels of an image in a horizontal direction is called horizontal flip (Fig. 5).

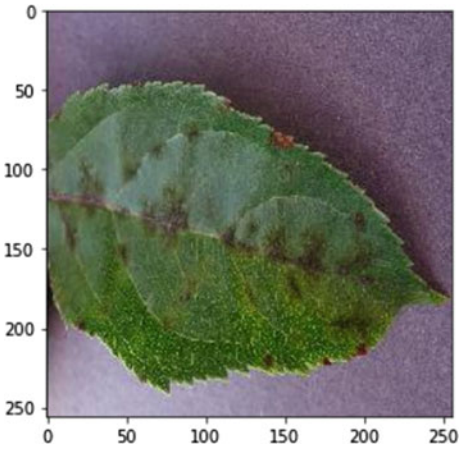


Fig. 2 Original image

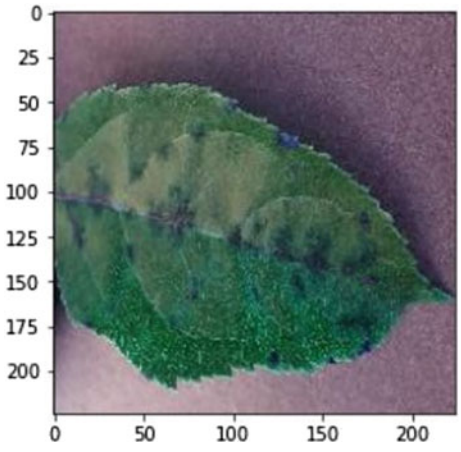


Fig. 3 Resized image

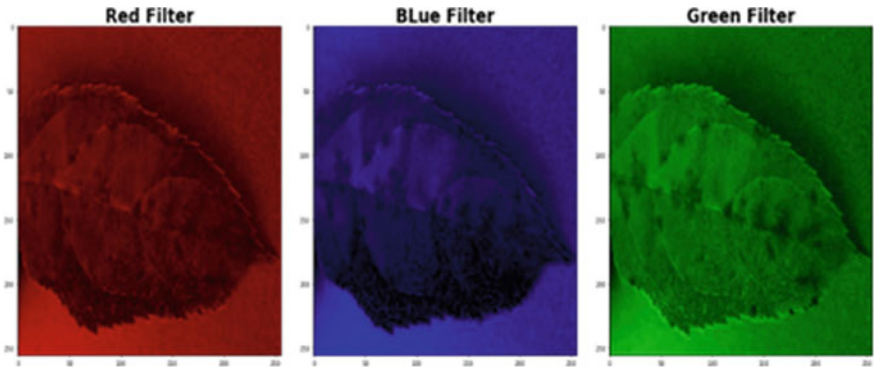


Fig. 4 RGB visualization

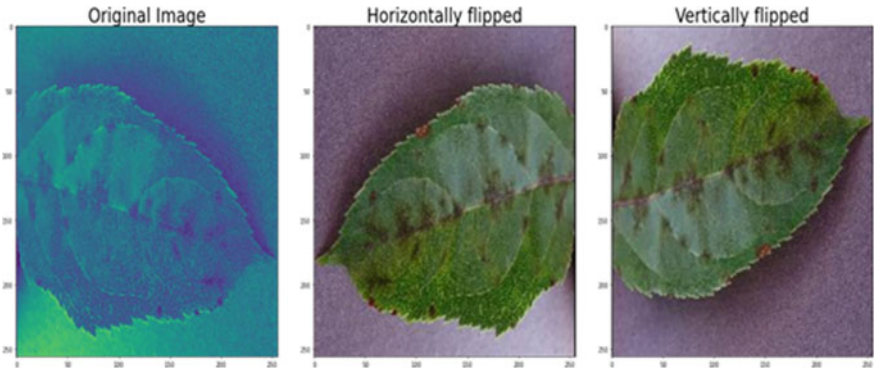


Fig. 5 Horizontal and Veritcal Flipped Images

4.3 Feature Extraction

The proposed work uses VGG16 model which is a convolutional neural network.

Predicting Percentage of Infected Region
Steps to find the percentage of leaf infected:

- **Color Transformation**: The input apple leaf image is transformed from RGB to HSI format. The hue, saturation, intensity (HSI) color model is based on human perception; hence, it is very popular form of color analysis and transformation.
- **Masking Green Pixels**: The green pixels in a leaf usually symbolize the healthy part of the leaf. Such green pixels are masked and later removed.
- **Segmentation**: The main aim is to extract the infected region of the leaf from the healthy leaf. This is done by segmenting the diseased part of the leaf from the other parts, i.e., with other similar colored parts (These similar parts can be a brown or a black colored region on the leaf that may look like the disease).

The masked out image is filtered and image processing is done over the region of interest.

- After the results are obtained, we analyze and evaluate the leaf features like percentage of leaf infected, area of leaf and the perimeter of the leaf. Different techniques used in this research are gaussian blur, contours and Canny edge detection.

5 Result Analysis

This section discusses about the results attained by training the entire dataset which consists of original and augmented images. For healthy or disease classification, XGBoost model performed better with the accuracy of 97.52%. Table 1 shows the classification result of XGBoost for the test images. Figure 6 shows the ROC curve for XGBoost for which the ROC curve area obtained is 0.99. For type of disease

Table 1 Classification result of XGBoost

	Precision	Recall	F1-score	Support
	0.98	0.98	0.98	169
	0.97	0.97	0.97	154
Accuracy			0.98	323
Macro average	0.98	0.98	0.98	323
Weighted average	0.98	0.98	0.98	323

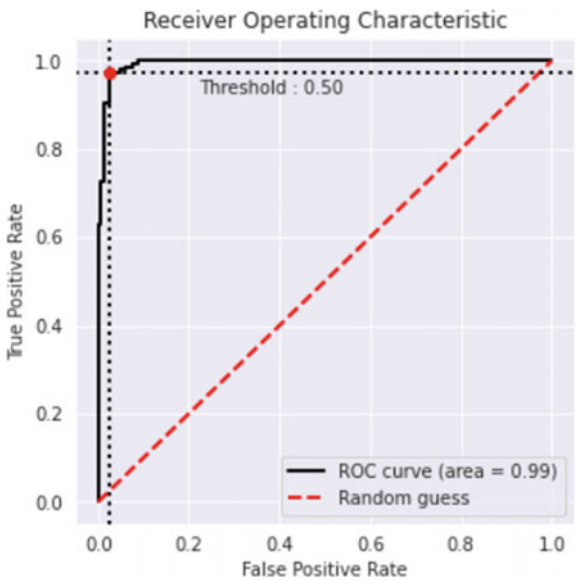


Fig. 6 ROC curve graph of XGBoost classifier

classification, VGG16 model performed better with the accuracy of 97.5% as shown in Figs. 7 and 8.

Figures 9 and 10 are the sample results obtained by XGBoost classifier.

The segmentation is based on two principles: discontinuity and similarity. In discontinuity, the regions with different properties like color, texture, intensity, etc.

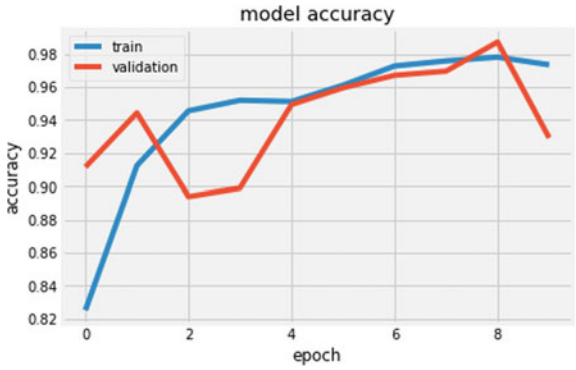


Fig. 7 Training and validation accuracy of VGG16 model

Fig. 8 Training and validation loss graph of VGG16 model

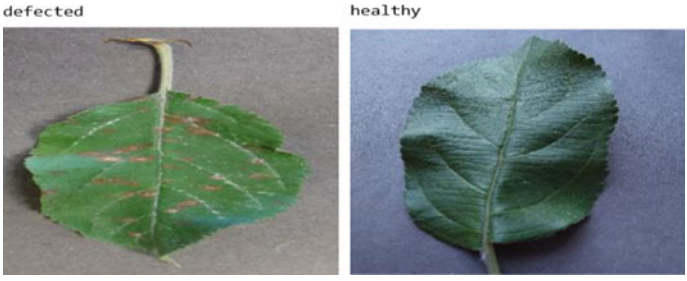


Fig. 9 Leaf is defected or healthy

Fig. 10 Different types of apple leaf defected result

are extracted. The identical image pixels are grouped with some predefined criteria. Based on similarity with the neighboring pixel, region-based algorithm is used. In leaf disease detection, segmentation is used to identify the infected area and features of a region are computed. For all the images, the following parameters like area of the leaf, percentage of the leaf infected and the perimeter are analyzed and evaluated using the above technique.

Following Fig. 11 represents the results of segmentation for extracting leaf features.

Below are some of the sample results obtained by predicting percentage of infected region which is shown in Table 2.

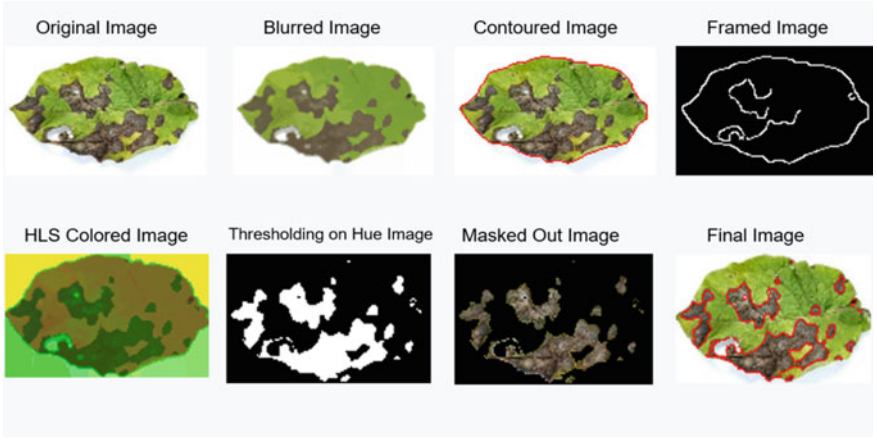


Fig. 11 Extracting leaf features

Table 2 Predicting percentage of infected region

Perimeter	335.46
Total area	6639.50
Infected area	2132.50
Percentage of infection region	32.12

6 Conclusion

This work is an attempt at improving apple leaf disease detection and classification. This research has contributed to design and develop a system that classifies the apple leaf is defected or healthy, and further, it detects the type of disease for a chosen input image. To train the model, gaussian blur and VGG 16 models are used to label the dataset. These methods are used to classify the apple leaves into healthy and defected and further into apple cedar rust, apple scab, apple black rot diseases. Also percentage of infected region for a leaf image is predicted. The model was trained with 4610 images of the leaf disease. Dataset considered for this research is from Kaggle source, which provides sufficient data to allow the system to generalize to most real-world situations. The model was evaluated using the testing chunk of the dataset, which gives better results, and also, the percentage of infected region is calculated for both healthy and defected leaves. This model has achieved an accuracy of around 97.5%.

References

1. Spoorti, S., Sourabh, D., Bhagyashri, N., Kanakaraddi, S.G.: Detection of stone weevil in mango fruit using nondestructive technique. Int. J. Emerg. Technol. Comput. Sci. Electron. (IJETCSE) **14**(2) (2015). ISSN: 0976-1353
2. Patil, A., Totad, S.G.: Non-invasive soya bean seed analysis using machine learning. Int. J. Recent Technol. Eng. **7**, 279–282 (2019). ISSN: 2277-3878
3. Divyashree, N., Ramsha, S., Shetti, S., Praveen, K., Kanakaraddi, S.G., Jahagirdhar, S.: Detection and quantification of disease in Cabbage using clustering and RGB colour features. Int. J. Emerg. Technol. Comput. Sci. Electron. (IJETCSE) **14**(2) (2015). ISSN: 0976-1353
4. Hampannavar, K., Totad, S.G.: Prediction of crop fertilizer consumption. In: 2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA), Pune, 16–18 Aug 2018
5. Baranwal, S., Khandelwal, S., Arora, A.: Deep learning convolutional neural network for apple leaves disease detection. SSRN Electron. J. (2019)
6. Dhaware, C.G., Wanjale, K.H.: A modern approach for plant leaf disease classification which depends on leaf image processing. In: 2017 International Conference on Computer Communication and Informatics (ICCCI) (2017)
7. Aravind, K.R.N.V.V.D., Prayla Shyry, S., Felix, Y.: Int. J. Innov. Technol. Explor. Eng. (IJITEE) **8**(12) (2019). ISSN: 2278-3075
8. Zhang, C., Zhang, S., Yang, J., Shi, Y., Chen, J.: Apple leaf disease identification using genetic algorithm and correlation based feature selection method (2019)
9. Sardogan, M., Tuncer, A., Ozen, Y.: Plant leaf disease detection and classification based on CNN with LVQ algorithm. In: 2018 3rd International Conference on Computer Science and Engineering (UBMK) (2018)
10. Puspha Annabel, L.S., Annapoorani, T., Deepalakshmi, P.: Machine learning for plant leaf disease detection and classification—a review. In: 2019 International Conference on Communication and Signal Processing (ICCSP) (2019)

11. Samajpati, B.J., Degadwala, S.D.: Hybrid approach for apple fruit diseases detection and classification using random forest classifier. In: 2016 International Conference on Communication and Signal Processing (ICCSP), pp. 1015–1019. IEEE, Apr 2016
12. Tulshan, A.S., Raul, N.: Plant leaf disease detection using machine learning. In: 2019 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT) (2019)

Secure Encrypted Word Proximity Retrieval Framework for Outsourced Relational Database Backed Applications

Kurra Mallaiah, Prateek Pande, Rishi Kumar Gandhi, and S. Ramachandram

Abstract Significant chunks of critical and sensitive data are being outsourced to untrusted public environments delivering commercial services. This is gradually becoming a serious threat for data security. Especially in these environments, it is usually necessary to safeguard sensitive data across its complete lifecycle (rest, transit and most crucial during-use). Mostly, data during-use still remains unsecured and hence most vulnerable. The enterprise applications generally use search queries for data retrieval from databases. These must be executed directly on encrypted data itself so as to achieve complete confidentiality of data throughout life cycle. In this paper, we present a framework to perform search based on proximity upon encrypted data while most research efforts are limited to only search based on exact keyword. We propose a novel framework for proximity-based retrieval on encrypted data for applications which are supported by relational databases at back ends. This framework comprises of a trusted proxy to search application's query over encrypted data hosted on outsourced database server while maintaining perfect secrecy by encryption algorithm involved. This is achieved by allocating words into domains, represented by their handles. The scheme presents two novelties. Firstly, the scheme supports security in all online transactional processing (OLTP) operations. This is evident in performance statistics observed. Secondly, it does not require any changes to existing stack of applications, thus it becomes practical toward real-world applications.

Keywords Cloud database security · Proximity searchable encryption · OLTP

1 Introduction

Cloud computing is emerging as quite practical for many organizations for outsourcing their databases over cloud environments. This minimizes capital cost and maintenance on IT resources. Governments and industries see a critical security concern in protecting the confidentiality of such databases hosted in an untrusted

K. Mallaiah (✉) · P. Pande · R. K. Gandhi · S. Ramachandram
Osmania University (OU), Hyderabad, India

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_77

service environment. Since only authorized entities should have access to data, hosting it after encrypting becomes quite obvious. It now requires data decryption so as to access data or perform a basic computation. But, this process demands encryption keys to be shared with cloud administrator. Thus, it can decrypt the encrypted data for usage/computation whenever required. This sharing of encryption keys with untrusted cloud administrator will destroy the purpose of data encryption. On the other hand, trying to keep the purpose by migrating entire encrypted database off cloud even for significant computations also remains highly infeasible. Thus, performing operations (e.g., search query) on encrypted data itself is a potential solution.

This solution addresses the threats arising due to rogue insiders and outsiders such as malicious database administrators and hackers, respectively. Complete confidentiality is maintained against threats from outsider and as well as insiders, as sensitive data always remains in encrypted form. This solution delivers security throughout life cycle of data, i.e., data at rest, in transition and also most importantly while in operation. Contemporary cryptographic protocols, techniques and algorithms adequately protect data at rest and transition. Data in operation can be secured by performing operations directly on encrypted data. Generally, enterprise level applications mostly use search operations to fetch required data from database servers. With the above scheme, these servers can also host encrypted data and facilitate search operations.

Defense Advanced Research Projects Agency (DARPA) has identified searchable encryption scheme as a potential solution to protect the sensitive databases in an outsourced environment. The various government regulations such as HIPAA, SOX, the Gramm–Leach–Bliley Act (GLBA) and PCI-DSS are mandating that companies have to encrypt their sensitive data to preserve the confidentiality of data from unauthorized entities. In this case also, need for a secure way of performing operations (searching) on encrypted data exists.

In past, organizations relied solely on a protected environment and strong access control for protecting confidentiality of their databases. The current use of cloud services offers organizations to rely solely upon crytopgraphic techniques to secure their data and confront threats such as business espionage [1]. Song et al. [2] addressed the problem of searching on encrypted data. This scheme supports users to create a trapdoor using symmetric key for a given search word and sends to the server where encrypted documents are stored. The server tests whether this encrypted search word is present in the document. Using public key concept for finding a keyword in the encrypted document at server side is addressed in [3]. CryptDB [4] addresses only full keyword searches for encrypted relational databases.

Various techniques in survey of literature support search based on exact keywords on encrypted documents. The proximity searching is very much useful in information retrieval (IR) systems to find the words that occur near each other [5]. In this paper, we propose a secure framework for retrieving the encrypted word on the basis of proximity to search keywords suitable for cloud-based relational databases. This is achieved by distributing the words into domains based on proximity with each other. The domains are organized as leaves of a binary tree. In this framework, advanced encryption standard (AES) and a trusted proxy (MySQL) are employed.

The remaining paper is arranged into three sections. In Sect. 2, related work discusses various research solutions available in the literature on searchable encryption. In Sect. 3, proposed scheme discusses secure encrypted word proximity retrieval framework for outsourced relational database backed applications, its performance evaluation and security analysis. Finally, Sect. 4 discusses conclusion and future scope.

2 Related Work

The most research efforts on searchable encryption [2, 3, 6–9] focus on searching of exact key word match on encrypted data. To the best of our knowledge, the phrase proximity search on encrypted documents described in [10] is only the work considering proximity search over encrypted data. The searchable encryption scheme proposed in [11] supports search on encrypted databases using exact matching of search word using modified version of Song, Wagner and Perrig scheme.

The fully homomorphic encryption schemes support all computations on encrypted data[12, 13] but are highly inefficient and thus highly impractical. Most of the research carried out on searchable encryption for retrieving the documents where the given search word matches exactly is meant for encrypted documents.

To the best of our knowledge, the proposed word proximity encryption search scheme is the first of its kind. The main contributions of this work are as follows:

i. A novel word proximity searchable encryption scheme is for cloud-based relational databases.
ii. The proposed scheme is practically usable in information retrieval systems for outsourced encrypted relational databases.
iii. Encryption, decryption and key management of words take place in the trusted MySQL Proxy which is under client's control. No encryption keys are shared with administrator of database server to search keyword on encrypted data.
iv. Ensures the confidentiality of sensitive data stored in the untrusted server from the malicious administrators by searching on encrypted data directly.
v. The proposed scheme is provably secure. The techniques used in this scheme provide provable secrecy of encrypted data, i.e., the untrusted database server can learn nothing about plain data given only the ciphertext.

2.1 Threat Model

i. *A malicious administrator cannot obtain the sensitive plain data from database server given only ciphertext.*
ii. *An adversary who compromises database server using a vulnerability in operating system or application cannot access sensitive plaintext data.*

These two threats are mitigated with proposed scheme by not sharing or storing the cryptographic keys in database server and supporting the search operation on encrypted data directly without decrypting in the database server while performing the search operation.

3 Proposed Scheme: Secure Word Proximity Retrieval Framework

3.1 Architecture Overview

This section presents the proposed word proximity searchable encryption scheme using the trusted proxy for the cloud-based relational databases. For simplicity, we assume that the length of all words in user data is constant. This scheme is also applicable to scenarios with different length words if an upper bound to proximity threshold is set equal to length of smallest word in the considered set. The scheme comprises of five procedures, viz. prepare, encryption, search trapdoor, search and decryption. A client, trusted proxy and untrusted database server are three main logical entities present in architecture. This is illustrated in Fig. 1.

i. The client initially sends data in plain text to store in database server and subsequently performs search over it. All such operations are scripted in SQL queries.
ii. The proxy receives plain data from client and segregates its each word into different domains as per proximity with other words. This is done using pro-

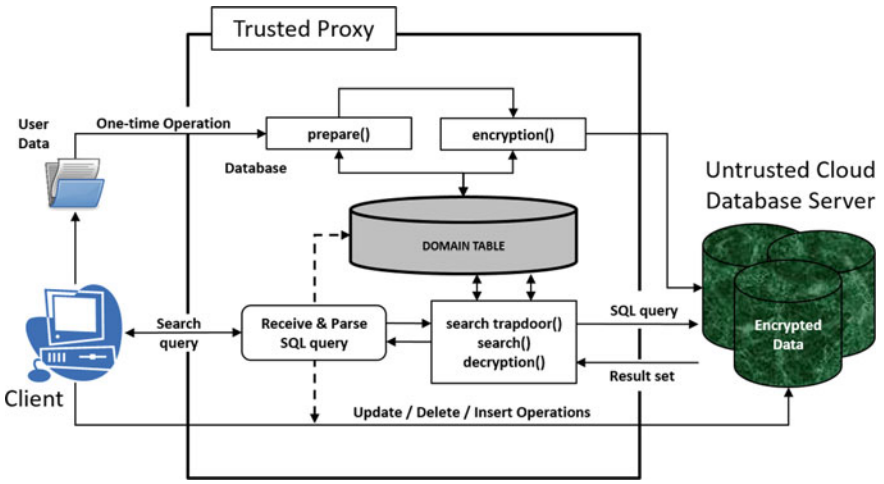


Fig. 1 Proposed secure encrypted word proximity retrieval framework for outsourced relational database backed applications architecture

cedure 'prepare'. The words belonging to each domain have a common prefix of bits which is referred to as handle. Once the words are assigned to different domains, words are encrypted in the proxy using a domain-specific key pair. The proxy sends these encrypted words to the database server for storing. Similarly, proxy encrypts the search keyword using 'search–trap' procedure to prepare fresh search query for execution on database server.

iii. Database server executes search query received from trusted proxy and returns the result set as per matched predicate back to trusted proxy.

iv. Proxy decrypts the result set using domain-specific key pair using 'decrypt procedure' and sends the search result back to the client.

3.2 Proposed Scheme Procedures

***prepare()* in proxy** The responsibility of prepare procedure is to organize plain text data into different domains. Each domain contains all those words which have same similarity level. For any two given words, the similarity level is the number of most significant bits which are same in both the given words. Higher the number, th more is proximity or similarity. Therefore, a domain contains the words which have same sequence of most significant bits. This sequence of most significant bits is referred to as domain handle of the concerned domain. Its length gives the minimal similarity between two words of that domain (though for some words, similarity can be higher than the length of their handle).

We create a binary tree which contains above explained domains represented by its leaves and domain handle by the path traversed from root node to leaf node. Initially, all the words are assumed to be residing at root node. Sequentially, each word is selected and propagated down the binary tree according to most significant bit of the word. For a given word at root node, if most significant bit is 0, then word is sent to left subtree otherwise right subtree. This process is recursively carried out at other nodes at each level in the binary tree. Putting up words in this manner results in a set of words at leaf nodes of the tree. All the words at a leaf node have same traversal path, thus have same domain handle. Therefore, each leaf node represents one domain.

For instance in Fig. 2, a user has following eight words each of length 1 byte to be hosted on the database. Each word is put in to binary tree through above mechanism.

$$\mathbf{w1} = 11001010 \quad \mathbf{w2} = 01110001 \quad \mathbf{w3} = 10011000 \quad \mathbf{w4} = 00010001$$
$$\mathbf{w5} = 01011111 \quad \mathbf{w6} = 11101010 \quad \mathbf{w7} = 01000001 \quad \mathbf{w8} = 00010011$$

Let us set similarity threshold (p) to 3. So, binary tree have three levels and eight (2^3) leaves with each leaf representing a domain. In this case, all the words are now assumed to be residing at the root node of binary tree. Starting from first word up to last word, most significant bit of word is read. If it happens to be 0, then word is put

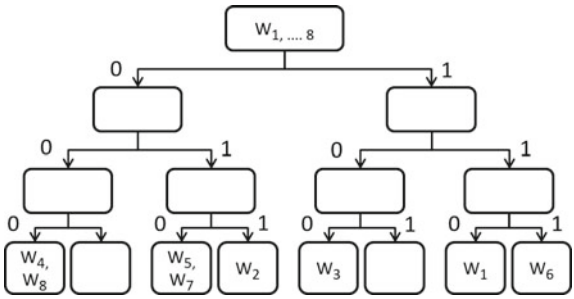


Fig. 2 Domains are represented by leaves of a tree structure stored at proxy

Domain Number D_m	Hash Digest of Domain Handle	Domain Key $k_{1\text{-}p}$	Domain Key $k_{p\text{-}l}$
0	handle_hash$_0$	.....	.....
1	handle_hash$_1$	.....	.....
2	handle_hash$_2$	.....	.....
...			
...			
g-1	handle_hash$_{g\text{-}1}$	.....	.....

Fig. 3 Domain keys and SHA256 digest of handle

to left subtree else to right subtree for further comparisons. After p comparisons, all words get propagated to leaves of tree. As in the illustration above, w4 reaches first node, whereas each of w5 and w7 reaches third leaf node and so on. So, now each leaf node represents a domain. Here, w5 and w7 form a single domain, and both have three most significant bits. This sequence of most significant bits is referred to as *handle* of that domain.

Once the given words are put in domains at leaves of binary tree, further procedures for encryption and search may be taken up. All the words are placed in the domains ranging from D_0 to D_{g-1}. A SHA256 hash digest of handle and two cryptographic keys are maintained for each domain in a table at proxy as shown in Fig. 3. Indexing in to this table gives details of corresponding domain.

The words are assumed to be of length l bits. Each word is divided into two parts. First part w_{1-p} comprises handle (most significant p bits), and second part w_{p-l} consists of remaining $(l - p)$ bits. Keys k_{1-p} and k_{p-l} are used to encrypt first and second parts, respectively. The hash digest of handle and keys for domains which are found to be empty need not be stored at proxy.

The algorithm for prepare() can be outlined as follows:

i. Let $W = \{w_0, w_1, ..., w_{n-1}\}$ be set of n words.
ii. Also, $|w_i| = l$, where $i = 0$ to $n - 1$ and $w_i \in$ W.
iii. Similarity threshold (length of domain handle) be p
iv. Let various domains be denoted as set D_0, D_1,....., D_{g-1}, where g = 2^p and $D_i = null$ for $i = 0$ to $g - 1$.

v. $handle = 0$
/* $handle$ is initialized to handle of first domain i.e. a vector having zero $p - bits$ */

vi. **/* for each word */**

```
for(j = 0 ; j < (n-1) ; j++) {
    /* for each domain */
    for(m = 0 ; m < (g-1) ; m++) {
        if(w_j<<(l − p) == handle) {
            {w_j} ∪ D_m
            break;
        }
        else
            handle++;
    }
}
```

For the sake of simplicity, linear search method is used above. It may be replaced by a binary search method for optimisation purpose.

Key Management Schemes in Proxy This framework requires encryption of words in two parts for which a range of schemes pertaining to key management can be adopted. Trade-off between desired level of privacy, computing power and storage capacity at trusted proxy is the basis to select a particular scheme of key management in the range. In this paper, the scheme chosen lies at higher end of this range. For each domain, a unique pair of keys and hash digest of handles are maintained at trusted proxy. Digest of a handle is used to obtain these keys to encrypt respective parts of words during various operations.

A scheme with a common pair of keys for all domains or a single key for both parts of a word in all domains also achieves reasonable functional security. Generally to select an appropriate scheme, user environment of implementation plays final critical role.

***encryption()* in proxy** The organization will set up the trusted proxy between their employees and cloud database server, and all the requests from organization to the database server are routed through the trusted proxy. NIST standard block cipher advanced encryption standard(AES) encryption algorithm is used to encrypt both parts of a word using keys k_{1-p} and k_{p-l}, respectively. This procedure is depicted in Fig. 4.

i. $CT_1 = \mathrm{E}(k_{1-p}, w_{1-p})$
ii. $CT_2 = \mathrm{E}(k_{p-l}, w_{p-l})$
iii. C = concatenate(CT_1, CT_2).

The ciphertext CT_1 and CT_2 are produced after encryption of each part, respectively. These are concatenated to obtain final ciphertext for a given word w. Similarly, each word in each domain(D_m) is encrypted in the proxy using domain-specific keys, and the resulting ciphertexts (C) will be sent to the database server to be stored.

***search Trapdoor()* in proxy** This procedure prepares a token for proximity-based search in the proxy up on receiving a search query from client/application. For a

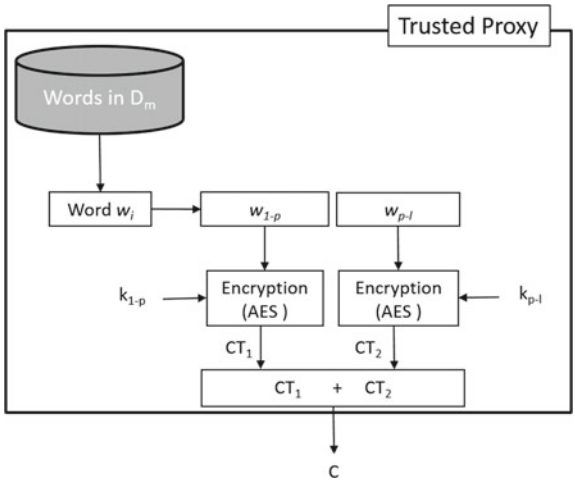


Fig. 4 Encryption procedure

given search word, the proxy determines its domain by comparing the hash digest of most significant p bits in the word against the table column having hash digest of various domain handles. Index value at which this match occurs indicates the domain number. Keys for the corresponding domain can also be obtained using this index value. Only first part, i.e., $(1 - p)$ bits in given search word is encrypted using the corresponding key.

For instance, the search word is SW, and hash digest of most significant $(1 - p)$ bits is generated and compared against the table column having hash digest of various domain handles.

i. $hash\_digest = \text{hash}(SW_{1-p})$
ii. $m = \text{compare}\ (hash\_digest\ , handle\_hash_i)$
where $i = 0$ to $(g - 1)$
This compares the $hash\_digest$ with all various domain handles and returns domain number m of the domain search word SW belongs to.
iii. Proxy uses first key of mth domain k_{1-p} to encrypt SW_{1-p} using AES algorithm. Search token (STK) is obtained as a result.
$\text{STK} = \text{E}(k_{1-p}, SW_{1-p})$

STK here becomes the new search word. Proxy prepares new search query using this STK and sends the same to the database server. The search trapdoor procedure of proposed scheme is depicted in Fig. 5.

***search()* in database server** This procedure executes the proxy's search query on encrypted databases and returns all the words which are within the specified proximity of search word. In order to make this search possible, encrypted words C_i for $i = 0$ to $(n - 1)$ are stored in the following manner. For every encrypted word in each domain, CT_1 and CT_2 are stored in a way such that a query can fetch all CT_2 s on the basis of a given CT_1. One simple implementation for such a scheme is to store

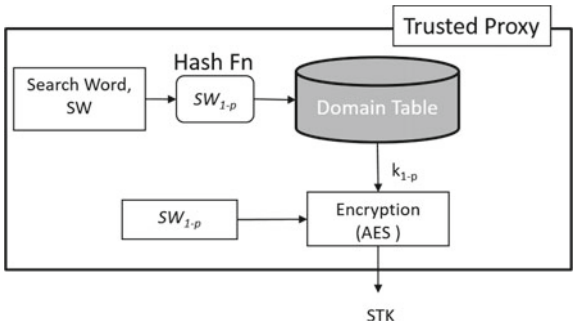


Fig. 5 Search trapdoor procedure

CT_1 and CT_2 in two separate columns of a database table. So that search query can be easily framed to fetch all CT_2 *s* for a given CT_1.

An alternative implementation could be to store encrypted words C_i for $i = 0$ to $(n - 1)$ in a single column, and corresponding search query may be created at proxy.

Up on receiving a search token (STK), it is compared with each CT_1. If a match occurs, corresponding pair $< CT_1 , CT_2 >$ is put into search result set. Once these comparisions are done, result set is in turn returned back to proxy. Thus, in other words, if STK matches for handle of domain D_m, then ciphertexts for all words in domain D_m are returned as search result.

***decryption()* in proxy** This procedure decrypts the search result set received from the database server in the proxy. Using the keys k_{1-p} and k_{p-l} for domain D_m, parts in each word of result set are decrypted, respectively.

i. For all ciphertexts in results set,
   $res_1 = \mathrm{D}(k_{1-p} , CT_1)$
   $res_2 = \mathrm{D}(k_{p-l} , CT_2)$
ii. $res = \text{concatenate}(res_1, res_2)$
iii. Decrypted result set is sent to client/application.

This scheme is designed in modular and stateless fashion. So both synchronous and asynchronous implementations are possible depending on behavioral requirements of the system. While in case of asynchronous implementation, the domain number *m* identified during *search trapdoor()* can be recovered in *decryption()* also. Otherwise in synchronous, domain number *m* is shared with all procedures till *decryption()*.

3.3 Secure Insert, Update and Delete Operations

Secure Insert Upon receiving a query from application to insert a word in the database server, proxy first determines the domain to which concerned word belongs to. Based on match, domain number *m* is obtained. Proxy encrypts the parts of word using respective cryptographic keys for domain D_m. The concatenated ciphertext is then sent to database server. The secure insert procedure is depicted in Fig. 6.

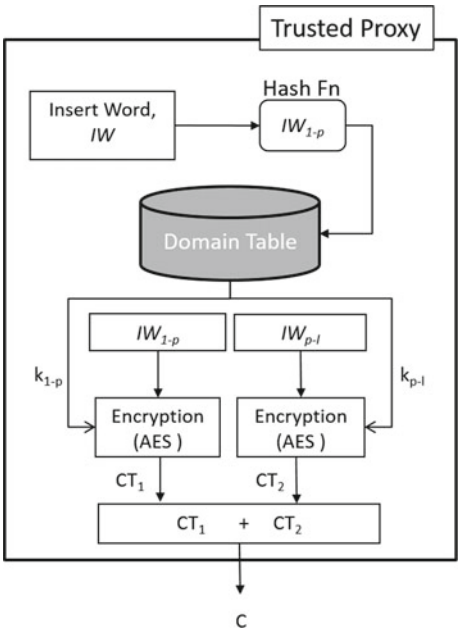


Fig. 6 Secure insert the word procedure

Let us consider IW as a word to be inserted in database. Hash digest of first p bits of IW is compared with stored hash digest of various domain handles in proxy. Once domain number m is determined, corresponding keys are used to encrypt respective parts IW_{1-p} and IW_{p-l}. Finally, concatenated ciphertext is sent to the database server.

i. $hash\_digest = \text{hash}(IW_{1-p})$
ii. $m = \text{compare}(hash\_digest, handle\_hash_i)$ where i = 0 to (k-1)
iii. Using keys for domain D_m for encryption.
$CT_1 = \text{E}(k_{1-p}, IW_{1-p})$
$CT_2 = \text{E}(k_{p-l}, IW_{p-l})$
iv. C = concatenate(CT_1, CT_2)
C is sent to database server for insertion.
Depending on implementation, C is stored either in a single column or is split in to CT_1, CT_2 to be stored in two columns.

Secure Delete Upon receiving a query from application to delete a word from the database server, proxy first determines the domain to which concerned word belongs to. Based on match, domain number m is obtained. Proxy encrypts the parts of word using respective cryptographic keys for domain D_m. The concatenated ciphertext is then sent to database server for deletion. The secure delete procedure is depicted in Fig. 7.

Let us consider DW as a word to be deleted from database. Hash digest of first p bits of DW is compared with stored hash digest of various domain handles in proxy. Once

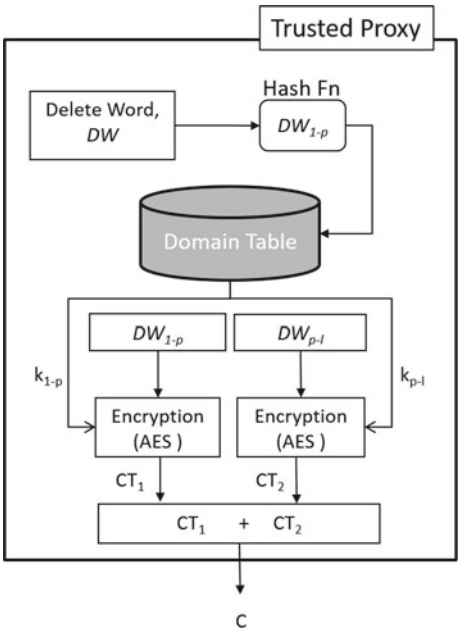


Fig. 7 Secure deleting a word procedure

domain number m is determined, corresponding keys are used to encrypt respective parts DW_{1-p} and DW_{p-l}. Finally, concatenated ciphertext is sent to the database server for deletion.

i. $hash\_digest = \text{hash}(DW_{1-p})$
ii. $m = \text{compare}(hash\_digest, handle\_hash_i)$ where $i = 0$ to $(k-1)$
iii. Using keys for domain D_m for encryption.
$CT_1 = \text{E}(k_{1-p}, DW_{1-p})$
$CT_2 = \text{E}(k_{p-l}, DW_{p-l})$
iv. C = concatenate(CT_1, CT_2)
C is sent to database server for deletion.
Also, if only CT_1 is used by proxy for framing delete query, proximity-based deletion can also be achieved.

Secure Update In principle, any update operation is equivalent to a delete and then an insert operation. Thus, procedures for secure insert and delete mentioned above can be used to formulate procedure for secure update. The procedure hence formulated can perform update over the *old word* even if it belongs to either same domain as *new word* or a completely different domain.

3.4 Performance Analysis of Proposed Scheme

The proposed scheme is implemented using C language with database *MySQL* and *OpenSSL* library. In order to measure performance statistics, a high-end computing facility with below configurations is used to host the trusted proxy, database server and client application.

Hardware Specifications:
Processor: Intel Xeon E52676 (Haswell)—2.4 GHz—16 vCPU; Memory: 64 GiB
Software Specifications:
Enterprise Linux 2.6.32–431.el6.x86_64; Database: MySQL v4.14; Library: OpenSSL-1.0.1.

For purpose of testing, 1 lakh (0.1 million) 8-bit words, with words with random content are created and fed to framework. Threshold value p is set to 3 bits. Two modes of storage for encrypted words are adopted during tests. In the first mode, all the words which are encrypted in *prepare()* are sent in memory in form of a data structure. In the other mode, a relational database MySQL is used to store encrypted words. In order to reflect marginal overhead due to encryption in the framework, all operations are also implemented with plain data words (i.e., without any encryption). For both storage mode, time elapsed for each operation in case of encrypted words and plain data words are shown in given Tables 1, 2, 3 and 4. Slight overhead due to encryption of words in first mode of storage is evident when Tables 1 and 3 are compared. The same is evident for second storage mode when Tables 2 and 4 are compared. This is illustrated in Figs. 8 and 9.

Everytime, *prepare()* remains a one time operation. Remaining operations (viz. *search(), insert(), delete(), update()*) take negligible time. This way entire set of operations can be considered as practically feasible. It is clearly evident that the proposed framework for word proximity-based search incurs marginal overhead. Besides, it is suitable to all real-time scenarios with response times of real world using a data set of large scale. The proposed scheme is meant for applications which have lists of keywords with considerable leading or trailing similarities. For instance, list of loan defaulters for a bank, list of patients with particular disease in a hospital or list of criminals for an investigation agency appears to be appropriate scenarios. In all these, sensitive data may be encrypted hosted in untrusted service environments yet available for all above demonstrated operations.

3.5 Security Analysis of Proposed Scheme

Security of the above scheme can be proved. This means that an adversary having access to ciphertext in database server or while executing a search query cannot derives plain data. AES (an NIST standard encryption algorithm) is used to encrypt data and search words in the proposed scheme. It protects confidentiality of sensitive data. Security of AES is already proven. Also, at any point of time, all the used

Table 1 Time elapsed in various operations on encrypted words stored in system memory

Iterations	1	2	3	4	Avg time elapsed (s)
Prepare	20.004339	20.040442	20.026257	20.00672	20.0194395
Search	0.001009	0.000925	0.000843	0.000703	0.00087
Insert	0.000029	0.000019	0.000006	0.000006	0.000015
Delete	0.001032	0.000961	0.000896	0.000728	0.00090425
Update	0.000968	0.000947	0.000901	0.000998	0.0009535

Table 2 Time elapsed in various operations on encrypted words stored in database

Iterations	1	2	3	4	Avg time elapsed (s)
Prepare	24.628572	24.637995	24.693247	24.653103	24.65322925
Search	0.012021	0.012039	0.012393	0.022339	0.014698
Insert	0.000119	0.00012	0.000119	0.000188	0.0001365
Delete	0.011419	0.011337	0.011185	0.011392	0.01133325
Update	0.054615	0.054567	0.054844	0.054776	0.0547005

Table 3 Time elapsed in various operations on plain words stored in system memory

Iterations	1	2	3	4	Avg time elapsed (s)
Prepare	19.63068	19.632608	19.597034	19.630104	19.6226065
Search	0.000774	0.000767	0.000767	0.000774	0.0007705
Insert	0.000016	0.000003	0.000016	0.000004	0.00000975
Delete	0.000731	0.000732	0.000831	0.000728	0.0007555
Update	0.000803	0.000806	0.000623	0.0009	0.000783

Table 4 Time elapsed in various operations on plain words stored in database

Iterations	1	2	3	4	Avg time elapsed (s)
Prepare	22.89455	22.940052	22.830717	22.958183	22.9058755
Search	0.008847	0.008475	0.008421	0.008425	0.008542
Insert	0.000062	0.000047	0.000048	0.000047	0.000051
Delete	0.007972	0.007808	0.007742	0.00783	0.007838
Update	0.035037	0.034907	0.034826	0.034853	0.03490575

Tables presented above provide time taken for various operations averaged over four independent iterations. As evident in Tables 1 and 3, *prepare()* takes marginally more time in case of words with encryption than in case of words with no encryption (Tables 2 and 4). When database is used to store words, the same trend is again evident

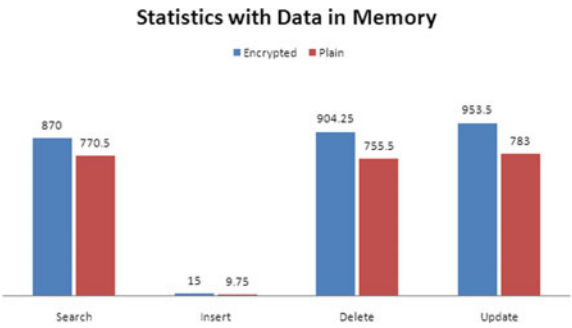


Fig. 8 Comparision of operation times in system memory (*in* μs)

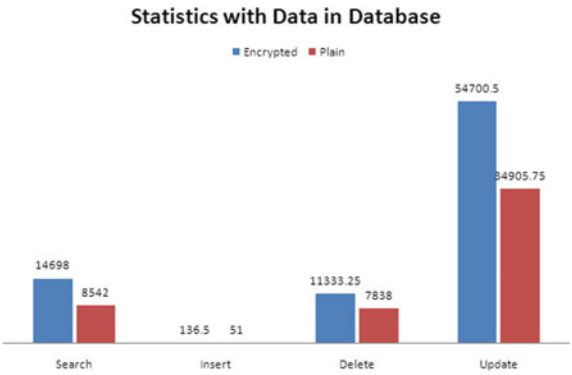


Fig. 9 Comparision of operation times in database (*in* μs)

cryptographic keys are always stored external to the untrusted database server. This validates the claim of provable security for the scheme hereby proposed in this paper.

Proof [14] AES is a block cipher algorithm. A block cipher is a function

$$E : K \times \{0, 1\}^n \rightarrow \{0, 1\}^n$$

such that $E_K(\cdot) = E(K, \cdot)$ is a permutation on $\{0, 1\}^n$. The inverse to the block cipher E is $D = E^{-1}$ defined by $D_K(Y)$ being the unique X ∈ $\{0, 1\}^n$ such that $E_K(X) = Y$. The way to quantify security of a block cipher $E : K \times \{0, 1\}^n \rightarrow \{0, 1\}^n$ works as follows.

Choose a random $K \leftarrow k$ and a random permutation π on n-bits. An adversary A is provided with black box access either to E_K or to π. The adversary attempts to make out which kind of object it has. We let

$$\begin{aligned} Adv_E^{prp}(A) = &\, Pr[K \leftarrow k : A^{E_K}(\cdot) \Rightarrow 1] \\ &- Pr[\pi \leftarrow Perm(n) : A^{\pi(\cdot)} \Rightarrow 1] \end{aligned}$$

where $Perm(n)$ indicates all n-bit string permutations. Alternatively, comparing a against a random function also sets a measure for security for block cipher. So,

$$Adv_E^{prf}(A) = Pr[K \leftarrow k : A^{E_K}(\cdot) \Rightarrow 1] - Pr[\rho \leftarrow Func(n,n) : A^{\rho(\cdot)} \Rightarrow 1]$$

where $Func(n, n)$ is set of all n-bit string to n-bit string functions . It is a standard result that $Pr[A^{\pi} \Rightarrow 1] - Pr[A^{\rho} \Rightarrow 1] \leq q^2/2^{n+1}$ for any adversary A that asks at most q queries. This makes the PRP and PRF notions of advantage close: $|\ \mathbf{Adv}_E^{prp}(A) - \mathbf{Adv}_E^{prf}(A)\ | \leq q^2/2^{n+1}$ if A asks q or fewer queries.

Resemblance of random permutations to forward and backward adversary oracles is also an important and strong notion of PRP security. This is equivalent to launching chosen ciphertext attack along with chosen plaintext attack. This advantage may be defined as

$$\mathbf{Adv}_E^{\pm prp}(A) = Pr[K \leftarrow k : A^{E_K(\cdot), E_K^{-1}(\cdot)} \Rightarrow 1] - Pr[\pi \leftarrow Perm(n) : A^{\pi(\cdot), \pi^{-1}} \Rightarrow 1]$$

Blockcipher like E = AES is widely considered good in PRP and 'reasonable' adversaries posses only "small" value $\mathbf{Adv}_E^{prp}(A)$ or $Adv_E^{\pm prp}(A)$. Thus, security of AES algorithm can be considered secure.

Besides, another security advantage in this scheme is that the keys required for all cryptographic operations are always restricted within user-controlled confines and are never allowed to be accessed/shared with administrator/database server. Therefore, it can be considered that the scheme of search based on proximity on encrypted data is *provably secure*.

4 Conclusion and Future Scope

In this paper, authors propose a secure encrypted word proximity retrieval framework for outsourced relational database backed applications using trusted proxy and a block cipher AES. The proposed scheme supports proximity search of words on encrypted relational databases. The scheme uses a trusted proxy to segregate all the words into domains depending on their similarity before encryption. In order to encrypt data and searching keywords, AES is used in trusted proxy. The security of the scheme proposed is provable and standard cryptographic definitions, and random oracle model is used to present proof for the same. All cryptographic operations are performed within trusted proxy which maintains all the cryptographic keys. The scheme protects confidentiality of sensitive data from untrusted database administrators by not sharing the cryptographic keys and performing all types of operations on encrypted data directly. It requires no alterations to client or server side environments. We have implemented and presented the scheme performance statistics. Besides, it is suitable to all real-time scenarios with response times of real world using a data set of large scale. The scheme is thereby practical and effective.

References

1. Bösch, C., Hartel, P., Jonker, W., Peter, A.: A survey of provably secure searchable encryption. ACM Comput. Surv. (CSUR) **47**(2), 18 (2015)
2. Song, D., Wagner, D., Perrig, A.: Practical techniques for searches on encrypted data. In: IEEE Symposium on Security and Privacy, pp. 44–55 (2000)
3. Boneh, D., Crescenzo, G.D., Ostrovsky, R., Persiano, G.: Public key encryption with keyword search. Proc. EuroCRYPT **2004**, 506–522 (2004)
4. Popa, R.A., Redfield, C.M.S., Zeldovich, N., Balakrishnan, H.: CryptDB: protecting confidentiality with encrypted query processing. In: SOSP'11, Cascais, Portugal, 23–26 Oct 2011
5. Salton, G.: Automatic text processing: the transformation, analysis, and retrieval of information by computer. Addison-Wesley (1989)
6. Goh, E.-J.: Secure indexes. Cryptology ePrint Archive, Report 2003/216
7. Bellare, M., Boldyreva, A., O'Neill, A.: Deterministic and efficiently searchable encryption. In: CRYPTO , pp 535–552 (2007)
8. Curtmola, R., Garay, J., Kamara, S., Ostrovsky, R.: Searchable symmetric encryption: improved definitions and efficient constructions. In: CCS, pp. 79–88 (2006)
9. Bao, F., Deng, R.H., Ding, X., Yang, Y.: Private query on encrypted data in multi-user settings. In: ISPEC, pp. 71–85 (2008)
10. Zittrower, S., Zou, C.: Encrypted phrase searching in the cloud. In: Proceedings of GLOBECOM (2012)
11. Popa, R.A., Redfield, C., Zeldovich, N., Balakrishnan, H. (2011) CryptDB: protecting confidentiality with encrypted query processing. In: SOSP, pp. 85–100 (2011)
12. Gentry, C.: Fully homomorphic encryption using ideal lattices. In: STOC, pp. 69–178 (2009)
13. Gentry, G., Halevi, S.: Implementing Gentry's fullyhomomorphic encryption scheme. In: Advances in Cryptology—EUROCRYPT, pp. 129–148 (2011)
14. Rogaway, P.: Evaluation of some Blockcipher modes of operation. http://www.cs.ucdavis.edu/~rogaway, 10 Feb 2011

Institutionalizing Individual Development Through Child-Centred Mechanisms in Indian Residential Childcare NGOs

Yashraj Wadalkar, Nishant Bagmar, and Wricha Mishra

Abstract This paper works towards understanding the layered nature of problems surrounding the domain of residential childcare service delivery and working towards solutions. We approach this research keeping in mind that we want to be able to empower the individual who is stuck in the clutches of a state regulated, institutionally operated residential childcare system. A focused study is made into understanding the problem areas that can be addressed through the usage of computing-led interventions. Through the conducted research, we arrive at a juncture, where we recommend a framework that can be implemented, based on a systems approach. This piece argues for the need of a deeper understanding of the problems in the domain, and not to try to follow a singular, blanket solution.

Keywords Child rights-based approach · Institutionalization · Individual

1 Introduction

India has the largest children population in the world (430 million) and along with that also has to handle one of the largest vulnerable child populations (40% of 430 million) [1]. Often in the popular discourse, they are referred to as "assets", [2] "resource" and everything else but individuals.

The following work examines deeply the contemporary research available in the domain of residential childcare and works towards establishing an outlook focused on self-discovery that can counter institutionalization. In order to tackle institutionalization in an Indian context, that aims at building resilient, congenial, peaceful secure childhoods, it is imperative that the organizations adopt a child rights based, interoperable, collaborative, action-based mechanism. It is the only effective approach that has a rigorous focus on empowering an individual child towards discovery of self, and building an identity, which uses technical interventions as a necessary aid. Therefore, it can deliver quantifiable, effective results.

Y. Wadalkar (✉) · N. Bagmar · W. Mishra
MIT Institute of Design, MIT-ADT University, Pune, Maharashtra 412201, India

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_78

We believe that the methods and insights discussed here are a step towards building a nurturing environment and an effort towards providing peaceful childhoods. This being the primary motivation towards how we approached our work. First, we try to understand the scale and nature of the problem. Section 2 works towards the same. As we move towards a solution, Sect. 3 highlights the recommendations and interventions towards the same. Section 4 shows a small model that we propose, which incorporates the interventions mentioned earlier. Section 5 displays the reception towards the solutions mentioned after consulting with organizations and highlights the critique of the same.

2 Literature Review

2.1 Lost Childhoods—The Tribulations Endured

An ever-changing developing society, reeling into the neo-liberal, neo-colonial, neo-conservative realities of the world, is all but a ripe atmosphere for communal, class, societal, gender disparities, distress and atrocities. The breakdown of traditional support systems [2] coupled with everyday survival against daunting odds, invisible clutches of poverty, distressed relationships, loss of hope, create distraught realities that an organism of the societal structure has to endure. This begets a population of children, who lead lives of quiet desperation. For no fault of their own, they are made to face the realities of a culture and upbringing that consider their existence as an "Unsustainable-Surplus" [3]. To quote Modi K., et al., "The situation adversely impacts the process of children's primary and secondary socialization, as the absence of a proper learning environment at home or in the community, inadequate parental emphasis on education, exposure to domestic violence and deviant practices, peer pressure and non-existent or inadequate coping mechanisms, financial hardship and emotional stress are all palpable phenomena in slums" [3].

It is at this critical juncture; institutions appear to be the only respite and last resort for the distressed parents and children. According to Malik, R. et al. "When education can be obtained for the children of poor families along with board and lodgings in a residential home, school or centre, there are big incentives for families to seek such opportunities for their children" [3]. It would only be fitting for these families to prefer NGOs as they themselves are not in a position to identify and support the emotional, physical growth of a child.

2.2 Institutionalisation—What Does It Bring with It?

Institutions, identified to be the rehabilitative refuge for the children to escape the tribulations of everyday survival, cause an irreplaceable harm. According to Islam,

T. et al., "Residential child and youth care involve living and learning environments that operate with 24 h, life space activity-based care and education—whether on-site or nearby, 7 days a week for designated periods of time measured by cohort, semester, term, season or year" [3]. This system is only best described by Browne [4] "residential care implies an organized, routine and impersonal structure to the living arrangements for children (e.g. all children sleep, eat and toilet at the same time) and a professional relationship, rather than parental relationship, between the adults and children". Malik, R. et al., substantiated the approach followed as "more administration centred then child centred". Thus, a machinery is created that breeds cohorts of troubled beings, at the mercy of the very system that vows to be a saviour for them.

Children who are a part of this infrastructure face severe problems in their holistic wellbeing. They tend to be emotionally withdrawn and experience loneliness [5, 6]. Padmaja, G. et al. explained the extremely poor levels of physical health on a comparative basis, displaying how institutionalization affects every aspect of living [7]. Multiple studies show the prevalence of hyperactivity well into adulthoods [8, 9]. Thus, this kind of care has high levels of psycho-social, emotional problems. Problems that affect on a daily as well as prolonged basis.

Reintegration of the child into mainstream society is difficult. The child does not feel that he is equipped to be able to be a part of the society. A mindset that is bereft with stigma and alienation towards the outside world, [7] and a skillset ill-equipped to carve a career ensues that this spiral of stigma, distress and trauma continues.

2.3 How Do We Walk Towards a Better Tomorrow?

Islam, T. et al. said that there is a lack of attention dedicated to the different psycho-social, intellectual, cultural, emotional aspects of the child development [3]. As is correctly identified by Malik, R. et al. (Residential care of orphans in Indian Kashmir; Residential child and youth care in a developing world 2018), there is a need for an overhaul in the childcare domain that is rooted in different aspects of wellbeing, and not merely provisions of infrastructure and residential services [3].

3 The Systemic Approach

It is clearly evident that a child rights-based perspective is necessary to be adopted. We have proposed a framework that enlists the high-level goals that need to be achieved and then supplement it with methodologies to be implemented on the ground.

3.1 High-Level Goals

Systems Approach. A systems approach is followed in order to weave together the different parts of the childcare delivery stakeholders together to work towards achieving inclusive, sustainable childcare mechanisms [2]. They further say that how creating such an approach would be able to implement comprehensive mechanisms to build protected environments, hence being able to address issues of the most vulnerable groups [2].

Child-Rights Approach. The soul of the whole structure works towards achieving an exaggerated, and vehement focuses towards individual child development. It is the soul of the framework and employs the need to work and design everything with children in mind.

Resilience-based Approach. Structuring the psyche of the children to push them towards believing and actualizing their potentials and aspirations. It is necessary to provide them with the resources and measures that would make them realize the need to get out of the spirals of risk and pain [7]. It is necessary to identify and engage with the children on an individual level that leads to resolving conflicts and a nudge towards betterment and improvement. Such an approach is important to push and provide every child with the resources to own up to their lives.

Tailor-Made Mechanisms. It is important to note that longing for a blanket solution that is able "deinstitutionalize" is not the answer. Throughout, we have talked about providing power to the individual, in order to be able to tackle institutionalization. As discussed by Padmaja, G. et al., it is imperative "to identify unmet psycho-social needs and enhancing the strengths, while helping the children gain control over the identified weaknesses" [7]. The tailor-made approach is to be integrated with a curriculum that focuses towards providing children mechanisms allowing them to build their identities around things that they have an aptitude for, and are good at, and assure them of their upbringing.

Empowering Caregivers. Empowering caregivers, like social workers, helpers, teachers with digital aid, would ensure an effective care relationship between the kid and them and creation of a fallback mechanism. As identified, social connections "play a crucial role in developing or reducing psychological problems" [7].

3.2 Implementation on the Ground

ICT-based infrastructure. Digitization and building up of an ICT infrastructure is extremely necessary to provide robust mechanisms to ensure implementation of education mechanisms. Integrating learning management systems with learning

analytic tools to empower caregivers to gauge growth and development of the child. Multisensory environments that provide support to empower themselves.

Interoperability Measures. The change that we propose to bring is only possible through following collaborative efforts aimed towards a common goal. The departments should be able to follow a free exchange of data, practices that can and have affected childcare.

Oversight and Accountability. There is a very high necessity of creating feedback mechanisms and grievance redressal structures in order to create transparent environments. The systems should have accountability measures to oversee the functioning of different plans and curriculums.

4 Proposed Solution Model

The following portion examines how the above-explained methodologies and practices could be tied together and implemented on the ground level.

4.1 Educational Curriculums Based on Traditional and Cultural Roots

To preserve the culture of the region, the students should also be taught the traditional art or other practices, in school, and inculcates, in each student, the values of their tribe. The school should also teach the children to uphold these virtues, live up to them and to promote them even outside their tribe, wherever they go.

The idea is to not only provide normal schooling to the children, but also make them culturally comfortable without feeling "left out". The schools should stress on the traditional values along with their academic curriculum. The essence of the tribe should not be lost in the hustle of studying at a school. It is important to follow these practices in order to look out for the bests for the children.

4.2 Mechanisms to Increase Visibility and Identity Building

Apart from a good academic module to brush up their cognitive skills, students should also be exposed to sports, the basics of finance, arts and so on, which will help enhance the personality of each child and improve their development. Such activities will also act as stress busters and open floodgates of opportunities for those students who do not ace in the classroom.

4.3 Volunteer Distant eLearning Programs

As we have discussed earlier, it is an imperative that the education and learning that the kids receive is the backbone of them having lives that are of value. What we propose is a portal where volunteers, or people who want to contribute towards the education of the kids can develop lectures, modules and teach, at their own discretion.

The idea here being the kids would be able to experience education that is in touch with the realities of the world out there.

4.4 Integrated Learning and Talent Management Systems

The technical backbone of this system rests on creating an integrated learning and talent management system. This is the common thread that runs through the different parts of the systems we have talked about earlier. The talent management system is used to, as written earlier, monitor, understand and evaluate the growth of the child throughout his entire journey.

Evaluation of aptitude and DNA analysis of the kids is done. The profile of the child is developed and updated after different psychometric and personality assessments. The learning management portal is used to maintain a singular database of the development and growth of the child, and how he/she has progressed. It includes learning analytics that should be used by the teachers, caregivers and volunteer faculties. We could employ these to monitor and plan education delivery according to the needs of the kids.

5 Reception and Results

We conveyed the solution developed to the following organizations. We prepared a questionnaire to help us evaluate how the system would fare in a real-world context. Throughout the course of the study, we had kept them in the loop for assessing the problem through contextual inquiry. The general response was fairly positive, with a few recommendations and advice.

1. Punarutthan-Samarsata-Gurukulam, Chinchwadgaon, Maharashtra [10].
2. Nachiket-Balgram, Ravet, Pune, Maharashtra [11].
3. Schools run by the Krantiveer Chaphekar Smarak Samitee, Chinchwad, Pune, Maharashtra [12].
4. Vanavasi Kalyan Ashram, Jashpurnagar, Chhattisgarh [13].

The results are presented here (Figs. 1, 2 and 3).

The Punarutthan-Samarsata-Gurukulm, Vanvasi Kalyan Ashram and Krantiveer Chapekar Samiti all are headed by the parent organization of the Rashtriya Swayamsevak Sangh (RSS) [14]. They showed a considerable interest towards implementing

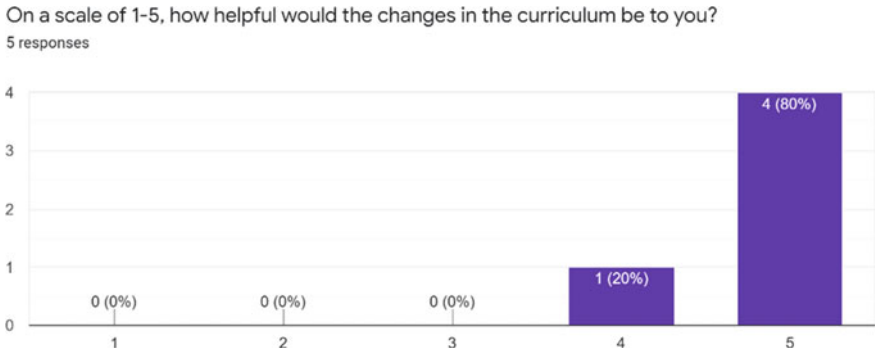


Fig. 1 Following graph depicts how useful the organizations find the changes recommended by the framework. The changes included are the tailor-made curriculums based on traditional and cultural roots

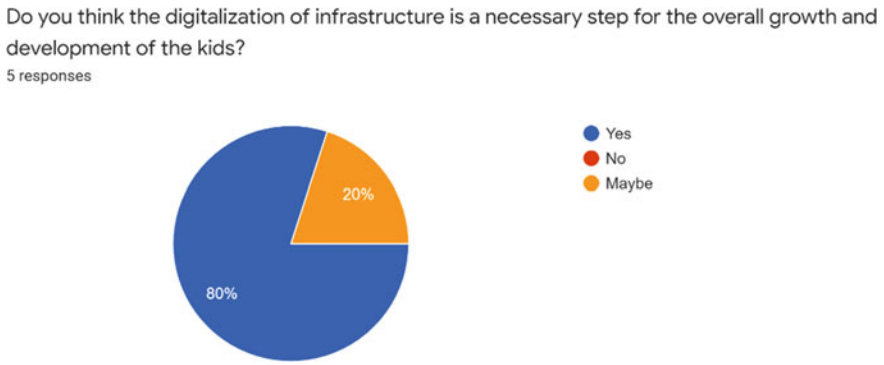


Fig. 2 Considerable amount of organizations surveyed agree that the introduction and employment of ICT-tools is necessary to implement the changes

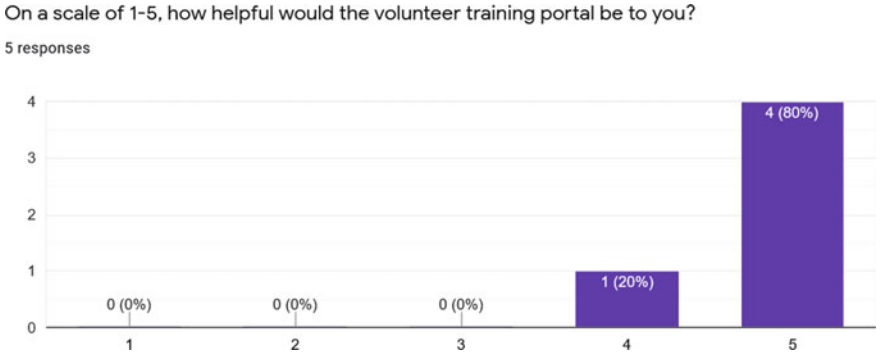


Fig. 3 Considerable number of organizations finds that the volunteer training portal would be helpful to very helpful

the solutions. We are in the process of working towards implementing the solution first, in the state of Maharashtra. Depending on the results, towards all the schools and institutions under the parent body of the RSS, throughout the nation. Their response, as shown before, was overwhelmingly positive.

Some concerns were expressed in the implementation. Another was the language of communication to be used. We were encouraged to work out the details and test out. We are ready to take the next step and go the last mile.

6 Conclusion

We have tried our best to identify, map and provide frameworks to approach problems relating to the Indian residential childcare environment. It is observed that a child rights-based approach has the power to subvert the mechanisms of institutionalization. What we have presented is a framework that could be implemented on the systems level. The real potential for the framework lies in HCI led interventions, designer-led solutions, social engineers, social workers and caregivers. We have mentioned a small model that could be one of the right steps towards the problem. Further, we have clearly demarcated and displayed the full possibilities of the systemic framework. This could act as a primary guideline while working on the problems of this nature. As we have said, the focus should be on the individual, and their place in the system. The goal is to provide happy childhoods to these vulnerable individuals. This acts as our step towards the same.

References

1. Census of India: Office of the Registrar General & Census Commissioner, India. Censusindia.gov.in. https://censusindia.gov.in/2011-common/ (2011)
2. Naaz, S., Meenai, Z.: Alternative care in India: issues and prospects. Rajagiri J. Soc. Dev. **11** (2019)
3. Islam, T., Fulcher, L.: Residential child and youth care in a developing world. Pretext Publishing, Cape Town, South Africa (2018)
4. Browne, K.: The risk of harm to young children in institutional care. The Save Children Fund, London. References—Scientific Research Publishing. Scirp.org, 2020. https://www.scirp.org/reference/ReferencesPapers.aspx?ReferenceID=1568787 (2009)
5. Han, E., Choi, N.: Korean institutionalized adolescents' attributions of success and failure in interpersonal relations and perceived loneliness. Child. Youth Serv. Rev. **28**(5), 535–547 (2006)
6. Zeanah, C., Smyke, A., Koga, S., Carlson, E.: Attachment in institutionalized and community children in romania. Child Dev. **76**(5), 1015–1028 (2005)
7. Padmaja, G., Sushma, B., Agarwal, S.: Psychosocial problems and wellbeing in institutionalized and non-institutionalized children. IOSR J. Human. Soc. Sci. **19**(10), 59–64 (2014)
8. Egelund, T., Lausten, M.: Prevalence of mental health problems among children placed in out-of-home care in Denmark. Child Fam. Soc. Work **14**(2), 156–165 (2009)
9. Wiik, K., Loman, M., Van Ryzin, M., et al.: Behavioral and emotional symptoms of post-institutionalized children in middle childhood. J. Child Psychol. Psychiatr. **52**(1), 56–63 (2010)

10. Punarutthan Samarasta Gurukulam: https://gurukulamonline.org/
11. Nachiket Balgram: https://nachiketbalgramsite.wordpress.com/
12. New Page 1: https://krantiveerchapekar.org
13. Akhil Bhartiya Kalyan Ashram: https://kalyanashram.org/
14. Rashtriya Swayamsevak Sangh: https://www.rss.org/

A Comprehensive Survey on Handwritten Gujarati Character and Its Modifier Recognition Methods

Priyank D. Doshi and Pratik A. Vanjara

Abstract In India, handwritten character recognition is becoming necessity region-alwise due to new education policy 2020. Various technologies are applied to solve the problem in this area like statistical or probability model, support vector machine, Bayes probability model, deterministic finite automaton (DFA), hidden Markov model, and many more which are used. Due to the advancement in machine learning, convolutional neural network is a good solution of HCR which gives more promising results but any new algorithm in machine learning that depends on training data, mathematical function, loss function, and method of evaluation of model. Focusing on past research of handwritten Gujarati character recognition is found that sufficient research is required for modifier level called "Barakshari". Results obtained in past are limited to character level only. In this paper, our effort is to analyze and summarize previous contributions in the handwritten character recognition for several Indian languages.

Keywords Support vector machine · Bayes probability model · Deterministic finite automaton · Hidden Markov model · Convolutional neural network

1 Introduction

The regional languages processing is still in the growing phase in India where various approaches and techniques contributed by the researchers and many more required. Internet is widely used as a platform for multiple Indian languages in the area of education, e-commerce, documentation, information sharing, etc. Therefore, regional language processing is becoming a challenge day by day. In this context, demand of offline handwritten character recognition research for Indian regional languages is rising rapidly. Focusing on Gujarati and other Indian languages problem

P. D. Doshi (✉)
Atmiya University, Rajkot, Gujarat 360005, India

P. A. Vanjara
Shree M. & N. Virani Science College, Rajkot, Gujarat 360005, India

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_79

of handwritten character recognition in the previous work is reviewed in this paper, and the challenges to solve Gujarati character recognition and its similar works in other languages are discussed. In this paper, we included HCR and OCR methods for Indian and foreign Languages to promote the work of Gujarati language handwritten character and vowels recognition. Offline handwritten character recognition for Indian regional languages is very much beneficial in number of ways.

Possible Advantages. Data should be taken from source so subsequently it avoids errors and unwanted changes which can be generated by retyping. Additionally, National Education Policy 2020 of India that depicts the medium of instruction will be student's mother tongue or regional language at least up to Grade 5. So Gujarati handwritten character recognition with vowels (character Modifiers) can be more advantageous to students and teachers. HCR can digitize old handwritten documents, postal address, and many more improvisation needed till it becomes easy and popular to use.

This paper is organized in five sections starting with general model (Table 1) which has mainly six different stages of handwritten character recognition given as

Table 1 Stages and task used in text recognition system

Stage	(Tasks)—description
Image acquirement	(Digitalization resizing, compression)-all images are converted into required format and size to prepare dataset
Preprocessing	(Noise removal, filtering, skew, thinning, edge detection and correction, morphological operation)-processing images to remove unwanted background to get region of interest. Image used to detect canny edge, bounded box, etc. Skewness detected and corrected
Segmentation	(Character based, word based, sentence based)-from the image, we required to separate lines with upper and lower boundaries. Then words are separated and then characters. A character made up of one or more connected components of pixels. Pixels intensity value can be used
Feature extraction	(Statistical—geometrical features)-curves and corners of connected components of pixels can be found to form characters. This is the area which is helpful in next stage of classification
Character recognition (classification)	The segmented content is fed to the classifier which gives result that input image corresponds to which character and modifiers
Post-processing	(Confusion matrix, contextual approaches dictionary-based approaches)—grouping characters and modifiers based on their location. Finding errors and correction. One to one correspondence between characters and unicode dictionary letters. Output obtained as a text file

under. Subsequently, it includes challenges motivation, efforts reviewed areas of their work, discussion findings, and conclusion in the last.

1.1 General Methodology

Scanned handwritten images are processed to get text file in target language. As shown in Table 1, processing of it divided into the phases like image acquirement, preprocessing, segmentation, feature extraction, classification, and post-processing.

2 Challenges and Motivation

Gujarati language character modifier recognition is required more efforts as it would be one step ahead since character recognition almost done so vowels recognition requires more research work. Upper line as it is in Devanagari language is missing in Gujarati, and many half characters are joined with another which is displayed differently. Skewed characters and modifiers are also giving variety to dataset. A character connected with another half character (conjunct consonants) shown as under and in Table 2 (counting them we get 12 so it is called "Barakshari") is also a challenging task. Following type of characters is varied very frequently person to person when written by their hands so it becomes very difficult to recognize by machine [1].

ક	ખ્ર	ગ્ર	ઘ્ર	ચ્ર	જ્ઞ
ચ્ય	જ્ય	ખ્ય	ઘ્ય	ચ્ય	જ્ઞ્ય
હ્ય	ધ્ય	ત્ત	હ્ન	ષ્મ	શ્ચ

Basically word formation is not possible without character modifiers. They are vowels attached with consonants. Even if character can be recognized, its combination with vowels (as shown in first column of Table 2) becomes difficult to recognize. Sometimes characters written by person may be vertically straight, forward slant, or backward slant. So it produces difficulty in character recognition.

3 Efforts Reviewed and Their Areas of Work

Various surveys in this context lead us to understand various [2] merits and demerits of techniques involved in each stage of text recognition. Language processing techniques [3] like chatbot, text-to-speech conversion, language identification, spell

Table 2 Basic shapes of geometry useful to recognize character modifier (vowels)

Character modifier	Detail of basic geometrical shapes observed in consonant modifier
અ–ક	No modifier present
આ–કા	Vertical line in right zone
ઈ – કિ	Vertical line in left zone, oval arc or circular arc in upper zone, connected or disconnected components
ઈ – કી	Vertical line in right zone, oval arc or circular arc in upper zone, connected or disconnected components
ઉ - કુ	Oval arc or circular arc in lower zone, connected components
ઊ - કૂ	Slanting line in lower zone, oval arc or circular arc in lower zone, connected components
એ - કે	Slanting line in upper zone
ઐ - કૈ	Two slanting line upper zone
ઓ - કો	Vertical line in right zone, slanting line in upper zone, disconnected components
ઔ - કૌ	Vertical line in right zone, Two Slanting line in upper zone, disconnected Components
અં - કં	One dot in upper zone
અઃ - કઃ	Two disconnected dots in right zone

check, medical record processing for real-time needs focused. We also found conclusion that [4] the selection of the classification as well as the feature extraction techniques needs to be proper in order to attain good rate in recognizing the character. Segmentation work in Gujarati character modifier still remains left. Over fitting problem in deep learning occurs when result is good fit on our model on the training data, but it is not generalized well enough on new data. It means model is very specific to the training data and inappropriate for other data. We can solve over fitting problem by adding more training data. So in deep learning [5] for image data augmentation result found that over fitting problem can overcome by improving the size and quality of training datasets. Augmentation like flipping, rotation, and zoom in and zoom out can be used to enrich the dataset in terms of size and quality. Comparatively it is found that offline character recognition is high accuracy and reliable system required [6], and also fuzzy membership functions could significantly outperform standard zoning methods [7]. For general stage of text processing having different techniques especially in classification stage including neural network, it is found that feature extraction and classification technique play an important role [8]

in the performance of character recognition rate and including neural network and data mining concepts merits and demerits [9] which are given for each stage.

3.1 Segmentation and Preprocessing

As there are six steps in general model of offline character recognition, and in each step, variety of scientific techniques of models can be applied segmentation and preprocessing can include histogram [10] projection and equalization technique [11], sliding window method, Hough transform technique [12] with preprocessing like dilation, erosion, and thinning. Also strip-based projection [13] and mixture of smearing and contour tracing for line segmentation are used in offline Gurumukhi language character recognition. Research has been done for segmenting modifiers as intact not subdividing further [14] from printed Bangla Text. They continued their work for overlapping, touching, and compound characters. Segmenting the upper and lower modifiers with characters is done by fuzzy functions in Hindi language [15].

3.2 Neural Network Used as a Classifier

Nowadays using deep learning [16] offline handwritten recognition system proposed based on convolutional neural networks (CNN) for Gujarati language with different accuracy and diversified data. Similarly by analyzing character's shape [17] and then neural network applied to recognize characters. For English language [18] applying neural network and KBS built for character recognition. Using ANN [19] an online multilanguage handwriting recognition system developed based on Bézier curves for 102 language including Gujarati using IAM-OnDB dataset. Paired adversarial learning (PAL) [20] method along with DNN is applied to recognize handwritten mathematical expression using CNN and RNN-based feature extractor.

Similarly for digit recognition [21] CNN and pooling layer for data reduction and fully connected layer and output layer applied as a classifier. For Urdu [22], HCR and digit recognition pioneer dataset were prepared with CNN. For HCR in English language shallow network based on the Fukunaga Koontz transform (FKT) [23] model used with neural network and to recognize characters from ancient Geez document [24] a deep CNN used with dataset. A model for English character recognition developed using DNN as classifier [25] and feature extraction by local Gabor pattern, Haar wavelet transform, histogram-oriented gradients, and grid level. Similarly, chain code and image centroid extraction method use along with feedforward ANN [26] as a classifier.

For classification of handwritten Bengali numerals [27], a model proposed with extra layers like zero padding, dropout, and max-pooling, and the number of filters

enhanced using CNN and for Gujarati numeral recognition [28] and multilayer feed-forward neural network and naive Bayes classifier used for handwritten Gurumukhi numeral [29] using backpropagation neural network with wavelength transform.

3.3 Support Vector Machine as a Classifier (SVM)

Using SVM [30–34] as a classifier many research work recorded for different languages like Gujarati and Devanagari. Along with SVM hybrid feature spaces like aspect ratio, image subdivision, polynomial kernel, Gaussian kernel are used. Also for Gujarati numeral recognition creating four features set of various size preprocessing and segmentation has been done. Distance profile, gradient profile, and wavelet form are used. Fourier descriptors as feature vectors and lexicon were used for post-processing.

3.4 K-Nearest Neighbor as a Classifier

K-nearest neighbor classifier [35] with distance transform method for zone identification and for segmentation projection profile and morphological operations used. Comparative analysis between KNN and principal component analysis (PCA) for Gujarati numerical classifier [2] developed. Noise removal, and thinning as preprocessing and stroke-based directional feature [36] also applied with this classifier. With this [37], normalization and interpolation used in preprocessing.

3.5 Other Method Used as Classifier

Decision tree approach for HCR for Gujarati characters [38] focuses on feature extraction of three types, i.e., connected or disconnected component, number of end point and number of closed loop. Similar concept like adjacent pixel connectivity and curvature-based pattern matching and classification used [39] in HCR method for Sinhala language.

For Gujarati language combined approach of Freeman Chain Code [40], Hu' Invariant Moment (4 order), center of mass applied with gradient feature. A deterministic finite automaton (DFA) [41] and fuzzy system [42] for intelligent word recognition also introduced in past using a regular grammar by labeling each pixel as vertical or as horizontal to group strokes for feature extraction.

A model using weighted integral image method, Bayes classifier [43] and statistical-structural features overcoming drawbacks of classic binarization method. For machine printed, Devanagari (Nepali) has been introduced in past.

Maximum mutual information [44], composite image and block-based PCA methods for character recognition are used in HCR. For English, writer independent character recognition HMM [45] used for each character with global and local processing features of images. For upper and lower modifiers and half characters are ignored [46] and topological features, heuristics for middle zone are used.

3.6 Dataset Creation

Including Gujarati language work noticed in dataset creation like page level handwritten document image dataset "PHDIndic_11" [47] of 11 official Indic scripts. Dataset for English language alphabets and numerals [48] is publically available and provides isolated characters and digits free of cost. A novel database [49] is consisting of 26,000 images of Hindi handwritten characters, and modifiers for offline recognition by segmentation and augmentation process were developed. Such dataset creation for Gujarati language is required to create for effective implementation HCR system.

4 Discussions and Findings

Due to emerging trends of machine learning, especially convolutional neural network can solve this problem more promisingly. As we can categorize machine learning algorithm in supervised or unsupervised in terms of label is present or absent. Label corresponds to prediction (output) which is based on features in training dataset. Further depending on labels, our problem of HCR lies in discrete classification problem. Along with sufficient training dataset, any new algorithm in machine learning depends on mathematical function like logistic, linear regression, loss function like mean square error (MSE), mean absolute error (MAE), and other classifications losses between predictions and actual observations. Methods of model evaluation also affect on this. So possibilities of new directions and improvements are always expected.

Unavailability or scarcity of dataset in HCR for Gujarati character modifiers (vowels) recognition is one of the necessary areas of work. Here we need to apply a holistic approach in which all stages should be integrated such a way that they should contribute effectively for central objective of the system. So successful integration of all phases of machine learning starting with features extraction (topological or geometrical) to classification is very important. Vowels sometimes are extremely cursive or artistic in nature. Using convolutional neural network, we can extract shapes (Table 2) by feature extraction and co-relate it with character to print and fed to classifier.

5 Conclusions

Offline handwritten Gujarati character modifier recognition ("Barakshari") is an interesting and challenging area of research. There are several classification and feature extraction techniques for handwritten character recognition problem. Various techniques can be applicable to complete the task in each steps of general model of recognition. Due to advancement in technology, machine learning attracts us toward selection of appropriate neural network model and it will be more effective if we can include good training dataset in case of Gujarati character modifier recognition. Any new algorithm in machine learning depends on how model has been trained with batch or mini batchwise of data, type of model or mathematical function has been used, its loss function, and method of evaluation. We can adjoin the effectiveness of each steps of model with good feature extraction of characters and modifiers using good augmentation techniques.

References

1. Patel, C., Desai, A.: Zone identification for Gujarati handwritten word. In: Proceedings of the 2nd International Conference on Emerging Applications of Information Technology. EAIT 2011, pp. 194–197 (2011)
2. Mj, B., Kv, K., Me, J.: Comparison of classifiers for gujarati numeral recognition. Int. J. Mach. Intell. **3**, 160–163 (2011)
3. Harish, B.S., Rangan, R.K.: A comprehensive survey on Indian regional language processing. SN Appl. Sci. **2** (2020)
4. Purohit, A., Chauhan, S.S.: A literature survey on handwritten character recognition. IARJSET **7**, 1–5 (2016)
5. Shorten, C., Khoshgoftaar, T.M.: A survey on image data augmentation for deep learning. J. Big Data **6** (2019)
6. Priya, A., Mishra, S., Raj, S., Mandal, S., Datta, S.: Online and offline character recognition: a survey. In: International Conference on Communication and Signal Processing. ICCSP 2016, pp. 967–970 (2016)
7. Impedovo, D., Pirlo, G.: Zoning methods for handwritten character recognition: a survey. Pattern Recognit. **47**, 969–981 (2014)
8. Sahu, V.L., Kubde, B.: Offline handwritten character recognition techniques using neural network: a review. Int. J. Sci. Eng. Res. **1**, 87–94 (2013)
9. Muthuraman, V.: A study on text recognition using image processing with datamining techniques. Int. J. Comput. Sci. Eng. Open Access (2019)
10. Dave, N.: Segmentation methods for hand written character recognition. Int. J. Signal Process. Image Process. Pattern Recognit. **8**, 155–164 (2015)
11. Dixit, S., Suresh, H.N.: South Indian Tamil language handwritten document text line segmentation technique with aid of sliding window and skewing operations. J. Theor. Appl. Inf. Technol. **58**, 430–439 (2013)
12. Shah, L. et al.: Rotation estimation of Gujarati script document using hough transform (2014)
13. Kumar, M., Jindal, M.K., Sharma, R.K.: A novel technique for line segmentation in offline handwritten Gurmukhi script documents. Natl. Acad. Sci. Lett. **40**, 273–277 (2017)
14. Akter, N., Hossain, S., Islam, M.T., Sarwar, H.: An algorithm for segmenting modifiers from Bangla text. In: Proceedings of the 11th International Conference on Computer and Information Technology. ICCIT 2008, pp. 177–182 (2008)

15. Pramanik, R., Bag, S., Kumar, R.: A fuzzy and contour-based segmentation methodology for handwritten Hindi words in legal documents. In: Proceedings of the 4th IEEE International Conference on Recent Advances in Information Technology. RAIT 2018, pp. 1–6 (2018)
16. Pareek, J., Singhania, D., Kumari, R.R., Purohit, S.: Gujarati handwritten character recognition from text images. Proc. Comput. Sci. **171**, 514–523 (2020)
17. Prasad, J.R., Kulkarni, U.V., Prasad, R.S.: Offline handwritten character recognition of Gujrati script using pattern matching. In: 2009 3rd International Conference on Anti-Counterfeiting, Security, and Identification in Communication. ASID 2009 (2009)
18. Kasthuri, M., Sivaprasatham, V.: Self-learning based cognitive reading and character recognition in image processing techniques. SN Comput. Sci. **1**, 1–12 (2020)
19. Carbune, V., et al.: Fast multi-language LSTM-based online handwriting recognition. Int. J. Doc. Anal. Recognit. **23**, 89–102 (2020)
20. Wu, J.W., Yin, F., Zhang, Y.M., Zhang, X.Y., Liu, C.L.: Handwritten mathematical expression recognition via paired adversarial learning. Int. J. Comput. Vis. (2020)
21. Ali, S., et al.: An efficient and improved scheme for handwritten digit recognition based on convolutional neural network. SN Appl. Sci. **1**, 1–9 (2019)
22. Ali, H., Ullah, A., Iqbal, T., Khattak, S.: Pioneer dataset and automatic recognition of Urdu handwritten characters using a deep autoencoder and convolutional neural network. SN Appl. Sci. **2**, 1–12 (2020)
23. Gatto, B.B., dos Santos, E.M., Fukui, K., Júnior, W.S.S., dos Santos, K.V.: Fukunaga–Koontz convolutional network with applications on character classification. Neural Process. Lett. (2020)
24. Demilew, F.A., Sekeroglu, B.: Ancient Geez script recognition using deep learning. SN Appl. Sci. **1**, 1–7 (2019)
25. Liu, Z., Pan, X., Peng, Y.: Character recognition algorithm based on fusion probability model and deep learning. Comput. J. **00** (2020)
26. John, J., Pramod, K.V., Balakrishnan, K.: Offline handwritten Malayalam character recognition based on chain code histogram. In: 2011 International Conference on Emerging Trends in Electrical and Computer Technology. ICETECT 2011, pp. 736–741 (2011)
27. Rahman, M.M., Islam, M.S., Sassi, R., Aktaruzzaman, M.: Convolutional neural networks performance comparison for handwritten Bengali numerals recognition. SN Appl. Sci. **1**, 1–11 (2019)
28. Sharma, A., Thakkar, P., Adhyaru, D., Zaveri, T.: Features fusion based approach for handwritten Gujarati character recognition. Nirma Univ. J. Eng. Technol. (2017)
29. Singh, P., Budhiraja, S.: Offline handwritten Gurmukhi numeral recognition using wavelet transforms. Int. J. Mod. Educ. Comput. Sci. **4**, 34–39 (2012)
30. Desai, A.A.: Support vector machine for identification of handwritten Gujarati alphabets using hybrid feature space. CSI Trans. ICT **2**, 235–241 (2015)
31. Farkya, S., Surampudi, G., Kothari, A.: Hindi speech synthesis by concatenation of recognized hand written Devnagri script using support vector machines classifier. In: 2015 International Conference on Communication and Signal Processing. ICCSP 2015, pp. 893–898 (2015)
32. Gupta, A., Srivastava, M., Mahanta, C.: Offline handwritten character recognition using neural network. In: ICCAIE 2011—2011 IEEE Conference on Computer Applications and Industrial Electronics, pp. 102–107 (2011)
33. Maloo, M., Kale, K.V.: Support vector machine based Gujarati numeral recognition. Int. J. Comput. Sci. Eng. (IJCSE) **3**, 2595–2600 (2011). ISSN 0975-3397
34. Nagar, R., Mitra, S.K.: Feature extraction based on stroke orientation estimation technique for handwritten numeral. In: ICAPR 2015—2015 8th International Conference on Advances in Pattern Recognition (2015)
35. Desai, A.A.: Gujarati handwritten numeral optical character reorganization through neural network. Pattern Recognit. **43**, 2582–2589 (2010)
36. Goswami, M., Mitra, S.: Structural feature based classification of printed Gujarati characters. In: Lecture Notes in Computer Science (including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). LNCS, vol. 8251, pp. 82–87 (2013)

37. Gohel, C.C., Goswami, M.M., Prajapati, V.K.: On-line handwritten Gujarati character recognition using low level stroke. In: Proceedings of the 2015 3rd International Conference on Image Information Processing. ICIIP 2015, pp. 130–134 (2016)
38. Thaker, H.R., Kumbharana, C.K.: Structural feature extraction to recognize some of the offline isolated handwritten Gujarati characters using decision tree classifier. Int. J. Comput. Appl. **99**, 46–50 (2014)
39. Madushanka, P.T.C., Bandara, R., Ranathunga, L.: Sinhala handwritten character recognition by using enhanced thinning and curvature histogram based method. In: 2017 IEEE 2nd International Conference on Signal Image Processing. ICSIP 2017, Jan 2017, pp. 46–50 (2017)
40. Macwan, S.J., Vyas, A.N.: Classification of offline Gujarati handwritten characters. In: 2015 International Conference on Advances in Computing, Communications and Informatics. ICACCI 2015, pp. 1535–1541 (2015)
41. Álvarez, D., Fernández, R., Sánchez, L.: Stroke-based intelligent character recognition using a deterministic finite automaton. Log. J. IGPL **23**, 463–471 (2014)
42. Álvarez, D., Fernández, R.A., Sánchez, L.: Fuzzy system for intelligent word recognition using a regular grammar. J. Appl. Log. **24**, 45–53 (2017)
43. Joshi, V., Panday, S.P.: Character component segmentation and categorization of machine printed text in Devanagari (Nepali) script in digital image processing. In: Proceedings of the 2018 IEEE 3rd International Conference on Computing, Communication and Security. ICCCS 2018, pp. 191–198 (2018)
44. Nopsuwanchai, R., Povey, D.: Discriminative training for HMM-based offline handwritten 2. In: Maximum Mutual Information Training of Analysis (2003)
45. Das, R.L., Binod, I., Prasad, K., Sanyal, G.: HMM based offline handwritten writer independent english character recognition using global and local feature extraction. Int. J. Comput. Appl. **46**, 975–8887 (2012)
46. Garg, N.K., Kaur, L., Jndal, M.: Recognition of offline handwritten hindi text using middle zone of the words. In: 2015 IEEE/ACIS 14th International Conference on Computer and Information Science. ICIS 2015—Proceedings, pp. 325–328 (2015)
47. Obaidullah, S.M., Halder, C., Santosh, K.C., Das, N., Roy, K.: PHDIndic_11: page-level handwritten document image dataset of 11 official Indic scripts for script identification. Multimed. Tools Appl. **77**, 1643–1678 (2018)
48. Yousaf, A., Khan, M.J., Imran, M., Khurshid, K.: Benchmark dataset for offline handwritten character recognition. Proceedings of the 2017 13th International Conference on Emerging Technologies. ICET2017, Jan 2018, pp. 1–5 (2018)
49. Nehra, M.S., Nain, N., Ahmed, M.: Benchmarking of text segmentation in devnagari handwritten document. In: 2016 IEEE 7th Power India International Conference. PIICON 2016, pp. 0–3 (2017)

Thai Stock Price Prediction from Daily News Contents

Pavit Wattanakul and Tanasanee Phienthrakul

Abstract Stock market is a major place for investment. Many investors would like to accurately predict the stock price in order to get more profit. Stock price prediction by machine learning techniques is widely used according to an uncertainty movement of the stock price. However, this should be better if the related information is combined. Daily news is a source of information that reflects many factors. Both good and bad factors are indicated in the news. In this research, the news is used to forecast the stock price. Text processing is applied to extract word features. Then, word features are analyzed and selected. The selected features are used in prediction process. Three predictive models, random forest (RF), AdaBoost and support vector regression (SVR), are compared. SVR yields the best result with the lowest root mean square error (RMSE). The result shows that on 50 features from Spearman ranking, SVR is able to reduce the root mean-squared error by 8.63%.

Keywords Stock price prediction · Daily news · Feature ranking · AdaBoost · Random forest · Support vector regression

1 Introduction

Stock market is one of the most popular investments in the world. There are two major styles to invest in the stock market, which are technical and fundamental analysis styled. Technical analysis will focus on the price, indicator and graph. For the fundamental analysis, the company profit and company performance will be focused on. Although technical analysis can make more profit for many investors, there are also more risk to them. Only stock price may not be enough to give important information in the analysis. Daily news is a source of information that affects to the

P. Wattanakul (✉) · T. Phienthrakul
Department of Computer Engineering, Faculty of Engineering, Mahidol University, 25/25 Phutthamonthon 4 Rd., Salaya, Nakhon Pathom 73170, Thailand

T. Phienthrakul
e-mail: tanasanee.phi@mahidol.ac.th

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_80

stock price. The bad news can make the decreasing in stock price, while as the good news may increase the price of stock in some sectors.

There are many researches that attempted to predict the stock price using machine learning. In a research of Kim [1], he applied support vector machine using technical indicator and the stock price movement to predict the movement of price. Polynomial and Gaussian kernel were tested. Then he found that the support vector machine was suitable for time series prediction and providing the high performance. Moreover, Nahil et al. [2] proposed support vector regression with the new kernel, which was called kernel principal component analysis (KPCA) using stock price as the input. The result showed that the performance of KPCA with support vector regression is better than the ordinary support vector regression.

Khaidem et al. [3] applied the random forest model which is an ensemble model to predict the stock movement by using technical indicator which is relative strength index (RSI) and stochastic oscillator (SO) as the input. The result shows that random forest model outperforms the other models, i.e., artificial neural network and logistic regression. Basak et al. [4] attempted to apply the indicator, which is RSI, SO, W%R, MACD, PROC and OBV to predict the stock price by random forest model, and it can overcome the other models, which is XGBoost, logistic regression, support vector machine and neural network model.

Kim and Han [5] proposed the enhanced random forest to find the stock movement from the closed price. Their model was adding more weight to the closed price when the closed price has significant changed. The case of the closed price was little in changing, and then weighting is small. The enhanced random forest performance was better than ordinary random forest model. AdaBoost model was one the ensemble methods by training the weak learner to improve the performance. Chang et al. [6] presented an improvement of AdaBoost model, which was joining the RSI indicator and average true range (ATR). Their technique was provided a remarkable performance. To improve the model efficiency, feature selection technique was applied. Huang et al. [7] proposed the prediction with feature selection by wrapper method and using 23 technical indicators as the input. The result showed that wrapper method can improve the prediction score of neural networks, support vector machine, *K*-nearest neighbor and decision tree models.

Regarding to the uncertainty movement of the stock price in the stock market, which has many factors affected to. One of the most affected factors is the news. When bad news comes, most of the stock price will be crumbled down. Nassirtoussi et al. [8] proposed a new technique to predict the stock movement from the news by using term frequency-inverse document frequency (TF-IDF) as a feature extraction and multilayer feature selection, which are heuristic–hypernyms feature selection, sentiment integration layer and synchronous targeted feature reduction. The result of the model was highly appreciated.

This research aims to predict the stock price by using TF-IDF as a feature extraction and find a model that can predict the stock price based on news with a low root mean-squared error. Next section will describe the background of this research, which is about feature selection method and predictive model analysis method. Then, the

research step and experimentation will be illustrated in Sects. 3 and 4, respectively. Lastly, the conclusion will be presented in Sect. 5.

2 Background

2.1 Feature Selection

An enormous of the vector is produced after the feature extraction method. Some variable is significant to the result but some not. In order to improve the prediction performance, some variables need to be cutting off. There are many techniques of the feature selection but in this research, it is focusing only two methods of the feature selection, which is expressed as below.

Spearman Ranking ($\boldsymbol{R_s}$). In order to reduce the number of variables by feature selection, correlation statistic is applied to handle it. Correlation is bivariate analysis. One of the best correlation techniques is Spearman rank correlation [9], which is used to find the relationship between two variables. This technique is called a filter method. In order to apply, the two variables need to be the independence. The formula can be expressed as below.

$$R_s = 1 - \frac{6\sum D^2}{N\left(N^2 - 1\right)}, \tag{1}$$

where R_s represents Spearman rank correlation, $\sum D^2$ is sum of the squared of the difference between the two ranks of each observation, and N is the number of observations.

Linear Regression Correlation ($\boldsymbol{L_s}$). This is the univariate regression analysis. The formula can be expressed as below.

$$L_s = \frac{\left(X - \overline{X}\right)\left(Y - \overline{Y}\right)}{sd_X - sd_Y}, \tag{2}$$

where L_s represents linear regression correlation, X and Y are variables, $\overline{X}$ is average of X, $\overline{Y}$ is average of Y, and sd_X and sd_Y are standard deviation of X and Y, respectively. The significance of these correlation can be confirmed by F-score.

2.2 Regression Analysis

Random Forest. Random forest model is one of the ensemble methods, which is built from hundreds of decision trees model, and the most voting of each decision tree will be chosen. This method was created by Leo Breiman, who is applying the random decision forest from Tin Kam Ho. In the regression task, multiple regression trees are created, and the results are combined by the averaging as shown in Fig. 1 [10].

Adaboost. AdaBoost is also one of the ensemble methods, but this method uses boosting technique to improve the weak learner to be the strong learner by adjusting the weight. Freund & Schapire proposed the AdaBoost model which is diminishing both of the bias and the variance [11]. Figure 2 shows process of AdaBoost to create the boosting tree.

Support vector regression. Support vector regression was created by Vladimir N. Vapnik. This model is modified from support vector machine to find a regression task but still keeps the same algorithm which is maximal margin as shown in Fig. 3. When applying the model to the high dimensional, kernel function will be used. Regular kernels are polynomial function, linear function and radial basis function [7].

3 Stock Price Prediction from News

This section shows the step to predict the stock price from the news. Two feature selection techniques are compared. Then, using three learning techniques, i.e., random

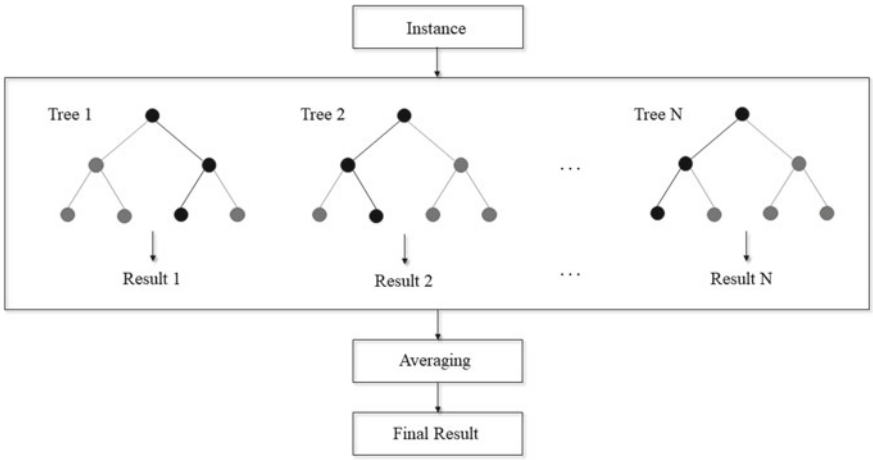


Fig. 1 Random forest model based on multiple trees

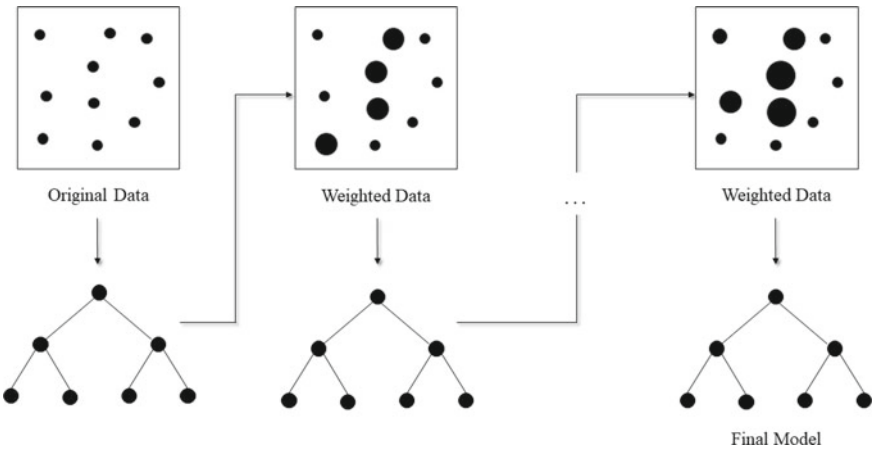


Fig. 2 Boosting tree

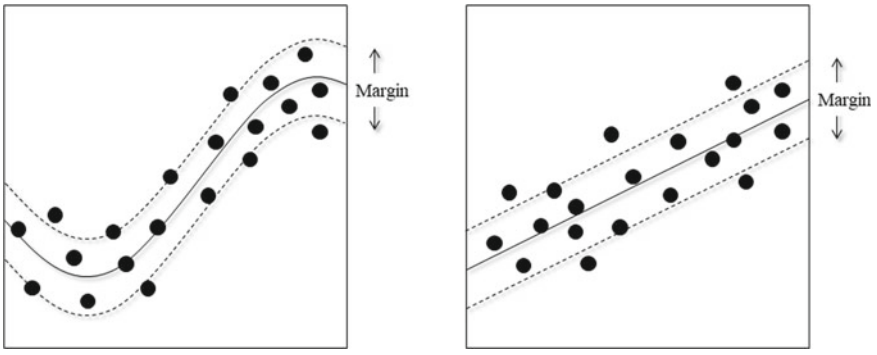


Fig. 3 Support vector regression

forest, AdaBoost and support vector regression are applied to get the prediction results. The overall process is presented in Fig. 4.

3.1 Data Collection

In order to predict the stock prices from the daily news, all related data must be collected. There are two parts of related information, i.e., stock prices and news. News in this research covers many topics such as finance, politic, sport and agriculture. However, the news is scoped by the source of them. Only a Web site is used as a source of data, to avoid the bias from news redundancy.

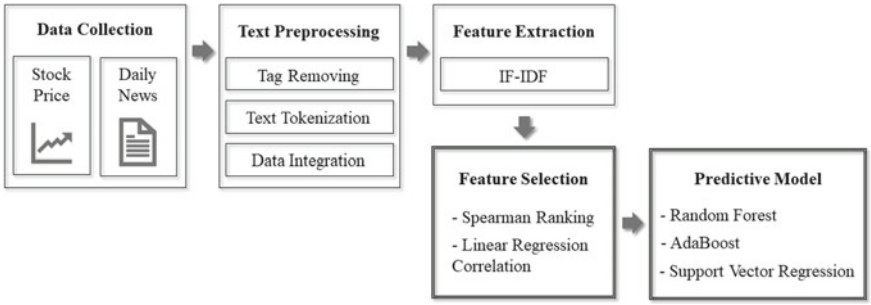


Fig. 4 Process of stock price prediction from daily news

3.2 Text Preprocessing

In this process, the raw text of news will be cleaned up by removing unnecessary tags. Since the news is collected from Internet source, HTML tag, XML tag and punctuation should be removed. Then, text will be separated into token, or it will be called tokenization. The maximum matching algorithm is applied in this step. Moreover, some tokens which have same meaning will be converted to their root form. In the lastly, frequencies of the tokens in a day are calculated, and these data are combined to the matched stock prices.

3.3 Feature Extraction

This part will be converting tokens to vector because in order to perform a predictive model, the input must be a vector or numerical. So far, there have many techniques to convert the tokens into the vector. But this research will be using term frequency-inverse document frequency (TF-IDF). TF-IDF method is a technique to find the importance of each word in the document. TF-IDF that can be expressed as below.

$$\mathrm{TF}(i, j) = \frac{\text{Term Frequncy of } i \text{ in Document } j}{\text{Total Words in Document } j}, \tag{3}$$

$$\mathrm{IDF}(i, \text{All Document}) = \log_2\left(\frac{N}{\mathrm{d}f(i)}\right), \quad \text{and} \tag{4}$$

$$\mathrm{TFIDF} = \mathrm{TF}(i, j) \times \mathrm{IDF}(i, \text{All Document}), \tag{5}$$

where N is the total number of document, and $\mathrm{d}f(i)$ is number of documents with term i.

3.4 Feature Selection

In this process, some features from the feature extraction will be removed. Only the most importance features will be used. In this research, two techniques of feature selection will be conducted, which is Spearman ranking and linear regression correlation techniques. Both Spearman ranking and linear regression correlation are filtering techniques that use the concept of correlation to find the relationship between variables. Due to the feature selection technique, task is to decrease the number of the variable, and in this research, the remaining features are 50 and 100 features.

3.5 Predictive Model

To perform the prediction, three learning techniques are compared. Random forest (RF), AdaBoost and support vector regression (SVR) are used for training and testing. The parameters are set as 100 trees in RF, regression trees with linear loss function are used in AdaBoost, and radial basis function (RBF) is chosen as the kernel function in SVR. The comparison of three predictive models will be performed by root mean-squared error (RMSE). The result of the predictive model will be described in the next section.

4 Experimentation

This research is predicting the stock price from the news. Dataset is composed of two parts. The first part is the news, which is collected from Thai news Web site on 2017–2020. The second part is the stock prices, which is collected from Thai stock market. Stock price will be chosen only by five companies per sector, and there are eight sectors, i.e., AGRO, CONSUMP, FINCIAL, INDUS, PROPCON, RESOURC, SERVICE and TECH. Hence, the total stock prices come from 40 companies. The data from 2017 to 2019 will be used to be the training set, and this is 80% of the collected data. The data from 2019 to 2020 will be used for testing. The performance of each learning technique will be evaluated by the root mean-squared error (RMSE),

$$\text{RMSE} = \sqrt{\sum_{i=1}^{n} \frac{(f_i - o_i)^2}{n}}, \tag{6}$$

where f_i is the values of sample i which is predicted from the model, o_i is the actual value of sample i, and n is the total number of the value.

Table 1 RMSE for each sector on whole features

Stock sector	Random forest	AdaBoost	SVR
AGRO	18.12	16.62	15.8
CONSUMP	2.02	1.94	1.97
FINCIAL	26.32	27.20	25.81
INDUS	7.74	8.15	7.32
PROPCON	1.13	1.17	0.97
RESOURC	2.87	3.07	2.50
SERVICE	7.54	7.88	7.03
TECH	11.51	11.44	10.91
Average	9.66	9.68	**9.04**

In the first experiment, three learning algorithms are compared based on whole features that are extracted from TF-IDF. The average RMSE for each stock sector is presented in Table 1.

Table 1 shows the comparison results among three algorithms. Bold value shows the lowest RMSE. The RMSE of SVR is lower than other models. There are seven sectors that SVR yields the best RMSE. Then, the feature selection techniques will be tested. Spearman ranking (R_s) and linear regression correlation (L_s) are compared for each learning technique as shown in Tables 2, 3 and 4.

Bold values from Tables 1, 2, 3 and 4 are the lowest RMSE value when compare to the other algorithm. From the results, RMSE from Spearman ranking is better than that of linear regression correlation for the most cases. Especially, when 50 features are used, SVR still provides the best RMSE in all sectors. Hence, based on this dataset, this research suggests to use Spearman ranking with 50 features and SVR for stock price prediction. Figure 5 shows the graph of the average RMSE when they are compared on the different number of features.

Table 2 RMSE of the random forest model with feature selections

Stock sector	50 features		100 features	
	R_s	L_s	R_s	L_s
AGRO	17.92	18.74	17.97	18.91
CONSUMP	2.00	2.10	2.02	2.10
FINANCIAL	26.44	26.55	26.31	26.63
INDUSTRIAL	7.68	8.46	7.92	8.02
PROPCON	1.10	1.14	1.10	1.18
RESOURCE	2.76	2.95	2.82	2.99
SERVICE	7.54	8.16	7.49	8.15
TECH	11.38	13.34	11.41	13.76
Average	**9.60**	10.18	9.63	10.22

Table 3 RMSE of the AdaBoost model with feature selections

Stock sector	50 features		100 features	
	R_s	L_s	R_s	L_s
AGRO	19.01	20.50	19.50	19.74
CONSUMP	2.06	2.13	2.05	2.12
FINANCIAL	26.83	27.13	27.23	27.26
INDUSTRIAL	8.50	8.34	8.40	8.50
PROPCON	1.09	1.21	1.12	1.26
RESOURCE	2.66	2.98	2.78	3.14
SERVICE	7.69	8.23	7.85	8.27
TECH	11.29	13.53	11.60	13.89
Average	**9.89**	10.51	10.07	10.52

Table 4 RMSE of the SVR model with feature selections

Stock sector	50 features		100 features	
	R_s	L_s	R_s	L_s
AGRO	14.09	15.59	14.71	15.67
CONSUMP	1.83	1.93	1.85	1.89
FINANCIAL	25.89	27.80	25.59	27.69
INDUSTRIAL	6.65	8.65	6.75	8.68
PROPCON	0.73	0.94	0.80	0.98
RESOURCE	2.14	2.64	2.29	2.56
SERVICE	5.89	8.31	6.17	8.40
TECH	8.82	12.14	9.43	11.71
Average	**8.26**	9.75	8.45	9.70

However, when each stock is considered, there are some values that are higher than the other stock such as CHOTI, BBL, BAT-3 K and DELTA. Figure 6 shows RMSE of each stock.

5 Conclusion

This research aims to find the prediction of the stock market by using daily news. The result shows that support vector regression model gives the results that is better than random forest and AdaBoost models which are the ensemble method. After enhancing the prediction by feature selection technique, the prediction score is increasing significantly. Support vector regression model with feature selection by Spearman ranking on 50 features yields the lowest root mean-squared error, which

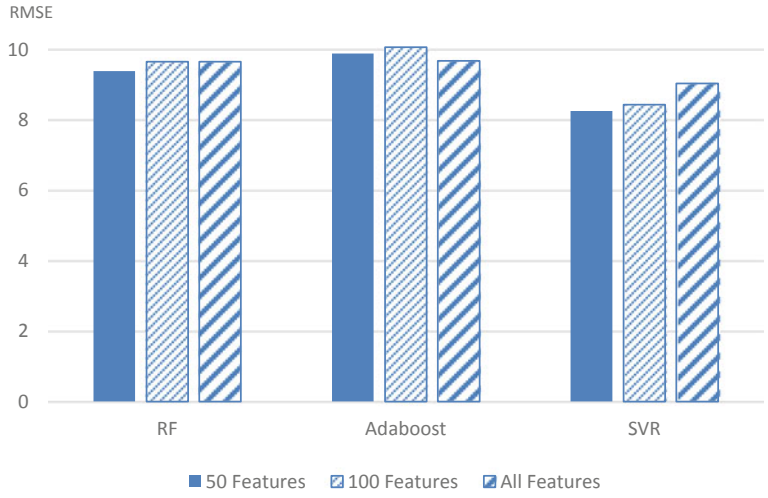


Fig. 5 RMSE on the different number of features

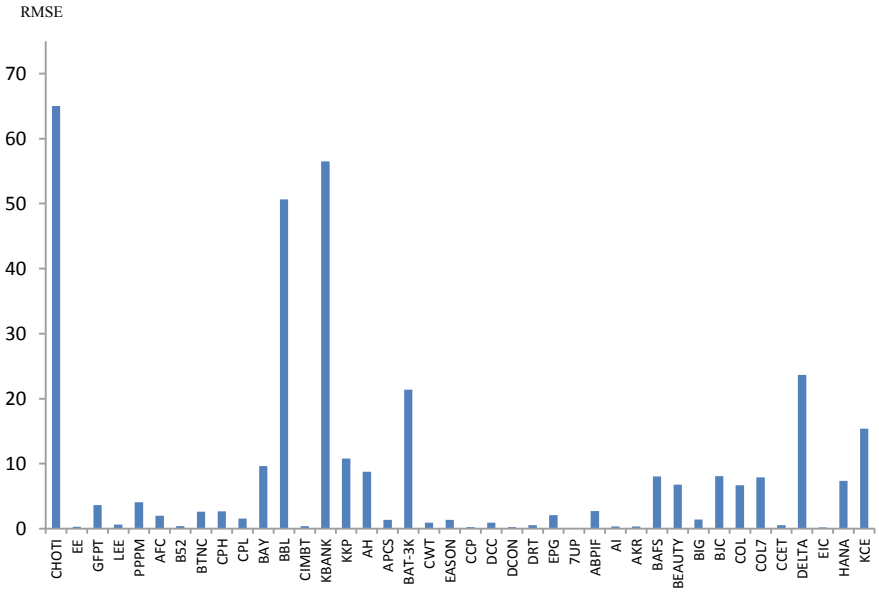


Fig. 6 RMSE of 40 stocks

can reduce the root mean-squared error around 8.63%, which is calculating from the prediction result of support vector regression without enhancement compared with the support vector regression enhancement with Spearman ranking with 50 features remaining.

References

1. Kim, K.: Financial time series forecasting using support vector machines. Neurocomputing **55**, 307–319 (2003)
2. Nahila, A., Lyhyaoui, A.: Short-term stock price forecasting using kernel principal component analysis and support vector machines: the case of Casablanca stock exchange. Proc. Comput. Sci. **127**, 161–169 (2018)
3. Khaidem, L., Saha, S., Roy Dey, S.: Predicting the direction of stock market prices using random forest. Appl. Math. Financ. **23**(20) (2016)
4. Basak, S., Saha, S., Khaidem, L., Dey, S.R.: Predicting the direction of stock market price using tree based classifiers. North Am. J. Econ. Financ. **47**, 552–567 (2019)
5. Kim, H., Han, S.T.: The enhanced classification for the stock index prediction. Proc. Comput. Sci. **91**, 284–286 (2016)
6. Chang, V., Li, T., Zeng, Z.: Towards an improved Adaboost algorithmic method for computational financial analysis. J. Parall. Distrib. Comput. **134**, 219–232 (2019)
7. Huang, C.-J., Yang, D.-X., Chuang, Y.-T.: Application of wrapper approach and composite classifier to the stock trend prediction. Expert Syst. Appl. **34**(4), 2870–2878 (2008)
8. Nassirtoussi, A.K., Aghabozorgi, S., Wah, T.E., Ngo, D.C.T.: Text mining of news-headlines for FOREX market prediction: a multi-layer dimension reduction algorithm with semantics and sentiment. Expert Syst. Appl. **42**, 306–324 (2015)
9. Wang, B., Wang, R., Wang, Y.: Compatible matrices of Spearman's rank correlation. Stat. Probab. Lett. **151**, 67–72 (2019)
10. Mei, J., He, D., Harley, R., Habetler, T., Qu, G.: A random forest method for real-time price forecasting in New York electricity market. In: 2014 IEEE PES General Meeting Conference, pp.1–5. National Harbor (2014)
11. Yutong, S., Zhao, H.: Stock selection model based on advanced AdaBoost algorithm. In: 2015 7th International Conference on Modelling, Identification and Control (ICMIC), pp. 1–7. Sousse (2015)

A Study on Deep Learning Predictive Models in Healthcare

H. N. Sowmya and S. Ajitha

Abstract In spite of large amount of data available today, the healthcare field is facing new challenges in order to automatically detect and diagnose diseases. Deep learning, a branch of artificial intelligence, is growing fast in computer science will provide various tools and techniques to address the challenges in the health care. The rapidly growing fields of predictive analytics and deep learning are playing a major role in the healthcare data practices and research. In this paper, we reviewed the benefits and risks associated with predictive analytics in health care. We studied various deep learning predictive models used in the health care as well as their applications and prominence in healthcare industry. We also summarized applications of different deep learning models and their results.

Keywords Deep learning (DL) · Artificial intelligence · Predictive analytics · Health informatics · Healthcare industry

1 Introduction

Artificial intelligence and deep learning are gaining a lot of attention and revolutionized several industries such as retail, finance, travel, manufacturing, health care, and so on. Health care is one such industry which is implementing these technologies to find solutions which are changing the shape of health care. AI and deep learning when combined together can provide a lot of benefits to the healthcare industry and changes the way medical professionals perform their roles and patient care. Healthcare organizations have developed an interest on how these technologies can support them to reduce cost and improve efficiency. Since health is a major priority, medical experts and professionals are trying to find ways to implement these technologies to provide best results.

H. N. Sowmya (✉)
Nitte Meenakshi Institute of Technology, Bengaluru, India

S. Ajitha
M. S. Ramaiah Institute of Technology, Bengaluru, India

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_81

Deep learning and predictive analytics are playing a vital role to aid researchers and medical professionals to discover the hidden features from the data and help in predicting future trends in the health care. Health care is benefiting tremendously from deep learning and predictive analytics and providing a lot of benefits to the health care. Deep learning, a branch of AI, will provide high-level solutions to medical technology. It helps doctors to analyze any disease accurately and helps them in treating patients in a better way, hence, resulting in better medical decisions.

In this paper, we will give a brief introduction to the variety of deep learning predictive models and their applications in health care. The following topics will be presented in the rest of the paper. In Sect. 2, we have discussed predictive analytics with its benefits and risks in health care. We outlined deep learning and popular deep learning models in health care in Sect. 3 and in Sect. 4 and we have presented the research carried out using deep learning models in health care for medical imaging, genome analysis, and disease diagnosis. We also presented a summary table showing applications of different deep learning models and their results. In Sect. 5, we provided the conclusion of the study with its future scope.

2 Predictive Analytics

Predictive analytics is a branch of statistics and advanced analytics that utilizes various techniques from modeling, statistics, machine learning, deep learning, and artificial intelligence. It has its roots in classical statistics which has the ability to predict the future events and behavior of variables by analyzing current and historical data. Predictive analytics can bring the technologies which have the ability to make future prediction and trends together. Predictive analytics is increasing its applications in manufacturing, marketing, cybersecurity, crime prevention, fraud detection, health informatics, and bioinformatics. The healthcare sector is becoming the key beneficiary of predictive analytics and has gained widespread adoption over the past few years. Deep learning techniques can be used for building predictive models due to their outstanding performance to manage large-scale datasets with noise.

Predictive Analytics in Health care

Today, health care is benefiting a lot from predictive analytics to perform business operations; to aid doctors and to improve the accuracy of prediction and diagnosis; and in assessing possible risk factors toward public health. There are also some risk factors associated with predictive analytics in health care. These include the fast pace of technology and impact on decision-making processes, moral hazard, and human involvement with the machine, algorithm bias, and lack of regulation and some privacy pressures. The following Fig. 1 demonstrates the advantages and risks with predictive analytics in health care.

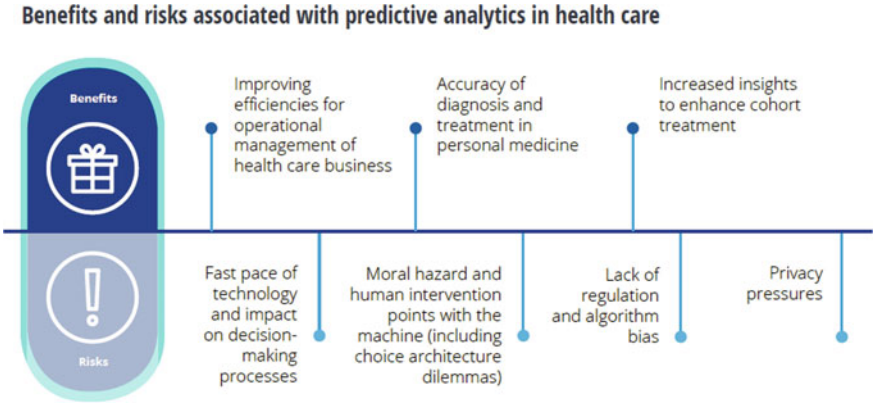


Fig. 1 Advantages and risks with predictive analytics in health care

3 Overview of Deep Learning Predictive Models

Deep learning (deep structured learning) is a new area emerging; a sub-field of machine learning is based on artificial neural networks. It has come to the foreground as a way to overcome the limitations of artificial neural networks to real-world problems. In these networks, data is passed through multiple layers, and each layer uses the output from its previous layer to generate results. Deep learning models can give accurate results as more data traverse multiple times. Deep learning models will learn from previous results to refine their ability to correlate and optimize the ultimate output. Deep learning models do not require processed data, since network itself takes care of preprocessing. Deep learning techniques can give accurate results very fastly, when compared to conventional machine learning techniques as more data becomes available. The workflow of deep learning is shown in Fig. 2.

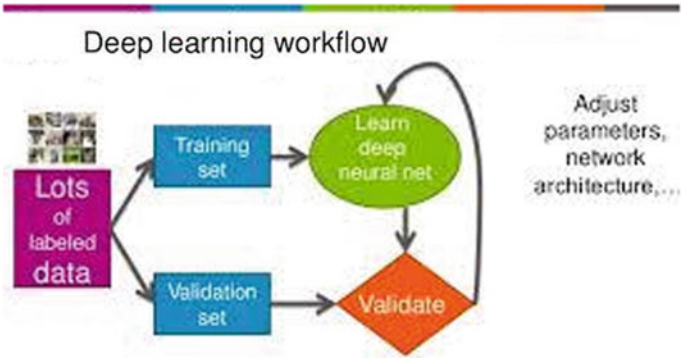


Fig. 2 Depicts deep learning workflow

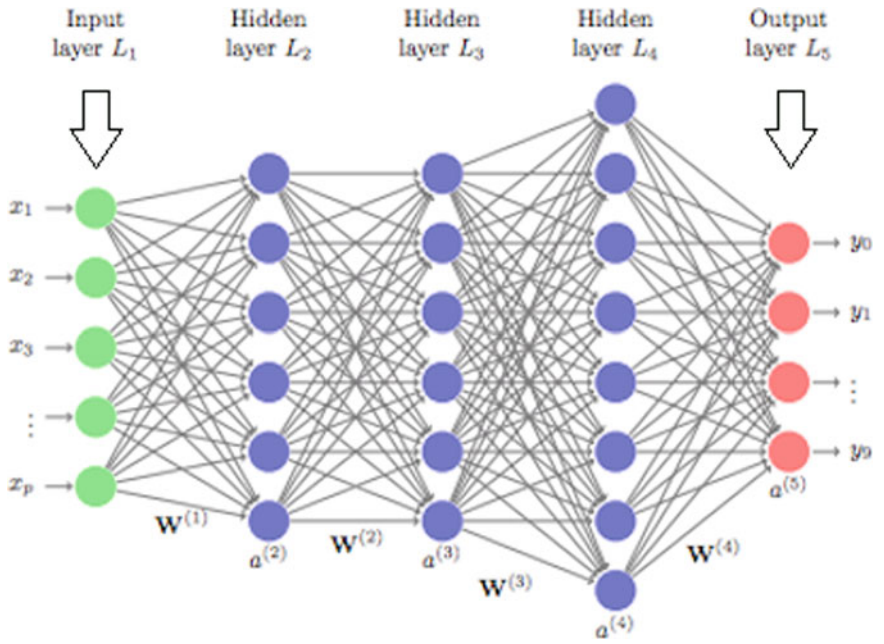


Fig. 3 Basic structure of a deep neural network

3.1 Deep Neural Network (DNN)

Deep neural network (DNN), a feedforward neural network has multiple hidden layers where each hidden layer has perceptrons contains multiple hidden layers. If the network has only three layers, then it is called shallow neural network, and when the network has more than three layers it is called deep neural network. The typical architecture of deep neural network is shown in Fig. 3.

There are various types of deep learning models applied in the healthcare industry including convolutional neural networks, recurrent neural networks, deep belief networks, deep autoencoders, and restricted Boltzmann machines.

3.2 Convolutional Neural Networks

Convolutional neural networks (ConvNet or CNNs) are deep neural networks which are used effectively for classification and image recognition. CNNs have been successful in areas like face recognition, object identification, and traffic signals. The convolutional neural network has many layers which are shown in Fig. 4.

The Convolutional Layer: This is the first layer whose purpose is to detect features present in the input images which is done by convolution filtering. Here, the window

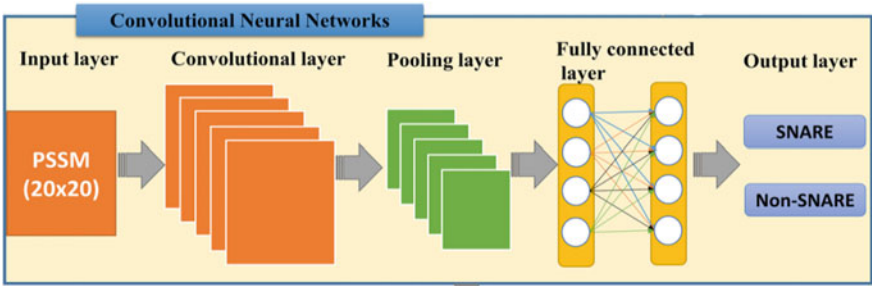


Fig. 4 Architecture of convolutional neural networks

containing the feature of image is dragged and a convolution product between the feature and image is calculated. This layer receives images and calculates the convolution of each of them.

The Pooling Layer: This layer is present between two convolution layers which applies pooling operation to each received feature map. Here, the size of the images is reduced preserving the image characteristics. It also reduces the number of parameters and calculations in the network thus improving efficiency of the network.

The Fully-Connected Layer: This layer is present at the last in the convolutional neural network which is not a characteristic of CNN. This layer receives input vector to which it applies an activation function and produces an output vector. This last layer is the one which classifies the input image which is received by the network.

3.3 Recurrent Neural Networks

Recurrent neural network (RNN) is a type of deep neural network which contains hidden units to analyze data streams. RNN's are well suited for applications including text analysis, DNA sequences, and speech recognition where the output of the network will be depending on previous computations. Usually, RNN's are provided with input samples having strong inter-dependencies and significant representations to maintain information of the previous or past steps. For traditional neural networks, there was a limitation that there inputs are not dependent on each other. RNNs usually deal with two input types, the present and the previous one, for providing output to the new data. RNN's also suffer from gradient problems even though they are powerful. Several variations of RNN's like long short-term memory units (LSTM's) and gated recurrent units (GRU's) were proposed to solve the problem of RNN's. These networks are more useful because they maintain long-term interrelations. RNNs have already successful in many applications like speech recognition, natural language processing, bioinformatics, and in stock market predictions. The architecture of RNN is shown in Fig. 5.

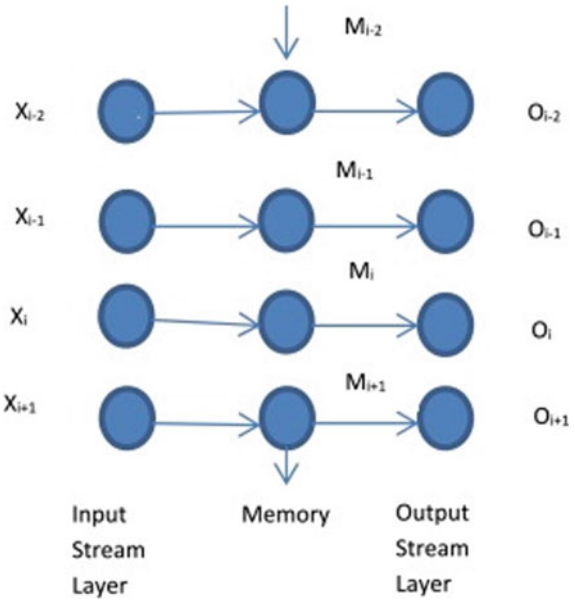


Fig. 5 Architecture of recurrent neural networks

3.4 Restricted Boltzmann Machine

A RBM's is a Boltzmann machine modeled using statistic units to learn probability distribution. RBM's were first proposed by Geoffrey Hinton which has two layers: Input layer and hidden layer are connected by symmetric bipartite graph. In this network, each node in the input layer is connected to each node in the hidden layer. Multiple RBM's can be stacked together to construct deep belief networks (DBN) and deep Boltzmann machine (DBM) which can be fine-tuned using gradient descent and backpropagation. RBM's are mainly used for classification, regression, and dimensionality reduction. The structure of RBM is shown in Fig. 6.

3.5 Deep Belief Network (DBN)

Deep belief networks are generative in nature and contain multiple layers including both directed and undirected layers. Here, direct connections exist between lower layer nodes and between upper layer nodes indirect connections will exist, where the nodes identify the correlations in the given data. These are the algorithms which are modeled to recognize patterns and process complex data. Deep belief networks are made up of many restricted Boltzmann machines (RBM's) but do not communicate within the layer like RBM's. The DBN's are trained in a layer-wise greedy manner using greedy learning algorithms by fine-tuning the generative weights to establish balance estimate of the learning probability. Since greedy learning algorithms are

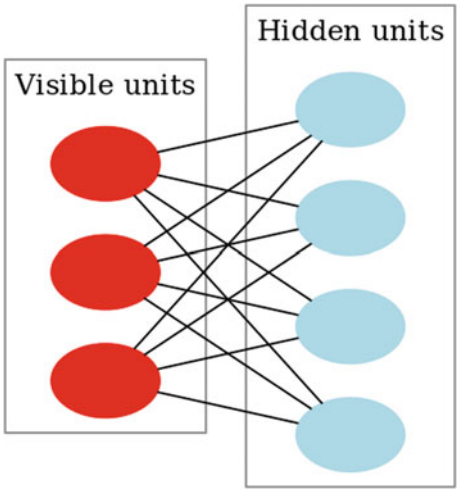


Fig. 6 Typical structure of restricted Boltzmann machine

quick and efficient, they are used for training deep belief networks. Deep belief networks are used in applications like image recognition, video recognition, predict time series prediction, and to capture motion data in various fields like medical diagnosis, financial market, home automation, security and health care, video game development, film making, and weather prediction. A typical architecture of DBN composed of RBM's is shown in Fig. 7.

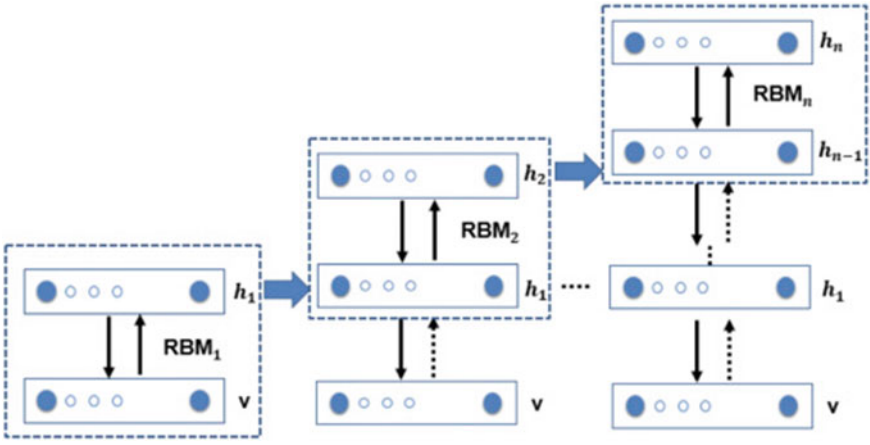


Fig. 7 Architecture of a deep belief network (DBN) with RBM's

3.6 Deep Autoencoders

Autoencoder is a deep neural network which has three layers including input layer, output layer, and hidden layer. The structure of autoencoder is shown in Fig. 8. In this network, both input layer and output layer will have same number of nodes, so that output layer can reconstruct its own inputs. Since there is no need of class label assigned to the input data, this method is therefore unsupervised. The autoencoders can be used for dimensionality reduction of the input data without any loss of data. For high-dimensional data, a single hidden layer will not be sufficient to represent the data. Therefore, there is a need for deep autoencoders.

Deep autoencoder is the extended version of autoencoder which has many hidden layers in order to represent the complex data. In this network, the first hidden layer may learn first-order features, and the second hidden layer learns second-order features and deeper layer may tend to learn even higher-order features. So put together, we can tell additional layers are required to handle complex data. During training, there are many variations of autoencoders are proposed including sparse autoencoders, denoising autoencoder, contractive autoencoder, and convolutional autoencoder.

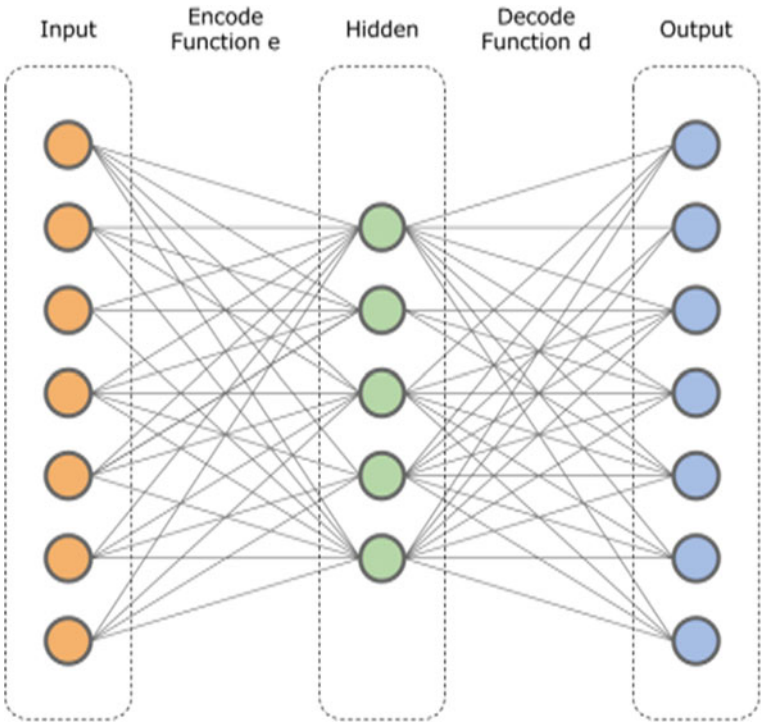


Fig. 8 Basic structure of autoencoder

4 Applications of Deep Learning in Health care

Health care is getting massive benefits from deep learning, since huge volume of data is being generated as well as the rapid increase of medical devices and digital record systems. Deep learning is playing a vital role in healthcare industry to assist medical professionals and researchers to discover the hidden knowledge in data. Deep learning is helping doctors and health professionals to predict and diagnose diseases like cancer, tuberculosis (TB), heart disease, brain tumor, diabetes, Alzheimer's, Parkinson's, and multiple sclerosis and aids in treating patients in a better way, thus resulting in taking better medical decisions. Some deep learning algorithms have already produced transformational outcomes in the healthcare industry. We searched many research papers, articles, and databases where we found various deep learning applications in health informatics. Here, we are going to outline few deep learning applications in health care and also we have done comparative study. Following are the few use cases of deep learning methods based on the relevant articles and databases we found:

4.1 Deep Learning in Medical Imaging

Medical imaging is a technique to produce optical representations of various parts of the body to predict, diagnose, and treat frightful diseases in health care. Magnetic resonance imaging (MRI), computed tomography (CT), electrocardiogram (ECG), X-ray, microscopy, ultrasound (US), mammography (MG), and optical images are some examples of medical images. Automatic analysis of medical images through interpretation is very much required in modern healthcare industry for earlier prediction of diseases. In modern medicine, various deep learning techniques are extremely applied in medical imaging research community because of its outstanding performance. Deep learning has shown promising results in automatic detection of diabetic retinopathy. In the approach [1], deep convolution neural network (DCNN) was applied to automatically detect diabetic retinopathy. Their work resulted in sensitivity of 97.5% and specificity of 93.4% on the eye picture archive communication system (EyePACS-1) dataset and achieved 96.1% sensitivity and 93.9% specificity on Messidor-1. The Inception-v3 architecture, a specific neural network was used in this work is shown in Fig. 9.

In [2] approach, deep convolutional neural network (DCNN) through spatial analysis was employed to identify the severity of the diabetic retinopathy disease. They deployed architecture with dropout layer techniques have claimed to achieve 94–96 percent accuracy. They used the image dataset from popular databases like STARE, DRIVE, and Kaggle fundus. The DCNN architecture used in this work is shown in Fig. 10.

Deep learning also showed excellent results in detecting breast cancer. In the approach [3], the researchers have applied DCNN on mammograms for the early

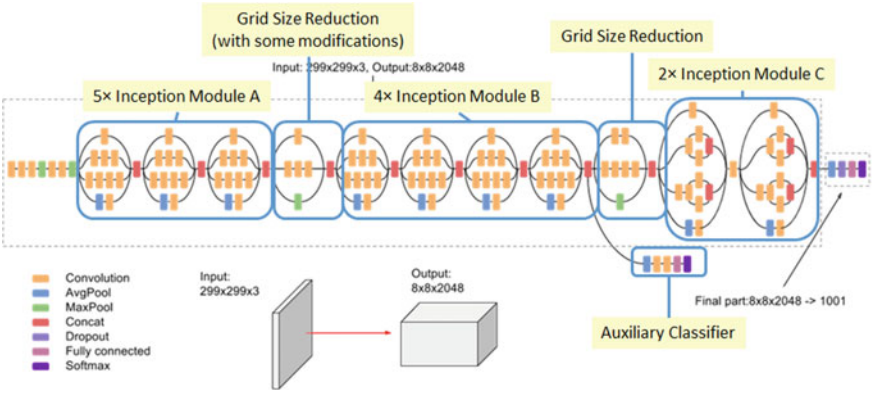


Fig. 9 Inception-v3 architecture used in the detection of diabetic retinopathy

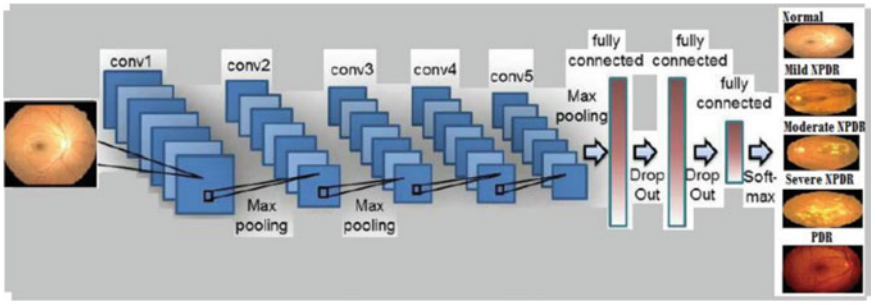


Fig. 10 DCCN architecture used in above research

detection of breast cancer. The convolutional network technique used by them has attained excellent results in comparison with previous methods for classifying screening mammograms. Their digital database for screening mammography (CBIS-DDSM) dataset of digitized film mammograms has claimed to achieve 86.1% sensitivity and 80.1% specificity. Another test dataset called full-field digital mammography (FFDM) images from the INbreast database has achieved a 86.7% sensitivity and 96.1% specificity. Parkinson's disease (PD) and Alzheimer's disease (AD) are the neurological disorder which medical industry facing significant challenges today. Deep learning models can be used to detect Alzheimer's and Parkinson's disease at the initial stage only. In [4] approach, convolutional network was used to distinguish Alzheimer's patients from healthy people brain. They employed a basic convolutional neural network (CNN) model called LeNet-5 to classify functional MRI data collected from Alzheimer's patients and normal controls where the trained data has reached an accuracy of 96.85%. In [5], deep convolutional neural networks (DCCN) were used to detect Parkinson's disease by automatically recognizing patterns in brain images that characterize the disease. They used 269 *DaTSCAN* images from

the database of Parkinson's progression markers initiative (*PPMI*) where 158 are Parkinson's brain images and remaining 111 are brain images from normal people. They also used isosurfaces to implement a classification system which resulted in 95.1% accuracy. In [6] approach, they applied sparse autoencoders which requires less labeled training set and less prior knowledge to predict Alzheimer's disease. The model was used to classify whether the condition is mild or prodromal. The algorithm resulted in 91.6% accuracy, 98.09% sensitivity, and 84.09% specificity. In paper [7], convolutional neural network (CNN) was implemented to classify different types of tuberculosis (TB) using lung CT images. In this work, they were successful in classifying five TB types out of six which was tested against 300 datasets. The accuracy value of 0.4067 was achieved in classifying TB types.

4.2 Deep Learning for Genome Analysis

There is a need for innovative approaches in genome science to overcome the challenges associated when conventional machine learning algorithms were used. One approach that is exciting and promising is deep learning made up of deep neural networks which can help doctors to diagnose and treat patients fat and accurate. Deep learning is applied genomics for cancer diagnosis, public health, and functional improvement. In [8] proposed a deep convolutional neural network (DCNN) for predicting the quantitative traits from single-nucleotide polymorphisms (SNPs). They got very good results in the prediction of quantitative phenotypes than current statistical methods. In approach [9], researchers proposed deep learning framework for modeling and predicting binding sites of RNA-binding proteins (RBPs). Their deep learning framework was able to automatically predict accurate RBP binding sites. In [10], deep learning approach called deep bind and an automatic standalone software tool to predict gene sequence specificities of DNA and RNA-binding proteins which outperformed than any other state-of-the-art methods. In [11] proposed convolutional neural network model for predicting DNA sequences and got excellent results than statistical methods. In [12], authors have proposed a hybrid framework composed of convolutional and bi-directional long short-term memory recurrent neural networks called DanQ, for predicting non-coding function from the sequence. In [13], stacked autoencoders were used for the prediction and classification of different cancer types based on gene expression data. The result has shown that the proposed method is better than previous methods. In [14] presented, a deep learning approach using stacked denoising autoencoder (SDAE) for cancer detection and to identify genes which are critical for breast cancer diagnosis. They found highly interactive gene as useful cancer biomarker for detecting breast cancer.

4.3 Deep Learning in Chronic Disease Diagnosis

Deep learning can be used for detection and diagnosis of chronic diseases like cancer, Alzheimer's, Parkinson's, and coronary heart disease. In [15], deep neural network (DNN) combined with mini-batch gradient descent (MBGD) was proposed to differentiate between PD patients and non-PD patients using vocal features. The result from the proposed method has shown that the model is better than the conventional methods. The accuracy achieved by the proposed method was 89.5%. In the approach, [16] proposed a deep neural network consisting of a softmax classifier and a stacked autoencoder for the diagnosis of Parkinson's disease. The results from the classifier were compared with traditional classification methods and the result showed the classifier as a very efficient classifier for Parkinson's disease diagnosis. [17] used convolutional neural network (CNN) to cope with PD automatic identification from handwritten dynamics. In [18], convolutional neural network-based framework called BC-DROID was proposed for automated breast cancer detection and diagnosis. Their framework was claimed to achieve 90% detection accuracy and 93.5% classification accuracy. In paper [19], an enhanced deep neural network (DNN) model was proposed to improve the accuracy in diagnosis of heart disease and medical condition in patients. The deep neural network model developed was resulted in 83.67% accuracy, 93.51% sensitivity, 72.86% specificity, and 79.12% precision value. In research work [20], proposed a residual neural network called ResNet-18 to perform classification of Alzheimer disease, and their model was significantly able to classify the subjects and got accuracy of 97.92% and 97.88%. In the approach, [21] proposed a two-stage method, where in the first stage, a two-layer neural network has been used to learn features from the data, and in the second stage, softmax regression was used. The proposed method was provided with dataset containing Alzheimer's brain images. The model has given excellent results and reduced the manual work and helps to diagnose intelligently. The summary of different deep learning applications in health care is presented in Table 1.

5 Conclusion

In this paper, we outlined different deep learning models and its applications in health care in different area. Even though deep learning is still in its early stages, it has shown substantial results compared to traditional machine learning techniques. It also showed better results for applications than healthcare professionals in some areas. In some areas, it helped doctors and patients in the early and fast detection and diagnosis of various chronic results. Since there is a limitation of data availability, the research in some areas is still in process. Deep learning techniques can be applied for more number of applications in every area of health care where human interpretation will be difficult. Finally, we want to conclude that with the help of deep learning

Table 1 Summary of applications of deep learning in health care

Application	Deep learning model	Results	References
Detection of diabetic retinopathy	Deep convolutional neural network (DCNN)	Sensitivity: 96.1% Specificity: 93.9%	[1, 2]
Detection of breast cancer and diagnosis	Deep convolutional neural network (DCNN)	Sensitivity: 86.1% Specificity: 80.1%	[3]
	Deep convolutional neural network (BC-DROID)	Detection accuracy: 90% Classification accuracy: 93.5%	[18]
Alzheimer's disease (AD)	Convolutional neural network (CNN)	Accuracy: 96.85%	[4]
	Residual neural network (ResNet-18) and resting-state functional magnetic imaging (rs-fMRI)	Average accuracy of 97.92% and 97.88%	[20]
	Sparse autoencoders, scale conjugate gradient (SCG), Stacked autoencoder	Sensitivity: 98.09% Specificity: 84.09% Accuracy: 91.6%	[6]
Parkinson's disease (PD)	Deep convolutional neural network (DCNN)	Accuracy: 95.1%	[5]
	Deep neural network (DNN) with mini-batch gradient descent (MBGD)	Accuracy: 89.5%	[15]
Tuberculosis (TB)	Convolutional neural network (CNN)	Accuracy: 0.4067	[7]
Heart disease diagnosis and prognosis	Deep neural network	Accuracy: 83.67%, Specificity: 72.86%, Sensitivity: 93.51%, Precision: 79.12%	[19]

predictive models, diagnosis of diseases can become smarter and faster. As future work, we would like to apply deep learning algorithms by considering a case study.

References

1. Gulshan, V., Peng, L., Coram, M., et al.: Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. JAMA **316**(22), 2402–2410 (2016). https://doi.org/10.1001/jama.2016.17216
2. Chandrakumar, T., Kathirvel, R.: Classifying diabetic retinopathy using deep learning architecture. Int. J. Eng. Res. Technol. (IJERT) **05**(06) (2016)
3. Shen, L., Margolies, L.R., Rothstein, J.H., et al.: Deep learning to improve breast cancer detection on screening mammography. Sci. Rep. **9**, 12495 (2019). https://doi.org/10.1038/s41598-019-48995-4

4. Sarraf, S., Tofighi, G.: Classification of Alzheimer's disease using fMRI data and deep learning convolutional neural networks (2016)
5. Ortiz, A., Munilla, J., Martínez-Ibañez, M., Górriz, J.M., Ramírez, J., Salas-Gonzalez, D.: Parkinson's disease detection using isosurfaces-based features and convolutional neural networks. Front. Neuroinform. **13**, 48 (2019). ISSN 1662-5196
6. Jha, D., Kwon, G.-R.: Alzheimer's disease detection using sparse autoencoder, scale conjugate gradient and Softmax output layer with fine tuning. IJMLC **7**, 13–17 (2017). https://doi.org/10.18178/ijmlc.2017.7.1.612
7. Gao, X.W., Yu, Q.: Application of deep learning neural network for classification of TB lung CT images based on patches. In: CLEF (2017)
8. Liu, Y., Wang, D.: Application of Deep Learning in Genomic Selection, pp. 2280–2280 (2017). https://doi.org/10.1109/BIBM.2017.8218025.
9. Zhang, S., Zhou, J., Hu, H., Gong, H., Chen, L., Cheng, C., et al.: A deep learning framework for modeling structural features of RNA-binding protein targets. Nucl. Acids Res. **44**, e32 (2016)
10. Alipanahi, B., Delong, A., Weirauch, M.T., Frey, B.J.: Predicting the sequence specificities of DNA- and RNA-binding proteins by deep learning. Nat Biotechnol **33**, 1–9 (2015)
11. Zeng, H., Edwards, M.D., Liu, G., Gifford D.K.: Convolutional neural network architectures for predicting DNA-protein binding. Bioinformatics, **32**, i121–i127 (2016)
12. Quang, D., Xie, X.: DanQ: a hybrid convolutional and recurrent deep neural network for quantifying the function of DNA sequences. Nucl. Acids Res. **44**, 11 (2016)
13. Fakoor, R., Ladhak, F., Nazi, A., Huber, M.: Using deep learning to enhance cancer diagnosis and classification. In: Proceedings of the 30th International Conference on Machine Learning, pp. 1–7 (2013)
14. Danaee, P., Ghaeini, R., Hendrix, D.A.: A deep learning approach for cancer detection and relevant gene identification. In: Pacific Symposium on Biocomputing **22**, 219–229 (2017). https://doi.org/10.1142/9789813207813_0022
15. Xu, Z., Wang, J., Zhang, Y., He, X.: Voiceprint recognition of Parkinson patients based on deep learning (2018)
16. Caliskan, A., Badem, H., Basturk, A., Yüksel, M.: Diagnosis of the Parkinson disease by using deep neural network classifier. Istanbul Univ. J. Electr. Electron. Eng. **17**, 3311–3318 (2017)
17. Pereira, C.R., Weber, S.A.T., Hook, C., Rosa, G.H., Papa, J.P.: Deep learning-aided Parkinson's disease diagnosis from handwritten dynamics. In: 2016 29th SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI), pp. 340–346. Sao Paulo (2016). https://doi.org/10.1109/SIBGRAPI.2016.054
18. Platania, R., Shams, S., Yang, S., Zhang, J., Lee, K., Park, S.-J.: Automated breast cancer diagnosis using deep learning and region of interest detection (BC-DROID), pp 536–543 (2017). https://doi.org/10.1145/3107411.3107484
19. Miao, K., Miao, J.:. Coronary heart disease diagnosis using deep neural networks. Int. J. Adv. Comput. Sci. Appl. **9**, 1–8 (2018). https://doi.org/10.14569/IJACSA.2018.091001
20. Ramzan, F., Khan, M.U.G., Rehmat, A. et al.: A deep learning approach for automated diagnosis and multi-class classification of Alzheimer's disease stages using resting-state fMRI and residual neural networks. J. Med. Syst. **44**, 37 (2020)
21. Salehi, A.: Alzheimer's disease diagnosis using deep learning techniques **9**, 874–880 (2020). https://doi.org/10.35940/ijeat.C5345.029320

Machine Learning to Detect the Toxicity of Web Textual Contents Against Gender-Based Cyberoffenses

K. Komathy

Abstract Internet serves as a knowledge resource medium for people on all walks-of-life. Especially children below 16 years are the current era's technology practitioners, who often refer the Web resources for their thirst on knowledge. The other side of the Internet also paves a medium for anonymity, which opens the channels for portraying women and children as sexual victims. Cyberbullying and cyberstalking are the most reported cases against girl children but the preventive measures except criminal action against victimization of girl children are yet to shape up. As a proactive measure, this paper proposes to design a machine learning model that detects the probability of toxicity in the textual contents displayed in Web sites and alarm the children so that the crimes may get averted. The model has two components namely text extraction and text classification. Long short-term memory (LSTM) algorithm based on the recurrent neural network (RNN), linear support vector classifier and the prediction model based on logistic regression are the primary machine learning algorithms that implement text recognition and text classification. The performance of the model proposed here achieves 90% accuracy due to the presence of multiclass texts in the data.

Keywords Gender-based cyberoffenses · Natural language processing · Text extraction · Text classifying algorithms · Multiclass classifier

1 Introduction

1.1 Gender-Based cyberoffenses

The Internet is a dynamic medium for new information and knowledge to search and reach. But, as per *Swedish Children's Welfare Foundation and Working Group for Cooperation on Children at Risk* states that the Internet also paves a way to people

K. Komathy (✉)
Academy of Maritime Education and Training, Chennai, India
e-mail: ithod@ametuniv.ac.in

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_82

who exploit children falsifying with their anonymity. Quayle et al. [1] reported that the medium is also opened up children with obscene, violent, and abusive materials and it also simplifies cybergrooming against children. Baines [2] in his report says that a collaborative research is vital at an international level to build a comprehensive and accurate assessment of the scale and nature of online offenses against children. Cyberbullying, another form of cybercrime using smartphone gadgets, is easy reach to children. According to a report [3] by the *World Health Organization* (WHO), *United Nations International Children's Emergency Fund* (UNICEF), and *United Nations Educational, Scientific and Cultural Organization* (UNESCO) on violence against children and the end violence partnership, approximately one billion children each year are affected by physical, sexual, or psychological violence, suffering injuries, disabilities, and death, because countries have failed to follow the established strategies to protect them.

The cyberviolence or bullying faced by boys and girls often differed by the dimension and the gender discrimination of children is one of the 12 overarching recommendations of the *Secretary-General*'s study [4]. Regarding the gender-based offenses [5, 6], the *United Nations Secretary-General* highlighted the structural and systemic causes of male violence toward women and girls, the risk factors that increase their vulnerability and the factors that shape their personal experiences of violence.

National Crime Records Bureau [7] records show that the cases of cyberstalking of women or children are higher by 36% in 2018 compared to 2017. However, the cases reported under cyberstalking are less in 2018 compared to 2017 due to cases being unreported. Sections of law in India, for example, Indian Penal Codes, IT Act, IR Women Act, TNPHW Act, POCSO Act, etc., in general, function in reactive mode of preventive measure while bringing the offenders under the court of law after the occurrence of the cyberviolence. Therefore, the objective of the paper is to emphasis a measure in a proactive mode to mitigate such crimes on line.

1.2 Necessity for Dynamic Patrolling and Vigilance of Web Resources

Cyberpornographers, who get access to children's chat rooms and then lure the children to meet offline to engage in illegal sexual behaviors, are those offenders who go unrecorded or unpenalized unless the child reports to police. Research is required to assess the forms and patterns of cyberstalking, the implementing strategies of law enforcement and at the same time protecting the civil liberties. FBI statistics [8] say that there are about a hundred to two hundred cases of children reported on kidnapping by strangers every year. To assess the offenders' modalities, the Web resources need a continuous and vigorous patrolling and vigilance of Web sites. Therefore, a dynamic and automatic monitoring and prediction tool is mandatory for 24×7 policing. However, freedom of speech or expression and jurisdiction beyond geographical boundaries are the barriers, which keeps the cyberoffenders still alive

and active. The suggestion is to create an exclusive *Cyber Information Act* from the perspective of seamless cybercrimes.

2 Related Works

A study on violence against girls [9, 10] reported that research should identify the modes that can victimize the girls from Internet abuses and the preventive measures such as awareness programs at schools. The report also recommends more stringent measures by *Internet service providers* (ISP) for filtering the Web site contents which uses abusive language and pictures. Sethi et al. [11] define the *cybersexual harassment* (CSH) as the use of the Internet to make uninvited advances, communications, or interactions with another person or entity. Their research study included the awareness of the gender-based CSH and developed a three steps model to identify, block, and report before taking legal action. UNESCO in its report [12] has highlighted a comprehensive solution involving all levels of society, government, legal authorities such as digital literacy and awareness on the safety of cyberspace; technology solutions to protect girls' and women's privacy; gender discrimination data collection with the use of technology.

Sathyanarayana et al. [13] made a call to the researchers to help the lawmakers in formulating policies and regulatory laws that will help to identify as well as curb the menace. Automatic extraction of sentence patterns is proposed [14] to detect the malign contents for dishonoring the girls and women in Japanese language. The authors developed a brute-force search algorithm to automatically extract sentence patterns and used those patterns in text classification of cyberbullying entries.

Researchers [15] conducted a systematic review and content analysis of bullying measures employed in bullying. The authors identified about forty attributes of school bullies involved inside or outside schools. The authors found difficulty comparing these measures due to inconsistency. A sentence is weighted using the frequency of repeated patterns, the length, and the unknown patterns.

3 Empirical Model to Assess the Toxicity of Web Contents Using Machine Learning Techniques

3.1 Extract Textual Contents from Web Pages

For extracting the textual contents from the Web, OCR-LSTM-based package in Python [16] is used. *Optical recognition system* (OCR) converts an image of text bounded with a dimension of length × width, into a printable text from images of jpeg, tiff, gif, png, tiff, and others. To achieve a better text recognition accuracy, the OCR engine, which embeds a *long short-term memory* (LSTM) [17] algorithm based

on the *recurrent neural network* (RNN) is preferred. Figure 1 shows the framework for image-to-text conversion using LSTM-based OCR engine. The framework, firstly, invokes a screen reader, which captures the screen contents that comprises texts as image. For preparing the image to extract the textual contents precisely, preprocessing processes such as gray scaling, noise removal, thresholding, and Canny edge detection are put in order. The training data uses the DDI-100 dataset [18] containing Web source pages and Web images.

Output image from the preprocessing is fed into a text detector, which employs the deep learning approach to mark the occurrence of text in an image. The text detector draws a bounding box with x and y coordinates around the text called region of interest (ROI) and outputs two arrays namely score and geometry. Score is an array identifying the probability of the text region. Score also maps the geometry that represents the bounding box locations in the input image. Non-maximum suppression (NMS) technique is used to filter out the regions that have the high score on intersection over union (IOU) than the threshold value with respect to all the regions identified. Extracted ROIs are then passed to LSTM deep learning text recognition algorithm.

The output of this process will undergo text analysis. The page segmentation mode namely *fully automatic page segmentation but no orientation and script detection* (OSD) mode is used by default. It is also found that the other available modes such as a *single uniform block of text treat the image as a single text line* also helpful to attain the text analysis. Scaling and padding are also added to balance the bounding box to fit the text rightly. Text lines are sequenced through blobs and are broken into words as per character spacing. A two-pass process is applied for text recognition. In the first pass, an attempt is made to recognize each word in turn. Word found satisfactory is then passed to an adaptive classifier as training data. The adaptive classifier is to recognize the text more accurately. The above process of text recognition is found not applicable for the text if skewed or rotated.

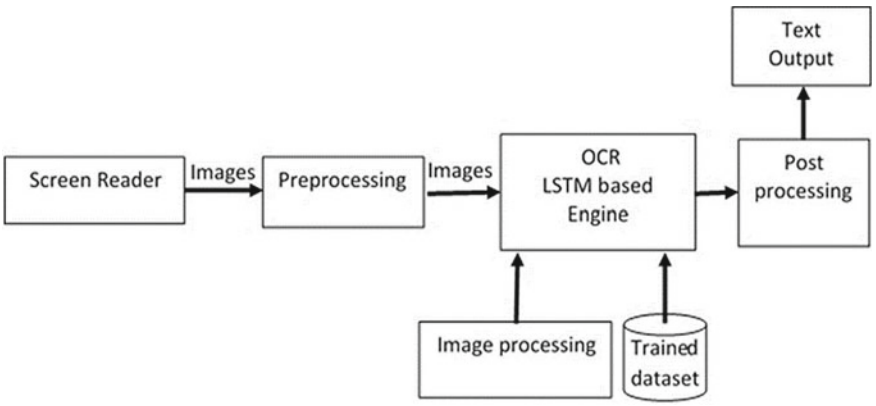


Fig. 1 Framework of OCR-based image processing

Table 1 Comparison of performance of CNN and LSTM

Method	Accuracy	Precision	Recall	$F1$ score	ROC_AUC
CNN	0.9819	0.9470	0.8667	0.9025	0.9306
LSTM	0.9824	0.9473	0.8670	0.9112	0.9346

Deep neural network-based architecture, such as *convolutional neural network* (CNN), and *long short-term memory* (LSTM) are compared for their performance on text classification. LSTM approach has given a better balance when was compared the number of iterations between precision and recall. This evaluation is performed on images on scanned documents or book pages. Five parameters to compare the performance of the techniques namely accuracy, precision, recall, $F1$-Score, and ROC_AUC are brought out between CNN and LSTM. Table 1 declares that LSTM-based text classifier gives a better performance.

3.2 Assess the Toxicity of Textual Contents from Web Sites

Second part of the framework of this paper is to evaluate the screen contents into normal and toxic. *Natural language processing* (NLP) provides a good source of solutions to text classification, which involves classifying text strings or documents into different categories, depending upon the contents of the strings. This model is based on supervised machine learning algorithm. Text classification tags the text data according to its content. *Python framework, PyCharm* [19] was helpful for experimenting the model with the addition of packages such as pickle, pandas, numpy, and nltk. Other packages of machine learning in Python are also utilized for predictive data analysis.

Figure 2 shows the text classification processes involved. Preprocessing of this data set such as HTML decoding, removal of stop words, convert text to lower case, removal of punctuation, anonymous characters are handled. Training data with 160,000 comments are used for training the classifier with labels. Training the classifier model is done where the training data is fed to the NLP algorithm with feature vectors and tag to produce the classification model. Identifying the harmful contents may sometimes involve multiple classes. To find an apt solution, linear support vector classifier (SVC), a text classification technique, is applied to perform the multiclass classification on the dataset [20]. Once the machine learning model is trained with dataset, the NLP model starts learning to apply precise prediction. Logistic regression performs on the document features to predict the outcome as 0 and 1. Figure 3 displays the output of the sample test data where a single label and multiple labels are identified by the model.

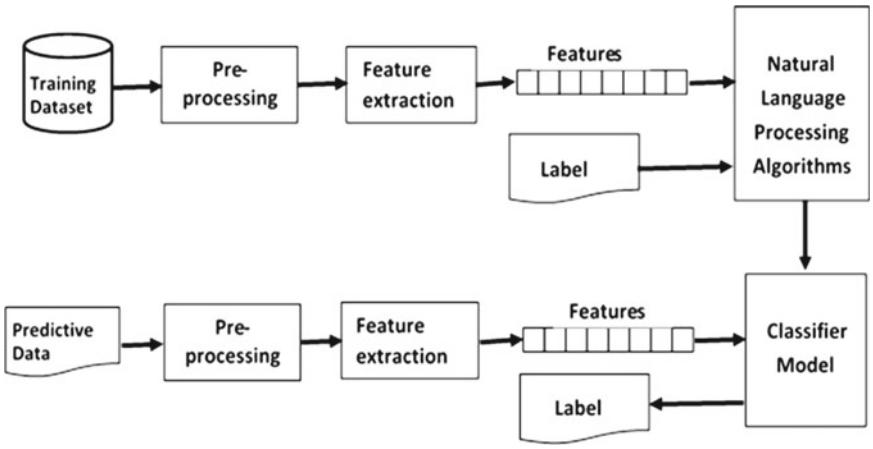


Fig. 2 NLP model for text classification

```
I am a good person.:[('Normal',)]
But I am rude with you.:[('toxic',)]
Sorry for hurting you.:[('Normal',)]
But you are a jerk.:[('insult', 'obscene', 'toxic')]
```

Fig. 3 A sample of classified output from testing data with single and multiple classes

4 Results and Performance Analysis

Combining the models discussed separately in Sects. 3.1 and 3.2, the resultant framework on identifying the toxicity level of the Web textual content is now analyzed for its prediction accuracy. Figure 4 shows the resulting classes as normal, sever toxic, obscene, threat, and insult when tested on another Web site contents.

To evaluate the performance of the classification model, the confusion matrix and $F1$ score are used. Precision is the ratio of true positive and total predicted positive whereas the recall is the ratio of total positive and the total actual positive. Precision shows how accurate the model behaves. To seek a balance between precision and recall, F1-score is used as given in Eq. (1).

$$F1 = 2 * (\text{Precision} * \text{Recall})/(\text{Precision} + \text{Recall}) \tag{1}$$

Table 2 shows the statistic results drawn out of the test data. It is inferred that the overall precision, recall, and F1-score accounted to 0.90, 0.885, 0.8922 are due to the fact that the text data are chosen with multiclass.

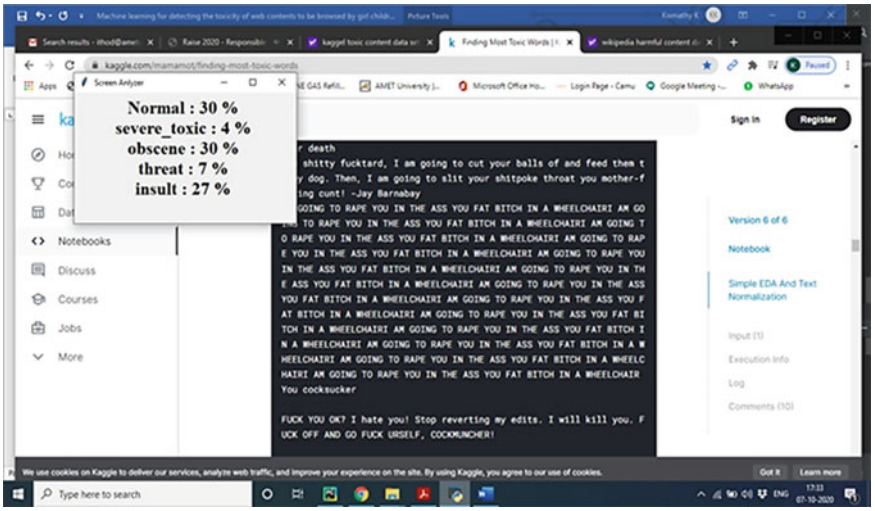


Fig. 4 Live prediction of the classes shows normal, severe toxic, obscene, threat, insult when applied on the Web site contents

Table 2 Statistic on precision, recall, and $F1$-scores

Class	Precision	Recall	$F1$-score
1	0.87	0.91	0.890
2	0.92	0.89	0.905
3	0.91	0.87	0.890
4	0.92	0.89	0.905
5	0.89	0.88	0.885
6	0.89	0.87	0.880

5 Conclusion

Referring to the objective of the paper, this paper attempts to emphasis a proactive measure to mitigate cybercrimes against girl children on line. As a proactive preventive measure, the law enforcement agencies, Internet service providers, information technologists, and data creators should join-in-hands with research to share their data and information sources and experiences to devise an authentication model, which can apprehend the toxicity level of the Web contents uploaded. The paper reviewed the situations when offenses are unreported and the victims are, repetitively, harassed by the cyberoffenders. Girl children, once they are alarmed of nonsensical, absurd, and toxic contents, they can simply withdraw from browsing the Web sites. The design of the dynamic model for detecting the toxic contents while browsing the Web resources by the girl children is experimented with a large trainer dataset. The classifier model is able to predict the class of the text with utmost precision. As a

future work, the performance of multiclass classification will be improved by an enhanced deep learning model.

References

1. Quayle, E., Loof, L., Palmer, T.: Child pornography and sexual exploitation of children online. World Congress III against Sexual Exploitation of Children and Adolescents. ECPAT International, Bangkok (2008)
2. Baines, D.V.: Online child sexual abuse: the law enforcement response. In: World Congress III Against the Sexual Exploitation of Children and Adolescents. ECPAT International, Rio de Janeiro, Brazil
3. Global status report on preventing violence against children. World Health Organization, Geneva (2020). https://www.who.int/teams/social-determinants-of-health/violence-prevention/global-status-report-on-violence-against-children-2020. Last accessed 2019/06/14
4. Report of the Independent Expert for the United Nations study on violence against children. United Nations, New York (2006). https://violenceagainstchildren.un.org/content/a61299. Last accessed 2020/06/20
5. In-depth study on all forms of violence against women: Report of the Secretary-General, A/61/122/Add.1. United Nations, New York (2006). https://www.un.org/womenwatch/daw/vaw/v-sg-study.htm. Last accessed on 2020/09/30
6. Pinheiro, P.S.: World Report on Violence against Children. United Nations Secretary-General's Study on Violence against Children, Geneva (2006)
7. Cybercrime Cases: National Crime Records Bureau (NCRB). https://data.gov.in/catalog/crime-india-2018?filters%5Bfield_catalog_reference%5D=6750431&format=json& offset=0&limit=6&sort%5B_score%5D=desc&query=Cyber+crime. Last accessed 2020/06/20
8. Cyberstalking: Stalking and violence: new patterns of trauma and obsession. Springer Kluwer Academic Publishers, Boston (2003)
9. A study on violence against girls. In: International Girl Child Conference. The United Nations Children's Fund (UNICEF), The Hague (2009)
10. Promoting gender equality through UNICEF-supported programming in child protection. UNICEF, New York (2011). https://www.unicef.org/gender/files/Protection_Layout_Web.pdf. Last accessed 2020/09/26
11. Sethi, D., Ghata, S.: Mitigating cyber sexual harassment, an insight from India. Asian Themes Soc. Sci. Res. **1**(2), 34–43 (2018)
12. Cyber Violence against Women and Girls: A Report by the UN Broadband Commission for Digital Development Working Group on Broadband and Gender. https://en.unesco.org/sites/default/files/genderreport2015final.pdf. Last accessed 2019/06/14
13. Sathyanarayana Rao, T.S., Bansal, D., Chandran, S.: Cyberbullying: a virtual offense with real consequences. Indian J. Psychiatry **60**(1), 3–5 (2018)
14. Ptaszynski, M., Masui, F., Kimura, Y., Rzepka, R.: Automatic extraction of harmful sentence patterns with application in cyberbullying detection. Human Language Technology. Challenges for Computer Science and Linguistics. LNCS, vol. 10930, pp. 349–362. Springer, Cham (2018)
15. Vivolo-Kantor, A.M., Martell, B.N., Holland, K.M., Westby, R.: A systematic review and content analysis of bullying and cyber-bullying measurement strategies. Aggress. Violent Behav. **19**(4), 423–434 (2014)
16. Smith, R.: An overview of the Tesseract OCR engine. In: IEEE Ninth International Conference on Document Analysis and Recognition ICDAR. IEEE Computer Society, vol. 02, pp. 629–633 (2007)s
17. Hochreiter, S., Schmidhuber, J.: Long short-term memory. Neural Comput. **9**(8), 1735–1780 (1997)

18. Zharikov, I., Nikitin, F., Vasiliev, I., Dokholyan, V.: DDI-100: dataset for text detection and recognition. https://github.com/machineintelligence-laboratory/DDI-100. Last accessed 2020/04/20
19. Pycharm Community Edition Version: 2020.2.3. https://www.jetbrains.com/pycharm/download/#section=windows. Last accessed 2020/04/20
20. Korba, K.A., Arbaoui, F.: SVM multi-classification of induction machine's bearings defects using vibratory analysis based on empirical mode decomposition. Int. J. Appl. Eng. Res. **13**(9), 6579–6586 (2018)

Machine Learning-Based Security Authentication for Wireless Multimedia Network

S. K. Gautham and Anjan K. Koundinya

Abstract For supporting various services in wireless multimedia networks, security is one of the major concerns. The traditional upper-layer authentication system and conventional security solution lack efficiency and does not solve the emerging security problems and ignores physical layer protection. To overcome these challenges, this paper presents a lightweight physical layer authentication system developed by understanding certain physical layer attributes in the wireless multimedia network through supervised machine learning algorithm. Support vector classifier (SVC) is used for classification. Classification can also be performed by other supervised algorithm and select based on each detection rate. Experimental analysis proved that the proposed system could protect wireless multimedia networks.

Keywords Security · Machine learning · Supervised algorithm · Wireless multimedia networks

1 Introduction

The new generation of connected society will be advancement of technologies like 5G and Internet of things [1–3]. The convergence of AI and multimedia system brings exciting services, and applications in the areas of smart home, health care, transportation, etc., were the applications that include training, monitoring, and entertaining [4].

As the complexity of multimedia system increases along with the number of multimedia sensors in the network system brings privacy and security challenges, for example, consider a scenario where a multimedia senor collects different types of data and a malicious node tricks the user by exploiting the interactivity and provides

S. K. Gautham · A. K. Koundinya (✉)
BMS Institute of Technology and Management, Bangalore, India
e-mail: anjank-cse@bmsit.in

S. K. Gautham
e-mail: 1by19scs02@bmsit.in

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_83

false information to the system [5]. This leads to security risks and causes life threatening consequences and leads to damages to the wireless network [6]. Therefore, to ensure safe communication transmission in wireless multimedia network, there is need for an effective protection system. However, new challenges arrive in security configurations if the network is large. And those multimedia devices which works on low-delay transmission will not support authentication techniques that will require high computational overhead, so the sensors to perform better communication will require a lightweight system. Therefore, in this paper, a new authentication system is been proposed that will defend against the attacks and challenges of conventional authentication systems.

1.1 Challenges of Conventional Security Systems

The main challenge for wireless multimedia networks based on conventional security system is high computing capabilities and resource constraints. Few challenges faced by conventional security systems are:

Unbalanced relation between low computational costs and high security

In traditional security systems, the security mechanism build on higher layers is difficult to handle the balance between the low computational cost and high security due to this malicious device gets an opportunity to steal sensitive data by forging public keys and bring serious damage to the authentication system. So, it would be beneficial if a machine learning-based lightweight authentication is available for multimedia environment.

Latency in real-time multimedia networks due to complex authentication steps

In traditional security systems, to increase the level of security, more efforts were given to extract complex features which lead to longer communication latency. So, whatever causes latency in a real-time system is always a problem. This latency also increases if the sensors in the multimedia network increases. Another problem is the manual selection of statistical characteristics which leads to time consumption and non-adaptive authentication process.

Therefore, there is need of lightweight authentication system in multimedia networks.

Inaccurate predictions during authentication

Traditional security system uses statistical hypothesis during testing which uses limited statistical properties which makes it difficult to establish an accurate detection model which leads to security threats in authentication systems. So, it is necessary to build an authentication system that does not depend on explicit programming.

In short, these challenges tell that there need for a new lightweight authentication system in wireless multimedia networks which can handle security issues where internet is everything.

1.2 Contributions

To address these challenges, we proposed a machine learning-based supervised algorithm to perform lightweight physical layer authentication, aiming to resolve security issues in multimedia devices. In summary, the contribution is as follows.

- Build a machine learning model based on supervised algorithm especially support vector machine (SVM) as a detection mechanism for wireless multimedia networks. The learning is based on exploiting the physical layer attributes. The reason for selecting physical layer attributes is because it is portable, efficient, and simple which does not cause any computational overhead or communication latency thus ensuring a lightweight authentication for the network.
- The performance and simulation results verify that the proposed system can significantly improve the authentication without much reduction in convergence.

2 Related Work

Security and machine learning are two in depended areas of research. But recently, interdisciplinary studies are going on between these two areas. The new networks intents to forsee new generation network architecture where there shall be support for all media types exchange where the privacy factor will be of paramount importance. A ML-based physical layer authentication is designed in [8] by exploiting multidimensional characteristics to defend against attacker in dynamic networks.

In [9], an ML-based authentication which exploits channel matrices and does binary classification which is threshold-free is proposed. To detect spoofers, [10] presents an algorithm that estimates the adaptive moment to train the authentication model and also ensures low latency. Reference [11] shows to analyze risk of unlabeled attributes based on the classified report, and the scheme is based on semi-supervised learning. In [12], a learning-based scheme is proposed for real-time authentication and identifying attacker from legitimate devices by exploiting the physical attributes. With better algorithm, better authentication system can be build using device and communication link characteristics and dependence between the channel [13, 14]. In [15], it is shown that lightweight authentication can also be build using mathematical models. The system presented [16] ensures secure wireless communication from adversarial attacks through feedforward network. The recent progress for security enhancement based on ML still has deficiencies to overcome the above challenges. More importantly, the conventional methods which depends on statistical properties leads inaccurate prediction which further leads to authentication failure, therefore focusing on providing a machine learning-assisted security scheme to overcome the challenges.

3 System Model

Figure 1 introduces to the system model of ML-based authentication system. Consider the scenario of wireless multimedia network, where several multimedia devices ($D_{1 \ldots n}$) are located at different location. The multimedia nodes collect the data and send it to the data collection center (DCC). The malicious node impersonates a legitimate device through eavesdrop and sends false messages to the data center for illegal activities. In data center, the model is trained in three phases.

- *Phase 1*: At first, the legitimate wireless devices are sent to the data center to get data about its corresponding physical layer attributes. Initially, the ML model is trained by training data like received signal strength (RSS) and its corresponding tags.
- In the data center, the model is trained by multiple physical layer attributes of different legitimate devices.

$$\boldsymbol{M}_{\text{legit}}(f) = [m_{\text{legit},0}(f), m_{\text{legit},1}(f), \ldots, m_{\text{legit},N-1}(t)]$$

where M_{legit} represents the estimated physical layer attributes, N represents the number of samples, and subscript legit means legitimate device.

- *Phase 2*: Using the estimated physical layer, the ML model is trained, and the trained model used is differentiated between the legitimate device and attackers during the authentication phase.

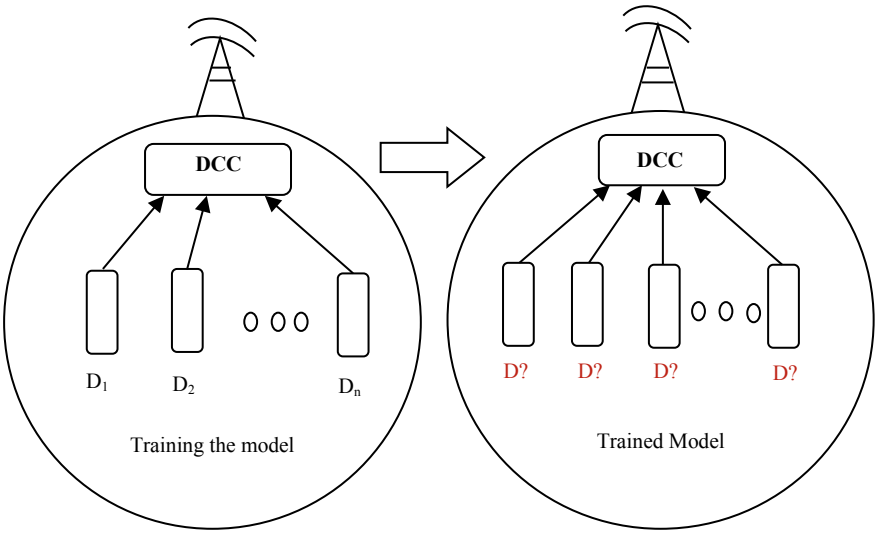


Fig. 1 System model of ML-based authentication scheme

$$\boldsymbol{M}_{\text{illegit}}(f+\tau) = \left[m_{\text{illegit}}(f+\tau), \ldots, m_{\text{illegit},N-1}(f+\tau)\right]$$

where M_{illegit} represents the attacker, τ represents the time interval between the two phases and N denotes the number of attributes.

- *Phase 3*: In this phase, the training data set used for modeling is updated. The training dataset is updated with the newly arrived physical layer attributes of legitimate devices. The updated dataset is again used to train the model and this cycle continues.

$$\boldsymbol{M}_{\text{legit}}(f) \leftarrow \boldsymbol{M}_{\text{legit}}(f+\tau)$$

With the trained model which is lightweight authentication system, the data collection center will be able to identify signals from illegitimate devices sent to the data center trying to mislead the behavior of the system and therefore the data collection center will reject the abnormal multimedia data.

4 Machine Learning-Based Scheme

This paper studies about a lightweight authentication system based on machine learning algorithm to meet the requirement of wireless multimedia networks such as low latency and higher accuracy. The ML model not only enhance the authentication but also provides retraining for the detecting model.

Security Authentication Based on SVM

The authentication model is trained based on support vector machine and multiple physical layer attributes. The data collection center on receiving data from the devices checks whether the characteristic of the channel, i.e., the attributes are legitimate or not and this can be done by the trained model, thereby the spoofing attackers can be detected by the trained model using the obtained authenticator. To improve the security, the ML-based algorithm, i.e., support vector machine generates an authenticator in the data collection center. To remove the problems of key distribution or management, the model ensures an endogenous security system by exploiting the channel characteristics. Due to proliferate and complex process in place for multimedia networks this shall impact the QoS parameters leading to higher laterncy. Therefore, a lightweight authentication scheme is placed in the data center. This method could avoid possible covert channel or hybrid covert channel [18] and thereby establishing secured communciation.

5 Performance Evaluation

To verify the security performance of the ML-based authentication, SVM-based authenticator is used. The process is based on using few physical layer attributes like received signal strength (RSS), distance between adjacent signal (DAS), and their corresponding Pearson correlation coefficient (PCC).

Figure 2 shows the confusion matrix of support vector classifier based on three attributes: RSS, DAS, and PCC. From the true positive and true negative, the authentication predication is right. It is also seen that with increase in the number of channel attributes it is expected that the detection performance increases. It is also observed that when the number of iterations increases, the rate of detection also increases.

Table 1 shows the classification report which gives information about the parameters like precision, recall, F1-score, and support of the SVM classifier. With the increase in the iteration and attributes, the detection rate increases. The detection

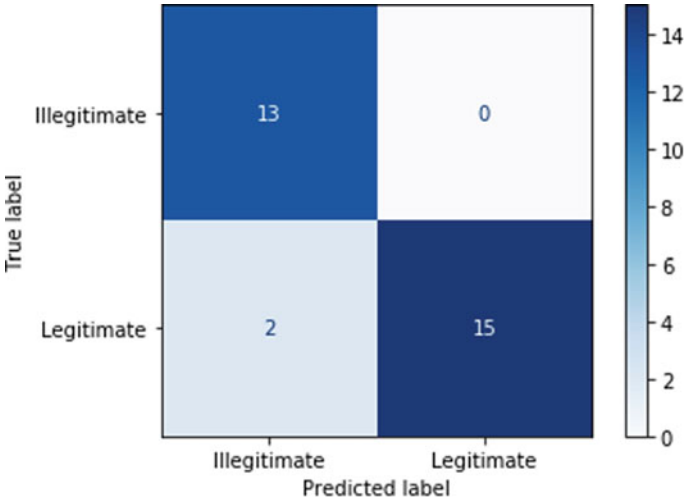


Fig. 2 Confusion matrix of SVC

Table 1 Classification report

	Precision	Recall	$F1$-score	Support
Illegitimate	0.87	1.00	0.93	13
Legitimate	1.00	0.88	0.94	17
Accuracy	–	–	0.93	30
Macro-avg	0.93	0.94	0.93	30
Weighted avg	0.94	0.93	0.93	30

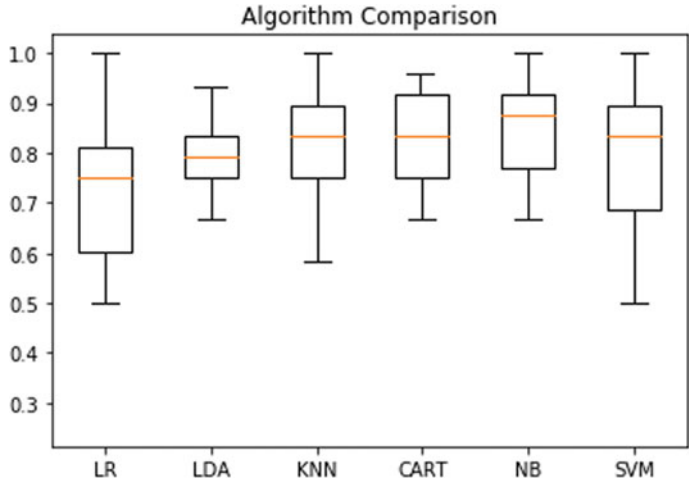


Fig. 3 Boxplot of algorithm comparison

rate of the SVM classifier is 94% which is confirmed by the $F1$-score in the classification report and confirms that it is a good authentication system and makes it difficult for the attackers. The time complexity of n authentications is $O(1)$.

Figure 3 shows the boxplot of different algorithm comparison. The different supervised ML algorithms (classifiers) used are logistic regression (LR), linear discriminant analysis (LDA), k-neighbors classifier (CART), Gaussian NB (NB), and support vector classifier with gamma value set to auto (SVM). Any of the classifier can be used to classify the authentication process. It is obvious to stick with SVM classifier, but other classifiers can also be checked because certain classifier shows better classification result when the number of attributes or attribute feature increases. There is also changes in detection rate with changes in iteration performed by different classifiers.

6 Future Scope

The challenges faced by convectional security schemes in Sect. 1 and discussed an authentication mechanism in Sect. 4, which gave an idea to study the physical layer attributes and based on the study use a ML-based security system to improve the detection rate, schemes that is based on statistical methods had problem in authentication especially by adversarial attacks, so the important aspect of authentication scheme is signal preprocessing. So, for the improvement in signal preprocessing, it is apt to select deep learning, which gives method for selecting physical layer attributes and also solves the problem of uncertainty in a wireless network. Deep learning tools

provide a data analytic-centric security scheme, where more importance is given to the data rather the model.

The conventional authentication technique is suitable for an ideal communication environment but not for a realistic interfered environment. So, from a practical perspective there is need for an adaptive authentication model. So, the changing communication environment and assumptions give direction for future research.

One direction for future research is blind feature learning, which develops a authentication scheme by exploring neural network to learn much more about the characteristics of the data collected at the data collection center.

7 Conclusion

The proposed machine learning-based authentication system for wireless multimedia networks has very good significance in real world. It not only provides privacy but also ensures a lightweight authentication system for multimedia networks. The model can be enhanced by increasing the number of attributes or features but it is time consuming for manually selecting it. The experimental analysis shows that the ML-based scheme effectively learns the physical layer attributes and provides a good detection rate and a short communication latency. It has also been found that with increase in the number physical layer attributes the authentication accuracy and the detection rate increase dramatically.

References

1. Wu, D., Zhang, Z., Wu, S., Yang, J., Wang, R.: Biologically inspired resource allocation for network slices in 5G-enabled internet of things. IEEE Internet Things J. **6**(6), 9266–9279 (2019). https://doi.org/10.1109/JIOT.2018.2888543
2. Li, Z., Liu, H., Wang, R.: Service benefits aware multi-task assignment strategy for mobile crowd sensing. Sensors **19**(12) (2019)
3. Li, Z., Jiang, Y., Gao, Y., Sang, L., Yang, D.: On buffer-constrained throughput of a wireless-powered communication system. IEEE J. Sel. Areas Commun. **37**(2), 283–297 (2019). https://doi.org/10.1109/JSAC.2018.2872374
4. Wu, D., Si, S., Wu, S., Wang, R.: Dynamic trust relationships aware data privacy protection in mobile crowd-sensing. IEEE Internet of Things J. **5**(4), 2958–2970 (2018). https://doi.org/10.1109/JIOT.2017.2768073
5. Wu, D., Deng, L., Wang, H., Liu, K., Wang, R.: Similarity aware safety multimedia data transmission mechanism for Internet of Vehicles. Future Gener. Compute. Syst. **99**, 609–623 (2019)
6. Fang, H., Qi, A., Wang, X.: Fast authentication and progressive authorization in large-scale IoT: how to leverage AI for security enhancement? (2019). arXiv:1907.12092
7. Qiu, X., Jiang, T., Wu, S., Hayes, M.: Physical layer authentication enhancement using a gaussian mixture model. IEEE Access **6**, 53583–53592 (2018). https://doi.org/10.1109/ACCESS.2018.2871514

8. Wang, N., Jiang, T., Lv, S., Xiao, L.: Physical-layer authentication based on extreme learning machine. IEEE Commun. Lett. **21**(7), 1557–1560 (2017). https://doi.org/10.1109/LCOMM.2017.2690437
9. Pan, F., et al.: Threshold-free physical layer authentication based on machine learning for industrial wireless CPS. IEEE Trans. Industr. Inf. **15**(12), 6481–6491 (2019). https://doi.org/10.1109/TII.2019.2925418
10. Liao, R.F., Wen, H., Wu, J., et al.: Deep-learning-based physical layer authentication for industrial wireless sensor networks. Sensors (Basel) **19**(11):2440. Published 28 May 2019. https://doi.org/10.3390/s19112440
11. Gan, H., Li, Z., Fan, Y., Luo, Z.: Dual learning-based safe semi-supervised learning. IEEE Access **6**, 2615–2621 (2018). https://doi.org/10.1109/ACCESS.2017.2784406
12. Chatterjee, B., Das, D., Maity, S., Sen, S.: RF-PUF: enhancing IoT security through authentication of wireless nodes using in-situ machine learning. IEEE Internet Things J. **6**(1), 388–398 (2019). https://doi.org/10.1109/JIOT.2018.2849324
13. Qiu, X., Jiang, T., Zou, W.: Physical layer security in simultaneous wireless information and power transfer networks. In: 2017 17th International Symposium on Communications and Information Technologies (ISCIT), Cairns, QLD, pp. 1–4 (2017). https://doi.org/10.1109/ISCIT.2017.8261231
14. Qiu, X., Jiang, T.: Safeguarding multiuser communication using full-duplex jamming receivers. In: 2017 IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC), Montreal, QC, pp. 1-5 (2017). https://doi.org/10.1109/PIMRC.2017.8292264
15. O'Shea, T., Hoydis, J.: An Introduction to deep learning for the physical layer. IEEE Trans. Cogn. Commun. Netw. **3**(4), 563–575 (2017). https://doi.org/10.1109/TCCN.2017.2758370
16. Qiu, X., Jiang, T., Wang, N.: Safeguarding multiuser communication using full-duplex jamming and Q-learning algorithm. In: IET Communications, vol 12, no. 15, pp. 1805–1811, 18 9 2018. https://doi.org/10.1049/iet-com.2018.0059
17. Wang, X., Cheng, W., Mohapatra, P., Abdelzaher, T.: Enabling reputation and trust in privacy-preserving mobile sensing. IEEE Trans. Mob. Comput. **13**(12), 2777–2790 (2014). https://doi.org/10.1109/TMC.2013.150
18. Koundinya, A.,Abraham, J.: Behavioral analysis of transport layer based hybrid covert channels, In Proceedings of 3rd CNSA, Springer Verlag published, June 2010 Chennai, India, pp. 83–92

Crime Data Analysis Using Association Rule Mining Technique

Vijendra Kamble, Aniket Muley, Nilesh Deshmukh, Parag Bhalchandra, Nitin Darkunde, Govind Kulkarni, and Pritam Tamsekar

Abstract Law authorization organizations utilize different crime analysis tool to find crimes. However, a major volume of crime data has made the way toward analyzing crime difficulties. Here, proposed technique is for improving the efficiency of the police personnel and police stations. The proposed methodology is used to advancement in the crime type data analysis with apriori algorithm in association rule mining. Techniques should be additionally valuable for police personals proposed a user-oriented crime analysis in Nanded city. The rule obtained in this study helps to identify the location and types wise crime associated cases. In future, it will help to deal with and minimize the most frequent cases.

Keywords Crime analysis · Association rule mining · Data mining · Discretization · Apriori

1 Introduction

Data mining is the way toward discovering Patterns in enormous data collections including strategies at the convergence of machine learning, statistics, and database frameworks. Data mining is an interdisciplinary subfield of programming designing and statistics with an overall target to isolate information from a dataset and change the information into a comprehensible structure for additional utilization [1]. Data mining is the investigation adventure of the knowledge discovery in database system or KDD. Next to the unrefined analysis step, it furthermore incorporates database

V. Kamble (✉) · N. Deshmukh · P. Bhalchandra · P. Tamsekar
School of Computational Sciences, Swami Ramanand Teerth Marathwada University, Nanded 431606, India

A. Muley · N. Darkunde
School of Mathematical Sciences, Swami Ramanand Teerth Marathwada University, Nanded 431606, India

G. Kulkarni
Department of Computer Science, Late Laxmibai Deshmukh Mahila Mahavidyalaya, Parli-Vaijjnath 431515, India

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_84

and data execute points, data pre-dealing with, model and determination considering, fascinating quality measurements, unconventionality, post-gathering of discovered structures, representation and online updating. The data mining is only the way that the goal is the extraction of patterns and data from huge amount of information, not just the mining of data [2]. It moreover is constantly applied to a huge scope data preparing similarly as any utilization of decision support system, including computerized analysis and business information. practical machine learning tools and procedures with Java was at first to be named just simply machine learning and the term data mining was included for displaying reasons. The real data mining task is the self-loader or customized analysis of huge amount of data to separate already unknown, exciting patterns, viz., social information records, and conditions [3]. This typically includes utilizing database techniques, for example, spatial records. These patterns would then be able to be viewed as a sort of summary of the input data, and might be utilized in further analysis or, For instance, the data mining step may distinguish various collections in the data, which would then be able to be utilized to acquire increasingly exact prediction results by a decision support system. The dissimilarity between data analysis and data mining is that data analysis is utilized to test models and theories on the dataset. Data mining utilizes machine learning and statistical models to reveal covert or hidden patterns in an enormous volume of data.

The objectives of this study are as following:

- To study and analyze the patterns of crime.
- To develop statistical queries and analysis tool.
- To develop crime analyzing tool in accordance with demographic features.
- To identify and locate hotspot location.

In subsequent sections, methodology, result and discussion, and conclusion are elaborated in detail.

2 Methodology

2.1 Study Area

The Geographical area of Nanded city situated in Maharashtra (Fig. 1) and co-ordinate of Nanded city somewhere in the range of 18° 0.15 and 19° 0.55′ North scope and 77° 0.7′ to 78° 0.15′ east longitudes (Figs. 2 and 3). Zone of Nanded city goes under Nanded Waghala Municipal Corporation (NWMC) surroundings is close to 51.76 km^2. Nanded city is distinct in two parts as: old Nanded which covers (25.00 km^2) situated on the north bank of Godavari River. The present investigation has been in old Nanded are blended and considered as a solitary region for this analysis. Four police stations have been chosen which incorporate north Nanded zone for the analysis according to the data accessibility.

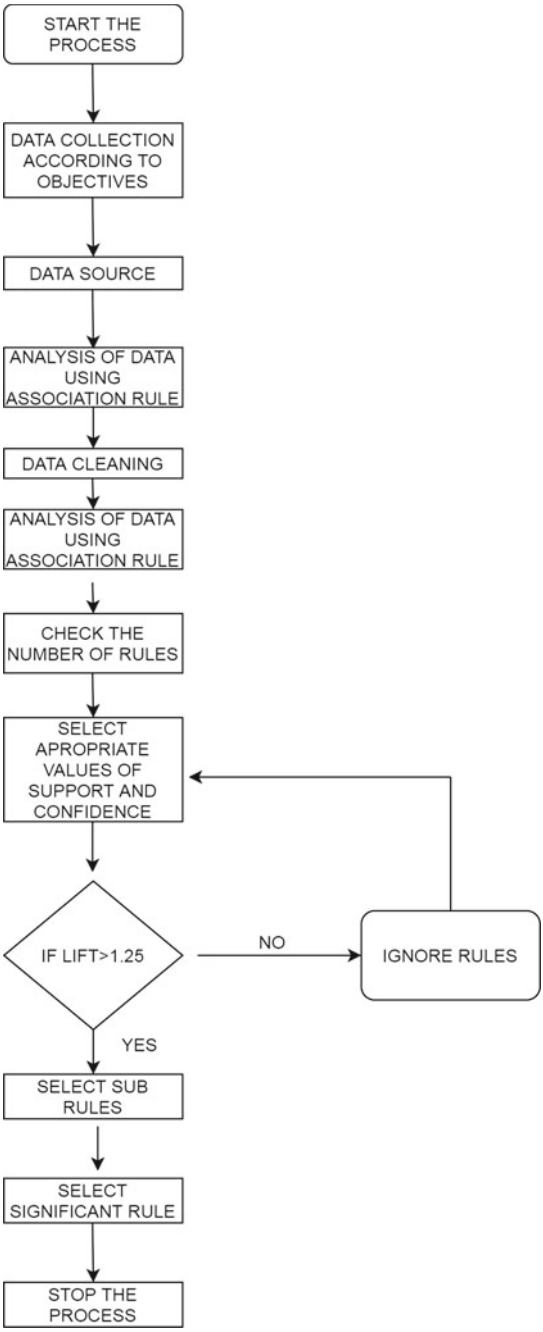


Fig. 1 Workflow of study

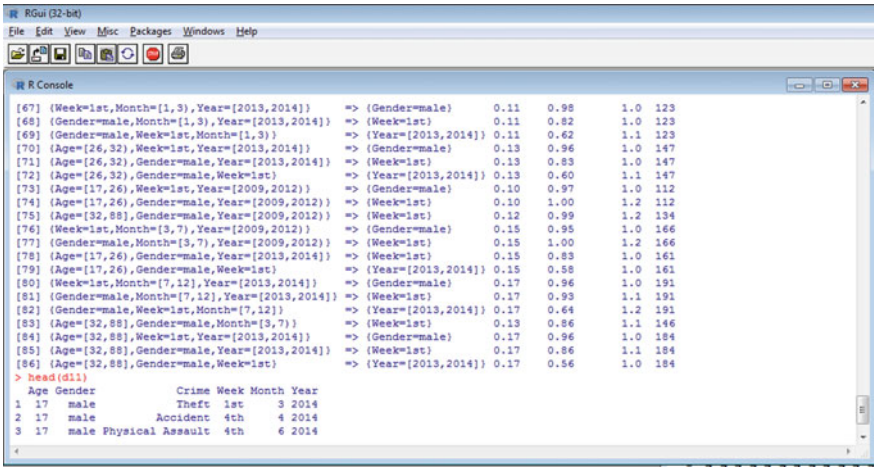


Fig. 2 Screenshot of association rule generated by *R*

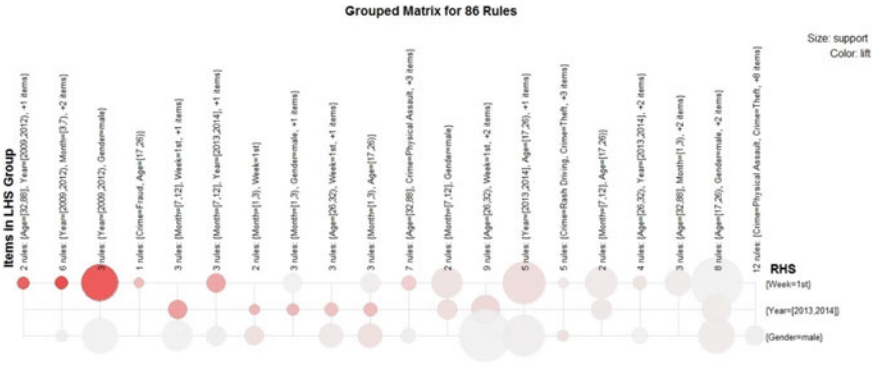


Fig. 3 Group matrix of generated associated rule

2.2 Data Collection

Crime data is the fundamental needs of a crime analysis which give the spot, date, and sort of crime which are in charge of investigation of crime. For the standard analysis of crime each police station gives multiyear FIR data. Those FIR which enlisted in police station can exceptionally make reference to what number of crime are occurred in those four police stations. The crime type was distinguished based on benchmarks techniques. Crime type is recognized from various site of police station locale region with arbitrary strategy. Crime parameters in such a way that to incorporate all kind of crime which is broke down according to standard strategy [4]. The investigation of this crime type is significant which make solid impact on the analysis of crime. From Bhagyanagar police station, we have gathered three years

of FIR data, limit maps of the police station and police faculty subtleties with their assignment. Bhagyanagar police station has around 4.67 square kilometer region. There is 107 police staff in Bhagyanagar police station. Also, having thickness of police workforce is about 8.11 per km^2 region. From Shivajinagar police station we have gathered two years of FIR data, limit of boundary line of the police station and police staff subtleties with their assignment. Shivajinagar police station has around 4.67 km^2 region. There are 106 police workforce in Shivajinagar police station. Also, having thickness of police staff is about 22.71 km^2 region. From Itwara police station we have gathered two years of FIR data, limit map of the police station and police workforce subtleties with their designation. Itwara police station has around 3.16 km^2 region. There are 110 police workforce in Itwara police station. What's more, having thickness of police workforce is about 34.82 per km^2 zone.

The primary source of the data used to hold this research was criminal authentic data taken from Nanded city from four police stations. The criminal database of the Nanded city contains in excess of 1500 records physically. As the record official displayed, this is a direct result of the deficiency of computer talented police persons and there are no database managers available in the police station. Subsequently, all FIR details were previously encoded in an Excel sheet. After the data was encoded, the whole dataset is placed in one document having numerous records. Each record associates with most legal data of one criminal [5]. Next, pre-handling strategies were applied to make it fitting for mining purpose. Table 1 represents list of attributes under study.

As analyzed before in data mining there are various types of standard techniques used for representing problems, however for this analysis, apriori algorithm under association rule mining is used, in request to produce huge data mining models for an analysis of crime activities [6]. Likewise, over the span of this analysis, the idea used to give an organized way to deal with networking a data mining project, its adaptability and its convenience when utilizing study to take care of critical diverse

Table 1 List of attributes used in this study

S. No.	Attribute	Possible value	Description
1	Age	15–20, 20–30, 30–40, 40–50, 50–60	Age of offender
2	Gender	Male and female	Sex of offender
3	Crime type	26	List of crimes
4	Time	(AM, PM)	Time of crime done in PM or AM
5	Week	Four or five week of the month	Weekly crime of the month
6	Date	Date wise crime	Crime date
7	Month	Every month of the year	Monthly crime of the year
8	Spot	Crime hotspot	Spot where actual crime happened
9	Year	Year wise crime	Year wise crime rate
10	Police station	4 Police station	Per police station wise crime

association rule mining. For the most part, this section manages the distinguishing proof of source of data, data cleaning and preprocessing of the data utilized in this analysis. Endeavors applied and procedures utilized during the time spent data structure and model design is additionally displayed. In addition, it presents consequences of the experiment and its interpretation.

2.3 Apriori Algorithm

Apriori is an algorithm for frequent item set mining and association rule learning over relational data. It proceeds by recognizing the consistent individual items in the database and extending them to greater and greater item sets as long as those item sets show up routinely in the database. Apriori uses a base-up approach, where subsets are extended every items thus a phase known as candidate generation, and assortments are attempted against the database. The algorithms end when no further powerful expansions are found. Apriori uses broadness first inquiry and a Hash tree structure to check up-and-comer item sets proficiently. It produces contender item sets of length from item sets of length. At that point, it disagrees with the candidate which has clashing sub patterns. As demonstrated by the descending lemma, the candidate set contains all progressive length item sets. Starting their forward, it patterns the exchange database to choose visit item sets among the contenders. The pseudo code for the algorithm is given underneath for an exchange database, and an assistance item set hypothetical documentation is used; anyway note that is multisets is the candidate set for level. At every progression, the algorithm is acknowledged to deliver the candidate sets from the tremendous item sets of the prior level; focusing on the diving descending conclusion lemma gets to a field of the data structure that addresses candidate set, which is at first thought to be zero. Numerous subtleties are redundant underneath, more often than not the most significant piece of the execution is the data structure utilized for putting away the competitor sets, and checking their frequencies [7].

Apriori is an algorithm for frequent item set mining and association rule learning over crime databases. It proceeds by recognizing the frequent individual items in the database and extending them to greater and greater item sets as long as those item sets show up oftentimes in the dataset. The progressive item sets dictated by Apriori can be used to choose association rules which highlight general patterns in the database. This has applications in area, for instance, crime analysis [1].

- Data Preprocessing: A data preprocessing is a procedure that comprises of data cleaning, data integration and data transformation which is typically prepared by a personal computer program. It means to shrink a few noises, incomplete data. The result from preprocessing step can be continuing by data mining algorithms.
- Data cleaning: fill in lost character, soft noisy data distinguish or evacuate exception, and determine irregularities there are numerous unassuming recommendations for satisfying absent character.

Diverse preprocessing procedures were utilized to get clean data, these contain:

- Removing irregularity: a portion of the records in the crime, criminal and geo-crime datasets speak to exceptions and can't be remembered for the analysis algorithm and techniques, so these data records were erased from the set.
- Insertion missing data: criminal age, crime name, and area name were not referenced in tables, normal and most generally utilized values were utilized to alternate these missing value.
- Data Integration: using different databases, data files as our data is collected from considerable police stations in this way, the data are coordinated to build three datasets.
- That by consistent the selected attributes and uniforms the type of significant attributes.
- Data transformation: standardization and accumulation in our research where the association rule is the pre-owned technique for mining the attributes are altogether changed over to letters appears if agree with attributes condition and disappear if not agree.
- Data reduction: reduction the volume however producing the equivalent or comparative analytical result in this research diminishing completely in two faces these are: face of excluding some unwanted character, viz., criminal wearing which cloths the criminal could change it. The later face is discarding whole records since they have in excess of three missing qualities so the filling will cause noisy.
- Data discretization: some portion of data reduction, replacing numerical attributes with supposed ones in this research the attributes of age as opposed to filling it as number will be written by letter, for example, A, where A will show up if criminal age over 40 years and vanish if age smaller [8].

2.4 Proposed Algorithm

Association rules investigation is a system to reveal how things are related to one another. There are three basic approaches to determine association.

Support: This says how famous an item set is, as estimated by the extent of exchanges wherein a crime type shows up.

$$\text{Support } \{\text{item}\} = \frac{\text{Occuance of item}}{\text{Total number of trasaction}}$$

Confidence: This says how likely robbery is happened under Shivajinagar police station, expressed as {robbery → shivajinagar police station}. This is measured by the proportion of cases of robbery, in which happened under Shivajinagar police station appears.

$$\text{Confidence}\{\text{robbery} \rightarrow \text{shivajinagar PS}\} = \frac{\text{Support}\{\text{Robbery,Shivajinagar PS}\}}{\text{Support }\{\text{Robbery}\}}$$

One disadvantage of the confidence is that it might misrepresent the importance of an association between the items. This is because it only record for how many times robbery cases happened, but not police station. If Shivajinagar police station having more robbery cases in general, there will be a higher chance that a rule containing robbery crime will also contain Shivajinagar police station, thus inflating the confidence measure.

Lift: says how likely when we considered Shivajinagar police station at the same time robbery cases are highly registered, while controlling for how shivajinagar police station associated with robbery crime type is. The lift of {attempt to murder → itwara PS} is 1, which displays no association between items. A lift value greater than 1 means that police station is likely to be associate if particular crime is registered, while a value less than 1 means that police station is unlikely to be associate if crime is not registered.

$$\text{Lift }\{\text{type of crime} \rightarrow \text{police station}\} = \frac{\text{Support}\{\text{type of crime, police station}\}}{\text{Support}\{\text{type of crime}\} \times \text{Support}\{\text{police station}\}}$$

Association principle mining is an alluring theme of research in the field of data mining. However, association rule mining is still in a phase of investigation and improvement. There are still some fundamental issues that should be read for recognizing valuable association rules.

Steps of algorithm are as follows:

k = 1.
Fk = { ili ∈ I ∧ σ({i}) ≥ N × minsup}.{Find all frequent 1-itemsets}
repeat
k = k + 1.
Ck = candidate-gen(Fk−1).{Generate candidate item sets.}
Ck = candidate-prune(Ck, Fk−1). {Prune candidate item sets.}
for each transaction t ∈ T do
Ct = subset (Ck, t). {Identify all candidates that belong to t.}
for each candidate itemset c ∈ Ct do
σ(c) = σ(c) + 1. {Increment support count.}
end for
end for
Fk = {clc ∈ Ck ∧ σ(c) ≥ N × minsup}.{Extract the frequent k-item sets.}
until Fk = Ø
Result = U Fk

The workflow our study is represented in Fig. 1. It elaborates the direction of mining the data analysis.

3 Results and Discussion

It is observed that mostly all types of crime done by the male. It is observed that Fraud and theft crime mostly done in 1st week of every month. Rash driving crime mostly happened at Shrinagar, Bhagyanagar which comes under Bhagyanagar police station area. Mostly male offender has found in rash driving cases and they were below 35 year age group. In the year 2012 and 2013, Bhagyanagar police station had registered 133 different cases in their FIR book. Theft is crime type which is mostly done by the male. Theft crime happened at the time of afternoon and evening. It is done in the first second and third week of every month in Shivajinagar as well as Bhagyanagar police station and the cased count is 179.

In the year 2013 and 2014, there are most of the cases registered in the month of January and February. Station and it is observed that mostly theft crime was happened. Mostly Physical assault crime done by the male and the age group of the offender is in between 15 and 86 years. In age group of 13 to 25, male offender has found mostly, rash driving and almost all types of crime done by this age group. The cases are mostly recorded in Bhagyanagar police station in the year 2009, 2010, and 2011. In the year 2003 and 2012 there are 366 cases are recorded in Bhagyanagar police station. In age group of 25 to 32, there are 374 cases are recorded and done by the male among that most of the cases physical assault and fraud cases were found. In the year 2013 and 2014, 592 cases are recorded and their offender was male only. This period is election period. In Bhagyanagar police station 818 cases are recorded and have been found to be the highest crime recorded area. Theft crime mostly happened in the Bhagyanagar police station the count is 139. Figures 3 and 4 represents the rules generated and bubble graph of rules, respectively.

4 Conclusion

Most of the crimes happened under Bhagyanagar police station. It means at the time of 12:00 pm or in afternoon patrolling time should be increased. These areas are full of coaching classes that's why the age of the offender is teenager. They are attracting someone by speedy driving there bike. Bhagyanagar area having dense population and number of private coaching classes are due to nearby colleges in that area. So that's why the age of the offender is teenager. They are newly started to drive the vehicle so they are not able to control their speed of the vehicle and as per their age they driving their bike speedily. Most of the crimes were recorded in the month of January to May. Further, it is observed that theft and fraud cases are recorded consistently. Theft crime happened at the time of afternoon and evening. It is done in the first second and third week of every month in Shivajinagar as well as Bhagyanagar police station. It is observed that attempt to murder cases are mostly done by male and were recorded under Bhagyanagar area and the incident mostly happened in the evening. Also, it is observed that majorly it happened between 3rd

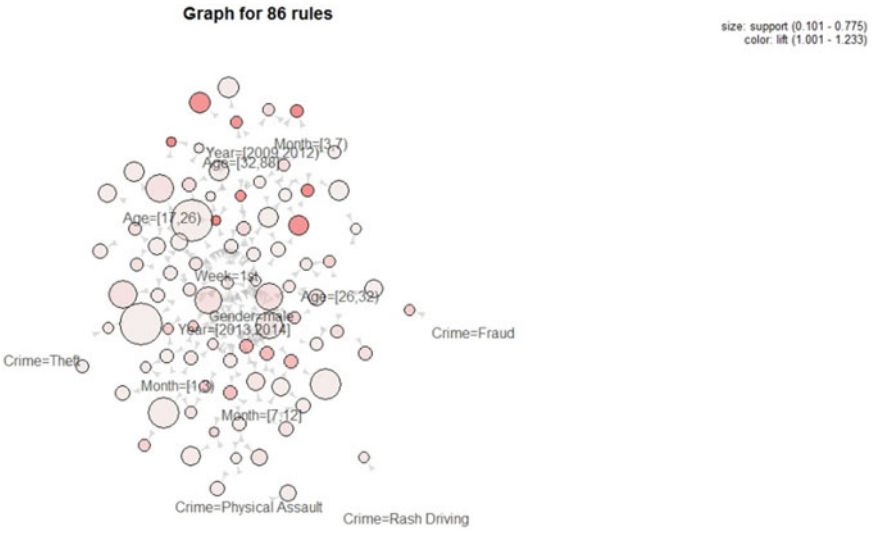


Fig. 4 Bubble graph of generated associated rule

and 5th weeks of every month. Theft crime type of crime is mostly happened in the Bhagyanagar police station. In the month of first and third of 2013 and 2014, it is observed that more than 300 cases are recorded in Bhagyanagar police station. Mostly Physical assault crime done by the male and in age group of 13–25 male offenders has found mostly. The cases are mostly recorded in Bhagyanagar police station in the year 2009, 2010, and 2011. In most of the cases, it is observed that 25 to 32 age group men were mostly involved in physical assault and fraud cases were found. In the year 2013 and 2014, 592 cases are recorded and their offender was male only and at the time of election. In Bhagyanagar police station, it has been found to be the highest crime recorded area. In future, it will help to deal with and minimize the most frequent cases in the respective areas.

References

1. Agarwal, A., Chougule, D., Agarwal, A., Chimote, D.: Application for analysis and prediction of crime data using data mining. In: Proceedings of IRF-IEEE forum International Conference, pp. 35–38 (2016)
2. Bharathi, A., Shilpa, R.: A survey on crime data analysis of data mining using clustering techniques. Int. J. Adv. Res. Comput. Sci. Manag. Stud. **2**(8), 9–13 (2014)
3. Jani, V.H.: Criminal pattern identification using genetic algorithm. Int. J. Sci. Res. Dev. **3**(4), 1345–1348 (2015)
4. Kumar, M.V., Chandrasekar, C.: Crime Hotspot detection using spatial Clustering: a literature review of related work. Int. J. Adv. Res. Comput. Sci. **1**(3), 415–417 (2010)
5. Sadasivam, V.R., Duraisamy, K., Bharathi, R.M.: Association rule mining and frequent pattern mining applications on crime pattern mining: a comprehensive survey. Int. J. Innov. Res. Sci.

Eng. Technol. **4**(6), 228–236 (2015)
6. Agrawal, S., Sejwar, V.: Frequent pattern model for crime recognition. Int. J. Comput. Intell. Res. **13**(6), 1405–1417 (2017)
7. Jabar, E.K., Hashem, S.H., Enas, M.: Propose data mining AR-GA model to advance crime analysis. IOSR J. Comput. Eng. **1**(5), 38–45 (2013)
8. Saeed, U., Sarim, M., Usmani, A., Mukhtar, A., Shaikh, A.B., Raffat, S.K.: Application of machine learning algorithms in crime classification and classification rule mining. Res. J. Recent Sci. **4**(3), 106–114 (2015)

Impact of Service Quality Factors on Customer Word of Mouth at Commercial Banks in the Highlands, Vietnam: The Mediating Role of Customer Satisfaction

Van Thuy Nguyen, Thi Huyen Trang Nguyen, and Thi Xuan Binh Ngo

Abstract This study analyzes the influence of service quality factors on word of mouth (WOM) on commercial banks in the Highland. Data used in the study was collected from 299 customers who are using commercial bank services in Central highland, Viet Nam. The research results show that when customers have empathy, reliability and the Bank's service capacity, they will be more pleased with the service quality, from which they share their experiences about use of the Bank's products and services to friends or relatives. The element of empathy, service capacity and reliability have a strong effect on customer satisfaction and the customer satisfaction factor also has a positive effect on word of mouth. In particular, empathy has a stronger impact on word-of-mouth activity than the other two elements.

Keywords Service quality · Customer satisfaction · Word of mouth · Highland · Vietnam

1 Introduction

Word of mouth is seen as a means of transmitting and storing messages from generation to generation. Word of mouth does not only convey good values, but it also makes it more expensive to build an image for a person or something. Today's word of mouth is no stranger to people, and especially to businesses and organizations who use word of mouth to build their image.

Quality is likely to have a crucial role in the service sector, especially in the retail sector, because service quality is critical to the organization's survival and profitability. Nowadays, quality of service and customer satisfaction are very important in almost all industries' service. Service quality has an impact on customer satisfaction that has a beneficial effect on word of mouth. Customers who use the company's products and services will share positive information about their businesses with their friends, family and relatives.

V. T. Nguyen (✉) · T. H. T. Nguyen · T. X. B. Ngo
Banking University Ho Chi Minh City, Ho Chi Minh City, Viet Nam
e-mail: thuynv@buh.edu.vn

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_85

Ioannis and Constantine [1] had explored the influences of service quality on WOM and satisfaction to mothers in Greece. The authors use the SME to analyze data collected from 1000 mothers who gave birth in 5 years. They indicate that, apart from satisfaction factor, empathy personnel is the single element which straightly has influences on word of mouth. The remaining factors indirectly affect word of mouth through satisfaction. With the topic "Service quality, corporate image and word of mouth in the domestic aviation industry of Vietnam," it was studied by Cong and Trang [2]. This study aims to investigate the elements effecting the quality of aviation services to the oral information through the image of domestic aviation enterprises in Vietnam. The author surveyed 298 passengers using domestic air services and selected 4 independent variables: convenience, reliability, customer service and in-flight service. The analytical results of the above 4 independent variables do not directly affect customers' word of mouth but indirectly through the image of the business. The article has shown that the relationship between service quality and customer satisfaction is always an urgent topic in the research. This study aims to evaluate the factors of service quality to customer satisfaction and word of mouth by commercial banks in the Central Highlands, Vietnam.

2 Literature Review

Service quality is a focused analysis that reflects customer perceptions of specific aspects of service: reliability, responsiveness, assurance, empathy, and tangible. Based on evaluations of the quality of services provided to customers, businesses can quickly identify problems, improve their services, and better evaluate customer expectations. The concept of service quality and debated by many authors such as Parasuraman et al. [3] is that it is the gap between expectation and perception of customer and Zeithaml [4] suggest that service quantity is the customer's evaluation of an entity's superiority and general excellence. The scope of the service quality concept is broad, complex and multidimensional [5].

According to Kotler and Keller [6], satisfaction is perceived by the customer through a rating from their past communication experiences with the service provider until the moment when this rating and rating is received by the customer. Use to predict future experiences. Bachelet [7], customer satisfaction is the feedback of the consumer, the comments on the features of the product, service or the product or service itself. These reflections will show different levels of satisfaction when they consume products or services. Post-consumer satisfaction or disappointment is defined as the customer's response to an assessment by perceiving the difference between expectations before consumption and the actual perception of a product after consumption [8]. The psychology of customer satisfaction is perceived by the comparison between product perceptions and customer needs and social expectations related to the product [9]. Customer satisfaction is achieved when the company's performance can be appropriate for their desires. If performance is below expectations, customers will not be satisfied; if performance is equal to expectations,

customers will be satisfied and if performance is higher than expected, customers will be very satisfied [10]. Satisfied customers will have a significant impact on the life of the organization now and in the future.

2.1 The Service Quality Factors and Customer Satisfaction

According to the research results of Ioannis and Constantine [1], tangibles variables have the strongest influence on customer satisfaction. The index of this variable relates to the facilities and surrounding space of the Banks, along with the friendliness and comfort from the environment, the cleanliness at the reception place, the equipment is always updated, rooms are always comfortable and clean. In addition, in the research of Masound et al. [11], physical variables have a relationship with customer satisfaction. So for commercial banks in Highland, Viet Nam, how does the facility factor affect customer satisfaction? The hypothesis is stated:

H_1:*Tangibles have a positive effect on customer satisfaction.*

Customers who have credibility with a brand will be more satisfied when using its brand's products and services [1]. The coefficient of this variable represents the ability to deliver the promised service reliably and accurately, the reliability variable is combined from good organization, capturing trust from customers, always keeping as promised to customers, the services are always done right and Banks always show interest in customers when problems occur. Besides, Parasuraman et al. [3] stated that reliability has a certain influence on customer satisfaction. So for commercial banks in Highland, Viet Nam, how does the reliability factor affect customer satisfaction? The hypothesis is stated:

H_2: *Reliability has a positive effect on customer satisfaction.*

Masound et al. [11] demonstrate that the relationship between responsiveness and satisfaction is quite positive. The index of this variable is related to the willingness to provide service to customers and provide fast service. This variable is a combination of customer service, staff ready to meet customer needs, staffs taking time to answer each customer's questions, quick response staff, and Information about translation products is provided. Similarly, research by Ioannis and Constantine [1] also demonstrated that responsiveness has the same positive impact on satisfaction factor. So for the commercial banks in Highland, Viet Nam, how does the ability of Responsibility affect customer satisfaction? The hypothesis is stated:

H_3: *Responsiveness has a positive effect on customer satisfaction.*

Assurances have a certain influence on customer satisfaction when they use products and services [1]. The indicators of this variable are related to their knowledge, etiquette, ability to convey their beliefs and trust. It is a combination of sub-variables: knowledgeable and experienced staff, friendly and courteous staff, staff who take customers seriously and respectfully, staffs who explain problems to customers and convenient operating hours for all customers. In addition, Masound et al. [11] also

suggest that there is a link between assurance and satisfaction. So for commercial banks in Highland, Viet Nam, how does the assurance factor affect customer satisfaction? The hypothesis is stated:

H_4: *Assurance has a positive effect on customer satisfaction.*

If a person has a sympathy for a product or service of a brand then they tend to be satisfied with that brand [1]. The index of this variable is related to the personal care and care that the organization provides to its customers. It is the resonance of staffs who understand the needs of customers, staffs who can give sincere attention, kind staffs to customers, staffs who seek the best for customers and staffs predict the needs of customers. According to Masound et al. [11], empathy has an impact on customer satisfaction. So for commercial banks in highland—Vietnam, the factor empathy affects how customer satisfaction. The hypothesis is stated as follows:

H_5: *Empathy has a positive effect on customer satisfaction.*

2.2 The Relationship Between Customer Satisfaction and Word of Mouth

Previous studies have demonstrated that customer satisfaction has a direct impact on word of mouth [1, 11]. When customers are meeting their needs and expectations, they tend to pass on those things to their relatives, friends, and people around them. So for the commercial banks in Highland, Viet Nam, how is the satisfaction factor affecting word of mouth of customers. The conceptual model and hypothesis in Fig. 1:

H_6: *Customer satisfaction has a positive impact on word of mouth.*

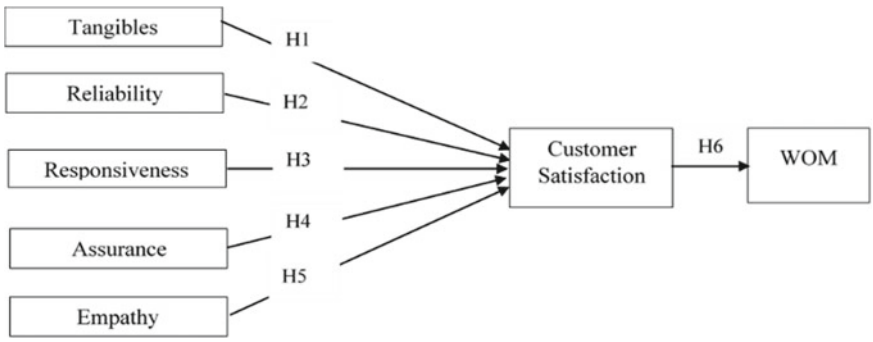


Fig. 1 Conceptual model

Table 1 The origin of the research scale

Dimensions	Number of items	Origin of the scale
Tangibles	5	Ioannis and Constantine [1], Shaham [12]
Reliability	5	Ioannis and Constantine [1], Samra Shaham [12]
Responsiveness	5	Ioannis and Constantine [1], Shaham [12]
Assurance	5	Ioannis and Constantine [1], Shaham [12]
Empathy	5	Ioannis and Constantine [1], Nimit [13]
Cus. satisfaction	5	Shaham [12]
Word of mouth	4	Herr et al. [14]

Source Summaries for study

3 Methodology

3.1 Scale Development

Based on an overview of the literature, the author decided to examine the relationships between service quality aspects of satisfaction and positive word-of-mouth activity on commercial banks in Central Highland, Viet Nam. In the conceptual model, it was used the study of Parasuraman et al. [3] related to SERVQUAL and Fuentes [5] to evaluate service quality affecting satisfaction and positive word-of-mouth activities (Table 1).

3.2 Sample and Data Collection

Data collected through detailed questionnaires for customers of 17 joint-stock commercial banks in the Central Highlands, Vietnam. A total of 299 responses were used in the analysis. Data used was encoded and analyzed using SPSS 25.0 software. Table 2 show the sample characteristics.

3.3 Data analysis

After encryption and data cleaning, the analysis was conducted following these steps: The scale reliability was evaluated by Cronbach's alpha and factor analysis. Cronbach's Alpha coefficient is used to assess reliability and eliminate variable trash. The variables correlated with the total variable less than 0.3 will be rejected while those with Cronbach's alpha coefficient greater than 0.6 will be accepted [15]. Exploratory factor analysis ensures that variables with transmission coefficient of less than 0.5 and difference in two-factor less than 0.3 will be disqualified. Eigenvalues (representing

Table 2 Sample characteristics

Sample information		Gender		Age					Total
		Male	Female	18–25	26–35	36–45	46–55	>55	
Time to use the service of commercial banks	<1 year	10	31	29	12	0	0	0	41
	1–3 years	40	81	92	25	4	0	0	121
	3–5 years	42	55	81	8	7	1	0	97
	>5 years	15	25	6	20	6	7	1	40
Total		107	192	208	65	17	8	1	299

Source Result data analysis, 2020

the variability explained by each factor) are greater than 1 and the total variance is greater than 50%. The Kaiser—Meyer—Olkin (KMO) test and Bartlett's test were used to evaluate the validity of the data. Multiple regression analysis is used. F-test for regression analysis is used to test the hypothesis about the suitability of linear regression model (Sig < 0.005).

4 Results and Discussion

The results of the reliability analysis show that Cronbach's alpha coefficient of all scales is over 0.8. The lowest is the responsiveness scale with $\alpha = 0.803$, and the highest is the customer satisfaction scale with $\alpha = 0.884$. Results of assessing the reliability of the scales are presented in detail in Table 3. The total number of observed variables was initially 34 and all of them met the reliability coefficient condition. Therefore, all 34 observations were included in the exploratory factor analysis in order to detect the structure and evaluate the convergence of the observed variables according to the ingredient.

The result of factor analysis with independent variables showed that KMO = 0.902 which satisfies the conditions and elements ensure reliability. Bartlett's test has Sig = 0.000 (<0.05) which indicates that the variables are correlated with each other and the factors ensure statistical significance. The load factor is greater than 0.50 (from 0.512 to 0.818). Table 4 shows Eigenvalues = 1139 (>1), the total extracted variance is 62,225% (>50%), from 25 items divided into 4 factors.

The factor analysis with dependent variables result showed that KMO = 0.883 (>0.5) satisfies the conditions and reliability assurance factors. Bartlett's test of that scale has Sig = 0.000 (<0.05). The extracted variance is 68,288% greater than 50% which means that the variation of the analyzed factors explains 68,288% of the measurement variation of the data originally surveyed. Eigenvalues of 1313 (>1) indicate that the convergence of the analysis stops at the second factor, this means that the survey data has two extracted factors. In addition, each observed variable

Table 3 Cronbach's alpha reliability result

No.	The scale	Number of items	Cronbach's alpha coefficients
1	TAN	5	0.805
2	REL	5	0.813
3	RES	5	0.843
4	ASS	5	0.803
5	EMP	5	0.846
6	CS	5	0.884
7	WOM	4	0.819

Source Results of data analysis, 2020

Table 4 EFA results of independent variables

No. of items	Component			
	EMP	SER	REL	TAN
1	0.805	0.818	0.816	0.767
2	0.719	0.712	0.732	0.749
3	0.718	0.610	0.653	0.704
4	0.692	0.582	0.598	0.650
5		0.537		
6		0.525		
7		0.512		
Eigenvalues	7.544	1.947	1.204	1.139
Cumulative (%)	39.65	49.896	56.231	62.225
Cronbach's alpha	0.812	0.863	0.807	0.775

Source Results of data analysis, 2020

has a factor load factor greater than 0.5, so all variables affect the factors that the variables represent in Table 5.

The regression analysis results showed that the adjusted $R^2 = 0.572$, the F-test for significance was Sig = 0,000. Therefore, the regression model is suitable to analyze the variation of the dependent variable explained by the independent variables, or in other words, the independent variables explained are 57.2% of the variance of the dependent variables. Customer satisfaction is explained by 4 independent factors. The Durbin-Watson coefficient = 2.060 showed that there is no autocorrelation between variables.

The results of regression analysis showed that empathy has the biggest impact on customer satisfaction ($\beta = 0.323$), followed by the service capacity variable (β

Table 5 EFA results of dependent variables

Scales	Component	
	CS	WOM
1	0.844	0.823
2	0.806	0.801
3	0.791	0.761
4	0.78	0.638
5	0.769	
KMO	0.883	
Bartlett's Test (Sig)	0.000	
Cumulative (%)	68.288	
Eigenvalues	1.313	

Source Results of data analysis, 2020

Table 6 Results of the model estimation

Model	Unstandardized coefficients		Standardized coefficients	*t*	Sig.	Collinearity statistics	
	β	Std. error	*β*			Tolerance	VIF
(Constant)	0.630	0.179		3.518	0.001		
EMP	0.323	0.046	0.356	7.066	0.000	0.564	1.772
SER	0.314	0.057	0.318	5.478	0.000	0.426	2.349
REL	0.198	0.051	0.199	3.885	0.000	0.55	1.819
TAN	0.021	0.045	0.021	0.460	0.646	0.683	1.464

Source Results of data analysis, 2020

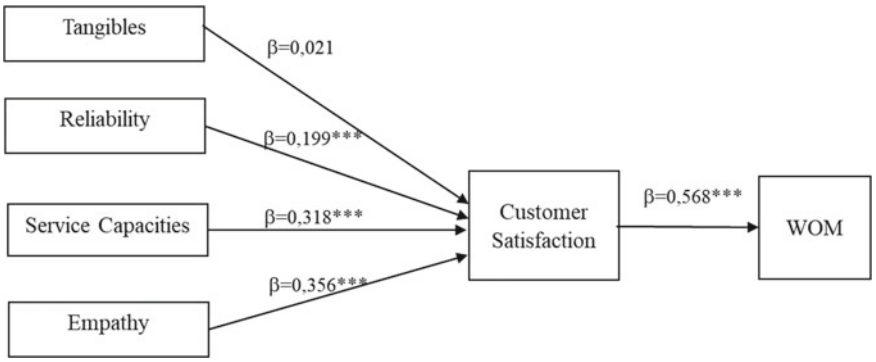


Fig. 2 Result of model estimation research

= 0.314) and finally the reliability ($\beta = 0.198$). The above independent variables have positive Beta coefficients, this proves that the independent variables are positively related to the dependent variables. As these independent variables increase, the dependent variable increases and vice versa (Table 6).

The result of satisfaction and word-of-mouth relationship shown Adjust $R^2 = 0.321$, which shows the degree of dependent variation explained by the independent variables. $F = 141.746$, Sig. = 0.000 (<0.05). Customer satisfaction has impact on word of mouth with $\beta = 0.568$. Result of the model estimation in Fig. 2.

5 Conclusion and Management Implications

Word of mouth is a form of advertising that is no stranger to businesses and organizations. Along with the continuous development of the socio-economy, the promotion of the image of businesses or organizations to customers is becoming more and more popular. Especially, in the field of Banking, commercial banks want to build a better

image and more prestige, the satisfaction of customers is very important. When customers are satisfied with the quality of service provided by the Bank, they are willing to share their experiences when dealing at the Bank to friends and relatives, which not only makes the image of Banks better in the eyes of potential customers but also create the reputation and professionalism of the Bank to customers before using the Bank's services.

The aims of this research is to explore the factors affecting the quality of service on word of mouth and customer satisfaction with commercial banks in highland, Vietnam. In addition, the research examines the extent of the impact of service quality factors on customer satisfaction and word of mouth. The research clearly indicated that the object of the research is the service quality that affects customer satisfaction and word of mouth for commercial banks and the survey is the customers using the services of Commercial banks. Quantitative research was conducted with the target customers who are using the services of commercial banks with $n = 299$. After checking Cronbach's alpha reliability, EFA factor analysis and regression analysis, there are 3 factors (Empathy, Service Capacity, Reliability) that positively impact satisfaction and word of mouth. Positively toward the Bank in Highland, Viet Nam. The development of these factor groups contributes to improving the quality of services, thereby improving customer satisfaction and positive word of mouth for the Bank in Highland, Viet Nam. Through the research, the implication are as follows:

Firstly, the Bank needs to consider devising a few strategies in order to find out the real needs of customers. Secondly, the Bank also needs to open more training courses for staffs not only hard on professional skills but also on soft skills. Thirdly, for the external image, the Bank needs to show its clients professional manner in working with a well-organized team.

The research gives an overview of word of mouth and customer satisfaction to commercial banks in Highland, Viet Nam and provides management solutions to promote positive word of mouth. It is not only building a business image but also increasing sales through improving service quality to increase customer satisfaction with the Bank. However, because the research was only limited to the scope of customer survey in Highland, Viet Nam. Therefore, it is necessary to expand the scope of the research for the whole country and make comparisons among regions, which will increase the meaning and the more general.

References

1. Ioannis, ECh., Constantine, L.: Service quality effect on satisfaction and word of mouth in the health care industry. Manag. Serv. Qual. **19**(2), 229–242 (2009)
2. Cong, L.T., Trang, N.T.M.: Service quality, corporate image and word-of-mouth in Vietnam domestic airline service. Sci. Technol. Dev. **16**(2), 107–116 (2013)
3. Parasuraman, A., Zeithaml, V.A., Berry, L.L.: A conceptual model of service quality and its implications for future research. J. Mark. **49**(Fall), 41–50 (1985)
4. Zeithaml, V. A.: Consumer perceptions of price, quality, and value: a means-end model and synthesis of evidence. J. Mark 2–22 (1988)

5. Fuentes, C.M.: Measuring hospital service quality: a methodological study. Manag. Serv. Qual. **9**(4), 230–239 (1999)
6. Kotler, P., Keller, K.: Marketing management, 12th edn. Prentice Hall, Upper Saddle River (2006)
7. Bachelet, D.: Measuring satisfaction; or the chain, the tree, and the nest. In: Brooker, R. (ed.) Customer Satisfaction Research. Emosar (1995)
8. Oliver, R.: Cognitive, affective and attribute bases of the satisfaction response. J. Consum. Res. **20**, 418–430 (1992)
9. Kavoussi, M., Saghaei, A.: Methods of measuring customer satisfaction. Sbzan Press (2005)
10. Kotler, P., Armstrong, G.: Principles of Marketing, 15th Global Edition. Pearson Prentice Hall
11. Masound, P., Mehrdad, G., Samin, Y., Shiva, M.: Service quality effect on satisfaction and word of mouth in insurance industry. Manag. Sci. Lett. 1773–1780 (2014)
12. Samra, S.: The effect of service quality on customer satisfaction, loyalty and word of mouth; evidence from tourists in Antalya, Turkey. Retrieved from https://i-rep.emu.edu.tr:8080/jspui/bitstream/11129/3589/1/shahamsamra.pdf (2016)
13. Nimit, S.: The effect of service quality on word-of-mouth in hotel industry: the mediating role of overall satisfaction. In: Proceedings of the International Conference on Marketing, Tourism & Hospitality, pp. 1–10, New York-USA (2017)
14. Herr, P.M., Kardes, F.R., Kim, J.: Effects of word-of-mouth and product-attribute information on persuasion: an accessibility-diagnosticity perspective. J. Consum. Res. **17**(4), 454–462 (1991)
15. Hair, J.J.F., Anderson, R.E., Tatham, R.L., Black, W.C.: Multivariate data analysis, 7th edn. Pearson Prentice Hall, Upper Saddle River (2010)

Design and Implementation of Conversion of Gesture to Voice Using OpenCV and Convolution Neural Network

Aditi Tiwari, Avani Gangrade, Ayushi Tiwari, and Kailash Chandra Bandhu

Abstract Communication is the most important aspect of society that people are dependent upon. Today everyone can communicate with the help of their native languages but there is a small part of society who cannot speak and faces a lot of problems in communicating. These people are the one who are suffering from speaking impairment. Such people are also members of this society and have equal right to enjoy their life and communicate with the external environment in an easy and professional manner. In today's world, technologies have made life easy. Everyday there is a new innovation, and a new device is ready. But there is a certain section of the community who still cannot take benefit of these technologies and gadgets for his or her communication. These people not able to speak and cannot convey their message and society calls them mute. This work will help such people to speak easily without the involvement of any third person or mediator. Mute people do the action ahead of the screen of the device, since this device is going to be user-friendly, an individual can feed action and its meaning into it. Person feeds his own action and gestures alongside its name or title, and therefore, the camera will detect the gesture, then print it on the screen and convert it into audio output. People working during a professional environment can cash in on it.

Keywords Convolution neural network (CNN) · Red green blue (RGB) · Open-source computer vision library (OpenCV) · Numerical python (NumPy) · TensorFlow and blue green red (BGR)

1 Introduction

In today's fast paced life, all are surrounded by technologies and everyday there is a new innovation; technologies have proved itself as a blessing on the human race. In our society, people of almost all sectors today have included technology in their day-to-day work. But the one who is deaf and the one who cannot speak still uses the

A. Tiwari (✉) · A. Gangrade · A. Tiwari · K. C. Bandhu
Shivajirao Kadam Institute of Technology and Management—Technical Campus, Indore, India

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_86

ancient method of sign language or takes the help of a mediator. Even after having equal mental capability, they are unable to get involved in society because of their physical disability and get neglected most of the time. There is a huge research going on in the field of artificial intelligence, machine learning, deep learning, and many more. Human race is all set and geared for new innovations, and now there is no turning back. Technology has proved to be a blessing for society. These technologies can be merged with sign language, and a new concept of communication can be generated. This research paper helps to overcome their physical disabilities, and they will be able to communicate with ease. There will be no more interference from any mediator. Since this project is going to be user-friendly, the user can feed his own actions and names in it; in short, a person can create his own dataset. It will eliminate the gap between reality and the virtual world.

In this work, the dataset is entered by the user. He initiates the file, and images are captured. The number of images captured is directly proportional to the accuracy. The images are converted into gray images. Those images are then stored in folders and are labeled accordingly and using the libraries, the images are processed. The data is considered as a dataset and with the help of its machine gets trained. Now as soon as the user runs the application and shows some hand gesture in front of the camera, those gestures are scanned by the webcam and the final message is printed on the screen and converted into voice. Our system gives freedom to the user to make his own dataset on that particular instance.

The paper organization is as follows: The Literature Survey of some hand gesture recognition system is included in Sect. 2, Proposed solution of design and implementation of conversion of gesture to voice using OpenCV and Convolutional Neural Network is well-explained in Sect. 3, Graphs and Calculation help in explaining accuracies in Sect. 4, Summary of results in Sect. 5, Conclusion of the system in Sect. 6, Future Scope is mentioned in Sect. 7, and at last the References is mentioned.

2 Literature Survey

Many researchers worked on artificial intelligence, machine learning, and deep learning and applied it to solve many real-life problems but few works were done on conversion of gestures into voice using neural networks.

Prajapati et al. [1] worked on feature extraction and classification of an image hand gesture can be converted. After image capturing, the image gets recognized using the PCA algorithm and classification of an image is done through KNN and SVM algorithm.

Vaghela et al. [2] worked on camera captures images in RGB format and then the image is scaled and transformed accordingly. The system works on convolutional neural networks (CNN). The convolution filters are the first layers that receive an input signal.

Krishi et al. [3] discussed a dataset that consists of different types of hand gestures that are used to train the system based on which the system performs the actions. The system is built using OpenCV and TensorFlow object detectors.

Patidar et al. [4] worked on recognition of letters from the hand language which is taken as per the American Sign Language. The detection is done by using the techniques of contour analysis and feature extraction.

Kumar et al. [5] worked on capturing images that are converted into binary data using the NumPy package and processed by using the H5py package.

Sood et al. [6] worked on communication system for deaf and dumb.

BGR: It is an additive color model where the different intensities of blue, green, and red give different shades of color [7].

RGB: Red green blue is a 3-D array of numbers, most image processing needs to operate on RGB images with three channels [7].

NumPy: Numerical python is a scientific computing package for python programming language [8].

TensorFlow: It is used for machine learning applications, e.g., neural networks. It is open-source and free software library for dataflow (CNN): Convolution neural network is one of the main categories to do image recognition, image classification, object detection, recognition faces, etc. [9, 10].

OpenCV: Open-source computer vision library is a machine learning software, consisting of optimized algorithms to perform various functions [11, 12].

3 Proposed Solution

Using this work people who are deaf–mute can communicate with others very easily. This application will detect their gestures and interpret it into voice and thereby make it easy for them to communicate with others who are unable to understand their sign language. This application will capture the images of their gestures and train the software so that it can detect their gesture and can convert into voice. And the camera detects the images and classifies it using CNN TensorFlow.

This system captures a hand gesture performed by the user in front of the camera and then proceeds to detect the valid gesture through CNN TensorFlow and execute the corresponding operation as shown in Fig. 1.

3.1 Algorithm

The overall implementation of system is described as follows:

Algorithm for Real Time Application:

Step 1	Start
Step 2	Input directory name

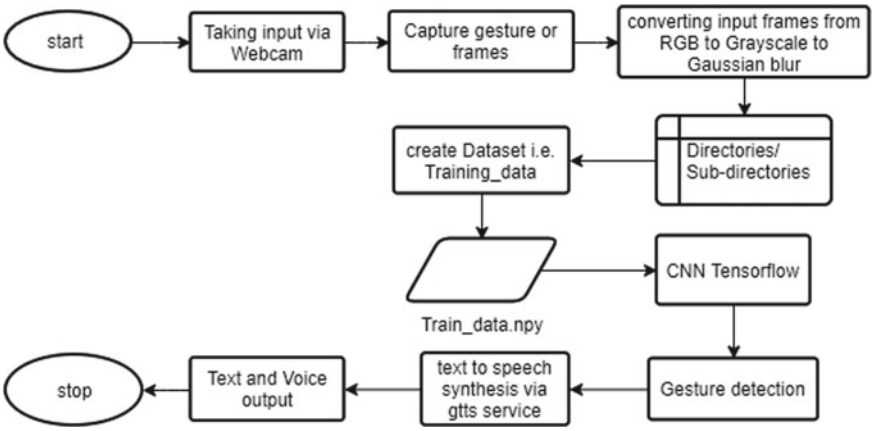


Fig. 1 Flowchart of system

Step 3	Path set
Step 4	Camera is called or Webcam is accessed
Step 5	While Loop
	Folder name as input is given
	Try (save to directory)
	Except (folder already exist)
Step 6	While Loop
	Camera is read.
	Image is processed
Step 7	Create training_datasets
	set path.
	shuffle and save
	create_train_data
Step 8	Train CNN model using the training_dataset
	create_training_model
	test model with split_test_train model
Step 9	List is created, and names of gestures are stored in it
	Functions of OpenCV are implemented
	Text converted into speech
Step 10	Camera Released

Fig. 2 Input RGB image

3.2 System Description

3.2.1 Capture Gesture

Capture the gesture through webcam (by default), the captured gestures will undergo processing using OpenCV; here, processing of gesture is done by performing operations on the video frame (frames can be treated as similar to an image, in a single second we have multiple frames).

One of the most popular computer vision libraries is OpenCV. It involves:

a. Reading images: OpenCV arranges the channels in BGR order.
b. Extracting the RGB (3-D array) values of a pixel.
c. Extracting the ROI: We select ROI from the frame, which is just a square region inside the frame image which can select a frame matrix.
d. Smoothing-now convert the RGB (3-D array) into grayscale (2-D array) and blur, it through Gaussian blur.
e. Rotating the image—It involves mainly: calculation of center image, i.e., image: center = (w/2, h/2) generating rotation 2-D matrix of three arguments-center, angle, scale.
f. It returns 2 * 3 matrices consisting of values derived from alpha and beta.
Alpha = scale * cos (angle).
Beta = scale * sin (angle)

$$\begin{bmatrix} \alpha & \beta & (1-\alpha)\cdot \text{center}\cdot x - \beta \cdot \text{center}\cdot y \\ -\beta & \alpha & \beta \cdot \text{center}\cdot x + (1-\alpha)\cdot \text{center}\cdot y \end{bmatrix}$$

g. Drawing a rectangle: It takes five arguments, i.e., image, Top-left corner coordinates, bottom-right corner coordinates, color (BGR), line width as shown in Figs. 2 and 3.

3.2.2 Create Training Dataset

It involves

Fig. 3 Grayscale image

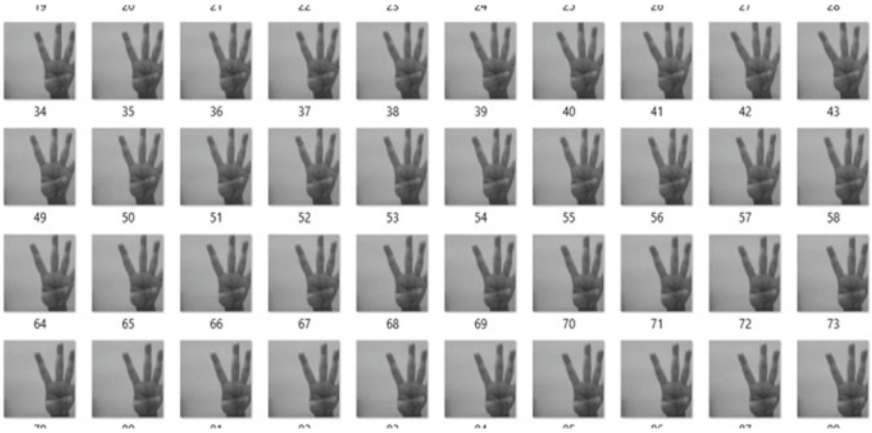


Fig. 4 Training dataset

a. Reading and resizing of gesture image to image_size 96.
b. Creation of train_data followed by converting it into np array.
c. Dealing with directories, sub-directories and file names.
d. Shuffle the train data and save the file under train_data.npy name. After the creation of train_data.npy file that will be our dataset and this dataset we can give to the neural network to train data.
e. Resizing the image: Here we have image_size 96, so we get image_size of 96/96 as shown in Fig. 4.

3.2.3 Classification Through CNN TensorFlow

CNN compiles different layers before making a prediction. A neural network has:

a. A convolution layer—It applys different filters on a subregion of a picture.
b. ReLU Activation function—It adds nonlinearity.
c. Pooling layer: It reduces the dimensionality of the feature maps.
d. Densely connected layer: All these layers extract essential information from images.

At last:

- The model was trained using 96 * 96 grayscale images.
- Here tensor boards in the final five hidden layers are used for increasing accuracy of the model.

3.2.4 Text to Speech

After whole processing, the meaning of the gestures gets printed on the screen. Using gTTS (Google Text To Speech) we will convert the text into speech. This library will be imported. The array consisting of messages will be passed into the function.

4 Graphs and Calculation

4.1 Accuracy with Time

Accuracy helps avoiding considerable time, money as it mitigates the errors through optimization and undue stress by preventing false or inaccurate models. Training graph is created by calculating accuracy on a particular time with a particular number of datasets. The accuracy of any model increases with the number of datasets since the model gets trained. And it speeds up the decision making of the model.

In this graph, accuracy was calculated using.evaluate() by passing testing data in it and was measured on the time interval of every 20 s. From the graph, it can be inferred that accuracy increases with time as shown in Fig. 5.

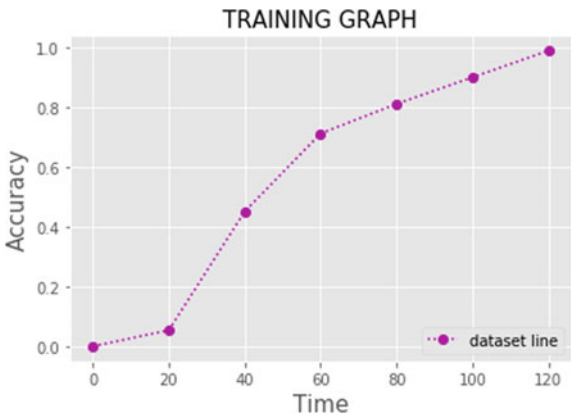


Fig. 5 Accuracy with respect to time

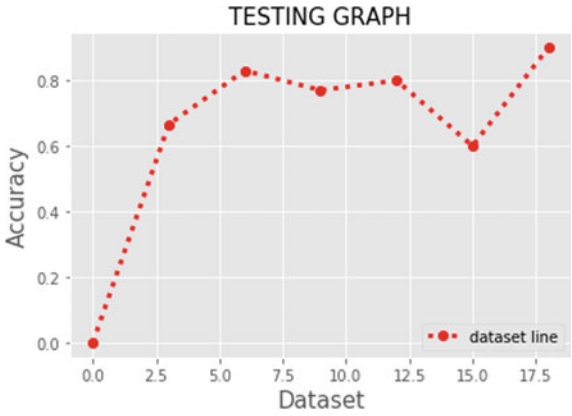


Fig. 6 Accuracy with respect to dataset

4.2 Accuracy with Dataset

Testing graph is created by calculating accuracy of dataset. The calculation here was done by identifying whether the module is recognizing the inserted dataset accurately. For this, the formula that was used is:

$$\text{Accuracy} = \text{No. of correct predictions/Total number of predictions}$$

Here the accuracy was calculated using the above formula, and the datasets were increased after each testing. From the graph, it can be inferred that the accuracy increases with time in datasets as shown in Fig. 6.

4.3 Accuracy of Gestures

The accuracy that was calculated from the above graphs is represented in the pie chart. The gesture having highest and lowest accuracy can be easily identified through this graph. First import plt from the matplotlib module with the line import matplotlib.pyplot [13] as plt. Then the method plt.pie() to create a plot is used. Matplotlib uses the pie() function to support the pie charts. Each value of a particular gesture is calculated as per the accuracy comes in the testing phase as shown in Fig. 7.

5 Results

Finally, it obtains a system that can read the values for a particular gesture done by the user, predicts the output for the gesture, displays in text on the laptop screen, and provides an audio output via a desktop GUI as shown in Fig. 8.

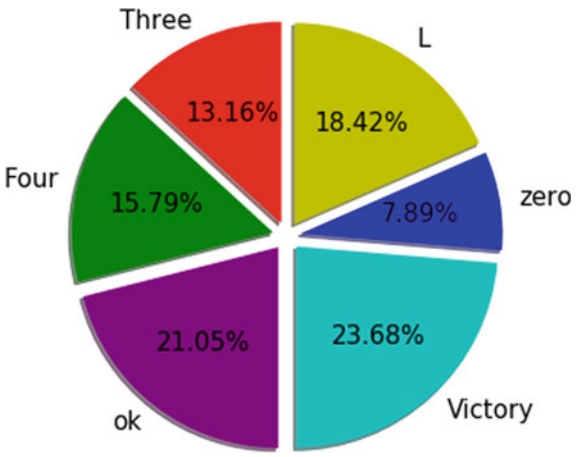


Fig. 7 Accuracy of gestures

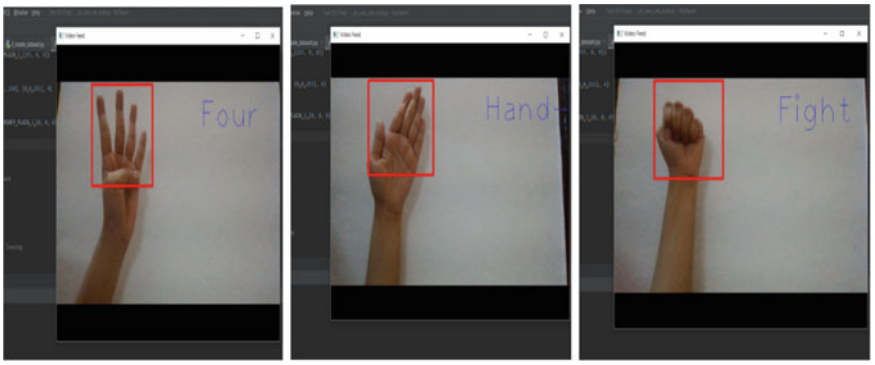


Fig. 8 Gestures to text and voice

6 Conclusion

The mute people use their standard sign language to communicate which is not easily understandable by common people without the knowledge of the sign language. Thus, this system allows them to capture the gestures of their own, through webcam followed by processing of images through OpenCV which results in conversion of BGR to grayscale image and will detect gestures in a video stream and save it to disk using OpenCV under train_data.npy file as dataset, which we provide it for classification through neural network to train data, CNN compiles different layers before making a prediction; hence, tensorboard in final five hidden layers is used for increasing accuracy of the model. After the whole processing, conversion of gestures (the sign language) into text will be printed on screen along with the conversion of text to audio using gTTs.

This system provides us with a high gesture recognition rate with accuracy 80% within minimal computation time. Sign language is a useful tool to ease the communication between the mute person and normal person. The performance of the software is depending on the algorithm used to develop the model and detecting sign images; also, the performance is greatly varying due to lighting conditions and background noises that can be reduced by using the webcam of high performance. As compared to the other system, this concept not only focuses on the gesture to word display but also on the speech synthesis.

7 Future Scope

The proposed system can also be implemented for entire video calls of some continuous gestures recognized and then converting these continuous gestures into corresponding sentences and also voice speech.

References

1. Prajapati, R., Pandey, V., Jamindar, N., Yadav, N., Phadnis, N.: Hand gesture recognition and voice conversion for deaf and dumb. Int Res J Eng Technol (IRJET) **05**(04) (2018). e-ISSN: 2395-0056, p-ISSN: 2395-0072
2. Vaghela, D., Vaity, A., Makwana, R., Shaikh, A., Chouhan, C.: Hand gesture recognition to speech conversion. IOSR JEN 21-25 (2015). ISSN (e): 2250-3021, ISSN (p): 2278-8719
3. Abhishek, B., Krishi, K., Meghana, M., Daaniyaal, M., Anupama, H.S.: Hand gesture recognition using machine learning algorithms. Int. J. Recent Technol. Eng. (IJRTE) 8(1) (2020). ISSN: 2277-3878
4. Patidar, A., Manikandan, K., Walia, P., Roy, A.B.: Hand gesture detection and conversion to speech and text. Int J Pure Appl Math **120** (2018). ISSN: 1314-3395
5. Kumar, P., Sharma, T., Rawat, S., Bhagat, S.: Hand gesture recognition for deaf and mute. IJITEE **9**(2) (2019). ISSN: 2278-3075
6. Sood, A., Mishra, A.: AAWAAZ: a communication system for deaf and dumb. In: 2016 5th international conference on reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO). IEEE (2016)
7. https://www.geeksforgeeks.org/color-spaces-in-opencv-python/
8. https://indianaiproduction.com/python-numpy-tutorial
9. https://www.toptal.com/machine-learning/tensorflow-machine-learning-tutorial
10. https://medium.com/@RaghavPrabhu/understanding-of-convolutional-neural-network-cnn-deep-learning-99760835f148
11. https://opencv.org/about
12. https://www.geeksforgeeks.org/opencv-python-tutorial
13. https://pythonspot.com/matplotlib-pie-chart/

Learners' Performance Evaluation Measurement Using Learning Analytics in Moodle

T. Sheeba Justin, Reshmy Krishnan, Sarachandran Nair, and Baby Sam Samuel

Abstract In recent years, the development of online environments has been rising exponentially and educators and learners are moving toward online learning systems. These online learning systems are open-source applications that have their advantages and disadvantages. Moodle is one of the widely used open-source learning platforms used by most institutions all over the world. Even though moodle provides a good framework for learning, it is static with minimal functionalities. The need for student preferences and their contexts is required for understanding and optimizing learning environments in a better way. The paper presents an approach to collect and retrieve student behaviors from the log files and table of moodle and classify learning preferences using the standard Naïve Bayes classifier based on the standard Felder Silverman learning style model. The retrieved learning preference based on students' behaviors and actions supports educators to view the preferences of students and to improve and enhance their teaching.

Keywords Learner analytics · Learning management system · Learning style · Moodle · Naïve Bayes classifier

T. S. Justin (✉) · R. Krishnan · S. Nair
Muscat College, Muscat, Sultanate of Oman
e-mail: sheeba@muscatcollege.edu.om

R. Krishnan
e-mail: reshmy@muscatcollege.edu.om

S. Nair
e-mail: saran@muscatcollege.edu.om

B. S. Samuel
InterTech.LLC, Muscat, Sultanate of Oman
e-mail: mail@babysam.biz

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_87

1 Introduction

Learning analytics collects and analyzes the data which is generated by learners during the interactions in both online and offline learning environments. It acts as a decision support tool in teaching, learning, and educational management.

In this paper, we describe how learning analytics can be provided as pointers of students' performance in the course. Students' activities related to assessments are considered as the key measurements. An overview of the emerging field of learning analytics is described initially, followed by a description of the case study which is done in moodle with the help of the learning analytics tool in moodle. The summary of the analysis helps us to conclude students' performance in that particular course as well as semester performance.

It covers three levels of analytics such as personal level, course level, and department level. The personal level does analytics on the performance of individual students, learning resources, and other students' study habits. Intelligent curriculum, conceptual development, and social networks are referred to in course-level analytics. The department level of learning analytics can identify the pattern of result (success/failure) and do predictive modeling which can be taken as the corrective measurement in quality teaching and students' performance [1].

2 Literature Review

Online educational technologies are adopted in all educational institutions during this pandemic. How students are interacting with these environments and technologies are still under research and learning analytics is the emerging approach to address this. It includes a variety of data gathering tools and analytical techniques that are used to examine students, interaction, activities, and performance on tasks [1]. These tools can support researchers to identify and analyze the diversity of learning behaviors that can be taken place in these environments [2].

Learning analytics can be utilized for improving the quality of teaching, curriculum, and assessment. Usage of logs from e-learning applications was not utilized properly in e-learning research. The term learning analytics refers to these data which is used to analyze the study behavior of the students [3].

The terminology around 'analytics' in education is evolving. Despite calls for its use in the early 1990s, usage logs from e-learning applications have been under-utilized in e-learning research [4]. The use of this automatically captured data, which records who accessed what and when to study student behavior, is now termed learning analytics [5]. This can be contrasted with academic analytics, which considers similar data at an institutional level [6].

Moodle is the world's largest online learning management system (LMS). It has a very user-friendly interface that is regularly reviewed improved and can be

customized as per the user's requirements [7]. To increase usability, moodle has introduced features of learning analytics [8].

Moodle is categorized as four services such as descriptive, predictive, diagnostic, and prescriptive. Descriptive tools give an idea about what happened and predictive tools provide the clue that what will happen next. The reason for the result is evaluated by diagnostic tools and suggestions for improvement are provided by prescriptive tools. There are several reports, blocks, and other plugins for moodle that provide learning analytics.

Learning analytics tools are one that can support the user to improve learning outcomes. The application of big data and analytics in education is the representation in learning analytics. There are many plugins like reports, blocks, and other plugins that are available in moodle to provide learning analytics. Some of the plugins are described below [9].

Logs are important plugins which are activity reports and available at the site and course level. A teacher as an administrator can generate a log of activity from the report option. A combination of fields of students such as date, activity, actions and level are selected for generating the log. Logs can be filtered as per levels like teaching level and participating level [10]. There are other plugins like activities, live logs, feedback, quiz statistics, course participation, survey, etc., in the moodle to support to analyze student performance [11]. Various systems like moodle activity viewer, Intelliboard, etc., are integrated with moodle to do learning analytics externally.

3 Methodology

3.1 Moodle-Based E-learning Portal

The online courses created in Moodle LMS have rich course topics represented in different component forms such as texts, power points, pdf, word, exams, hyperlinks, and assignments. The interaction of students is represented by using forums discussions and chats where the students can post topics and reply to others and view others' opinions. Any interaction of the students performed in moodle is stored in the log files and also in the database. As the portal is available on the Web site of the institution, the students with Internet access can access the course materials at any time.

3.2 Felder Silverman Learning Style Model (FSLSM)

Each student has their own learning style. This paper focuses on student learning style behavior patterns extraction based on the standard model of the FSLSM. FSLSM is a standard learning style model tested in various education systems. According to

Table 1 Felder Silverman's learning dimensions and learner characteristics

Learning dimensions	Learner characteristics	
Processing	**Active**: Do something physical by involving in the discussion, applying, or explaining to others	**Reflective**: Prefer observation and tend to think about information rather than doing active experimentation
Perception	**Sensitive**: Tend to do more practical and careful with details. Like to solve problems using standard methods and doing hands-on work	**Intuitive**: More innovative and tend to discover possibilities and relationships. Prefer abstractions and mathematical formulations and good at grasping new concepts
Input	**Visual**: Prefer visual representations such as diagrams, charts, graphics, charts, etc., to best remember the information	**Verbal**: Prefer verbal information such as lectures or written documents
Understanding	**Sequential**: Prefer to learn each topic in well-structured information sequentially in step by step	**Global**: Prefer to learn topics randomly in large chunks without seeing the relationship between them. Easily grasp the big picture and able to solve complex problems quickly

FSLSM [9, 11], learning style preferences are determined based on four different dimensions: processing, perception, input, and understanding which are grouped into various learner characteristics as sensitive/intuitive, visual/verbal, active/reflective, and sequential/global as shown in Table 1.

3.3 System Architecture Overview

The proposed architecture (Fig. 1) shows the entire process performed on the e-learning platform moodle for creating student reports.

When students enroll and take the online course on moodle, the students' interaction is being captured in log files and stored into the moodle database on the WAMP server. Further, the captured data on these files and tables have been analyzed to extract the behaviors of the students. Finally, the classification of student behaviors is done using a standard Naïve Bayes classifier to detect the learning style based on the standard FSLSM. The output can be used by educators to improve the learning of the students by making the course content based on their preferences.

Data Collection. In the context of students' performance, the information can be acquired explicitly (by collecting students' profile data) or implicitly (by monitoring students' behavior, such as visits to teaching courses, documents downloaded). The explicit method includes the personal data of the students such as identification

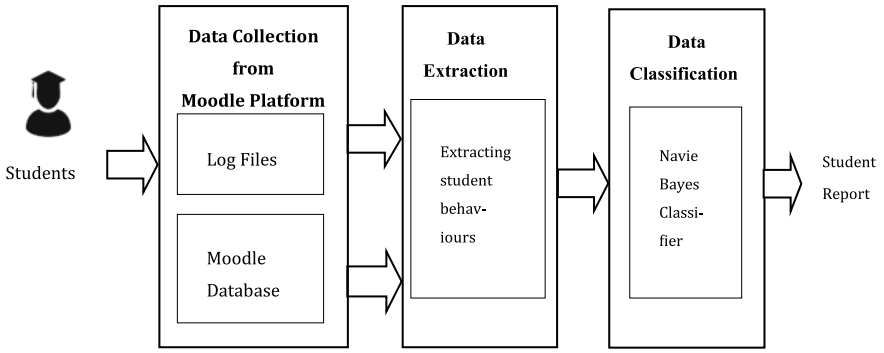


Fig. 1 Overall system architecture

number, name, mode of attendance, university number, date of birth, date of registration, e-mail, phone number, education, profession, address, place of birth, and nationality. The implicit method acquires the student behavior (the type of files, the time spent with reading files, the number of times a particular file is accessed, the number of visits/postings in forum/chat, number of visits and time spent on assignments, amount of time spent on a test, number of performed tests, number of visits and time spent on assignments, etc.) automatically from the log events and tables stored in the database of Moodle LMS. The proposed system focuses on both the explicit and implicit methods to collect students' data.

Capturing student profile data. In the proposed system, the explicit data of students registered for the online courses in Moodle LMS are obtained using the software that manages student data in the institution. The static data such as identification number, name, mode of attendance, university number, date of birth, date of registration, e-mail, phone number, education, profession, address, place of birth, and nationality, etc., are obtained and stored in the comma-separated values (CSV) file for further processing.

Capturing Log Event. When the student logs in for the first time and enrolls for the course, all the course contents are made available and provide access rights to various facilities such as formative assessment, self-assessment quizzes, news forums, discussion groups, chat rooms, and easy authoring tools for creating course contents such as videos, PowerPoint presentation, pdf, word, and other files including hyperlink insertion created by the instructor. The students access the material as per their preferences.

While accessing the courses and access other facilities of moodle, their daily interaction is obtained by observing the behaviors that are recorded in the log files automatically. The instructors can access the reports that record which course contents and activities of a course have been accessed, when, and by whom. One of the reports 'Logs' shown in Table 2 generates a filtered report showing information about the student activities. The information stored is 'Time', 'User full name', 'Event context',

Table 2 Sample log file data

Time	User full name	Event context	Component	Event name	Description	Origin	IP address
5/08/20, 16:21	alaa alfahdi	File: Chapter 1-Networking	File	Course viewed	The user with id '523′ viewed the 'resource' activity with course module id '2358'	Web	5.21.253.78
5/08/20, 16:19	Alaa	Quiz: Chapter 1	Quiz	Quiz attempt viewed	The user with id '509′ has viewed the attempt with id '313' for the quiz with course module id '3243'	Web	5.21.253.78
4/07/20, 15:38	Sofia	Quiz: Chapter 2	Quiz	Quiz attempt reviewed	The user with id '539' has had their attempt with id '434' for the quiz with course module id '3243'	Web	5.162.128.223
4/07/20, 15:15	Hasna	Forum: Networking	Forum	Discussion viewed	The user with id '540' has viewed the forum discussion for the forum with course module id '3243'	Web	5.162.128.223
4/07/20, 15:14	Hasna	Assignment: Chapter 1	Assignment	Submission form viewed	The user with id '349′ viewed their submission for the assignment with course module id '3846'	Web	5.162.128.223

'Component', 'Event name', 'Description', 'Origin', and 'IP address'. The log events that are captured are stored in the database in the WAMP Server.

The details of the log files header are given below:

- **Time**: The date and time on which the student accesses the page while using the LMS.
- **User full name**: This identifies the name of the student who visited the LMS.
- **Event Context**: Involves the name of the Event accessed by the student.
- **Component**: This includes the component such as quiz, assignment, forum, chat, and resources accessed by the student.
- **Event name**: This involves the behavior and actions of the student who access the LMS.
- **Description**: This gives a detailed description of the events performed with user id and course id.
- **Origin**: Represents the origin of all messages.
- **IP address**: This is the temporary address assigned to the computer to identify the student by the Internet service provider (ISP). This address is also used to identify revisited students.

Capturing Moodle Tables. The moodle database has more than 200 tables that are created for each course. The data in these tables keep track of students' significant activities regarding that particular course. From these tables, the important data required for student interaction and other related actions can be obtained using Structured Query Language (SQL) queries.

Data Extraction. Both log events and moodle tables are used to extract student behavior and actions required for detecting learning styles. Table 3 shows the relationship between the learning styles of FSLSM dimensions, learning styles, behavior patterns, and learning contents. Based on this mapping, the corresponding behaviors from the log files and moodle database can be extracted for each student to determine their learning style as the final output. This output supports educators to recommend learning content to the students based on their learning style and also helps students to understand the concepts in a better way.

Extracting student behaviors and actions from log events. Log events are used to monitor the activity of each student. From the log events, the most important actions for each activity shown in Table 4 such as the type of files (PowerPoint presentation, video, graphics, pdf, word, hyperlinks) the student accesses, the time spent with reading files, the number of times the student access a particular file, the number of visits/postings in forum/chat, number of visits and time spent on assignments, amount of time spent on a test, number of performed tests, number of visits and time spent on assignments are accessed for each student and these data are concatenated and saved in a single CSV file which is used to determine the students' learning style.

Extracting student behaviors and actions from moodle tables. The moodle tables are used to extract the behavior patterns of students such as number of visits/postings in forum/chat, number of visits and time spent on assignments, amount of time spent on a test, number of performed tests, number of visits, and time spent on assignments

Table 3 Relationship between FSLSM dimensions, learning styles, behavior patterns, and learning content category

FSLSM dimensions	Learning styles	behavior patterns	Learning content category
Perception	Sensitive	Amount of time spent on a test, no. of revisions before handing in a test, no. of performed tests, no. of visits and time spent on examples, performance on questions regarding facts	Facts, case studies
	Intuitive	Performance on questions regarding theories, no. of visits and time spent on exercises	Theoretical text
Input	Visual	Amount of time spent on contents with graphics, performance in questions related to graphics	Image, diagram, charts, video, and others
	Verbal	No. of visits/postings in forum/chat	Audio, text, and others
Processing	Active	No. of visits/postings in forum/chat, no. of visits and time spent on exercises	Example, practical exercise, activity, discussion, experimental, problem solving, and others
	Reflective	Amount of time dealt with reading material	Question, examples, links, readings, and others

Table 4 Activities versus actions of WebLog files

Activities	Actions
Quiz	Quiz attempt viewed, quiz attempt reviewed, quiz attempt submitted, quiz attempt started
Assignments	Submission form viewed, course module viewed, submission created, a file has been uploaded, status of submission has been viewed
Chats	Message sent, chat viewed
Forums	Discussion viewed, post created
Word	Course viewed
PDF	Course viewed
Powerpoints	Course viewed
Videos, graphics	Course viewed
Hyperlinks	Course viewed

Table 5 Set of attributes selected per student in moodle courses

Name	Description
id_student	Identification number of the student
id_course	Identification number of the course
num_sessions	Number of sessions
num_assigment	Number of assignments done
num_quiz	Number of quizzes taken
a_scr_quiz	Average score on quizzes
num_posts	Number of messages sent to the forum
num_read	Number of messages read on the forum
t_time	Total time used on moodle
t_assignment	Total time used on assignments
t_quiz	Total time used oil quizzes
t_forum	Total time used on forum

are obtained using SQL queries. Tables 5 and 6 show the summary table which integrates the most important information about the online activities in the courses.

Data Classification using Naïve Bayes Classifier. The demographic characteristics of the students and their behaviors and actions extracted from log events and moodle tables in some of the tasks such as chats, forums, tests, assignments, and files accessed are collected in three separate CSV files for each dimension and are analyzed and classified using the Naïve Bayes classifier to identify the learning style of students.

The Naïve Bayes classifier algorithm is the most straightforward and fast classification algorithm successfully applied for various applications such as text classification, sentiment analysis, and recommender systems. Initially, the train_test_split() method is applied to split the data set as 70 and 30% for the training and testing, respectively. Then the training is done by calling the method GaussianNB() and then fit() method used to fit the training data.

Table 6 Online activities of students in the moodle courses

Name	Description
course	Identification number of the course
n_assigment	Number of assignments done
n_quiz	Number of quizzes taken
n_quiz_a	Number of quizzes passed
n_quiz_s	Number of quizzes failed
n_posts	Number of messages sent to the forum
n_read	Number or messages read on the forum
total_time_assignment	Total time spent on assignments
total_time_quiz	Total time spent on quizzes
total_time_forum	Total time spent on forum

4 Experimentation and Results

The proposed approach is experimented on Moodle LMS by collecting the log events and moodle tables used by approximately 100 students studying one semester in computing science course in the institution.

The performance is evaluated based on various parameters such as accuracy, precision, and recall. Table 7 contains the results of the performance of the Naïve Bayes classifier, showing a classification report that shows the values of precision, recall, $F1$-score, and accuracy, classified to the three dimensions of FSLSM. The overall classification accuracy obtained is 85%.

The final results are used to cluster the students that have the same learning style based on FSLSM three dimensions. Clustering is a technique of grouping data into a smaller number of clusters so that the records in the group are very similar to each other based on the attributes of the data compared. Figure 2 shows the cluster of sample students populations based on learning style according to FSLSM dimensions.

5 Conclusion

The paper proposes an analysis of student data on the e-learning portal moodle to determine the learning style of each student based on FSLSM. The final result shows that the student learning style is classified precisely and accurately using the Naïve Bayes classifier based on three FSLSM dimensions thereby obtaining an overall classification accuracy of 85%. The research is restricted to a limited number of students and focused on the computer science domain. Future research could include more number of students under various domains. Moreover, further research could extend by extracting more student activities from the moodle tables for more accurate learning style classification and also attempt with other standard classifiers and use it for predictions. The final result supports educators to analyze the student preferences and design their learning contents based on their knowledge and requirements. The educators can either improve or change their approaches to working with students or to engage more with students. The educators can either change their approach for the specific student or opt for learning materials that better fit the student's needs.

Table 7 Classification report

Classifier	Performance measures	FSLSM dimensions				
		Processing				
		Extremely_reflective	Medium_reflective	Neutral	Medium_active	Extremely_active
Navie Bayes	Precision	1.00	0.78	1.00	0.64	1.00
	Recall	1.00	1.00	0.77	1.00	1.00
	*F*1-score	1.00	0.88	0.87	0.78	1.00
	Accuracy	87%				

FSLSM dimensions								Overall accuracy
Perception					Input			
Extremely_intuitive	Medium_intuitive	Neutral	Medium_sensitive	Extremely_sensitive	Extremely verbal	Extremely visual	Neutral	
1.00	1.00	1.00	0.19	0.00	1.00	1.00	1.00	
1.00	1.00	0.28	1.00	0.00	1.00	1.00	1.00	
1.00	1.00	0.43	0.33	0.00	1.00	1.00	1.00	
67%				100%				85%

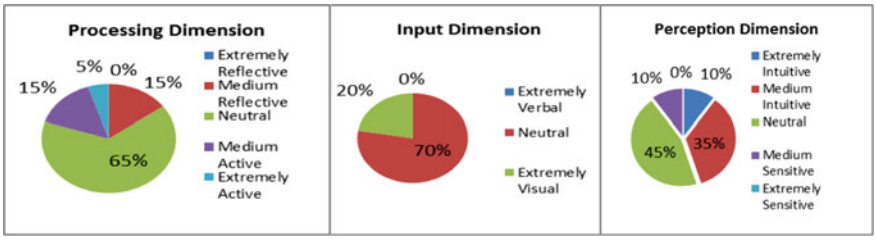


Fig. 2 Cluster of students populations based on learning style

Acknowledgements This paper is part of funded Project 'Construction and Retrieval of Ontology Based Semantic Learner Profile' from Research Council, Sultanate of Oman in call 2018.

Reference

1. Siemens, G., Gasevic, D., Haythornthwaite, C., Dawson, S., Buckingham Shum, S., Ferguson, R., Duval, E., Verbert, K., Baker, R.S.J.: Open learning analytics: an integrated & modularized platform proposal to design, implement and evaluate an open platform to integrate heterogeneous learning analytics techniques project overview (2011). www.solaresearch.org
2. Calvo, R.A., Markauskaite, L., Trigwell, K: Factors affecting students' experiences and satisfaction about teaching quality in engineering. Australasian J. Eng. Educ. **16**(2), 139–148 (2010). https://doi.org/10.1080/22054952.2010.11464049
3. (Tony) Bates, A.W., Sangra, A: Managing technology in higher education: Strategies for transforming teaching and learning. 1st edition, John Wiley & Sons Inc (2011)
4. Baker, R.S., Inventado, P.S.: Educational data mining and learning analytics. In: Larusson, J.A., White, B. (eds.) Learning analytics, pp. 61–75. Springer, New York (2014)
5. ERIC—EJ982673: Predictive modeling to forecast student outcomes and drive effective interventions in online community college courses. J. Asynchr. Learn. Netw. (2020). Retrieved 16 Aug 2020. https://eric.ed.gov/?id=EJ982673
6. Learning analytics—MoodleDocs (n.d.). Retrieved 13 Aug 2020, from https://docs.moodle.org/33/en/Learning_analytics
7. Moodle. Accessed 10.01.2018
8. Hasan, L.: The usefulness and usability of Moodle LMS as employed by Zarqa University in Jordan. J. Inf. Syst. Technol. Manag. **16**, 1–19 (2019). https://doi.org/10.4301/s1807-1775201916009
9. Ahmad, N.B., Shamsuddin, S.M., Abraham, A.: Granular mining of student's learning behavior in learning management system using rough set technique. In: Computational Intelligence for Technology Enhanced Learning, p. 273 (2010). https://doi.org/10.1007/978-3-642-11224-9_5
10. Moodle plugins directory. Available: https://moodle.org/plugins/. Accessed 29 Nov 2015
11. Kolekar, S.V., Pai, R.M., Manohara Pai, M.M.: Adaptive user interface for Moodle based e-learning system using learning styles. Procedia Comput. Sci. **135**, 606–615 (2018). https://doi.org/10.1016/j.procs.2018.08.226

Multiple Criteria Decision Making to Improve Retention Ratio of Employees by Identifying and Analyzing Critical Prevention Factor

Nilay M. Vaidya, Jaimin N. Undavia, and Atul Patel

Abstract One of the core assets of any organization is their employees. Retention and appraisal of them are one of most challenging tasks for the higher authorities, as replacement is easy but replacement in any ongoing project becomes costly affair in terms of time as well as money. The idea behind this study is to identify the critical preservation factor (CPF) of an employee that helps higher authorities in taking corrective steps in posing healthy culture. Here, we are proposing a model that helps in identifying the 360° critical effect measure that reflects the contributory factor and happiness factor of an employee in the organization. There are many factors which need to be taken while finding the critical factor. Many of the factors have known value; many of the factors have grey value and many have unknown value. The model uses the multiple criteria grey decision-making theory that tracks and uses these factors collected or derived from many stakeholders to find the critical factor. This critical factor helps not only to find the personality factor of an employee but also gives notification to the higher authorities in foreseeing the bonding between employees that reflects culture of the organization.

Keywords CPF · Critical effect measure · Critical factor · Grey value · Multiple criteria grey decision theory · Retention factor

1 Introduction

In the globalization and competitive market, many companies consider their primary asset as a great talent pool which allows the company to survive in the market.

N. M. Vaidya (✉) · J. N. Undavia · A. Patel
Charotar University of Science and Technology, Changa, Gujarat, India
e-mail: nilayvaidya.mca@charusat.ac.in

J. N. Undavia
e-mail: jaiminundavia.mca@charusat.ac.in

A. Patel
e-mail: atulpatel.mca@charusat.ac.in

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_88

The companies are having different appraisal and measurement schemes to evaluate and promote the employees. These approaches sometimes become very complex and uncertain that indulge into a greater challenge of retaining the employees in their organization. Retention of an employee is basically a term that denotes various measures being taken by the employer to motivate and encourage their employees in such a manner that (s)he remain with them for a longer period of time. The primary reason to retain them is to retain the talent [1]. Many of the departure factors not taken into consideration like in the study are: family migration, moving for the higher study, fired from the organization. The purpose of this study is to propose an approach in the organization that helps in evaluating the employee's performance, involvement, contribution, and work approach in such a way that generates a common indicator measure that plays an alarm and allows higher authorities to take better decisions and to take corrective actions in the environment to make it healthier environment.

It has been observed that, perceptual factors have a wide influence in the performance management system. Along with the healthy environment, the key factor which also is affecting the performance management system which is responsible in measuring the performance of the employee in the organization. Every organization appreciates the value of capitalizing in a healthy performance appraisal system. Organizations are aware that they gain advantages from the employees who have loyal, fruitful, and transparent performance evaluations by all the stakeholders within the organization.

Measuring the performance for the appraisal shall manage and capture qualitative as well as quantitative feedback that helps in tuning actionable insights within the organization as well as outside the organization.

2 Related Literatures

As many factors are affecting the performance of the employee in the educational institute, we here considered a few and tried to find the significance of these factors in the performance management system.

Internal politics in the environment makes the system dicey and ambiguous which plays a vital role in spoiling the mindset of all the employees [2].

Ethical environment and climate guides, strong bonding among employees, lower turnover rate, positive attitude, satisfaction in the job and feeling toward job find direct impact between climate of organization and satisfaction of the employees [3].

Many researchers have shaken their hands to address this critical issue in the growing corporates environment. Substantial number of papers and related literatures have been reviewed to propose the paper. Directly related and referred papers have described herewith. James and Mathew wrote an article in 2012 [1] to address the employee retention issue. They have described different strategies to retain employees in IT industries. They narrated such strategies in their paper titled "Employee Retention Strategies: IT Industry", and that was published in SCMS Journal of Indian Management. A paper titled "Perceived organizational politics

and employee silence: Supervisor trust as a moderator" was published in Journal of the Asia Pacific Economy in year 2016. Khalid and Ahmed [2] published that article and they have focused on the issue of employee retention where internal politics may be affecting their retention. Affecting parameters in employee retention were noted and published in the paper titled "Occupational values, work climate and demographic characteristics as determinants of job satisfaction in policing". Tomaževič et al. published that paper in 2018 [3] in the journal of Police Practice and Research—An International Journal. Employee retention and its affecting parameters were discussed in that article.

In month of August-2015, Vaidya et al. [4] published a paper titled "Evaluating learning effectiveness in collaborative learning environment by using Multi-Objective Grey Situation Decision Making Theory". Decision-making theory in collaborative leaning environment discussed in that paper which was been published in International Journal of Scientific and Engineering Research. Dazhi et al. [5] also discussed decision-making theory, in 2009 in their paper titled "Study on the Application of Multi-objective grey situation decision-making theory for transportation mode choice". The decision-making theory was discussed so cleanly and effectively, so as it was published by IEEE.

In the year 2016, Indian Journal of Science & Technology has published a research paper authored by Undavia et al. [6]. They have proposed a model through decision tree algorithm in their paper titled "Customized Prediction Model to Predict Post-Graduation Course for Graduating Students Using Decision Tree Classifier". A customized model was developed by them to generate classification results with high accuracy. A research paper titled "Determinants of Employee Retention" was published at IGI global platform by Melissa Rodrigues Ataide Silva, João Conrado de Amorim Carvalho, Alvaro Lopes Dias. They claimed that the study shows the reality in Brazil for the matter of employee retention. They have discussed issues, controversies, and problems of the employee retention problems. Mr. Tanmay Prakash Salunkhe defended his dissertation of Master of Science at Dubai and acknowledged the issues and affecting parameters in the employee retention domain. The dissertation was titled "Improving Employee Retention by Predicting Employee Attrition using Machine Learning Techniques" and suggested some valuable techniques for HR and project managers to retain their valuable employees in the organization.

Transparent and fair performance management system shall impact directly in building the positive attitude and behavior toward organization.

3 Objectives of the Study

An automated performance review system helps to identify, review, and act on the several known as well as unknown gaps between different stakeholders of an organization. The effective measure may help organization to

(1) Offer 360° performance appraisal

(2) Reward and recognition
(3) Betterment in performance plan and many more.

There are many misconceptions related to the appraisal process that inculcates all the stakeholders of the organizations connected.

(1) There are so many curves and long-lasting interconnected processes connects that makes the process complicated and complex.
(2) Many stakeholders are involved makes the process less user friendly.
(3) There are only weaknesses highlighted.
(4) The system is being used to promote the employees.
(5) System normally used once or twice in a year and many more …

The study is just to find and suggest a conceptual and perceptual model based on the certain criteria with their significance in constructing a performance management system. This study tries to map the criteria and performs the calibration with the interests of the employee in the educational field.

This study also tries to find the happiness and contributory factor of the employee in the institute/university. These factors help in making environment more competitive, live, and energetic, also to help higher authorities in preserving their major asset, i.e., their talented employees by foreseeing the consequent factor.

4 Methodology

The proposed approach to find the performance factor uses a performance cycle (Fig. 1):

(1) Setting up the standards by taking into the consideration the overall vision, mission, and objectives of the organization
(2) Fetching the day-to-day patterns as well as feedback on regular basis from all the interconnected and related stakeholders.
(3) Calibrating the real factor versus the critical factor and also comparing the same with the previous standards.

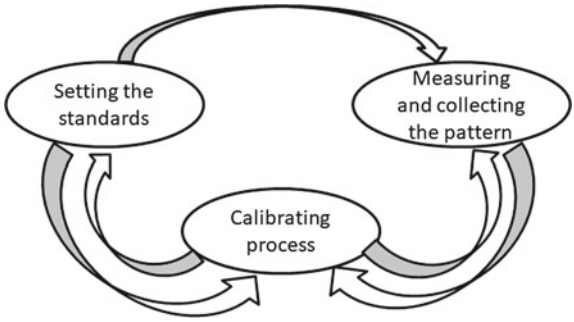


Fig. 1 Overall working of the model

Performance factor, also can be considered as evaluation factor in the process of appraisal, is an assessment process where the many internal as well as external factors reviews performance at work [6].

To find the overall contributory factor of employees working in the organization, we are coming up with some preliminary known details being collected and analyzed that give an effect measure which can then be compared with some grey criteria which has got some weights assigned depending upon the objective and significance in the given frame of scope and environment.

As decisionmakers are facing difficulty in such a complex problem, multiple criteria decision making (MCDM) is used to deal with it. As the fuzzy set approaches and human evaluation scheme are used as different input criteria to the system, these input to the system need weights being assigned that help in finding better meaningful outcome. Many of the details may have degree of uncertainty or may have dependency factor associate. Such details to the system are called grey variable/constant. These normal inputs to the system are having specific weights assigned to them as they behave normally in the system and are responsible in generating the known output. The degree of uncertainty in the criteria we call it grey criteria needs to be processed separately as they need different weights to assigned in the different environmental constraints (Fig. 2).

The technique uses multiple attributes as an input in the decision-making system. System performs series of actions on it in such a way that it generates the output variable as well as grey variables for each employee. It then creates a performance matrix consisting of each criterial measure of employees in the organization. This helps in finding the overall effect measure by using the multi-objective grey situation decision-making theory which helps in ranking and generating decision graph can be used by decisionmakers improve the decision-making process. The outcome then helps in understanding level of involvement of the employees in the organization that is capable of forecasting the involvement of employees using grey theory and multi-objective programming [4].

This theory includes basic four measuring elements: event, strategy, effect, and target. In our case, we considered the event and adaptive capability. For the same, we considered known input as basic profile and as grey as well as uncertain details

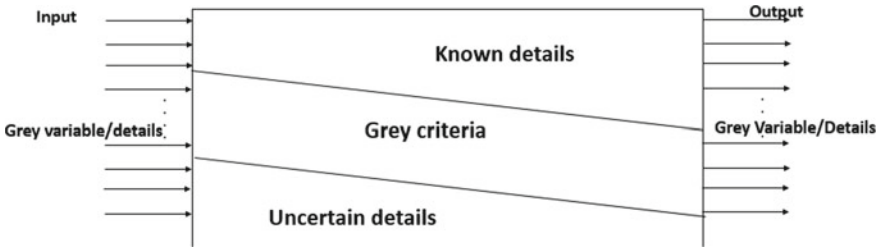


Fig. 2 Types of input to the model

we considered few surveillance details. We considered following criteria to measure the adaptively in the organization [7, 8]:

(1) Span of service in the current organization
(2) Involvement in the project (team leader/member/client communicator)
(3) Involvement in the organizational activities (other than routine)
(4) Initiatives taken in the organization
(5) Bonding and helping the peer (solving problems of peer via internal FAQ or Forum, etc.)
(6) Clearing or pursuing the higher studies during the job
(7) Clearing/attempted the competitive exams or certifications
(8) Frequency of writing or reading peer blogs
(9) Aggregate rating given by peer, authorities, and other stakeholder
(10) Involved in any unethical practices
(11) Missing or not meeting deadlines

Data collected for each of the criteria may have certain value, fuzzy value, or greyness. These fuzzy as well values contain greyness shall be processed in such a way that it can be used in calculating the effective measure. The crisp values shall be calculated and taken for the further processing [9]. Data processing is to find the contributory factor of an employee in the organization, using information systems by applying multi-objective grey situation decision-making theory. This makes the system more adaptive by flavoring the intelligent agent onto it. Each criterion used is classified into upper, lower, or central measures. The decision making during the measurement of the performance factor includes four major elements: (i) action, (ii) approach, (iii) result, and (iv) measure. In case of finding the critical measure factor, there are many criteria considered. Depending upon the deviation in the scale in the criteria, the effect measure is classified into either: (i) upper, (ii) lower, and (iii) center. The criteria use the effect measure that is depended upon the deviation in mass of the group, objective of the criteria, criticality defined for the criteria, and many other criteria can be taken into consideration to have the effect measure. To generate the effective measure, we classified the criteria in upper, lower, or central layer [4, 9].

For the upper level, it considers criteria value as higher the better and calculates the effect measure for the criteria using

$$r_{ij} = \frac{u_{ij}}{u_{\max}} \tag{1}$$

where u_{ij} is the actual effect measuring value and the $u_{\max}$ is the maximum data.

For the lower level, it considers criteria value as lower the better and calculates the effect measure for the criteria using

$$r_{ij} = \frac{u_{\min}}{u_{ij}} \tag{2}$$

where u_{ij} is the actual effect measuring value and the $u_{\min}$ is the minimum data for the criteria $u_{ij} \geq u_{\min}$: $r_{ij} \geq 1$ [5, 10].

There many such effect measures are calculated, then there is a need to have a comprehensive effect measure. Decision factor also needs to be generated to prepare the decision matrix which then helps in identifying comprehensive value. Employee effect measure calculated as sij and decision factor also needs to be calculated.

The decision matrix which is prepared as

$$\begin{bmatrix} \frac{r_{11}^{(k)}}{s_{11}} \\ \frac{r_{21}^{(k)}}{s_{21}} \\ \cdots \\ \frac{r_{n1}^{(k)}}{s_{n1}} \end{bmatrix} \tag{3}$$

Similarly, the comprehensive matrix then can be calculated as:

$$r_{ij}^{(\sum)} = \frac{1}{N} \sum_{k=1}^{n} r_{ij}^{(k)} \tag{4}$$

This comprehensive value is being used to find the contributory factor of an employees. Also, several such cluster-based factors help in finding the happiness factor of an employee in the different sectors. This gives an opportunity to higher authorities to introspect their routine procedures being followed and also gives an indicator to all the stakeholders how to improve the bondage and environment increase the productivity in the organization.

The comprehensive value gives an indication about the effectiveness of the employee, environment, and the happiness factor of an employee. The value indicator is shown in Fig. 3.

Problem solving by the multi-objective situation decision-making theory for sample criteria for four employees. Calculation to find the effective measure by considering sample criteria is shown in Table 1:

Defining the effect measure for different parameters will be:

$$r_{ij} = [7/7, 6/7, 2/7], r_{ij} = [6/8, 8/8, 8/8],$$

Fig. 3 Critical factor classification

Table 1 Sample data for three employees

Criteria with scale	Sample 1	Sample 2	Sample 3
Span of service in the current organization **Scale**: Upper **Range**: 0–10	$\mu_{11} = 7$	$\mu_{12} = 6$	$\mu_{13} = 2$
Involvement in the project (team leader/member/client communicator) **Scale**: upper **Range**: 0–10	$\mu_{21} = 6$	$\mu_{22} = 8$	$\mu_{23} = 8$
Involvement in the organizational activities (other than routine) **Scale**: Upper **Range**: 0–10	$\mu_{31} = 8$	$\mu_{32} = 6$	$\mu_{33} = 4$
Initiatives taken in the organization **Scale**: Upper Range: 0–10	$\mu_{41} = 5$	$\mu_{42} = 5$	$\mu_{43} = 5$
Bonding and **helping** the peer (solving problems of peer via internal FAQ or Forum, etc.) **Scale**: Upper **Range**: 0–10	$\mu_{51} = 9$	$\mu_{52} = 8$	$\mu_{53} = 7$
Clearing or pursuing the higher studies during the job **Scale**: Upper **Range**: 0–10	$\mu_{61} = 0$	$\mu_{62} = 5$	$\mu_{63} = 3$
Clearing/attempted the competitive exams or certifications **Scale**: Upper **Range**: 0–10	$\mu_{71} = 8$	$\mu_{72} = 6$	$\mu_{73} = 5$
Frequency of writing or reading peer blogs **Scale**: Upper **Range**: 0–10	$\mu_{81} = 9$	$\mu_{82} = 9$	$\mu_{83} = 7$
Aggregate rating given by peer, authorities, and stakeholders **Scale**: Upper **Range**: 0–10	$\mu_{91} = 7$	$\mu_{92} = 8$	$\mu_{93} = 9$
Involved in any unethical practices **Scale**: Lower **Range**: 0–10	$\mu_{101} = 2$	$\mu_{102} = 0$	$\mu_{103} = 0$
Missing or not meeting deadlines **Scale**: lower **Range**: 0–10	$\mu_{111} = 2$	$\mu_{112} = 1$	$\mu_{113} = 0$

$$
\begin{aligned}
r_{ij} &= [8/8, 6/8, 4/8], r_{ij} = [5/5, 5/5, 5/5] \\
r_{ij} &= [9/9, 8/9, 7/9], r_{ij} = [0/5, 5/5, 3/5], \\
r_{ij} &= [8/8, 6/8, 5/8], r_{ij} = [9/9, 9/9, 7/9] \\
r_{ij} &= [7/9, 8/9, 9/9], r_{ij} = [1/2, 0, 0], r_{ij} = [1/2, 0, 0] \\
r_{ij}^{(\sum)} &= \left[r_{11}^{(\sum)}, r_{12}^{(\sum)}, r_{13}^{(\sum)} \right] = [0.77, 0.73, 0.53]
\end{aligned}
$$

Here looking at the effective measure of three sample employees, it is clear that compared to sample 2 and 3, the first employee is having higher retention rate compare to the other two as the effect measure is nearer to 1.

5 Conclusion

The formal model is to provide and agent that keep observing the behavior and pattern of an employees in an organization. It also extracts and then processes using the multi-objective grey situation decision-making theory to find the retention factor for employees. These help in finding the happiness factor of employees. An aggregate factor also helps higher authorities in foreseeing the retention factor of employees. Clustering the criteria and finding the effect measure help the higher authority to identify whether or not any change is required in the current environment of an organization.

References

1. James, L., Mathew, L.: Employee retention strategies: IT industry. SCMS J. Indian Manag. **9**(3) (2012)
2. Khalid, J., Ahmed, J.: Perceived organizational politics and employee silence: supervisor trust as a moderator. J. Asia Pacific Econ. **21**(2):174–195 (2016). Available at: https://doi.org/10.1080/13547860.2015.1092279
3. Tomaževič, N., Seljak, J., Aristovnik, A.: Occupational values, work climate and demographic characteristics as determinants of job satisfaction in policing. Police Pract. Res. 1–18 (2018). Available at: https://doi.org/10.1080/15614263.2018.1500282
4. Vaidya, N., Sajja, P., Gor, D.: Evaluating learning effectiveness in collaborative learning environment by using multi-objective grey situation decision making theory. In: Int. J. Sci. Eng. Res. (IJSER) **6**(8):41–45. ISSN 2229-5518
5. Dazhi, J., et al.: Study on the application of multi-objective grey situation decision-making theory for transportation mode choice (2009). 978-0-7695-3583-8/09. IEEE. https://doi.org/10.1109/ICMTMA.2009
6. Dolia, P., Patel, A., Undavia, J.N.: Customized prediction model to predict post-graduation course for graduating students using decision tree classifier. Indian J. Sci. Technol. **9**(12):1–7 (2016)
7. https://kissflow.com/hr-process/performance-management/performance-appraisal-system/#:~:text=A%20performance%20appraisal%20system%20manages,turning%20them%20into%20actionable%20insights. Accessed on 25 Aug 2020
8. https://www.indeed.com/career-advice/career-development/performance-review-phrases. Accessed on 28 Aug 2020
9. Vaidya, N.M., Patel, K.K.: Learner performance and preference meter for better career guidance and holistic growth. 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_5
10. Wei, W., Xuewu, C.: Transportation Planning. The People's Transportation. Beijing (2007)

Preference-Oriented Password-Based Authentication

Amanpreet Kaur and Khurram Mustafa

Abstract Cyberspace security is largely authenticated by password-based schemes. Its usage is increasing day by day, and efficiency considerations are inevitable along with the effectiveness. Consideration of preference of users enhances the credential safety as well as the efficiency of an authentication mechanism. This paper presents a novel preference-based password authentication scheme (PPA). The scheme uses a recognition-based graphical design and offers greater resistance to shoulder surfing attacks. The trial involved the simulation of the basic PPA scheme, which provided users to choose their preferred method to register and log in. This freedom of choosing their authentication process conforms to the usability principle and can enhance the security of the process. The results obtained in the user trial indicate that the PPA is efficient and secure.

Keywords Graphical authentication · Efficiency · Security · Usability · Attacks · Passwords · Preference-based authentication · Recognition technique · PIN · Face

1 Introduction

During the last fifty years decades, textual passwords significantly affected authentication regardless of any security issues. Textual passwords are generally used, where a user can have a user-id and password known to the user only [1]. Generally, users tend to choose easily guessable passwords [2] and make it a habit to reuse the same passwords among multiple accounts for short-term benefits [3]. Many warnings and guidelines have been given to the users to avoid writing down their passwords to mitigate their misuse [4–6]. Hence, organizations should maintain a balance for

A. Kaur (✉) · K. Mustafa
Jamia Millia Islamia, New Delhi, Delhi, India
e-mail: iamanpreetkaurcs@gmail.com

K. Mustafa
e-mail: kmustafa@jmi.ac.in

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_89

regulating the password policies between memorability issues and security breaches [7].

Weak passwords are prone to security breaches, even then they are prevalent and common. This problem becomes catastrophic when the users are left to generate the passwords on their own. In an endeavor to conquer the shortcomings of traditional passwords, graphical password authentications were developed as a conceivable security improvement. A graphical password is an authentication framework that works by having users select pictures in a specific order provided in a graphical user interface (GUI).

Graphical passwords are an alternative as they could mitigate certain security attacks and curb the weakness of traditional passwords.

Studies related to user authentication show that usability is impacted by contextual attributes (human, technology, design) [8]. These findings support the need of adaptivity in user authentication design; the support of schemes intelligently adapts to the precise traits of each person depending upon the situational parameters or context of use. However, it is observed that some providers still follow 'one-size-fits-all' paradigm; the same password authentication scheme is deployed disregarding the fact that users have various characteristics that influence the usability of the task. Moreover, it is seen that the users' preferences are sometimes not seen by the experts and developers. Hence, it is desirable to focus on users demand without logging behind the security issues [9]. Hence, it is required to build authentication schemes that will bring utmost efficiency considering increasingly different users, their respective choices and situations.

2 Genesis

By the review of literature and developments during the study, a set of ideas appeared to be pertinent for a relatively efficient password scheme. Such ideas form the basis for the formation of the proposed scheme and include 'Preference', 'Recognition,' 'Design,' and context of use. The existing literature review on graphical passwords has driven us to follow perceptions that established the framework for this work.

2.1 Preference-Based Authentication

An authentication system in which users have freedom to choose their preferred method of authentication from the schemes available such as text passwords, object pass-tiles [10], persuasive cued click points [11], and persuasive text passwords [12]. The schemes display relative strength and usability for users.

The scheme needs special CYOA [12] module to be implemented for the collaboration of all the techniques in one system, so the resource required for the CYOA is increased. Moreover, such schemes may require special software for installation

such as Web browser, Adobe Flash, or Java. So, a system where users can choose a preference scheme by the virtue of familiarity is taken care of automatically would be a nice fit.

2.2 Recognition-Based Techniques

Passface algorithm [13] is largely based on the concept that human faces are recalled easily than any other picture. This scheme helps users to select their password by choosing four faces from their FACES database at the registration time. During login, a grid with nine human faces is presented, out of which a user recognizes and selects chosen face. The scheme is repeated four times till the user can identify all the correct faces.

According to Brostoff and Sasse [13], Passfaces are easy to memorize. The comparative study between text passwords and Passfaces showed that the login failure rate of Passfaces has decreased to one-third. Tari et al. [14] studied that Passfaces with keyboard are less vulnerable to guessing attacks as the attacker must look at two places at the same time. According to a study by Valentine [15] Passfaces are memorable over long period of time. Passfaces can easily be mapped to usability features of ISO standards such as easy to use, create, and remember [14]. According to the study conducted by Levin [16] on the Passfaces, there has been obvious pattern selection by users. It is more likely that a user will choose faces of people from the same race. In lieu of these problems, it can be said that Passfaces are vulnerable to guessing attacks [16].

2.3 Quick Access to Users

Primarily logging onto a system is not an important task, that to multiple times. Users want to speedy get beyond authentication. The traditional passwords take about eight o twenty seconds for authenticating, while PIN takes 5 to 10 s to enter. Biometrics takes the least time by taking less than 5 s. According to a study by Werner et al. [17], well-designed graphical passwords can reflect rapid access times only, slightly extra slowly than conventional password schemes. So, providing quick access will empower more users and hence successful authentication.

2.4 Context of Use

A situation where credentials can be ordered according to the level of risk involved, i.e., benefit lined up with risk analysis. The contextual data plays significant role for

taking access-related decisions. Access control decisions should be taken dynamically by considering contextual data (user behavior, time channel, etc.). Various frameworks were recommended that utilize context [18].

3 Preference-Oriented Password-Based Authentication

Graphical schemes have been developed due to the usability problems faced by textual passwords. Graphical passwords can be dual coded visually and semantically, hence can be remembered after long intervals [19]. However, they have been blamed for being susceptible to over the shoulder attack (OSA). To achieve this security goal, schemes have been designed to being less vulnerable to OSA [20].Common strategies used to decrease OSA are grouping targets with distracters [20], translating the targets mentally to different location [21, 22], and disguising targets by changing their appearance [23]. It is blending all numbers, characters, and pictures, hence strengthening the authentication scheme. The proposed scheme hereby tries to accommodate the following challenges:

- Decrease the user effort
  - Making the password easier to remember.
  - Giving users a choice to select between two authentication mechanisms for input. Based on their experience and preference, they can choose any scheme. Hence, their training cost is likely to be very low.
  - The length of the password is less so the user can quickly access their system by login fastly than another graphical authentication system, as two different entities are embedded together without compromising the security
- Provide over the shoulder resistance
  - Disguising appearance of targets
  - Using a keyboard to enter the password and not clicking with a mouse

We propose recognition-based graphical authentication which associates contextual parameters and user preference to authenticate. In this scheme, a user will be given a choice (depending upon its preference) to select its authentication scheme from pin-based authentication or face-based authentication given below, respectively. Once the user selects the scheme, initially at the first step users will select a random shape in which at the next step, the user-chosen pin or faces will be inserted.

3.1 Face-Based Authentication

Since Passfaces have high memorability and low login failure rate [13], the proposed scheme is highly likely to have improved memorability. Face-based authentication

takes the benefit of recognizing the faces and increases the security by embedding it in extra shape and typing the alphanumeric code as the password.

3.2 Pin-Based Authentication

Textual password and PIN are the widely used knowledge-based authentication schemes. According to a study by Weiss and De Luca [24], it has been observed that PINs are low on memorability and security. Hence, there was a move on graphical password authentication. The proposed scheme takes the advantages of both the graphical and PIN authentication scheme. The scheme requires users to remember the number of images in a sequential way such as PIN, which provides decreased speed and error rate.

4 Methodology

A study was conducted to examine the goals. Participants were approached to use the prototype and afterward answer a related survey. The following sections clarify the prototype and blueprint the means and techniques the participants needed to follow.

4.1 Software Prototype

The prototype of preference-based password authentication (PPA) is based on Navision interface. Navision is a business ERP application which is having a front end in.net and backend is controlled by SQL DB; the language used to control the business logic is CAL which is Client Application Language. To simulate and test, basic PPA was implemented to check out the feasibility of the design.

4.2 Procedure and Steps

All participants were contacted via e-mail or face-to-face conversation. In all 60 participants were targeted, however, only 40 made it through at the end. The software prototype saved the participant user-id, names, address, date of birth, shape, an order of images selected details of login attempts as well. The study considers short-term performances for the participants; hence no longitudinal study was accomplished. However, the long-term performance will be pursued as a future study enabling the comparison between the short-term and long-term performance.

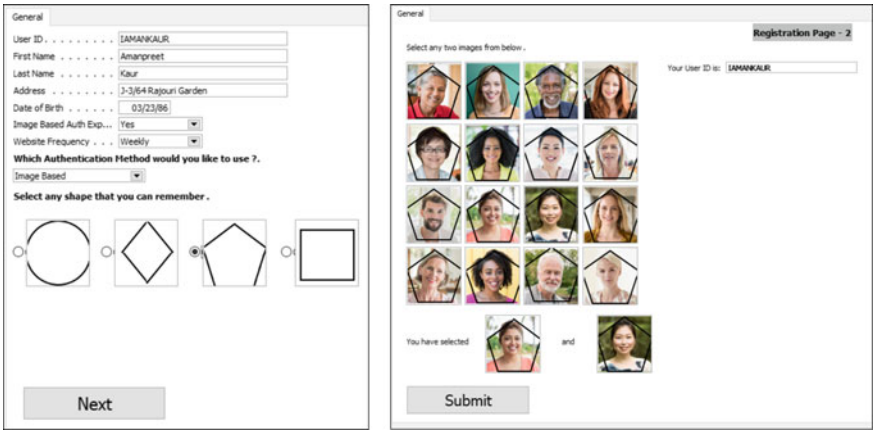


Fig. 1 Registration of Face-based authentication

Registration Time

At the time of registration to create a password, the user should choose 2 faces/pin from a grid of 16 images and a shape from eight defined shapes. The images will be embedded into a user-selected shape as shown in Fig. 1 (shows registration of face-based authentication, i.e., similar to pin-based authentication), respectively.

Login Time

At the time of login, a user must remember the faces/pin and the shape he/she chooses at the time of registration. The login page shows a grid of 16 faces (images) with different embedded shapes and alphanumeric characters shown in each image (Fig. 2 shows the login of pin-based authentication, i.e., similar to face-based authentication). The user remembers the shape and two faces/pin and types the alphanumeric code written in each image as a password for the scheme.

Sequencing of number images is a characteristic of the recall-based algorithm (passpoint) making the shoulder surfing attack less vulnerable. It also minimizes shoulder surfing attacks by using the WIW approach [25] of integrating similar sequencing and using alphanumeric codes.

5 Results and Analysis

Forty participants participated in the prototype testing. The minimum age of the participants was 19 years old and the maximum age was 73 years old. Most people are young, i.e., lie in the age group of 19–38. Most people, i.e., 55% of people, have experience with image-based authentication (Table 1).

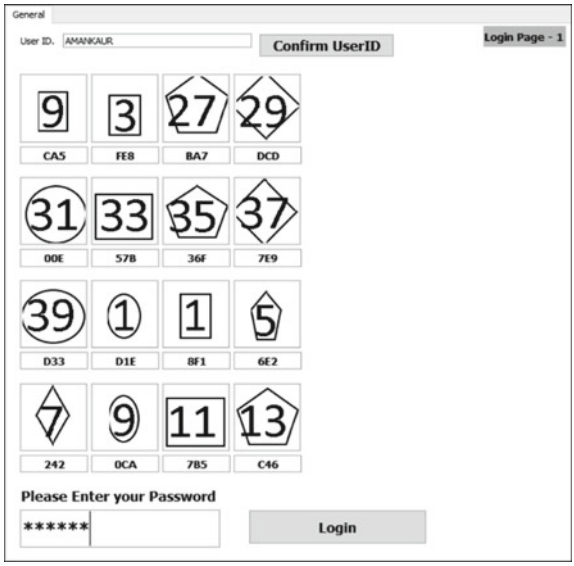


Fig. 2 Login pin-based authentication

Table 1 Participants characteristics for the main experiment

Age	
19–38	62.50%
39–58	25%
57–78	12.50%
Frequency of use	
Daily	42.50%
Weekly	37.50%
Monthly	20%
Experience for image-based authentication	
Yes	55%
No	45%

5.1 Login Time

The database of the simulated software recorded the time required to log in. The average time for both the authentication methods (PPIN, PPIA) was calculated and the mean differences show that there is not much difference between the two methods following their login time. The standard deviation for PPIN is marginally low by 0.07 from PPIA (Table 2). However, their mean time ranges around 11 s.

Table 2 Login time

	Authentication time	
	Mean	SD
PPIN	11.04	2.70
PPIA	11.12	2.77

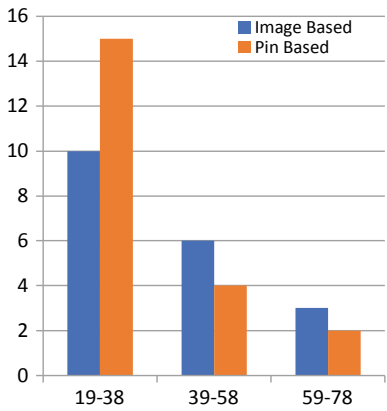


Fig. 3 Authentication method and age

5.2 *Authentication Method and Age*

It has been observed that young people (19–38) prefer to choose pin-based authentication over image-based authentication (Fig. 3). However, image-based authentication is a preferable option for older adults. This can be assumed from the fact that older people have memorability issues and images providers more memorability [26].

5.3 *Authentication Method and Frequency of Use*

It has been observed that people whose frequency of use is daily or weekly are generally opting for pin-based authentication while those whose frequency of use is limited to monthly are generally opting image-based authentication (Fig. 4). This also reveals that people assume that face-based authentication is suitable for operations where the authentication is not often and requires long-term recall.

5.4 *User Review Results*

The user experience of the proposed scheme PPA is calculated with the help of a user experience questionnaire (UEQ) [27]. The 27 parameters are summarized

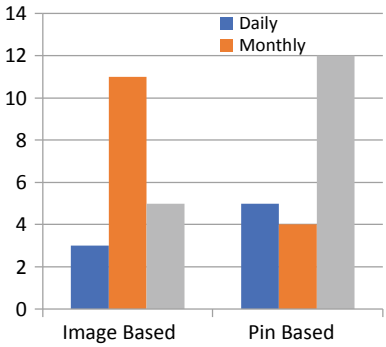


Fig. 4 Authentication method and frequency of use

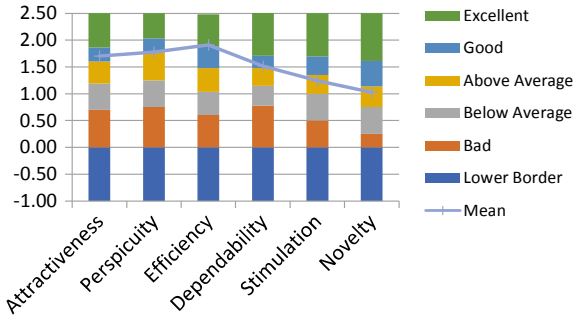


Fig. 5 Benchmarks for prototype resulting from UEQ

by the following six attributes: attractiveness, perspicuity, efficiency, dependability, stimulation, and novelty [27].

Out of the six attributes covered in the questionnaire, users found that the scheme was the most efficient with the mean of 1.92 (SD = 0.89). The users also found the authentication mechanism was easy to familiarize and straightforward as shown by the perspicuity value of 1.77 (SD = 1.35) and is quite attractive (mean = 1.70, SD = 1.40). Overall, most of the respondents are glad about the performance of our proposed prototype as all the attributes are forming above average level (Fig. 5) and that is a significant result.

6 Security Assurance and Attack Resilience

Security assurance provides trust in the security-related properties and functionalities, just as the activity and organization strategies, of a developed solution [28]. It is the reason for increasing legitimate confidence that an application will reliably show all the properties required to guarantee its reliable operation.

Table 3 PPA levels

Authentication scheme	Image chosen	Shapes chosen	Entropy	password length
Basic PPA	Two	Out of four	2 * (log 216 * 4) = 12	Six
Customized PPA	Four	Out of sixteen	4 * (log 225 * 16) = 34.58	Twelve

6.1 Security Assurance

The degree of assurance (LOA) required depends on the consequence of authentication error as well as the abuse of credentials. As the results of authentication error increments, the degree of assurance should also be incremented. Casual or low worth solicitations will require less stringent assurance [29]. Preference-based password authentication (PPA) offers adaptive architecture and hence offers two types of authentication levels:

- **Basic Preference-Based Password Authentication (PPA)**: It can be applied when the impact of the risk is low, i.e., level I.
- **Customized Preference-Based Password Authentication (PPA)**: It can be applied when the impact of the risk is moderate to high, i.e., level II and III. The user can choose from one of the two schemes when the risk level is moderate. However, for assurance level III, the user must register for both the schemes and one of the two can serve as fallback authentication. Moreover, in level III, the PPA can impose strict rules, with multigraded authentication.

Table 3 describes the different levels of preference-based authentication schemes with the password length.

6.2 Attack Resilience

Our foremost goal is to determine how resilient our password scheme is against a determined threat attack and identify gaps in our mechanism to respond to a targeted attack [30]. The following are the detailed description of attacks on the two available schemes. Table 4 describes different attacks and PPA response to it.

1. **Brute-force Attack**: The proposed scheme can be cracked by guessing two faces from a grid of 4 × 4. There are 240 possible patterns for selecting two faces out of 16 faces and 1560 possible patterns for selecting two numbers out of forty. Moreover, the addition of a defined shape will help in disguising the appearance of targets as the same face is available in more than one shape. This will also multiply upon the selection of patterns by a value of four. Moreover, the alphanumeric password will increase security by 36 × 36 × 36 (46,656). Furthermore, restricting login attempts will enhance the security of the scheme.

Table 4 Security attacks

Attack types	Textual passwords	Passfaces	Man et al.	Deja Vu	Triangle	PPA
Dictionary attack	Y	–	N	N	N	Y
Brute-force attack	Y	Y	Y	Y	Y	Y
Guessing attack	Y	Y	**N**	**Y**	Y	Y
Spyware attack	**Y**	**N**	Y	N	N	N
Shoulder surfing	**Y**	Y	N	Y	N	N
Social engineering	–	**N**	N	N	N	N

2. **Social Engineering Attack**: During login, the images are placed randomly in a grid with a unique code inserted inside it, making the password entry unique every time.
3. **Resistance over Shoulder Surfing Attack:**
   - The user-chosen shape embedded over a face/pin image will help in disguising the targets. Assuming a user is assigned a rhombus as its shape, the login page will show the user-selected faces with rhombus, as well as other shapes on the selected faces to disguise the target.
   - It is observed that there is a smaller number of guessing attacks when the keyboard is used instead of clicking of a mouse. The proposed scheme uses the keyboard to enter the alphanumeric keys associated with the faces.

7 Conclusion

A preference-based password authentication (PPA) is an efficient authentication scheme through which users can select from one of the available schemes depending upon the user preference for pin or faces. The user can choose to authenticate from pin or faces depending upon the familiarity or cognitive skills. PPA helps users to bring down their effort, as the schemes available are familiar bringing down their learnability and cognitive workload issues. The users must memorize only two image objects and a shape that is embedded over image objects (lesser than earlier graphical authentication schemes). It provides resistance to shoulder surfing attacks as the shape used is disguising the targets. The time needed for the authentication (as per survey results) on an average is 11 s, which means that the time effort by the authentication is also nominal. Moreover, no extra hardware or software is required to implement PPA as both the objects for authentication (pin and faces) are stored as images. So, graphical password authentication is implemented with minimal economic and resource consumption.

References

1. Kaur, A., Mustafa, K.: A critical appraisal on password based authentication. Int. J. Comput. Netw. Inf. Secur. **11**(1) (2019)
2. Bonneau, J., Herley, C., Oorschot, P.C.V., Stajano, F.: The quest to replace passwords: a framework for comparative evaluation of web authentication schemes. In: 2012 IEEE Symposium on Security and Privacy (2012)
3. Florencio, D., Herley, C.: A large-scale study of web password habits. In: Proceedings of the 16th international conference on World Wide Web, pp. 657–666 (2007)
4. Inglesant, P., Angela Sasse, M.: The true cost of unusable password policies: password use in the wild. In: Proceedings of the Sigchi Conference on Human Factors in Computing Systems, pp. 383–392 (2010)
5. Riley, S.: Password security: What users know and what they actually do. Usabil. News **8**(1), 2833–2836 (2006)
6. Zviran, M., Haga, W.J.: Password security: an empirical study. J. Manag. Inf. Syst. **15**(4), 161–185 (1999)
7. Wood, C.: Effective information system security with password controls. Comput. Secur. **2**(1), 5–10 (1983)
8. Katsini, C., Belk, M., Fidas, C., Avouris, N., Samaras, G.: Security and usability in knowledge-based user authentication: a review. In: Proceedings of the 20th Pan-Hellenic Conference on Informatics, pp. 1–6 (2016)
9. Kaur, A., Mustafa, K.: Efficiency—a layered model for authentication. In: IEEE 8th International Conference on Reliability, Infocom Technologies and Optimization (ICRITO' 2020), 4–5th June 2020. Accepted for Publication
10. Stobert, E.: Memorability of assigned random graphical passwords. Ph.D. Dissertation, Carleton University (2011)
11. Chiasson, S., Stobert, E., Forget, A., Biddle, R., Van Oorschot, P.C.: Persuasive cued click-points: Design, implementation, and evaluation of a knowledge-based authentication mechanism. IEEE Trans. Dependable Secure Comput. **9**(2), 222–235 (2011)
12. Forget, A.: A world with many authentication schemes. Ph.D. Dissertations. Carleton University (2013)
13. Brostoff, S., Angela Sasse, M.: Are Passfaces more usable than passwords? A field trial investigation. In: People and computers XIV—usability or else!. Springer, London, pp. 405–424 (2000)
14. Tari, F., Ozok, A.A., Holden, S.H.: A comparison of perceived and real shoulder-surfing risks between alphanumeric and graphical passwords. In: Proceedings of the Second Symposium on Usable Privacy and Security, pp. 56–66 (2006)
15. Valentine, T.: An evaluation of the Passface personal authentication system. Technical Report, Goldsmiths College (1998)
16. Levin, D.T.: Race as a visual feature: using visual search and perceptual discrimination tasks to understand face categories and the cross-race recognition deficit. J. Exp. Psychol. Gen. **129**(4), 559 (2000)
17. Werner, S., Hauck, C., Masingale, M.: Password entry times for recognition-based graphical passwords. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting, vol. 60, no. 1, pp. 755–759. Sage CA, SAGE Publications, Los Angeles, CA (2016)
18. Bardram, J.E., Kjær, R.E., Pedersen, M.Ø.: Context-aware user authentication—supporting proximity-based login in pervasive computing. In: International Conference on Ubiquitous Computing, pp. 107–123. Springer, Berlin (2003)
19. Liu, X., Gao, H., Wang, L., Chang, X.: An enhanced drawing reproduction graphical password strategy. J. Comput. Sci. Technol. **26**(6), 988–999 (2011)
20. Gao, H., Liu, X., Wang, S., Liu, H., Dai, R.: Design and analysis of a graphical password scheme. In: 2009 Fourth International Conference on Innovative Computing, Information and Control (ICICIC), pp. 675–678. IEEE (2009)

21. Kim, D., Dunphy, P., Briggs, P., Hook, J., Nicholson, J.W., Nicholson, J., Olivier, P.: Multi-touch authentication on tabletops. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 1093–1102 (2010)
22. Zangooei, T., Mansoori, M., Welch, I.: A hybrid recognition and recall based approach in graphical passwords. In: Proceedings of the 24th Australian Computer-Human Interaction Conference, pp. 665–673 (2012)
23. Lin, D., Dunphy, P., Olivier, P., Yan, J.: Graphical passwords & qualitative spatial relations. In: Proceedings of the 3rd Symposium on Usable Privacy and Security, pp. 161–162 (2007)
24. Weiss, R., Luca, A.D.: PassShapes: utilizing stroke-based authentication to increase password memorability. In: Proceedings of the 5th Nordic conference on human-computer interaction: building bridges, pp. 383–392 (2008)
25. Man, S., Hong, D., Matthews. M.M.: A shoulder-surfing resistant graphical password scheme-WIW. In: Security and management, pp. 105–111 (2003)
26. Nicholson, J., Coventry, L., Briggs, P.: Age-related performance issues for PIN and face-based authentication systems. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 323–332 (2013)
27. Laugwitz, B., Held, T., Schrepp, M.: Construction and evaluation of a user experience questionnaire. In: Symposium of the Austrian HCI and Usability Engineering Group, pp. 63–76. Springer, Berlin (2008)
28. Beznosov, K., Kruchten, P.: Towards agile security assurance. In: Proceedings of the 2004 Workshop on New Security Paradigms, pp. 47–54 (2004)
29. News, D. I. T.: User authentication and levels of assurance, 2 Oct 2015. Retrieved from https://it.wisc.edu/about/user-authentication-and-levels-of-assurance/
30. Kaur, A., Mustafa, K.: Qualitative assessment of authentication measures. In: 2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom), pp. 694–698. IEEE (2016)

Detection of Optimum Fumigation in Grain Storage Using FUMON

Ravi Sankar, Tarun Kanti Ghosh, Hena Ray, Alokesh Ghosh, Sangit Saha, Devdulal Ghosh, and Nabarun Bhattacharyya

Abstract Grains are generally stored in warehouses after harvesting often for months. These stored grains are distributed to consumers through distribution channels as and when required. During such long storage, grains are affected by pests/insects and there can be considerable amount of storage losses. Pest infestation is highly affected by the temperature and relative humidity of the warehouse. In order to control such loss, a process called fumigation is performed to kill the pests by means of pesticides. The operation is hazardous and is carried out in buildings, soil and grain warehouses. Fumigation is also done during the process of export and import of goods to prevent the transfer of unusual organisms (Saeung et al., 2018 5th International conference on industrial engineering and applications (ICIEA), Singapore, 2018, pp 179–183 (Saeung, P., Santalunai, S., Thosdeekoraphat, T., Thongsopa, C.: Improved efficiency of insect pest control system by SSPA. In: 2018 5th International Conference on Industrial Engineering and Applications (ICIEA), Singapore, pp. 179–183 (2018). 10.1109/IEA.2018.8387092)). Phosphine (PH_3) gas is mostly used presently to carry out fumigation in warehouses. This is important to know the performance of the fumigation process. But phosphine being highly toxic in nature, the warehouse is kept sealed during the entire fumigation period. Hence, the efficacy of fumigation during the process cannot be known. This paper describes an indirect way of checking the effectiveness of fumigation by measuring the concentration of fumigant in real time. A device named FUMON (fumigation monitoring system) is designed and developed for the said purpose. FUMON device automatically sniffs fumigant from the grain stack under fumigation at certain predefined interval of time. A PC/laptop connected to the FUMON device continuously plots the concentration (C) against time (T) and calculates the CT product value at every instance of time. The CT product value indicates the effectiveness of ongoing fumigation and seeing the value, the warehouse manager may take decision to lengthen or shorten the fumigation. Validation of the developed FUMON device is also done at Indian Institute of Food Processing Technology (IIFPT), Thanjavur.

R. Sankar (✉) · T. K. Ghosh · H. Ray · A. Ghosh · S. Saha · D. Ghosh · N. Bhattacharyya
Centre for Development of Advanced Computing, Kolkata 700091, India
e-mail: ravi.sankar@cdac.in

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_90

Keywords Fumigation · FUMON · Concentration–time (CT) curve · Phosphine (PH_3) · Parts per million (ppm)

1 Introduction

Food grains stored in warehouses are subjected to loss and/or damage due to pest attack. The total grain value lost, in terms of quantity and quality, between harvest and consumption in developing countries is between 25 and 50%. Survey shows that storage loss in India was about 60,000 ton between 2011 and 2012 and 2016 and 2017. Killing of pests, insects, and their eggs is an important activity that needs to be carried out by warehouse management in order to control or reduce this loss. Fumigation is a process of killing pests with the help of toxic gases and chemical smokes. Methyl bromide and phosphine are known as the most commonly used fumigants for controlling insects due to its rapid action and comparatively low cost [2]. The dose of fumigant for certain grain type and pest class is generally given by concentration-time (CT) product, which means that a fixed concentration (C) to be maintained for a fixed time (T), as shown in Fig. 1.

This CT product is nothing but the area under the curve obtained by plotting concentration of phosphine versus time. While fumigating using AlP tablets, the concentration of phosphine does not remain constant, but increases with time from zero, attains a maximum, and then slowly goes down over time, as shown in Fig. 2.

This is expected that the AlP tablets would release phosphine as per specification and the phosphine concentration under the polythene sheet would be maintained at the desired level for the entire fumigation period. Hence, there is a need to continuously

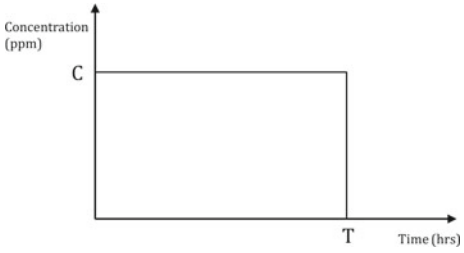


Fig. 1 Fumigation dose–concentration–time curve

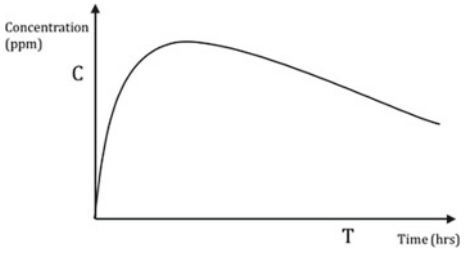


Fig. 2 Concentration–time curve in real situations

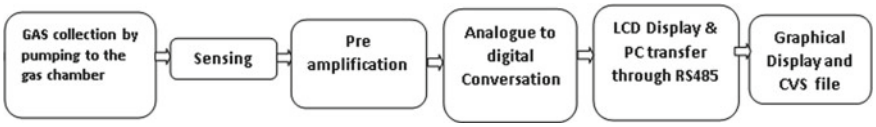


Fig. 3 Block diagram of FUMON

monitor the concentration of phosphine during the entire fumigation period in order to correlate that with the killing of pests. The real-time cumulative CT product value would help the warehouse manager to understand what is happening inside the stack under fumigation and would enable him to take corrective actions whenever required. Accordingly, the management would be able to take decision whether to lengthen or shorten the fumigation process or to start the process afresh abandoning the ongoing one with necessary corrective measures. Figure 3 describes the block diagram of the system.

2 Phosphine

The most commonly available fumigant is phosphine which is low cost. The real-time applications of phosphine are very simple and are accepted in the market. Phosphine scatters quickly inside the enclosed area and no separate fans are required [3]. The handling of stored grain and other robust commodities depends on the fumigation process with phosphine because of the international standard [4]. Phosphine (PH_3) is a colorless combustible gas in its actual form. Phosphine is dispersed very quickly in the open environment because the density of phosphine is almost similar to air [5]. The effectiveness of the fumigation process depends on the collective effect of the time and the concentration of phosphine [6].

3 Hardware Environment

The FUMON consists of three parts, namely the electronics circuit along with the sensor, the mechanical and electrical construction, and the software that includes the firmware functioning in the microcontroller and the PC software. The electronics circuit further consists of the sensor, the analog and digital circuit, and the microcontroller, the communication module. The electrical and mechanical construction consists of a micropump and an IP65 cabinet. The electronics portion of the system is responsible for command, sensing, pre-amplification, analog-to-digital conversation, display and communication. The system consists of a '7-PH3-2000′ electrochemical gas sensor which is capable of sensing 0 to 2000 ppm with 1 ppm resolution. Figure 4 shows the photographs of the developed FUMON device.

Fig. 4 Photographs of FUMON

The sensor is interfaced with a separate analog circuit. The analog circuit maintains the bias voltage of potentiostat between the reference sensing electrode and the current to voltage converter. The 0 V biasing voltage is maintained between the reference voltage and the sensing electrodes for PH_3 sensor. The current flow is being reduced in the sensing electrode by the oxidization of sensing electrodes. The current signal is also converted into voltage signal by using the transimpedence circuit. The digital circuit consists of a microcontroller which was inbuilt 10-bit ADC along with other peripheral. The system also consists of the alphanumeric LCD display that is being connected with the general-purpose input–output (GPIO) ports of the microcontroller. The LCD module is capable of displaying 16 × 2 numbers of characters or numeric value at a time. The MAX485 module (TTL to RS485) converts the serial data from the system to RS485 format and vice versa. The microcontroller on receiving the command from the PC starts the micropump and then samples the signal from the sensor to calculate the PPM concentration of PH_3 gas. The ADC of the microcontroller has been used for the signal sampling and digitization. The microcontroller further processes the data and displays in the LCD. The microcontroller simultaneously sends the data to the PC. A 12 V micropump is used for suction of the air from the stack that contains PH_3. A separate dust filter is used to prevent the dust particles. Customized Teflon made gas chamber is used for channeling the flow of air containing high concentration of PH_3. The chamber ensures the gas does not leak. The sensor is placed inside the gas chamber which is fitted to the PCB with screw. A brass made connecter is fitted with the gas chamber for silicon pipe connection. Figure 5 represents the schematic view of the system.

4 Software Environment

The software of FUMON (eFUMON) can detect the concentration of fumigant gas at any instant of time and calculates and presents the cumulative CT product value at that moment. Developed software sends command in regular interval of 30 min to the sensing system through RS485. The command tells the microcontroller to sniff for a particular time of 30 s/1 min/2 min. The testing was done by choosing the sniffing time

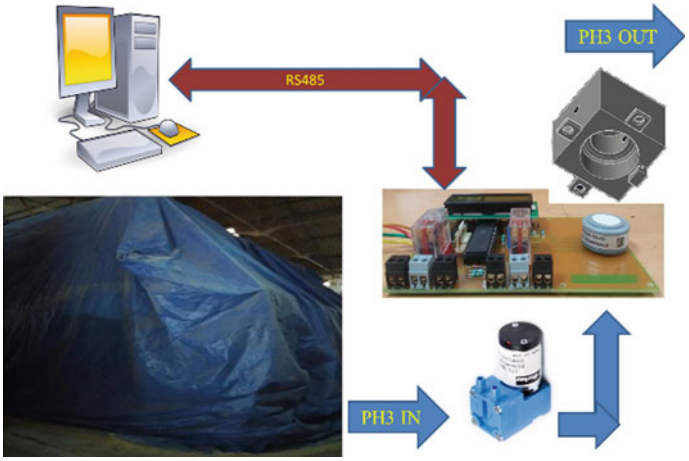
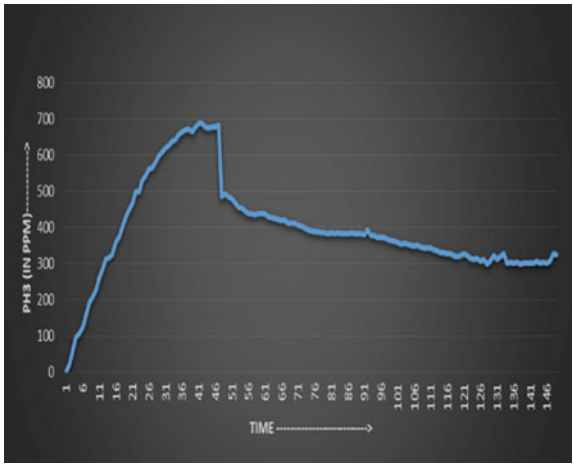


Fig. 5 Schematic view of working principle of FUMON

as 1 min and the pump starts at 0 s and runs up to 30 s and takes reading in 1 s interval for 60 s. The software shows the concentration (C) versus time (T) plot during entire fumigation period and simultaneously shows the maximum concentration achieved so far. The software development was done in LabVIEW environment. The multiple testing of FUMON device was done at Food Corporation of India (FCI) Kolkata and was deployed at FCI Raipur. Figures 6 and 7 show the different test results of FUMON device under various climatic conditions. The software flow diagram of FUMON is shown in Fig. 8.

Fig. 6 C-T curve obtained at FCI



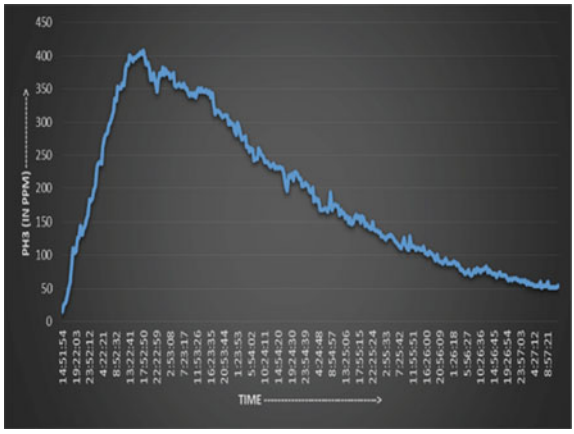


Fig. 7 C-T curve obtained at FCI, Kolkata

5 Validation of FUMON Device

The performance of FUMON has been validated by Indian Institute of Food Processing Technology (IIFPT), Thanjavur. A 14 L fumigation chamber was designed and developed for performing laboratory-scale fumigation using AlP tablets. Phosphine gas liberated within the chamber was pumped by an inbuilt pump into FUMON and it was given back through the outlet, to maintain the concentration and reduce the gas loss. The concentration of phosphine gas in parts per million (ppm) measured by FUMON was compared with properly calibrated UNIPHOS 250 PM phosphine monitor. Experiments were conducted with AlP tablets of 3.0, 1.0, 0.5, and 0.25 g with and without rice in chamber at ambient conditions of 30 ± 1 °C and 75 ± 5% RH. FUMON was analyzed for sensing range, accuracy, repeatability, and response time. The experimental setup is shown in Fig. 9.

FUMON was verified under different conditions for its performance characteristics and limitations. The detection range was found to be 0 to 3000 ppm and the values were repeatable under similar conditions. The difference in concentration of phosphine gas is negligible as shown in Figs. 10, 11 and 12. Response time for the whole system was observed as 0.5 min. While analyzing the plots of fumigation with rice, as shown in Fig. 13, the maximum concentration decreases to 1600 ppm due to the absorption of phosphine by rice grains. When the same grains were fumigated again, a greater concentration of phosphine gas was observed. This might be because of the saturation of molecular sites. A 100% mortality was obtained in all treatments for red flour beetle in rice.

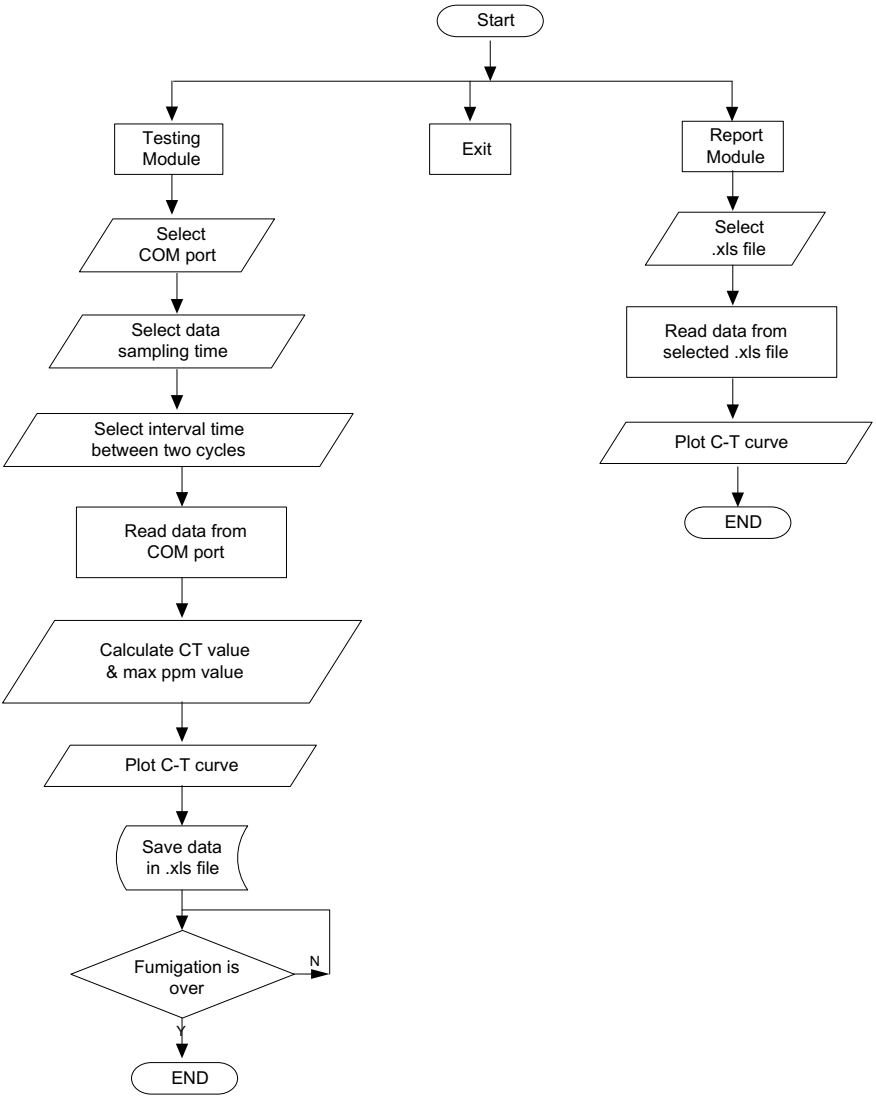


Fig. 8 Software flow diagram of FUMON

6 Conclusion

In this paper, we presented a system which is capable of measuring the concentration of phosphine gas with respect to time. The results of fumigation play a valuable role in many pest control operations. With the advancement of science and technology, various researches have been done on fumigants. After getting sufficient evidence on fumigants like phosphine, an electronic system has been developed to find out

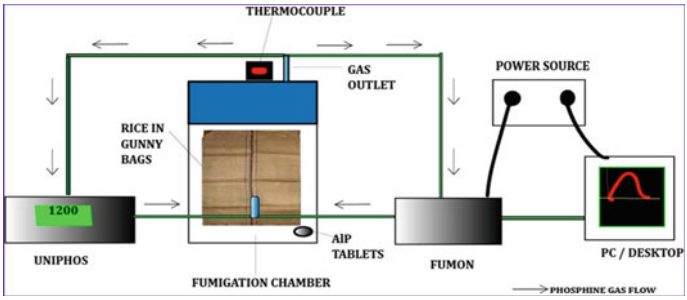


Fig. 9 Schematic of FUMON validation setup

Fig. 10 Trial #1

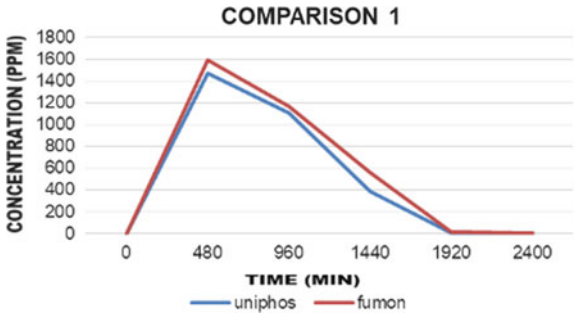


Fig. 11 Trial #2

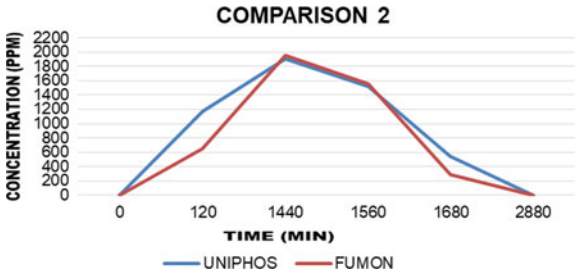
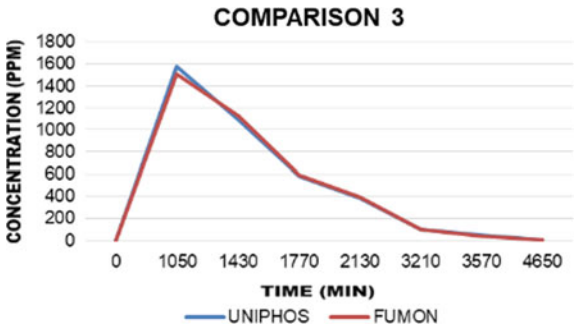


Fig. 12 Trial #3



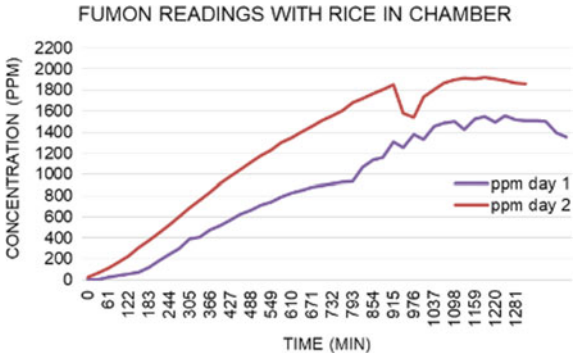


Fig. 13 Fumigation trial with rice

the concentration of the fumigant under various climatic conditions. C-T plot for different food grains under different operating conditions showed varied results. This kind of device is not readily available in market where a user can view the online C-T plot along with cumulative CT product value in real time over the entire fumigation process. Third-party validation of the developed FUMON device shows good accuracy. This can become a potential tool for fumigation. In future, we may add the environment monitoring sensors (e.g., temperature and relative humidity) to the device and may study the effect of these environmental parameters on phosphine fumigation. In addition, the warehouse managers may study a detailed study on phosphine fumigation for various grains under different climatic conditions.

References

1. Saeung, P., Santalunai, S., Thosdeekoraphat, T., Thongsopa, C.: Improved efficiency of insect pest control system by SSPA. In: 2018 5th International Conference on Industrial Engineering and Applications (ICIEA), Singapore, pp. 179–183 (2018). https://doi.org/10.1109/IEA.2018.8387092
2. Rajendran, S.: Status of fumigation in stored grains in India. J. Grain Storage Res. (2016). https://doi.org/10.5958/0974-8172.2016.00022.5
3. Wu, L., Dong, M.: Modeling and simulation of imported wood fumigation process based on TCPN and Flexsim. In: 2011 Fourth International Symposium on Computational Intelligence and Design, Hangzhou, pp. 250–253 (2011). https://doi.org/10.1109/ISCID.2011.70
4. Chadda, I.C.: Fumigation with phosphine—a perspective. J. Grain Storage Res. (2016). https://doi.org/10.5958/0974-8172.2016.00023.7
5. Kopacki, M., Pawlat, J., Terebun, P., Kwiatkowski, M., Starek, A., Kiczorowski, P.: Efficacy of non-thermal plasma fumigation to control fungi occurring on onion seeds. In: International Conference on Electromagnetic Devices and Processes in Environment Protection with Seminar Applications of Superconductors (ELMECO &AoS), Lublin, pp. 1–4 (2017). https://doi.org/10.1109/ELMECO.2017.8267746
6. Chadda, I.C., Alice, J., Sujeetha, R.P., Alagusundram, K.: Improved procedures for fumigating grain storages with phosphine in Indian warehouses. In : Proceedings of 10th International Conference on Controlled Atmosphere and Fumigation in Stored Products (2016)

Empirical Analysis on Parameters for Adoption of Cloud-Based e-learning in Indian Higher Education System: A User's Perspective

Nidhi Gupta, Neeraj Sharma, and Sanjay Sood

Abstract E-learning provides a highly reliable and effective teaching–learning environment. A cloud-based e-learning system enhances the quality and performance of education, especially higher education. Cloud computing as a whole has a great impact on society and everybody's day-to-day life. The institutions which are adopting cloud computing in their learning process have been successful in providing enhanced and global standard education. Many developing countries like Afghanistan, India, Iran, Nepal, etc., have started taking the benefits of these technologies for the effectual learning system. Cloud computing delivers benefits like easy accessibility of e-learning contents, conducive teaching–learning platform, elasticity, scalability, reduced IT skills, standardization, etc. This paper primarily investigates the end users' perspective in concern to the adoption of cloud-based e-learning platform by the colleges and universities of India, particularly Chandigarh and Punjab. The paper also discusses the benefits attained from the cloud-based e-learning system in higher education and explores the various challenges being faced by the different stakeholders in the adoption of this system.

Keywords Adoption factors · Cloud computing · Cloud-based e-learning system · Issues in cloud computing · User's perception

1 Introduction

Education is a principal requirement for the growth and survival of every human being. Education not only helps the individual to grow but also nurtures the nation. The two major concerns of the educational institutions today are to provide quality

N. Gupta (✉)
I. K. Gujral Punjab Technical University, Kapurthala, Punjab 144603, India

N. Sharma
Gian Jyoti Institute of Management and Technology, Mohali, Punjab 160055, India

S. Sood
C-DAC Mohali, Punjab 160071, India

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_91

education and state-of-the-art infrastructural facilities. The proliferation of ICT and other technological resources for learning, known as e-learning, has facilitated the institutes to achieve the goal of eminence learning. Both teachers and students are accepting the practice of using Web-based e-learning methods as a new tool for the teaching-learning process. Hiu Li et al. say that "'e' in e-learning not only represents electronic but also represents evolving, enhanced, extended, everywhere, everytime, and everybody" [1]. E-learning has the capability of providing a conducive learning environment to the learners and allows them to learn in different formats and at their ease. By acquiring the benefits of contemporary educational methods, e-learning thus delivers new methodologies of learning. However, most of the academic institutions are still using the asynchronous or Web-based method of e-learning because synchronous e-learning involves a lot of investments and maintenance of ICT resources. This limits the academician to take the comprehensive advantage of the latest learning technologies. The limitations of the e-learning can be overcome and said benefits can be increased manifolds with the amalgamation of cloud computing (CC) along with e-learning.

CC is the latest computing paradigm that offers a great set of customized, dynamic, and scalable hardware and software resources as a service to the users as per their requirements. S. B. Bele in his paper defines cloud computing as "Resource sharing is a pure plug and play model that dramatically simplifies infrastructure planning" [2]. CC assists in eliminating the limitations of Web-based e-learning system and provides an affordable and easy solution to educational institutions.

In this paper, we aim to parameterize a cloud-based e-learning system keeping in view the following aspects: deployment applications, impact on society, and challenges for adoption in Indian Higher Education Institutions (HEIs).

2 Theoretical Background

2.1 Cloud Computing

Cloud computing (CC) is one of the fastest growing areas in the field of data science and has emerged as the most sought technology because of its potential benefits. Cloud computing is an Internet-based computing model which ensures that the computing resources are available to the users as per their requirement. The users pay for the services according to their usage in a metered manner [3].

Gartner, one of the biggest IT research organizations, defines cloud computing as "a style of computing in which massively scalable IT-related capabilities are provided 'as a service' using Internet technologies to multiple external customers" [4].

The main aim of cloud computing is to run various services over the Internet in a dynamic and controlled manner. Cloud computing delivers services with a user's information, application software, and computation on an application programming

interface (API) over a network. The services include IT infrastructural integration, system and applications software provisions, data cloud, networking services, etc.

2.2 Cloud Computing and Education System

Over the years, the education system has changed and is not confined to the chalkboard classroom or course books. The learners are now connected to each other or to the teachers with the use of smart devices and enjoy the flexibility of learning.

The attributes like scalability and affordability have made cloud computing an impeccable solution for addressing the requirements and issues of the higher education system and have been widely accepted by academic institutions since 2018. The services of cloud computing have maximized the learning outcomes of an institution and allow teachers and students to acquire more benefits while using this latest technology.

In today's scenario, when everyone is trying their level best to fight the pandemic caused by COVID-19, CC has emerged as a boon for the education industry. It has completely transformed the practice of teaching–learning, from maintaining students' records to data storage, online education platforms to upgraded research analytics.

2.3 Impact of Cloud-Based E-learning System on the Education Sector and Society

HEIs have been swift to recognize the advantages of CC and have been enthusiastically adopting it for various reasons. A special thanks to different free and open online cloud platforms that have made the easy availability and accessibility of resources. Some of the obvious advantages of CC which have encouraged its adoption are as follows:

1. **Low-Cost Resources**: The need for technical resources can never be eliminated in today's technological world. Most of the resources which are expensive to buy are now affordable because of CC. Professional industries are already doing consumerization of the CC technologies and considering it as a cost-effective solution. The education sector should also follow a similar trend, which will both augment technology-enabled education and importantly, lessen the budget effect [5].
2. **Cultural Exposure**: The subscriptions on YouTube or the connections in LinkedIn are evidence of how people are interacting with each other and coalescing socially to make it a culturally diverse realm. From individual places to the global platform, an event can reach around the world with its exposure to cloud technologies.

3. **Virtual Classrooms**: Education today is not bounded to four walls of the classroom, however, it has been diversified. The teachers and the learners can now collaborate anytime, anywhere and provide a virtual classroom scenario. This also promotes self-paced and personalized learning.
4. **Eco-friendly**: The cloud technologies are considered to be more environmentally responsive and effective than the classic IT infrastructural system as fewer resources are required and shared among multiple users, thus saving the energy and reduces carbon emission [6]. Also, as most of the data is available online and can be accessed anytime, need to have the hardcopy of learning material has reduced, thus encourages afforestation.
5. **Backup and Recovery**: CC has simplified the process of backup and recovery. The data is backed up on the cloud data centers (mostly in a distributed manner) and not on physical storage devices, which have limited life [7]. In case of any disaster or data loss, the information can easily be recovered from cloud servers.
6. **On-Demand e-Content**: Apart from the routine teaching workload, the faculty members in HEIs are often laden by additional activities like development and upgradation of in-house learning resources, maintenance of ICT resources, etc. [6]. This hinders their personal knowledge growth and a very less percentage of academicians are actively involved in quality research and innovations. Thus, CC is an enthusiastically acceptable solution as it provides readily on-demand customized learning resources that supports their teaching needs.

2.4 Challenges in Implementing a Cloud-Based E-learning System

Though the adoption of CC in educational institutions has been increasing every year, its rate of acceptance especially in developing countries, particularly in India, is fairly low due to the following barriers and challenges.

1. **Dependence on network performance**: The backbone for successful implementation of the cloud-based e-learning platform is high-speed Internet connectivity. Though Prime Minister Shri. Narendra Modi Ji's initiative of *Digital India* has improved the state of ICT and Internet connectivity in India, its penetration rate in India is around 50% in 2020 [8]. It means half of the population still not have the access to the online platform or has low Internet connectivity, especially in rural areas.
2. **Security and Confidentiality**: Data security and privacy is the second biggest barrier in the adoption of CC for learning. Although several measures like multi-verification login, encryption techniques, data backup at distributed servers, restricted access, etc., are taken for the secure deployment and usage of data of cloud, there is always a risk of data preaching. Since the data lies externally at data centers and can be accessed by CSPs themselves [9], consumers are never be completely assured that their information will not be accessed by unauthorized people.

3. **Lack of Multilingual Learning Resources**: Most of the available e-content is in English language. Also, various tools which are used for the development of e-contents support a single first language. In countries like India, where we have cultural diversity, the lack of e-contents in different regional languages is a major concern.
4. **Resistance to the latest technologies**: Digital divide has always been an obstacle toward acceptance of any new technology. The transition from the traditional learning system to a new online learning system requires lots of great patience, understanding, and involvement [10]. Most of the time, teachers or students are resistant to accept new technologies and prefer to use the custom teaching–learning method.
5. **Compatibility**: Compatibility has constantly been an issue to be considered while adopting new technologies. The integration of new technologies or platforms with the existing one can be a cumbersome task for the CSPs. Most of the HEIs usually use their in-house software compatible with their structure. So customization of the cloud environment to collaborate with the prevailing system may be painstaking.
6. **Vender Lock-In**: Dependency on service providers is another issue while using CC. Sometimes, the services provided by CSPs do not suit user, then it becomes difficult to migrate to another CSP due to legal commitments or storage of huge data on CSPs servers [6]. This situation is known as "vender lock-in." Lack of standardized cloud implementation regulations makes situation even more problematic.

3 Literature Review

Salunkhe and Kelkar [11] carried out the analytical study to recognize the role of CC in education, its advantages, and risks involved. From the review of 08 research papers, the authors highlighted the importance and plausible usage of CC. A survey was conducted on the students of Management Institutions of Mumbai, India, to validate the outcomes of the reviewed study. From the survey, it was inferred that ***rate of adoption*** of CC by the management colleges is ***not adequate***. Those institutions which are using CC along with their traditional teaching methodology are also ***not satisfactorily using CC*** because of the ***risk issues*** like security, hacking, reliability, privacy, etc. [11].

Basak et al. [13] in their paper identified the main factors which affect the implementation of e-learning in African educational institutions. The study was a review of 40 research articles aiming at the said objective. After reviewing these articles, they proposed a framework of the factors that affect the adoption of e-learning as shown in Fig. 1. The framework was designed based on eight factors which were further categorized as: **"*close context, medium context, and larger context*"** [13].

Ashtari and Eydgahi [12] studied the students' perception to measure effectiveness of cloud applications in higher education. The study included the undergraduate

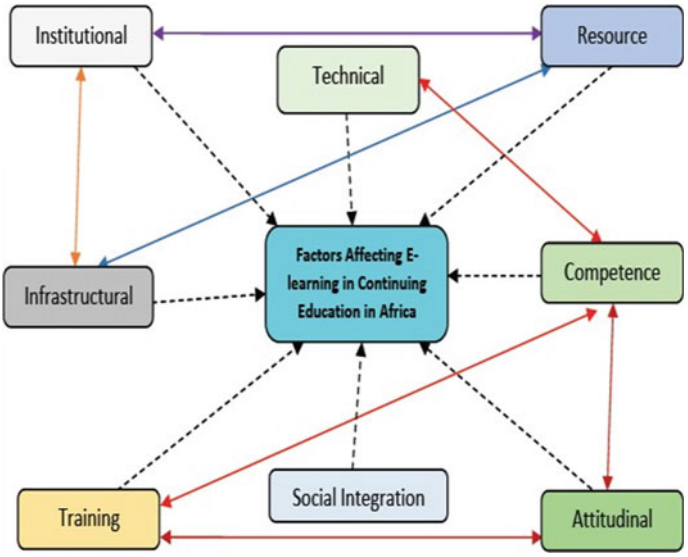


Fig. 1 A conceptual framework. *Source* [13]

students of the College of Technology at Eastern Michigan University. Based on initial study, a theoretical framework was proposed based with perceived utility of cloud applications as a major variable. From the results, authors believed that ***ease of use of applications*** is an important factor for usage of technology after its adoption. The authors also identified from the responses that ***self-efficacy*** and ***willingness to use the new technology*** are the driving factors for the adoption of new advancements [12].

Banu et al. (2018) in their study gave a SLR to explore the current status of implementation of CC in various HEIs. To achieve the said objective, a total of 20 research papers were included in SLR. From the review, they found that ***financial gain*** is one of the major reasons for CC adoption. However, authors also realized that though HEIs have shown interest to include CC in their present system, there is a lack of empirical studies in support of acceptance of CC by educational institutions [14].

Upadhyaya et al. [15] through their study emphasized on the quality of cloud services the users expect after embracing CC for the educational needs. They selected "Quality of Experience (QoE)" as the parameter for diagnostic. The survey data was obtained from the students of HEIs of India and USA. The questionnaire included questions on students' experience about various cloud-based software used for online admissions, assignments, projects, assessment, etc. It was observed during the study that ***Google Docs is the most preferred application***. The authors believed that due to the absence of any International Cloud Quality Standards, QoE indicators need to be ***defined and implemented***. Also, they found that there is a remarkable difference between the quality and service standards adopted by India and the USA [15].

4 Objective and Scope of the Study

The objective of this research study is manifold. The research aims to answer the following questions:

1. What are the perceptions of various stakeholders for the implementation of cloud-based e-learning solutions in colleges and universities?
2. What are the benefits attained through the cloud-based e-learning platform in colleges and universities?
3. What are the issues and challenges in implementing cloud-based e-learning platform in colleges and universities?

The scope of the study is restricted to the colleges and universities of Chandigarh and Punjab. In this study, responses from all the stakeholders like administrative staff, faculty members, IT managers, and students have been recorded.

5 Research Design

5.1 Data Collection and Sample Size

The research methodology comprises of the rigorous study and analysis of cloud-based e-learning systems in higher educational institutions and thus requires a study based on both primary as well as secondary data.

1. **Primary Data**: Since the scope of the study is being proposed to the colleges and universities situated in Chandigarh and Punjab, the primary data was collected from the stakeholders of various colleges and universities through a self-designed and pre-tested online questionnaire which was designed using Google Form.
2. **Secondary Data**: The secondary data for this research was collected from the survey of the literature, reports, internal records of the organization, and other published data in various journals, conference proceedings, etc.

The population of the study consists of 57 respondents from various colleges and universities of Chandigarh and Punjab. The respondents are the stakeholders or the end users who have a direct impact on the inclusion of CC in their respective academic institutions. The sample was selected using a random sampling technique.

5.2 Research Methodology

In this study, a research instrument called questionnaire was adopted which comprises of the 37 closed-ended questions on a five-point Likert scale ranging from 1 (strongly

disagree) to 5 (strongly agree) along with the demographic information of the respondent. The questionnaire was distributed to the sample with the help of e-mails, WhatsApp groups, and messaging services. Data was collected through Google Forms, and, thereafter, documented directly into an Excel file. Later on, SPSS software, *t*-test, and chi-square test were used for data analysis. The survey was conducted to take the viewpoint of sample of target population to be able to infer the perception of the entire population.

6 Data Interpretation

This study aims to investigate the willingness of the respondents to adopt the CC technologies. The study also elucidates several benefits and challenges of CC in HEIs.

Table 1 shows that the percentage of respondents from Chandigarh is 61% and from Punjab is 39%. It is evident from Table 2 that the number of respondents majorly includes faculty members and students, who are the eventual end users of the CC.

Cumulative Data Analysis of Research Question 1 Perceptions and knowledge of various stakeholders of colleges and universities for implementing the CC Services.

The *p*-value of test statistics in Table 3 is 0.31, which indicates that knowledge mean scores are not equal for respondents. The knowledge of the respondent, who are using CC Services, is more as compared to those who are not using it. It is also interpreted that demographic location of user does not affect their knowledge about CC services.

The *p*-value of the test in Table 4 is 0.002 which shows that the majority of respondents are in favor of the implementation of a cloud-based e-learning system.

Cumulative Data Analysis of Research Question 2 This analysis is done to understand the benefits attained by the respondents from cloud-based e-learning systems

Table 1 Demographic statistics of respondents

Location	No. of responses	%
Chandigarh	35	61
Punjab	22	39

Table 2 Distribution of profile of respondent

Current profile	No. of responses	%
Administrative	03	5
Faculty member	13	22
IT manager	1	1
Student	41	72

Table 3 Knowledge about cloud computing services

	Mean ± Standard error	Confidence interval	*t*-statistics (*p*-value)
Yes	8.60 ± 0.135	[−2.10, −0.89]	0.031
No	10.08 ± 0.263		

Table 4 Willingness to implement cloud-based e-learning system

	Mean ± Standard error	Confidence interval	*t*-statistics (*p*-value)
Yes	65.41 ± 0.705	[1.225, 5.790]	0.002
No	61.91 ± 0.920		

(Table 5). Table 5 also highlights those factors which encourage the respondents to adopt CC.

Survey scale: *1*—Strongly disagree, *2*—Disagree, *3*—Neither agree or disagree, *4*—Agree, *5*—Strongly agree

Cumulative Data Analysis of Research Question 3 This analysis is done to understand the challenges or issues that act as barriers for the adoption of CC and its services in their teaching–learning process by the respondents (Table 6).

7 Research Analysis and Discussion

The prime objective was to recognize the parameters affecting the adoption of CC by HEIs in India, particularly Chandigarh and Punjab. All vital determinants affecting the acceptance of the CC system have been considered, which have been categorized into five broad categories (based on the results obtained), viz. as follows:

1. ***Technological Determinants (TD)***: TD (Q1 to Q8) focus on those parameters which provide technological innovation advantages to the user in terms of compatibility with an existing system, better understanding, scalability, flexibility, etc. Among all of these factors, **self-paced learning using different formats** is the most preferred factor for adoption, followed by **compatibility and collaboration**.
2. ***Personal Determinants (PD)***: PD (Q9 to Q13) refer to knowledge and usage of CC in teaching–learning process, out of which **knowledge of CC and user's own interest** to attain more knowledge have been derived as the prominent parameters.
3. ***Organizational Determinants (OD)***: OD (Q14 to Q18) include various ICT facilities, learning resources, e-learning tools, and other support provided by the college/university that facilitate the adoption of cloud-based e-learning system. **Support and encouragement from the institution** to embrace CC has

Table 5 Percentage of responses for questions encouraging adoption

Frequency of responses (in percentage)						
	Statement	1	2	3	4	5
Q1	Using cloud-based e-learning for my course encourages me to continue learning on the Internet by myself	2	4	7	51	36
Q2	It is simple to access and share cloud-based e-learning resources	2	5	7	50	36
Q3	Cloud-based e-learning encourages me to actively participate in the teaching–learning activities	2	11	17	50	20
Q4	It is easy to take the backup and recovery of data using CC	2	6	17	47	26
Q5	Use of cloud-based e-learning is well-suited with my current work	5	6	14	51	24
Q6	I can understand my course content better using ppts, videos, etc.	2	2	10	54	32
Q7	It offers effective administrative management opportunities	5	8	29	42	16
Q8	It is easy to increase or decrease the resources using CC	2	10	25	47	16
Q9	CC helps to enrich my knowledge and learning ethics	2	6	22	46	24
Q10	I am aware of the usage of CC in education	2	8	19	56	15
Q11	It is easy to get the updates about my course using a cloud platform	2	6	14	56	22
Q12	I have a smart device along with Internet connectivity	5	5	6	43	41
Q13	I use computer and other smart devices for online learning and studying	5	7	16	43	29
Q14	My college/university knows the advantages of using a C	3	4	10	48	35
Q15	My institute is ready to adopt a cloud-based e-learning system	3	4	9	49	35
Q16	The hardware resources required for the adoption of cloud-based e-learning are provided by my institution	2	7	27	43	21
Q17	It reduces the expenses of purchasing costly ICT devices	6	15	20	39	20
Q18	It minimizes the need of procuring and renewing software licensing	4	9	35	37	15

(continued)

Table 5 (continued)

Frequency of responses (in percentage)						
	Statement	1	2	3	4	5
Q19	The laws and regulations laid by the administration allow the use of CC in educational institutions	4	7	21	49	19
Q20	The government has issued various instructions for the safe disposition of data on the cloud	2	6	25	48	19
Q21	Peer institutions are embracing cloud in their respective institution	3	9	20	48	20
Q22	E-learning or CC is the need of an hour to raise the learning principles	3	10	12	44	31
Q23	I am interested to incorporate a cloud-based e-learning system with my present system	4	5	20	48	23
Q24	Learners today are ready to use advanced learning approaches	6	4	14	48	28

come out to be the most crucial acceptance parameter out of all organizational determinants.

4. ***Environmental Determinants (ED)***: ED (Q19 to Q24) are associated with parameters like support from the government by implementing laws and policies, peer pressure, and consumer's behavior toward technological advancements. Among all of the environmental determinants, **government support and the user's behavioral aspiration** are two most preferable factors for adoption.
5. ***Barrier Determinants (BD)***: BD (Q25 to Q37) comprise of those parameters which can act as a barrier or challenge for the adoption of a cloud-based e-learning system. These include determinants like complexity in understanding or using the innovation, concern about data reliability, digital divide, or CSP switching restrictions. Out of 13 questions related to challenges in CC in the questionnaire, **Internet connectivity issues, data security, and CSP lock-in** are the most challenging parameters in implementing a cloud-based e-learning system. Table 7 shows the significance of the above five determinants in adoption of CC.

Data Reliability Analysis To measure the reliability of the determinants, Cronbach's alpha (α) analysis was done as shown in Table 8.

According to Viswanath et al., since all the determinants return α-coefficients >0.70 [16], a high degree of reliability is determined for all the determinants.

Table 6 Percentage of responses for questions which act as barriers for adoption

Frequency of responses (in percentage)						
	Statement	1	2	3	4	5
Q25	This system incurs huge structural and Internet costs	5	7	19	44	25
Q26	Online course contents for my course are not available	4	16	31	35	14
Q27	I have a lack of knowledge about cloud computing and e-learning perspective and custom	13	26	30	25	6
Q28	There is a threat of data malfunctioning while using CC	7	13	32	34	14
Q29	I am biased to use CC because of security issues	17	39	25	15	4
Q30	I am afraid of the unlawful use of data by external sources	7	24	22	35	12
Q31	E-learning resources available are mostly inconsistent	10	27	35	22	6
Q32	CC restraints us to adopt the system as per the conditions of cloud service providers	6	13	33	37	11
Q33	It is very difficult to shift CSP due to contract constraints	5	17	36	35	7
Q34	There are no worldwide standards to implement CC	7	12	44	30	7
Q35	I rarely use my smart device for learning and training	28	34	19	16	3
Q36	I am more comfortable in using traditional board teaching	10	16	30	27	17
Q37	It involves lots of self-study and time	15	28	32	19	6

8 Study Findings and Recommendations

The following **findings are deduced** from the research conducted to understand the opinion of the users about adoption of cloud-based e-learning systems in HEIs in India:

1. The study showed that approximately 84% of respondents are using smart gadgets like smartphones, laptops, tablet, etc., equipped with an Internet connection.
2. Anytime–anywhere–anyform assistance is one of the critical reasons for the adoption of CC services by 86% of the respondents.
3. In today's skill-oriented world, the results showed that the students (77% of the respondents) are interested to induct new technologies for improvised learning.

Table 7 Significance of different determinants

Determinants	Mean ± standard error	Confidence interval	*p*-value (*t*-statistics)
Technological determinants	31.09 ± 0.365	[0.516, 2.87]	0.005[a]
	29.40 ± 0.492		
Personal determinants	19.43 ± 0.241	[0.640, 2.20]	0.00[a]
	18.01 ± 0.328		
Organizational determinants	17.95 ± 0.250	[−0.0821, 1.479]	0.079
	17.25 ± 0.312		
Environmental determinants	22.77 ± 4.198	[0.243, 0.226]	0.013[a]
	21.64 ± 4.744		
Barrier determinants	39.77 ± 0.566	[−4.277, −0.881]	0.004[a]
	42.32 ± 0.676		

[a]Significant—It means that these determinants are effected depending on whether a respondent is using CC services or not. If a respondent is using CC, then he/she has a better understanding of these determinants

Table 8 Cronbach's alpha score

Determinant	Cronbach's alpha
Technological determinants	0.897
Personal determinants	0.84
Organizational determinants	0.81
Environmental determinants	0.87
Barrier determinants	0.86

4. The findings of the study revealed that the various stakeholders of an educational institution are keen to take the advantage of CC services; however, they have insufficient knowledge and expertise about its usage and deployment.
5. From the study, it was found that most of the HEIs support the cloud platform as it is compatible with their existing system.
6. The study showed that Google Drive (44%) is the most preferred and used cloud service by respondents followed by Microsoft Cloud (22%). This evidences that majority of respondents are using CC mainly for backup of data and information.
7. The empirical research also concludes that people, in general, are not ready to shift to a new cloud paradigm or to incorporate it with their routine tasks.

Based on the above findings, the following **recommendations are inferred** which can be adapted to improve the status of usage of CC in HEIs in India:

1. The colleges and universities must ensure good Internet connectivity for the successful implementation of a cloud-based e-learning system.
2. Before transiting to CC, the staff and students of colleges and universities must be given demonstrations about its usage in their teaching–learning process.

3. Standardized policies and regulations must be framed to free the CC consumer from the fear of vender lock-in.
4. The government must frame strict laws for secure and reliable deployment of data.
5. The public must be aware of the potential of CC and the benefits they can achieve after adopting it.

9 Limitation of the Study

Although the research has been successful in achieving the desired objectives, there are certain limitations of the study. The research was conducted on a small sample size of various stakeholders in HEIs. Also, there are inadequate empirical evidences about the adoption CC by HEIs, thus unlocks the scope of further research.

10 Conclusion

A report from UGC shows that there are more than 900 HEIs and 41,500 affiliated colleges in India, with around 35 million students enrolled, which is approximately 110% growth in the education sector in the last 10 years [17]. The need for integrating the latest technologies in the education section for achieving the international learning standards has increased. This is required to sustain and stand in today's fast-moving life. Learners not only demand quality education but also want complete course understanding and satisfaction.

Apart from an extensive literature review, an online questionnaire with 57 respondents was conducted to understand the perspective of the users toward CC adoption. In the study, it was found that the respondents understand the benefits of CC services and are willing to incorporate it. The study also showed that this system has both benefits and challenges for adoption. However, the advantages achieved by the adoption of CC overshadow its barriers and can be successfully implemented in HEIs to provide a sustainable blended learning environment. To completely realize the benefits and to reduce the issues in future, efforts must be done to widen and expand the adoption of CC services across the educational arena for its longer and enhanced sustainability.

Acknowledgements This work is a part of research work done as a part of a Ph.D. degree under I. K. Gujral Punjab Technical University, Kapurthala, Punjab. The authors would like to thank all the respondents who took the time to participate in this research.

References

1. Li, H., Masters, J.: E-learning and knowledge management in the early years: where are we and where should we go. Knowl. Manage. E-Learn.: Int. J. **1**(4), 245–250 (2009)
2. Bele, S.B.: An empirical study on 'cloud computing'. Int. J. Comput. Sci. Mob. Comput. **2**, 33–41 (2018)
3. Backialakshmi, B., Sumalatha, V.: A novel approach for cloud-based e-learning system. Int. J. Sci. Technol. Res. **9**(4), 1587–1591 (2020)
4. Prentice, B.: Cloud computing: what's it all about? In: Gartner Sservice Ooriented and Aapplication Ddevelopment and Integration Summit, Sydney, Australia, 28–29 July 2010 (2010)
5. Rao, K.S., Challa, R.K.: Adoption of cloud computing in education and learning. Int J Adv. Res. Comput. Commun. Eng. **2**(10), 160–4163 (2013)
6. Karim, F., Rampersad, G.: Cloud computing in education in developing countries. Comput. Inf. Sci. **10**(2), 87–96 (2017)
7. Omankwu, O., Ezeme, C., Ugwa, C.: Impacts of cloud computing in the society. Int. J. Eng. Sci. **5**(6), 86–97 (2016)
8. Online at Statista: https://www.statista.com/topics/2157/internet-usage-in-india/
9. Odeh, M., Garcia-Perez, A., Warwick, K.: Cloud computing adoption at higher education institutions in developing countries: a qualitative investigation of main enablers and barriers. Int. J. Inf. Educ. Technol. **7**(12), 921–927 (2017)
10. Kihara, T., Gichoya, D.: Use of cloud computing platform for e-learning in institutions of higher learning in Kenya. In: IEEE IST Conference Proceedings, Africa, pp. 1–6, May 2014. https://doi.org/10.1109/ISTAFRICA.2014.6880638
11. Salunkhe, U., Kelkar, S.: A study on the scope of cloud computing in management education. AIMA J. Manage. Res. **10**(2/4), 1–9 (2016)
12. Ashtari, S., Eydgahi, A.: Student perceptions of cloud applications effectiveness in higher education. J. Comput. Sci. **23**, 173–180 (2017)
13. Basak, S.K., Wotto, M., Bélanger, P.: Factors affecting to e-learning in continuing education in Africa: a review of literature. Int. J. Eng. Sci. Manage. Res. **4**(1), 86–97 (2017)
14. Ali, M.B., Wood, T., Mohamad, M.: Benefits and challenges of cloud computing adoption and usage in higher education: a systematic literature review. Int. J. Enterp. Inf. Syst. **14**(4), 64–77 (2018)
15. Upadhyaya, J., Ahuja, N.J., Sharma, K.D.: Evaluating user expectations and quality of service: a novel approach to understanding cloud services. Int. J. Recent Technol. Eng. **7**(5), 381–385 (2019)
16. Venkatesh, V., Morris, M.G., Davis, G.B., Davis, F.D.: User acceptance of information technology: toward a unified view. MIS Q. **25**(5), 425–478 (2003)
17. Online at UGC Aannual Rreport 2018–2019. https://www.ugc.ac.in/page/Annual-Report.aspx

Modeling Control Measure Score of COVID-19 Outbreak Using Fuzzy c-Means-Based Adaptive Neuro-Fuzzy Inference System

Abraham Varghese, Shajidmon Kolamban, Vinu Sherimon, Eduardo M. Lacap Jr, Saad Salman Ahmed, Jagath Prasad Sreedhar, Syed Rafeek Ahmed, Hasina Al Harthy, and Huda Salim Al Shuaily

Abstract The current outbreak of COVID-19 continues to threaten across the globe even after 9 months of its starting. The preventive measures and self-protecting measures are the suggested and viable solution to mitigate the spread of pandemic till new vaccine is discovered. The preventive measures are framed based on the detailed analysis of the current data of the ongoing pandemic. The correctness of the policies relies on the quality of the data and its analysis with a meaningful interpretation. In this paper, the present status of the epidemic across the world is analyzed using

A. Varghese (✉) · S. Kolamban · S. S. Ahmed
Faculty of Mathematics, University of Technology and Applied Sciences, Muscat, Sultanate of Oman
e-mail: abraham.v@hct.edu.om

S. Kolamban
e-mail: shajidmon.k@hct.edu.om

S. S. Ahmed
e-mail: saadretaj@hct.edu.om

V. Sherimon · H. S. Al Shuaily
Faculty of IT, University of Technology and Applied Sciences, Muscat, Sultanate of Oman
e-mail: vinusheri@hct.edu.om

H. S. Al Shuaily
e-mail: huda.alshuaily@hct.edu.om

E. M. Lacap Jr · J. P. Sreedhar
Faculty of Statistics, University of Technology and Applied Sciences, Muscat, Sultanate of Oman
e-mail: eduardo.jr@hct.edu.om

J. P. Sreedhar
e-mail: jagath.sreedhar@hct.edu.om

S. R. Ahmed
Faculty of Engineering, University of Technology and Applied Sciences, Muscat, Sultanate of Oman
e-mail: rafeek@hct.edu.om

H. Al Harthy
Department of Research, Royal Hospital, Muscat, Sultanate of Oman

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_92

a control measure score which is computed using adaptive neuro-fuzzy inference system (ANFIS) where partition of data is performed using fuzzy c-means clustering. We have used transmission rate and recovery rate independently obtained from the discrete version of the Susceptible-Infected-Recovered (SIR) model to predict the score. Since the transmission parameters depends on incubation period, recovery and death, the parameter has been calculated based on active cases where its mean value changes significantly. It is observed that the recovery rate is more than the transmission rate of all Gulf countries at present which shows that the outbreak reached its highest value. The current control measure score emphasizes on the normal working atmosphere keeping strict social distancing measures.

Keywords ANFIS · Fuzzy c means · Control measure score · COVID-19 · Adaptable neuro-fuzzy inference system · SARS-CoV-2 · FCM

1 Introduction

The novel coronavirus originated in December 2019 in a seafood wholesale market in Wuhan City, Hubei Province, China, and it was named as COVID-19 by WHO. Since the first patient was hospitalized in December 12, 2019, it has been reported a total of 41,042,579 cases, of which 30,628,417 have recovered and 1,129,591 deaths as on October 21, 2020 [1]. Due to its high contagion, the virus is still spreading fast, and the amounts of patients are still increasing. According to the International Health Regulation, WHO declared this current coronavirus (COVID-19) pandemic as a international concern with respect to Public Health Emergency. As this pandemic is becoming critical in many countries, it is the need of the hour to gain a high-level interpretation of the newly recognized virus which may successfully lead to the discovery of drugs, methods and treatments to contain the spread of this disease. It is very crucial to investigate about the epidemiological characteristics of the virus so that such accurate data may help the medical researchers to find a solution to contain the virus. As many countries are in the start of community transmission of this virus, it is highly crucial and critical to identify the causes and solutions.

Earlier studies prove that interdisciplinary research is very essential in any complex situations. Computers and computer technologies are used in medical field for several decades. There are many potential technologies like neural networks, fuzzy logic, deep learning, etc. that can be used in medical science. Researchers all over the world are trying their best to model the pandemic using various methodologies such as mathematical as well as machine learning model to recognize the nature of the disease which would help the authorities to take right decision at right time. Since all models are only approximations, but some models are useful, our aim is to come up with a model which is suitable for the available data. As the correctness of the model depends on the quality of the data, it is essential to have a quality data to predict the accurate nature of the disease. But due to the uncertainties in the number of tests conducted per day, the asymptomatic behavior of the disease and so many

other factors, there may be variations in the number of data and our model needs to accommodate such factors.

A neural network is used to gather information from the highly interconnected processing data where input–output data is properly arranged. But they [2] have limitations with their characterization of inferred knowledge. Fuzzy logic systems [3] are a good choice to generate models to accommodate uncertainties in the data, but it is subjective and heuristic. Also, in fuzzy systems, the rules and I/O are set by trial-and-error method. So, the whole process becomes a prolonged task. Therefore, a combination of both neural network and fuzzy logic may be a right choice to model the data in order to attain maximum information from the data.

The shortcomings of neural networks and fuzzy logic systems can be overthrown by blending fuzzy logic system with neural network technology, called adaptive neuro-fuzzy inference system (ANFIS) [4]. It is possible to train all the parameters of ANFIS similar to a neural network inside the structure of a fuzzy logic system. Sugeno-type fuzzy inference systems (FIS) use a hybrid learning algorithm to classify the membership function parameters of single output [4]. To model a specified set of input/output data, the parameters of the membership function are trained using backpropagation gradient descent methods and least square methods. ANFIS has attracted the attention of ability to produce intelligent methods. The benefit of the ANFIS method is that it combines the positive features and contributes a sounder program to improve systems performance for several design purposes [4]. In data clustering method fuzzy c-means (FCM), each data point is assigned to a particular cluster depending on the membership grade. This algorithm was introduced by Dunn in 1973, and later, it was modified by Bezdek in 1981 [4]. The objective of this research paper is to present an adaptive neuro-fuzzy inference model to compute the Control Measure Score (CMS), which determines the present situation of any country in terms of COVID-19 outbreak.

The rest of the paper is organized as follows: Sect. 2 includes SIR model. ANFIS model is explained in Sect. 3. Discussion and Results are illustrated in Sect. 4, and Conclusion is given in Sect. 5 followed by Acknowledgements and References.

2 Time-Dependent SIR Model

In Susceptible-Infected-Recovered (SIR) model, the entire population is separated into three classes. The susceptible (S) class is a fraction of population not infected now but having a risk of getting infected. The reported infected Individuals belong to the class Infected (I). The fraction of population which is either recovered or deceased included in the class Removed (R). At any moment, sum of these three classes constitutes the total population ($N = S(t) + I(t) + R(t)$). Also, it is assumed that the infected person develops immunity and not gets infected again due to suggested preventive measures. The rate of transmission between the compartments can be modeled using the differential Eqs. (1)–(3) at time t which is given by

$$S' = -\beta I(t)\, S(t) \tag{1}$$

$$I' = \beta I(t) S(t) - \gamma I(t) \tag{2}$$

$$R' = \gamma I(t) \tag{3}$$

$\beta(t)$ is the rate at which new persons are getting infected. The parameter $\gamma(t)$ is removal rate which means the rate at which infected individuals are removed in the form of recovery or death and are no longer infectious. The discrete version of SIR for implementation of COVID-19 pandemic is expressed as given in (4).

$$I(t+h) = I(t) + (\beta - \gamma) I(t) h \tag{4}$$

where h is the interval at the parameters are calculated. The infected individuals with unit time depend upon the difference of $\beta(t)$ and $\gamma(t)$, and these parameters $\beta(t)$ and $\gamma(t)$ can be computed from the formula given in (5).

$$\beta(t) = (I(t+h) - I(t))/I(t)h \quad \text{and} \quad \gamma(t) = (R(t+h) - R(t))/I(t)h \tag{5}$$

But the realistic value of transmission parameter $\beta(t)$ depends on delay distributions for infection (incubation period of SARS-CoV-2 is 2–14 days) recovery and death. In order to accommodate these factors, we calculate the parameters $\beta(t)$ and $\gamma(t)$ using the formula during the period where there is a significant change in the active case. This can be calculated based on the period where mean of active cases changes most abruptly. Figure 1 depicts the 10 periods where the mean of active cases of Oman changes abruptly from beginning till October 8, 2020.

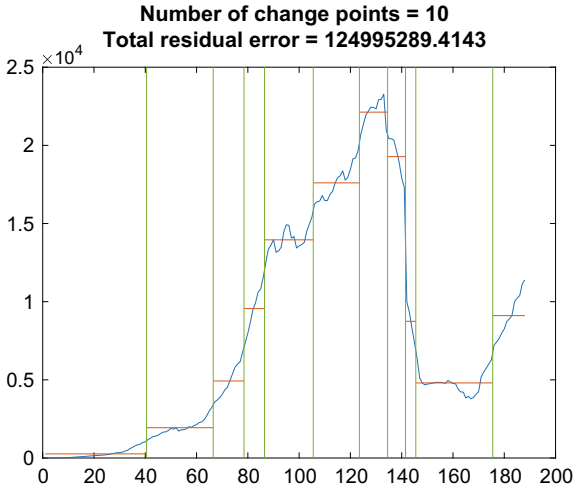


Fig. 1 Significant changes of mean of active cases in Oman

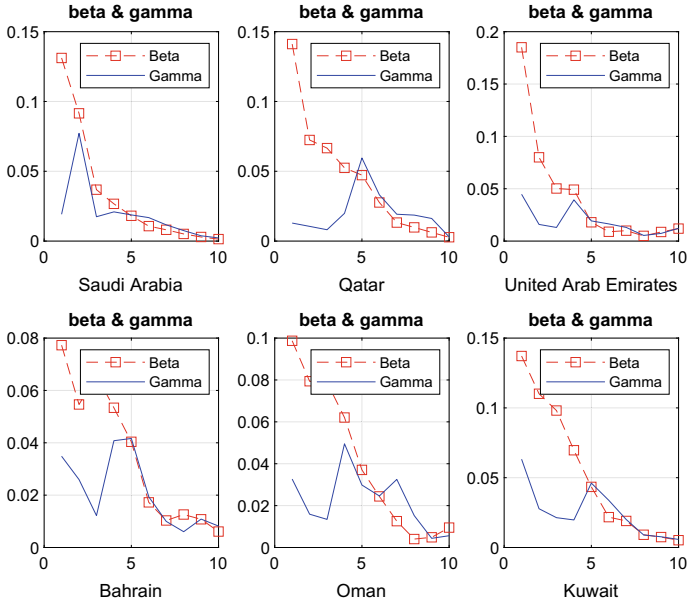


Fig. 2 Transmission and recovery rate of GCC countries

Figure 2 shows the beta and gamma of Gulf countries based on the period where mean of the active cases changes significantly. It is observed from the graph that the peak of the pandemic has been reached for all GCC countries because after initial days of higher rate, the transmission rate became less than the recovery rate.

We have taken the transmission rate and the recovery rate of 188 countries to model the control measure score for COVID-19 pandemic. In the absence of proper vaccination, the strategies to be adopted to control the epidemic depend on the growth rate of active cases. Therefore, the strictness level of the control measures (Strict, Moderate and Relaxed) mainly depends on two important parameters $\beta(t)$ and $\gamma(t)$. The impact of both parameters acts opposite to each other on the strictness level of control measures, i.e., when $\beta(t)$ increases the strictness level also increases, and while recovery rate $\gamma(t)$ increases the strictness level decreases. Therefore, optimizing these two parameters is inevitable to arrive at a meaningful strictness level score. Since it is difficult to arrive at a sharp boundary for these parameters while combining them, we use fuzzy logic which takes continuous value to model the parameters.

3 Adaptive Neuro-Fuzzy Inference System (ANFIS)

The first-order Sugeno fuzzy model with two inputs and one output having general rule: If x is A_i *and* y is B_i, then $f_i = p_i x + q_i y + r_i$. The final output is $f = \frac{w_1 f_1 + w_2 f_2}{w_1 + w_2}$, where w_1 and w_2 are the corresponding weights are used in the ANFIS architecture to model the input–output variable [4]. The ANFIS integrates the advantage of fuzzy logic and neural network as a single unit in a single system. The inherent capability of ANN to learn from experience and fuzzy logic ability to knowledge representation in terms of linguistic form helps to make ANFIS more realistic. The basic ANFIS structure comprises five distinct layers—fuzzification layer, rule base layer, membership functions (MFs) layer; defuzzification and summation layers, respectively. The whole ANFIS model design depends on the two sections: construction and training. In the first part, the number and type of membership functions are defined, and rules are generated accordingly. The input–output variable clustering is done through various methods such as subtractive clustering, grid partitioning and fuzzy c-means (FCM) [5]. FCM is used to systematically create the fuzzy MFs and fuzzy rule base for ANFIS as given in Fig. 3.

The input of the model is beta and gamma parameters. The output is modeled based on the average value of these two parameters after grouping the data into three clusters. The output is assigned based on the information in Table 1, and ANFIS structure with FCM is used to model it. The training data pairs are generated from the input–output data to form ANFIS model. The membership functions are changed

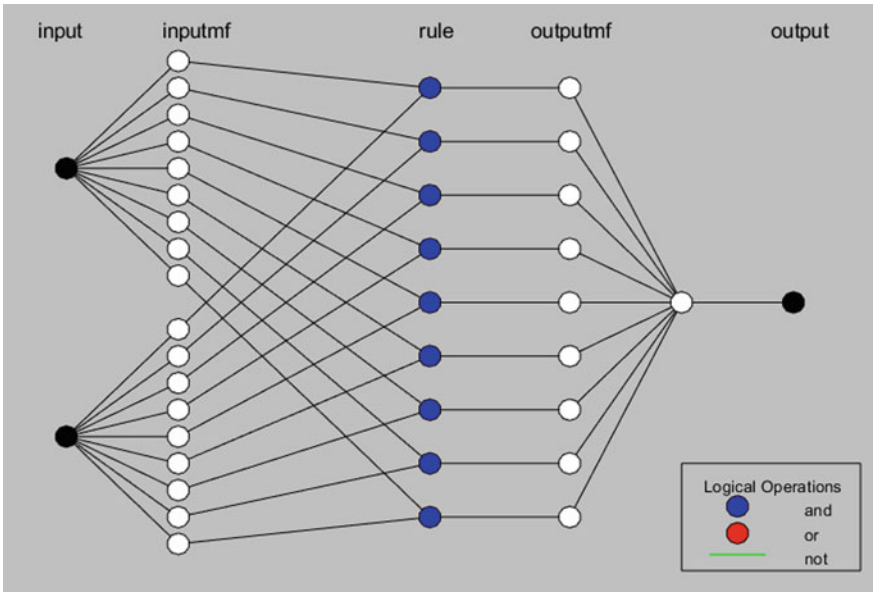


Fig. 3 ANFIS structure

Table 1 Output modeling

$\beta(t)/\gamma(t)$	Low (0–0.01)	Medium (0.01–0.13)	High (0.13–0.47)
Low (0–0.04)	Controlled (0.5)	Relaxed (0.3)	Relaxed (0.1)
Medium (0.04–0.29)	Strict (0.8)	Controlled (0.6)	Relaxed (0.4)
High (0.29–2.5)	Strict (1.1)	Strict (0.9)	Controlled (0.7)

Table 2 ANFIS model parameters with values

Parameters	Value
Number of nodes used	65
Number of nonlinear parameters involved	40
Number of linear parameters involved	30
Total number of parameters	70
Number of testing data pairs	422
Number of training data pairs	986
Number of fuzzy rules	10

to get a minimum error, and this is achieved through backpropagation learning. Table 2 depicts the ANFIS information.

4 Discussion and Results

The data for this study was collected from WHO dashboard [6]. The error of the training data (70%), test data (30%) and all data (100%) is given in Figs. 4, 5 and 6. The mean square error (MSE), root mean square error (RMSE) and histogram of error distribution are computed for training data, testing data as well as the whole data. It is observed that the MSE and RMSE values are more or less same for both training and testing data. The mean of the error is very close to zero which explains the robustness of the model.

The correlation coefficient value (R) is also computed for training as well as testing data, and it is observed that the value is 0.92 as shown in Fig. 7.

The control measure score is computed for GCC countries from the beginning of the confirmed cases till October 8, 2020, and the result is shown in Fig. 8. It is observed that the control measure score is high in the initial days and is going down. The current control measure score in Oman as on October 8, 2020, is 0.476, and it denotes the status of the country as control state now.

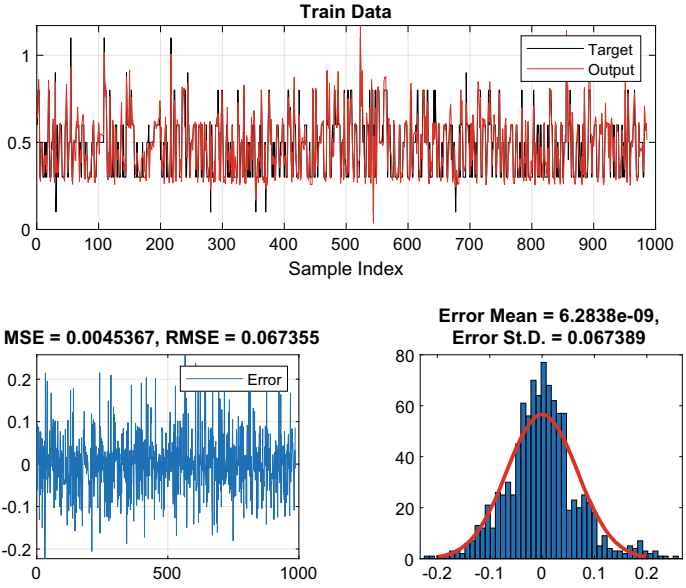


Fig. 4 Error information of training data

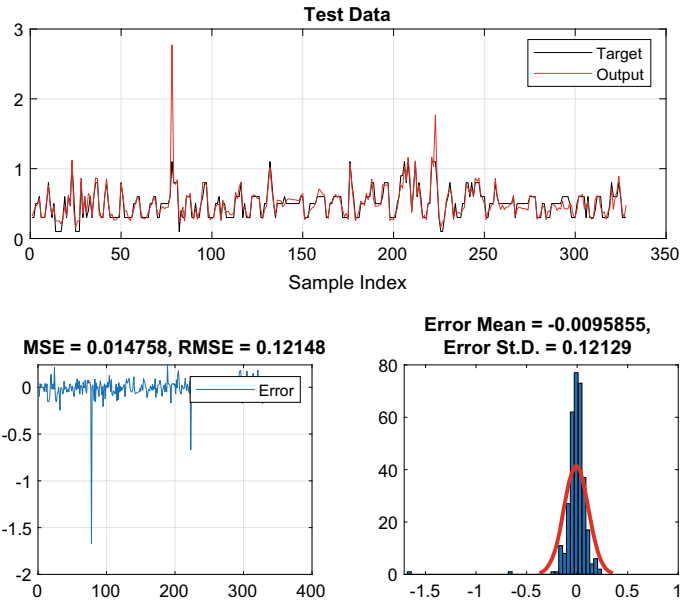


Fig. 5 Error information of testing data

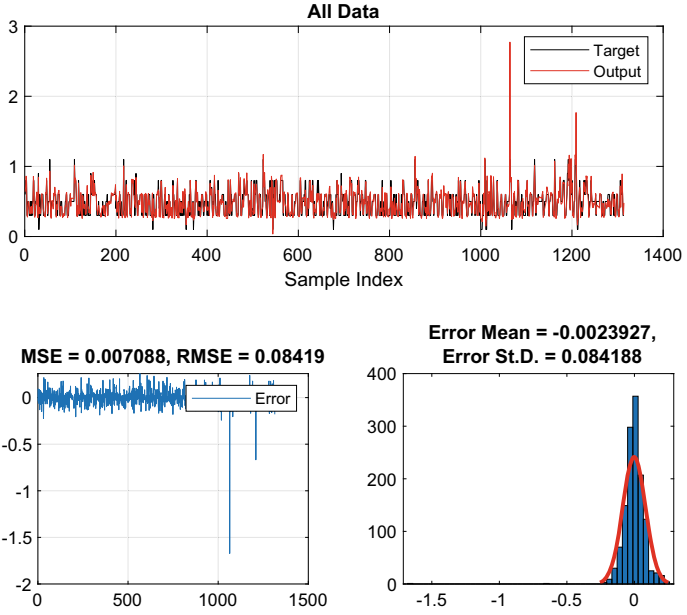


Fig. 6 Error information of whole data

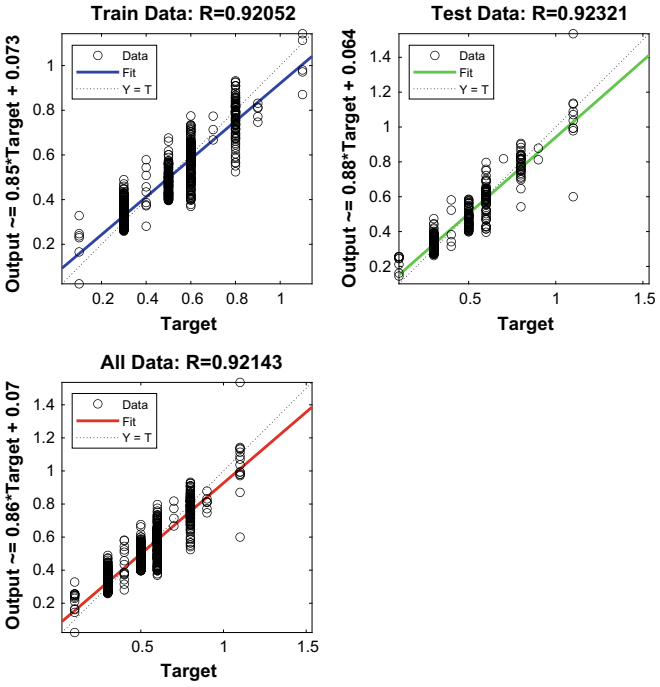


Fig. 7 Correlation coefficient (R) for training, testing as well as all the data

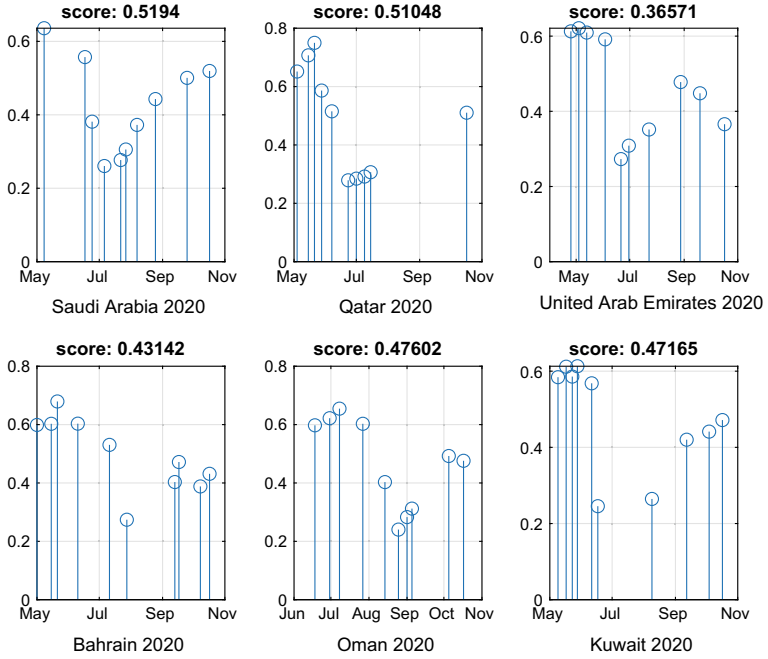


Fig. 8 Control measure score of GCC countries

5 Conclusion

In this paper, control measure score is predicted based on the recovery rate and the transmission rate using adaptive neuro-fuzzy inference system. The score is predicted by considering the transmission rate and recovery rate of all countries. These parameters are computed during the period at which the mean value of active case changes abruptly. This would help to accommodate the information on delay distribution of infection and removal rate in the transmission parameter. As the transmission and recovery rates are two deciding parameters to find the impact of the disease, we model the control measure based on these parameters using FCM (FCM-ANFIS). The error is calculated for both training and testing data; it is found that there is no much variance which ensure the robustness of the model. The machine learning technique is applied in time series data by taking the information on parameters from all countries which gives a meaningful information. Appropriate policies also can be framed based on the control measure score in order to lessen the spread of the disease.

Acknowledgements This research is a part of funded research project approved by The Research Council (TRC) of the Sultanate of Oman, under Commissioned Research Program, Contract No. TRC/CRP/HCT/COVID-19/20/06.

References

1. WHO Coronavirus Disease (COVID-19) Dashboard. https://covid19.who.int/. Last accessed 21 Oct 2020
2. Yegnanarayana, B.: Artificial Neural Networks. PHI Learning Pvt. Ltd., New Delhi (2009)
3. Mendel, J.M.: Fuzzy logic systems for engineering: a tutorial. Proc. IEEE **83**(3), 345–377 (1995)
4. Jang, J.S.: ANFIS: adaptive-network-based fuzzy inference system. IEEE Trans. Syst. Man Cybern. **23**(3), 665–685 (1993)
5. Guillaume, S.: Designing fuzzy inference systems from data: an interpretability-oriented review. IEEE Trans. Fuzzy Syst. **9**(3), 426–443 (2001)
6. COVID-19 Dataset. https://raw.githubusercontent.com/datasets/covid-19/master/data/time-series-19-covid-combined.csv. Last accessed 08 Aug 2020

A Study on Benefits of Big Data for Urban Flood Control in Surat City

Pankaj J. Gandhi and Prasit G. Agnihotri

Abstract Recently, across the globe use of big data has been increased to a great extent. Due to the use of computer, the generation of data in every domain has been increased to a great extent, due to the same a concept of big data has been crystallized among the researchers. Flood situation in India is very common, and due to the poor infrastructure, its control and management are difficult at the grassroot level. Surat is one of the cities in western India, which is highly prone to flooding situations. Urban flood control is one of the necessary needs in the context of smart urban planning and the said can be derived with the help of big data in today's era. Similar to most advanced and developed cities of countries like US, Europe, China, and Japan leveraging the benefits of big data for flood control and mitigation, Surat can also make use of the big data. In the present paper, we have started with the overview the need of early prediction of flood and its control across the urban areas of the globe as well as Surat city, followed by the common advantages of big data in flood control for coastal cities like Surat as well as field-specific uses and concluded with a broad framework portraying big data in context to flood control of Surat city.

Keywords Big data · E-governance · Surat · Urban flood control

1 Introduction: Overview of the Urban Flood and Its Control

Urban flooding is one of the most severe and costly natural hazards in the world. In absentia of sufficient data related to the urban flooding events, such as flood severity, occurrence of flood, and impact of flood on people, it is difficult to prepare

P. J. Gandhi (✉) · P. G. Agnihotri
Civil Engineering Department, Sardar Vallabhbhai National Institute of Technology, Surat, India

P. G. Agnihotri
e-mail: pga@ced.svnit.ac.in

P. J. Gandhi
School of Engineering, P. P. Savani University, Kosamba, Surat, India

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_93

resiliency plan and other disaster management activities [1]. Recent past of few years is a witness of devastating floods in major urban areas across the globe [2, 3]. Scholars have pointed out those major reasons for urban flooding are climate change, unplanned urban development [4], and urban administrative apathy [5]**.** In the post-globalization era, sociologist and economists have realized that for keeping the pace of economic development, and there is a need to develop urban centers as socio-economic engines to support the GDP of the nation [6].

The urban development in major Asian nations is a kind of haphazard or non-harmonized development to cope up with population explosion, resulted into urban sprawl in low terrain areas of the city in the form of slums [6, 7], such areas are highly vulnerable to urban flooding. The non-harmonized and unplanned urban development in major Indian cities has caused the problems of impervious pavements, filling of natural water bodies–for to urban need of residential/commercial/industrial development and slum type settlements in coastal areas, in nearby vicinity of river belt and in low-level terrain of the city where risk of flooding is high [6–13].

Some of the research scholars have identified that flooding in urban areas is mainly occurs due to the land use change, unsustainable use of natural resources, pollution, climate change, and soil erosion [4, 8, 14]. Dewan and Yamaguchi [10] have considered the factors like river flow, rainfall, flood plain of city, and inundated areas of the city in the context of different urban development scenarios for the prediction of flood risk [11]. Post-floods scenarios are a part of disaster management system where the major focus is to reduce the socio-economic damages [15, 16]. The cities which exhibit higher urban growth rate have higher chances of urban flooding and devastated impact of flood [11, 12]. Waghwala and Agnihotri [9] have analyzed the integration of the urbanization with flood risk, and accordingly, it can be said that there is a need to use past data of floods and its damages to the urban clusters in the form of big data [9].

Zhang et al. [3] in their study found that the flood risk increases with the unplanned urban development [4]. Surat is a witness of such unplanned urban development and is vulnerable to frequent urban flooding [9]. Therefore, it is necessary to design and develop a framework of big data analytics for controlling the urban flood. Witting and Kallee [17] have highlighted the need of big data use in urban flood hazards from the perspective of capital investments needed after the catastrophic flood events [17]. Keiza and Mary [18] have analyzed the risk of rapid flood as irremediable calamities from the large datasets of rainfall and MapReduce technique as a part of big data [18]. Wang [1] has highlighted the limitation of exiting geospatial techniques for the development of big data pool for urban flooding, mainly the cost aspects and involvement of advanced technology and suggested to use data mining of social media and crowd-sourced data in collaboration with mycoast.org—a mobile app company to collect thousands of photos from the general public and develop big data of flooding [1].

Wang [1] has labeled the flooding events photos of social media with respect to time and location and with the help of computer vision program and natural language processing (NLP) tool and taking out the details from Twitter, news reports, and traffic data to generate the big data pool for urban flooding [1]. Wu et al. [19]

have used Web crawler technology for the mining of flood-related information from Internet and social media and developed big data-based flood disaster vulnerability assessment model on analytic hierarchy process (AHP) with traditional statistical data and assess the rainstorm and flood disaster vulnerability based on geographic information system (GIS) for pluvial flooding in Zhengzhou city [19].

This era of computer and information technology has resulted into the generation of plenty of data, known as big data and emergence of a theory of machine learning algorithms. Both these concepts are combined by water engineers for the prediction of rainstorm, waterlogging, urban flooding, modeling urban coastal flood severity, hyper-resolution monitoring of urban flooding, and corrective steps initiation for mitigating the flood disasters [19–22]. Use of big data is very much important as abundant technologies now available for the collection of urban flooding data [19]. Due to the usage of Internet and social media, it is easy to predict that in near future, and the degree of information related to flood disasters will continue [19]. In the context of such development, the urban local body of Surat city has to develop big data that serves as a first step for urban flood disaster risk assessment. The objective of this study is to highlight the importance of using big data by city like Surat for the development of urban flood control strategies and flood disaster vulnerability assessment model.

2 About the Surat and Floods of Surat

Surat is one of the coastal cities located in South Gujarat having an average elevation of 13 m and has a population of about 44,67,797 and is one among the fastest developing urban cluster in Asia [23]. Tapi River is a lifeline of this city as it provides the water for domestic, irrigation, and industrial need. The city is having a higher flood risk due to the presence of Tapi River, and several major and minor creeks pass through the city. The average rainfall in lower Tapi Basin is around 1376 mm, which often results in a heavy discharge from the Ukai dam resulted into frequent flood incidences in Surat.

In India, the city of Surat is known as ‘art silk Manchester’ or ‘textile city,’ ‘diamond city,’ but due to the frequent flood incidences in pre-modern times, it is also known as ‘flood city.’ The city of Surat has a strategic location on DMIC and one of the important economic centers for diamond and textile industry [24]. The geographic location of Surat is in flood prone area and bearing two types of flood risk, such as river flood and Khadi flood, and past data indicates that the frequency of flood and subsequent flood risk has been much higher than the other river deltas of western India [25]. Past data of lower Tapi Basin flood indicates that the flood occurred in the years 1883, 1884, 1894, 1942, 1944, 1945, 1949, 1959, 1968, 1994, 1998, and 2006 [26]. The city has experienced the devastating nature of Tapi in 1968 and 2006 [27]; apart from this river flood, the city is a witness of the socio-economic damages from frequent Khadi floods, which has been increased due to the rapid urbanization [6, 25].

The Surat city is a financial capital of Gujarat state and considered as an economic nervous system of the state. Hence, it is essential to protect the city from flooding and use of big data can be advantageous to minimize the socio-economic damages.

3 Advantages of Big Data in Urban Flood Control

The use of big data can help to reduce the flood risk and losses from the damages and enhance the resiliency of the city. Starting with the discussion about the applicability of big data in flood control on a general basis, the major benefits are as follows:

- **Prediction of Flooding Scenarios**: With the use of big data server, the past data of flooding can be easily stored and retrieved for decision making.
- **Simulation Modeling**: Based on the various combinations of data, the worst scenario of flooding can be developed. Based on the simulated outcome, the users, mainly the government will get the idea of how, where, and when to mobilize the resources to minimize the damages from floods and establishment of faster resiliency in the area.
- **Development of the disaster management strategies**: Based on the processing of the data systematic classification can be developed to frames the rules, and to investigate their pursuance on the use of 'big data' for flood management and disaster control and reduction of flood hazards in the Surat Municipality Area.
- **Useful for other coastal cities**: The outcome of the usage of big data in Surat can be equally deployed in any coastal cities of the world as a case study where similar frequency of flood occurs.
- **Emergency response**: It will provide a valuable support for the development of emergency response for meeting the need of resiliency and minimum damages from flood.
- **A real-time detection**: With the help of big data, it is possible to develop real-time flood detection with TensorFlow by considering the development in social media. In this approach, there is no need of GIS and remote sensing data.
- **Cost of Model Development is low**: Most of the models in flood vulnerability assessment are based on statistical data recorded in literature and very few models are of real time. Big data sources, such as picture, video, and text data, are collected from the Internet and social media platforms which can be applied for the development of cost-effective models for flood control and its vulnerability [28, 29].
- **Fusion with earth Science**: Past few decades are the witness of utilization of the earth observation data through satellite images and other geospatial techniques, which can now be synchronized and combined with advances in machine learning algorithms to have more explorations in the field with big data to predict and analyze the rainstorm and flood disasters. Nowadays, big data are being extensively used in the rainstorms and flooding, such as urban runoff and water logging risk assessment [20], modeling urban coastal flood severity [29] hyper-resolution

monitoring of urban flooding [1], and real-time identification of urban rainstorm waterlogging disasters [22, 30].

- **Synchronization with Internet and Social media**: It is easy to use on hand multiple platforms for the collection of multi-spectrum data regarding the urban flooding through the use of Internet-connected devices and information related to flood disasters through social media. Use of both these sources will serve as a path breaking development of big data for the urban flood disaster risk assessment.
- **Web Crawler Technology**: With the help of this, a text data on Internet platform can be a source for the development of big data pool. Such type of text data can be collected from various social media platforms, such as Facebook, Whatsapp, Twitter, Internet Web sites, and many more and processed and stored for meeting the requirements of the flood control [20, 22, 31].
- **Low reliance on historic data**: Most of the earlier research studies related to the flood control and disaster management are heavily focused on the historical data, and because of the same, there are errors in predicted outcome, when used through the simulation or modeling. The said error can be minimized with the use of an emerging big data for the vulnerability assessment.

4 Model of Big Data for Urban Flood Control in Surat City

In this section, we discuss a model for Urban Flood Control of Surat City of Gujarat, India. The big data for flood control is to be developed from different resources as shown in Fig. 1. Before the usage of the data, there is a need to carry out preprocessing of data smoothing and cleaning to avoid the data duplicity. The processing of all the collected data is in the context of Surat and its surroundings, where some of the data are to be collected from government departments such as meteorological department, weather department, city planning office of Municipal Corporation, zone-wise population, and social media data of people. After the collection of the details of every data, their analysis to be carried out on in the context of real-time needs to derive the constructive and helpful information. The said information derived from these

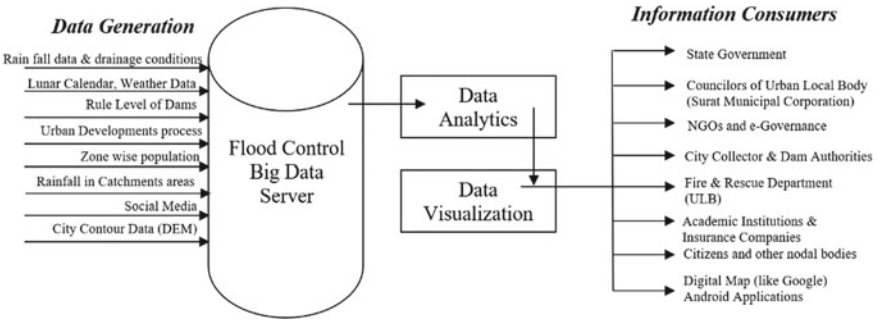


Fig. 1 Model of big data for urban flood control of Surat city of Gujarat, India

real-time data will be used by various consumers for their need. Also, the derived information will be useful for the development of the disaster management plan for the rescue of the people and also for the development of resilient strategies.

The details regarding how the gained information can be used by different information consumers are given below:

- **State Government**: The concerned authorities of the Gujarat state government can monitor the situation online on the computer monitor at the Gandhinagar, Gujarat and according to mobile the needed resources and support. Also, they can provide the support in decision making. The needed support from paramilitary forces can be deployed by them for the rescue operations. Based on the severity of the flood, the government can declare the emergency and extends the monetary help.
- **Councilors of Urban Local Body (Surat Municipal Corporation)**: The councilors of the respective ward of the city can provide support to the citizens for the distribution of grocery, food, medicines, and other needed items, mobilization of support resources, carryout of the rescue operations for sick and needy people.
- **E-Governance**: The E-Governance is an extended IT wing of Municipal Corporation and functions in coordination with the IT team, municipal commissioner, and Mayor of the city and provides a support for the decision making and resources mobilization. E-Governance team, after an approval of mayor, commissioner, and collector, will release the messages in the social media, radio, and television to control the rumors in the city.
- **NGOs**: They will functions in a similar way to municipal councilors. They coordinate with self-help groups (SHGs) for the supply of foods, medicines, and other needed items. Also, they provide a support to carry out rescue operations.
- **City Collector and Dam Authorities**: Based on the input data of the weather, rainfall, catchments area position, and lunar calendar, they will take decisions for maintaining the discharges of water from the dam and follow the rule level guidelines. Based on the information, the city collector will coordinate with his team and commissioner of Municipal Corporation for managing the level of flood in a city and initiation of actions for managing the flood disaster. Also, the collector will regularly update the dam authorities for the city position. Dam authorities can also observe the likely scenarios of the city and accordingly prepare their work plan. As a part of rescue, they can allocate the corporation school to the affected people for the temporary shelters.
- **Fire and Rescue Department (ULB)**: It is one of the most important pillars of the urban flood control activities. They will act and mobilize the needed resources for the disaster control. They are like a rapid action force team for the safety of the city at large.
- **Academic Institutions**: The data will be useful for carryout the further studies by the students, development of models, and carry out simulation runs. Also, based on the simulation study, they can develop a decision support system (DSS) for the apex body of the city.

- **Insurance companies**: Insurance companies are providing insurance for the protection of the properties, such as residential, commercial, and industrial for protecting them against the natural calamities, including the floods. In such cases, the amount of premium to be charged against the sum insured of property can be varied according to the location of the property in a city. For this, the insurance companies can use the big data, such as (1) water level in a nearby vicinity of the property to be insured, (2) number of hours water remains in the area, (3) water discharge from dam, (4) lunar calendar, and (5) rainfall. Accordingly, insurance companies can use big data to identify and prevent false claims as well as they can use big data for the premium to be paid.
- **Citizens and other nodal bodies**: They can also use the big data for their day-to-day need and travel plan. The nodal bodies like temple, mosques, churches, gurudwara, and community halls can allocate their premises for the use of flood-affected people.
- **Digital Map (like Google)—Android Applications**: This application is a kind of traffic density indication on Google Map. Based on the various inputs from social media, the system will generate the details of the water depth level of respective area, and accordingly, it will guide the people of the city for their movement.
- **Reliability of data**: The data collected through this big data framework (Fig. 1) are highly reliable in nature as they are collected from the actual agencies and are of real time, such as data of rainfall collected by the urban local body; lunar calendar is fixed; rule level of dam and water discharge from dam are provided by the irrigation depart of government; data pertaining to contours of the city is developed by the geomatics survey department of state government; population details are from the urban local body; hence, the data used in this big data system are robust in nature, highly reliable, and validated through various systems.

5 Conclusion

Urban flooding is one of the most serious issues across the globe, especially in the coastal regions. It is one of the serious natural disastrous occurrences which take place every year and damages a lot to the cities in terms of socio-economic-and-environmental aspects. It is a well-known fact that urban flood is a kind of natural disaster which cannot be avoided, but with the use of advance concepts of technology such as big data, we can minimize the losses happening due to flooding.

Based on the experience, it can be said that it is essential to have a proper flood planning which comprises of: (1) the estimation of flood extent, and (2) risk of flood hazard at different conditions of flow discharge, rainfall in city, and lunar conditions. As indicated in Fig. 1, the big data generated are of real time in nature, and accordingly with the help of data analytics and data visualization, they will provide a visual images of likely scenarios (simulated) of the flood. Based on the availability of real-time big data and utilization of data analytics and data visualization software, the authorities can predict the level of flood in a Surat city, and accordingly they can prepare the

flood evacuation strategies for the transfer of people and disaster management plan for the post-flood scenarios.

The said complexity of urban flooding arises due to the: (1) high intensity rainfall in a city area, (2) drainage conditions and flooding caused by overtopping in the channels or rivers, (3) flooding due to high tides—lunar calendar, (4) level of water discharge from the dam, and (5) rainfall in the catchments area. In coastal urban cities like Surat, mostly severe flood scenarios happened due to the combination of urban runoff, channel overtopping, and tidal flooding. For effective coastal urban flood management and mitigation plans, the possible flooding scenario is to be simulated with the help of big data for various combinations of the factors responsible for flooding, such as level of rainfall events and discharges of water from dam.

References

1. Wang, R.Q.: Big data of urban flooding: dance with social media, citizen science, and artificial intelligence. In: Geophysical Research Abstracts, vol. 20. EGU2018–404, 2018 EGU General Assembly 2018
2. Dhiman, R., Vishnu Radhan, R., Eldho, T.I., Inamdar, A.: Flood risk and adaptation in Indian coastal cities: recent scenarios. Appl. Water Sci. **9**(5), 1–16.
3. Zhang, H., Ma, W.C., Wang, X.R.: Rapid urbanization and implications for flood risk management in hinterland of the Pearl River Delta, China: the Foshan study. Sensors **8**(4), 2223–2239 (2008)
4. Bhat, G.K., Raghupathi, U., Rajasekar, U., Karanath, A.: Urbanization—Poverty—Climate change. A Ssynthesis Report—India, vol I. TARU Leading Edge Pvt. Ltd., Gurgaon, Haryana, India
5. Biswas, A.K., Saklani, U., Tortajada, C.: Truth about urban flooding: cities like Mumbai get inundated regularly due to administrative apathy, not climate change. Times of India, Mumbai Edition 31.08.2017 (2017)
6. Mohan, R., Dasgupta, S.: Urban development in India in the twenty first century: policies for accelerating urban growth. In: Fifth Annual Conference on Indian Economic Policy Reform at the Stanford Centre for International Development on 4–5 June 2004. Working Paper No. 231 (2004)
7. Ahmed, A., Dinye, R.D.: Urbanisation and the challenges of development controls in Ghana: a case study of WA Township. J. Sustain. Dev. Afr. **13**(7), 210–235 (2011)
8. Nukala, R.B., Mutz, D.: Strategic approach for sustainable land use in an emerging country—case of India. In: 2015 World Bank Conference on Land and Poverty. The World Bank, Washington DC, 23–27 Mar 2015 (2015)
9. Waghwala, R., Aagnihotri, P.G.: Flood risk assessment and resilience strategies for flood risk management: a case study of Surat City. Int. J Disaster Risk Reduct. **9**(5), 1–16 (2019)
10. Dewan, A.M., Yamaguchi Y: Effect of land cover changes on flooding: example from Greater Dhaka of Bangladesh. Int J Geoinf 4(1). https://creativecity.gscc.osaka-cu.ac.jp/IJG/article/view/604 (2008)
11. Hammond, M.J., Chen, A.S., Djordjević, S., Butler, D., Mark, O.: Urban flood impact assessment: A state-of-the-art review. Urban Water J **12**(1), 14–29 (2015). https://doi.org/10.1080/1573062X.2013.857421
12. Dewan, A.M., Islam, M.M., Kumamoto, T., Nishigaki, M.: Evaluating flood hazard for land-use planning in greater Dhaka of Bangladesh using remote sensing and GIS techniques. Water Resour. Manage. **21**, 1601–1612 (2007). https://doi.org/10.1007/s11269-006-9116-1

13. Dewan, A.M., Yamaguchi, Y.: Land use and land cover change in Greater Dhaka, Bangladesh: using remote sensing to promote sustainable urbanization. Appl. Geogr. **29**(3), 390–401 (2009)
14. Patel, K.A., Bhagat, S.S.: An overview of flood resilience: a case of Surat City. Int. J. Manage. Technol. Eng. **9**(1), 3043–3049 (2019)
15. Brody, S., Blessing, R., Sebastian, A., Bedient, P.: Examining the impact of land use/land cover characteristics on flood losses. J. Environ. Planning Manage. **57**(8), 1252–1265 (2014)
16. Thieken, A.H., Apel, H., Merz, B.: Assessing the probability of large-scale flood loss events: a case study for the river Rhine, Germany. J. Flood Risk Manage. **8**(3), 247–262 (2015)
17. Witting, A., Kallee, D.: Ruling big data in urban flood risk mitigation. In: Yearbook of Swiss Administrative Sciences, pp. 11–22 (2015)
18. Keiza, S.P., Mary, A.V.A.: Prediction of rapid floods from big data using map reduce technique. Glob. J. Pure Appl. Math. **12**(1), 369–373 (2016). ISSN 0973-1768
19. Wu, Z., Shen, Y., Wang, H.: Assessing urban areas' vulnerability to flood disaster based on text data: a case study in Zhengzhou City. Sustainability **11**(4548), 1–15 (2019). https://doi.org/10.3390/su11174548
20. Lin, T., Liu, X.F., Song, J.C., Zhang, G.Q., Jia, Y.Q., Tu, Z.Z., Zheng, Z.H., Liu, C.L.: Urban waterlogging risk assessment based on internet open data: a case study in China. Habitat Int. **71**, 88–96 (2018)
21. Sadler, J.M., Goodall, J.L., Morsy, M.M., Spencer, K.: Modeling urban coastal flood severity from crowd-sourced flood reports using Poisson regression and Random Forest. J. Hydrol. **559**, 43–55 (2018)
22. Yang, X., Beiqun, L., Zaiwu, G.: Real-time identification of urban rainstorm waterlogging disasters based on Weibo big data. Nat. Hazards **94**, 833–842 (2018)
23. City Mayor's Report. The world's fastest growing cities and urban area from 2006 to 2020. https://www.citymayors.com/statistics/urban_growth1.html. Accessed 07 Dec 2018
24. Singapuri, N.C., Bipinkumar: Two local industries of Surat under the muddled conditions of business environment–evaluation under the international factors. J. Adv. Scholar. Res. Allied Educ. (JASRAE). **15**(12), 783–801. E-ISSN 2230-7540. https://doi.org/10.29070/JASRAE.
25. Kale, V. S.: Long period fluctuations in monsoon floods in the Deccan peninsula, India. J. Geol. Soc. India **53**, 5–15 (1999). https://hdl.handle.net/11718/114
26. Agnihotri, P.G., Patel, J.N.: Improving carrying capacity of river Tapi (Surat, India) by channel modification. Int. J. Adv. Eng. Technol **II**(II), 231–238 (2011)
27. Mavalankar, D.: Lessons from massive floods of 2006 in Surat city: a framework for application of MS/OR techniques to improve dam management to prevent flood (2008)
28. Ahmad, A., Khan, M., Paul, A., Din, S., Rathore, M.M., Jeon, G., Chio, G.S.: Towards modeling and optimization of features selection in big data based social Internet of things. Future Gener. Comput. Syst. **82**, 715–726 (2018)
29. Aqib, M., Mehmood, R., Alzahrani, A., Katib, I., Albeshri, A., Altowaijri, S.M.: Rapid transit systems: smarter urban planning using big data, in-memory computing, deep learning, and GPUs. Sustainability **11**, 2736 (2019)
30. Wang, R.Q., Mao, H., Wang, Y., Rae, C., Shaw, W.: Hyper-resolution monitoring of urban flooding with social media and crowdsourcing data. Comput. Geosci. **111**, 139–147 (2018)
31. Eilander, D., Trambauer, P., Wagemaker, J., Van Loenen, A.: Harvesting social media for generation of near real-time flood maps. Procedia Eng. **154**, 176–183 (2016)

Multi-sink Routing Protocol for Underwater Sensor Networks

Vijayalaxmi R. Patil and Anita Kanavalli

Abstract A major issue of UWSNs is energy consumption of resources. UWSNs are equipped with batteries of limited power. Recharging or replacement of batteries is a difficult task when battery power is completely exhausted. Hence, it is very much essential to utilize power efficiently so as to enhance the network lifetime. To bring reduction in energy consumption of each node in the network, nodes are deployed layer wise and a sink is provided at each layer. The nodes belonging to particular layer send the sensed data to sink designated for that layer. This eliminates the participation of more number of nodes as intermediate nodes which indirectly helps in balancing the energy of network. The simulations are performed first by deploying a single sink on surface of water and then by deploying sink on each layer. The performance metrics are determined for both static single sink and static multi-sink architecture and are compared in this paper.

Keywords UWSNs · Energy consumption · Routing protocol · Multi-sink · Energy tax

1 Introduction

Water covers 70% of earth's area. UWSN is a network that is the main portion lies underwater. It mainly consists of underwater sensors, underwater sinks, surface sinks and autonomous underwater vehicles in order to monitor specific region. The UWSNs are characterized by small bandwidth, high delay, signal attenuation, mobile nodes and higher error probability. UWSNs may have one-dimensional, two-dimensional or three-dimensional network topology. Applications of UWSNs include collection of oceanographic data, monitoring of the presence of chemicals in water, oil and

V. R. Patil (✉)
Ambedkar Institute of Technology, Bengaluru, India
e-mail: vijumalghan@yahoo.co.in

A. Kanavalli
M S Ramaiah Institute of Technology, Bengaluru, India
e-mail: anithak@msrit.edu

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_94

gas exploration, disaster detection and prevention, navigation assistance (detection of rocks underwater), tactical surveillance, tsunami and seaquake detection.

2 Related Work

The VBF [1] protocol uses a concept of vector pipe for routing packets. Only the nodes which are lying within vector pipe are involved in forwarding packets. In HHVBF [2], each node has its own routing pipe. In NIR [3] protocol, the weight of neighbor node is calculated to select the next hop node. CRP [4] protocol addresses the reliability aspect due to high probability of error in UWSNs. Connectivity index of node is used to decide on next forwarding node. TCBR [5] protocol uses multi-sink architecture, ordinary and courier nodes. The sensor nodes are divided into clusters based on their locations. FBR [6] protocol addresses the issue of energy consumption. The transmitting node adjusts power level of transmission and only the nodes within cone angle become candidate relay nodes. The node near to sink is selected to be next relay node. DFR [7] protocol addresses packet reliability and void problem. DBR [8] protocol eliminates location information requirement. DSVBF [9] uses location information and residual energy as factors to determine route. H2DAB [10] protocol uses multi-sink architecture in which route creations are done by assigning dynamic hop ids. EEDBR [11] uses residual energy and depth of nodes for selecting next forwarding node. MRP [12] protocol uses supernodes having abundant energy.

3 Routing Mechanism

Let L, B, H represent length, breadth and height of water environment. Then total volume of water given by $V = L \times B \times H$ is divided into layers in which surface of water forms Layer-1. Layer-2, Layer-3 and Layer-4 are at depth D, $2 \times D$ and $3 \times D$, respectively, from water surface. Sensor nodes are deployed layer wise. Each layer has one sink. The nodes belonging to particular layer always send data to the sink of that layer. All the nodes and the sink nodes are stationary. Instead of sending data to surface sink always, the nodes route sensed data to sink designated for that layer. If sink is not within range of sender, it is routed via the intermediate node. The packet is forwarded only to node which belongs to same layer.

3.1 Network Architecture

The sensors are deployed in 3D with terrain of 1000 m $\times$ 1000 m $\times$ 1000 m. The sensor nodes are deployed layer wise. Each layer has the same number of nodes. There exist two different scenarios:

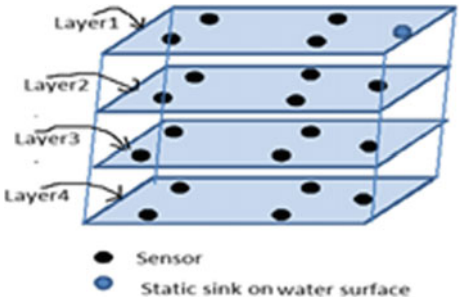


Fig. 1 Network model for single surface sink architecture

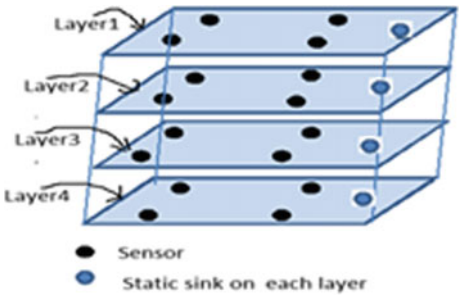


Fig. 2 Network model for layer wise multi-sink architecture

Scenario-1 Figure 1 shows the single static sink deployed on water surface. All the sensors deployed on water surface and all sensors deployed underwater send sensed data to only one sink deployed on surface of water.

Scenario-2 Multi-sink architecture given in Fig. 2 shows that each layer has a static sink. The sensed data is sent to the sink of that layer.

NS-2-based network simulator is used for simulation. The parameters used for simulation are listed in Table 1.

3.2 Performance Metrics

The various performance metrics considered for research are average energy consumption, energy tax, average residual energy, packet delivery ratio and throughput.

- Average Energy Consumption: It is ratio of total consumed energy in network to total nodes in network.
- Energy Tax: It is ratio of total energy consumed in network to total packets received.
- Average Residual Energy: It is defined as ratio of total remaining energy to network size.

Table 1 Parameters

Parameter	Value
Node deployment	3D
Terrain	1000 m × 1000 m × 1000 m
Initial energy of node	50 J
Transmission power	2 W
Receiving power	0.1 W
Idle power	0.001 W
Sleep power	0.0001 W
Packet size	1000 Bytes
Traffic type	CBR
Total simulation time	20 s

- Packet Delivery Ratio: It is ratio of packets delivered to packets sent.
- Throughput: It is count of packets received per second.

4 Results

4.1 Scenario-1

The results obtained for scenario-1 on simulation are given in Figs. 3, 4, 5, 6 and 7.
Average energy consumption decreases as network size increases.
Energy tax decreases as network size increases.
Average residual energy increases with network size.

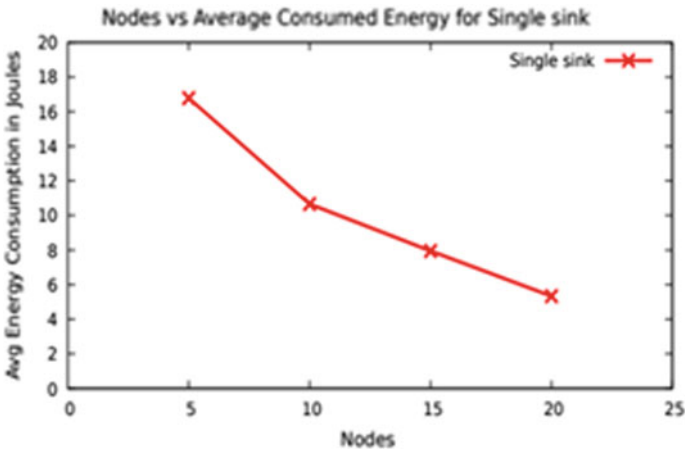


Fig. 3 Average energy consumption

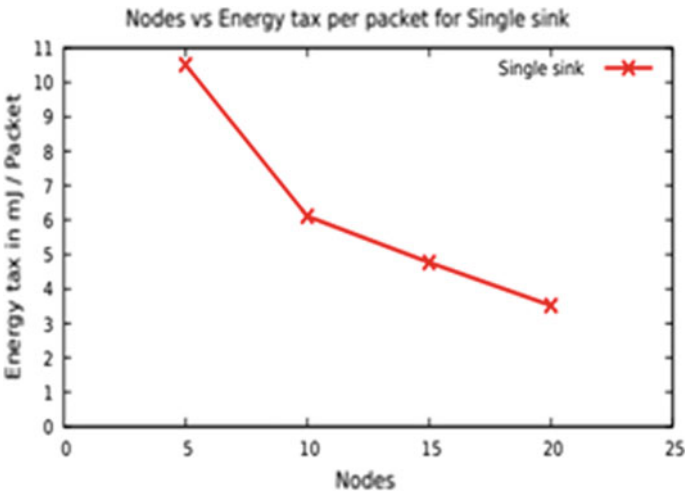


Fig. 4 Energy tax

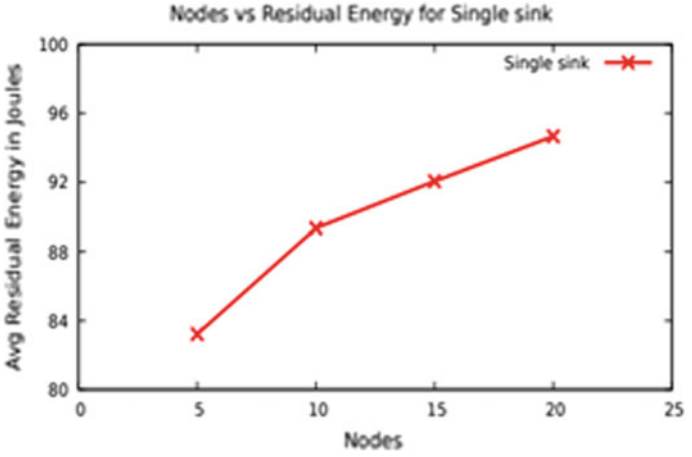


Fig. 5 Average residual energy

PDR is decreasing when network size is increased. This shows that packet drop has been increased due to increase in network size.

Throughput decreases with rise in network size.

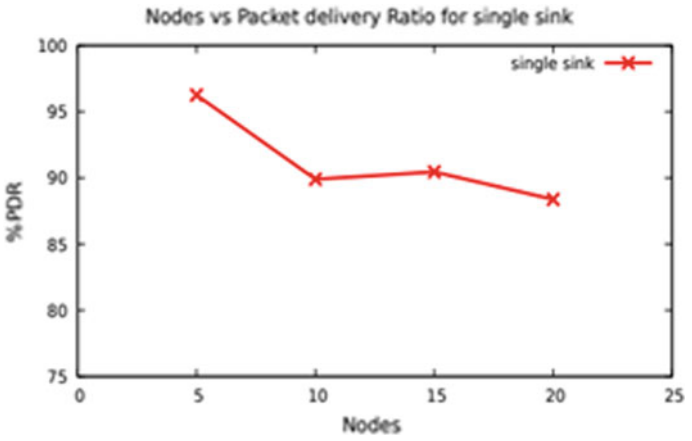


Fig. 6 Packet delivery ratio

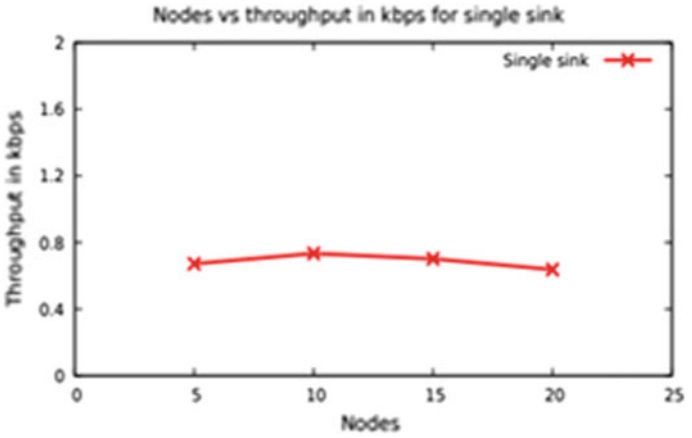


Fig. 7 Throughput

4.2 Scenario-2

The results for scenario-2 are shown in Figs. 8, 9, 10, 11 and 12. Average energy consumption decreases with increase in network size.

Energy tax decreases with increase in number of nodes.

Average residual energy increases with number of nodes.

PDR increases with increase in number of nodes ranging from 10 to 20.

Throughput increases for increase in number of nodes ranging from 10 to 20.

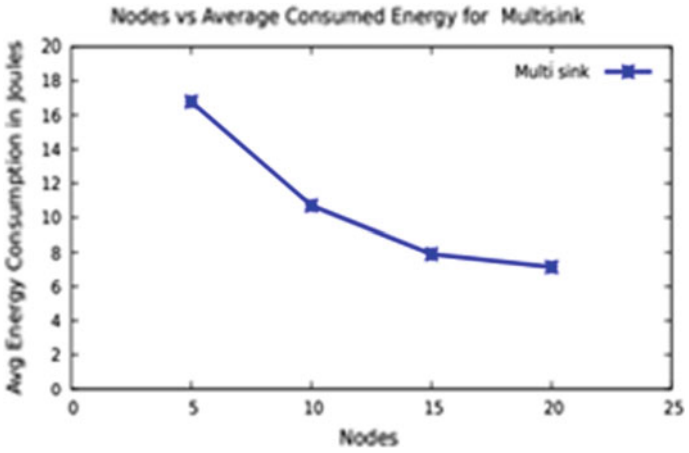


Fig. 8 Average energy consumption

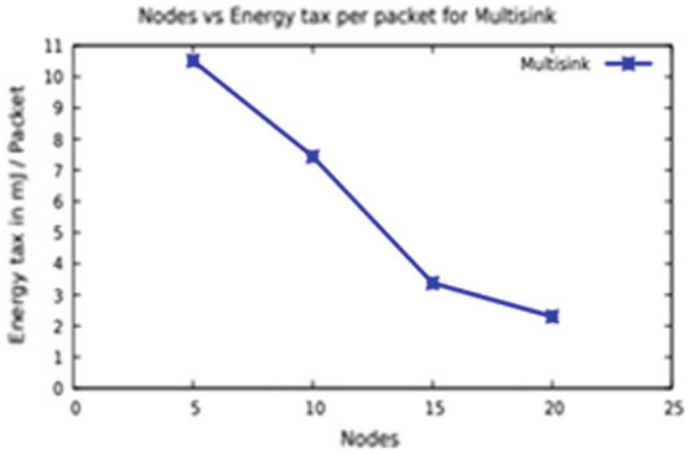


Fig. 9 Energy tax

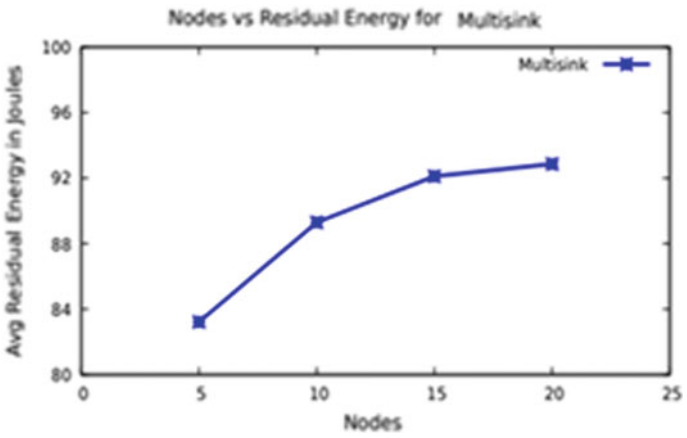


Fig. 10 Average residual energy

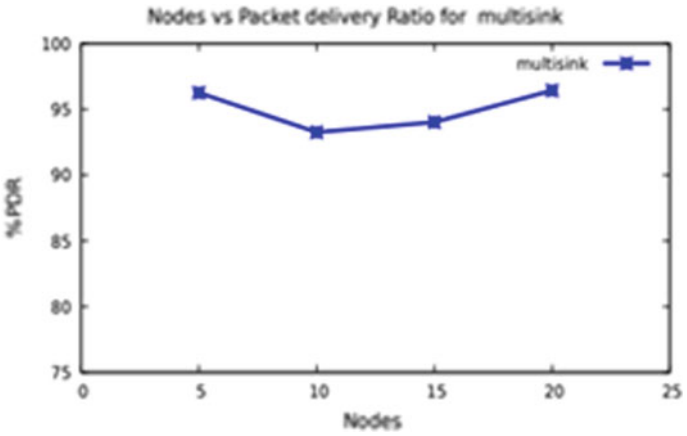


Fig. 11 Packet delivery ratio

4.3 Comparison of Results Obtained in Scenario-1 and Scenario-2

The results of comparison are depicted in Figs. 13, 14, 15, 16 and 17. The results show that the average energy consumption remains almost the same for 5, 10 and 15 nodes, but for 20 nodes, there is decrease in average consumption when single surface sink is used.

In terms of energy tax, multi-sink architecture shows a better performance for network of size 15 and 20.

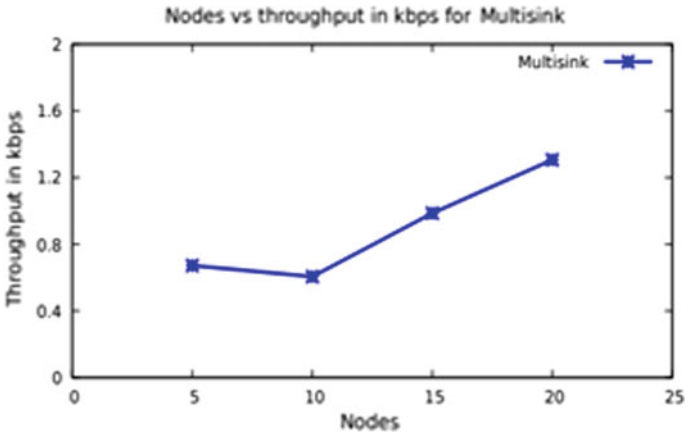


Fig. 12 Throughput

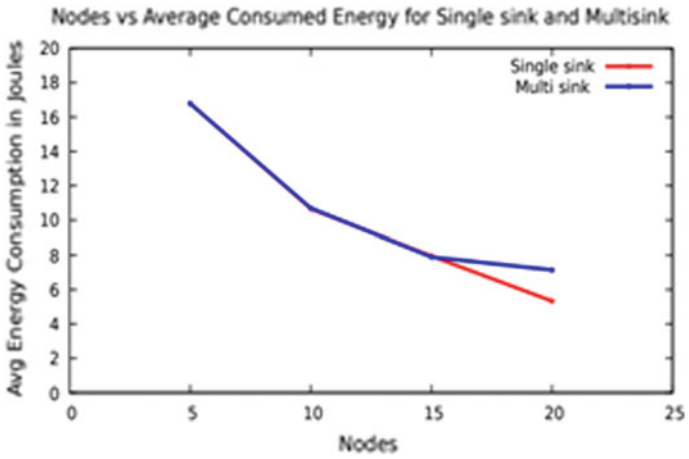


Fig. 13 Average energy consumption

Average residual energy remains almost the same for network of sizes 5, 10 and 15. But single sink architecture shows a better performance for network of size 20.

PDR is improved in multi-sink architecture compared to single sink architecture.

Throughput increases for nodes ranging from 10 to 20 in multi-sink, whereas it decreases in single static sink architecture.

Figures 18, 19, 20 and 21 show node wise residual energy for 20 s simulation time. For single layer of 5 nodes, there exist no change in residual energy when single sink or static multi-sink architecture are used.

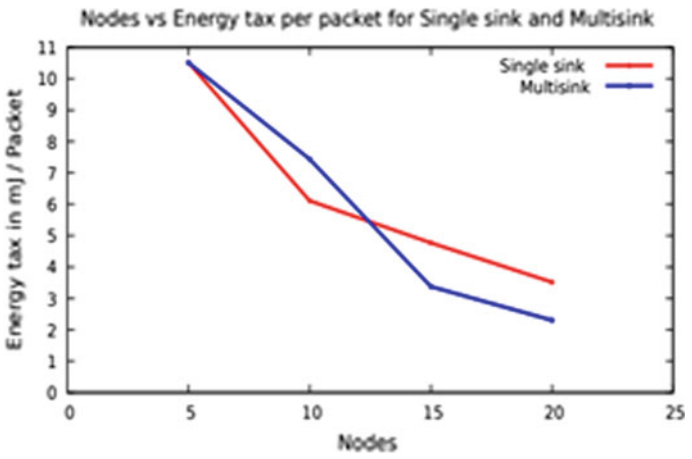


Fig. 14 Energy tax

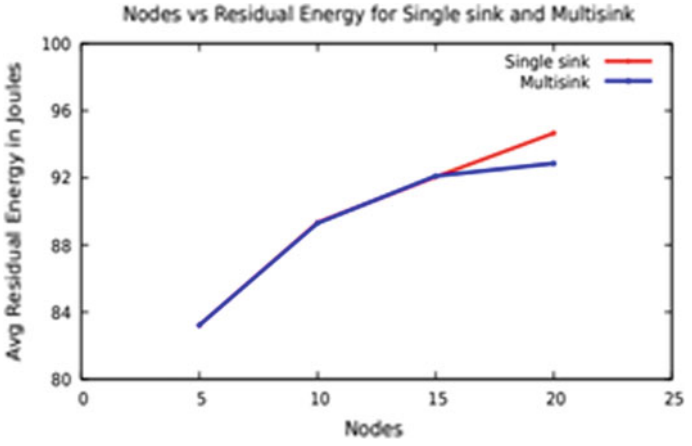


Fig. 15 Average residual energy

Instant throughput for single sink and multi-sink architectures are as depicted in Figs. 22, 23, 24 and 25.

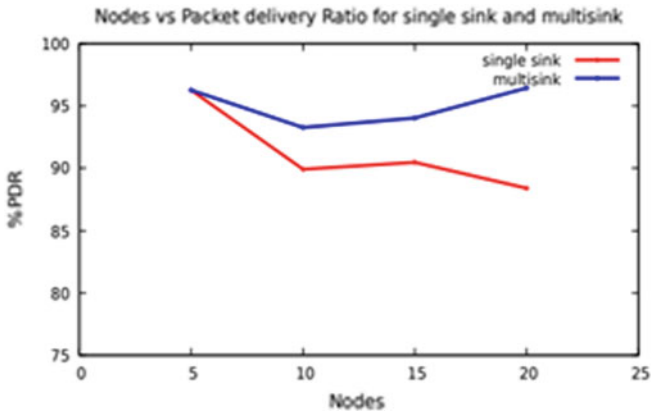


Fig. 16 Packet delivery ratio

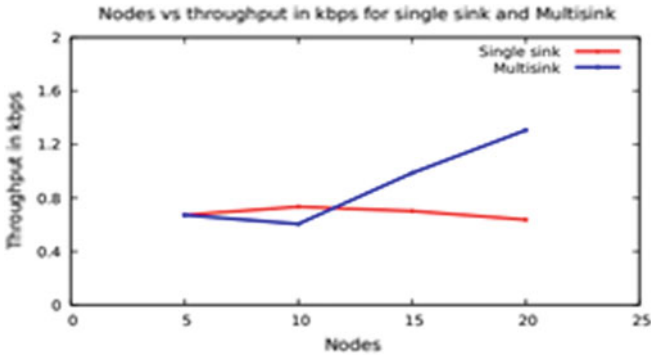


Fig. 17 Throughput

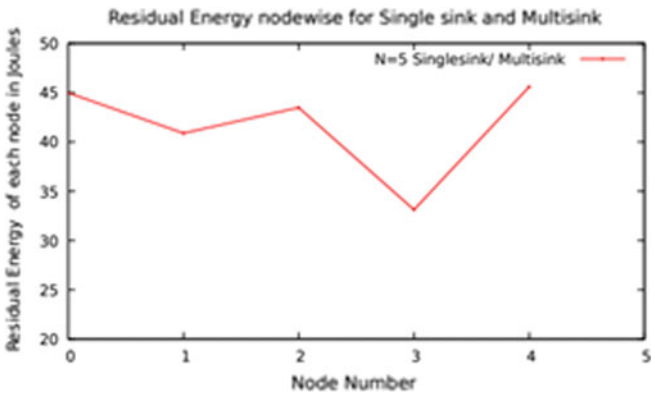


Fig. 18 Residual energy for 5 nodes

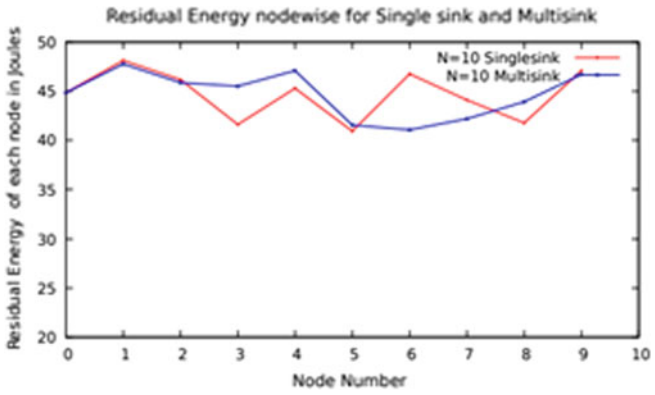


Fig. 19 Residual energy for 10 nodes

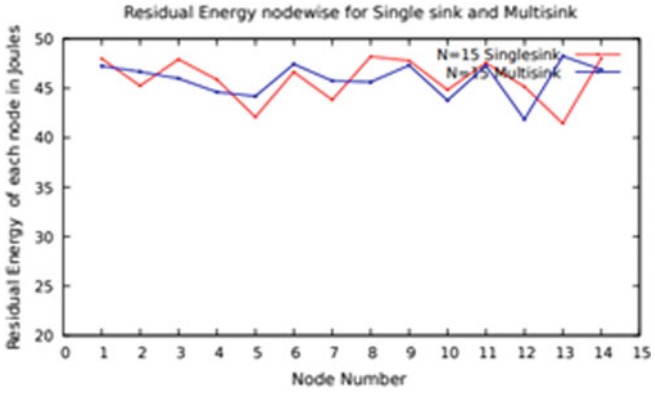


Fig. 20 Residual energy for 15 nodes

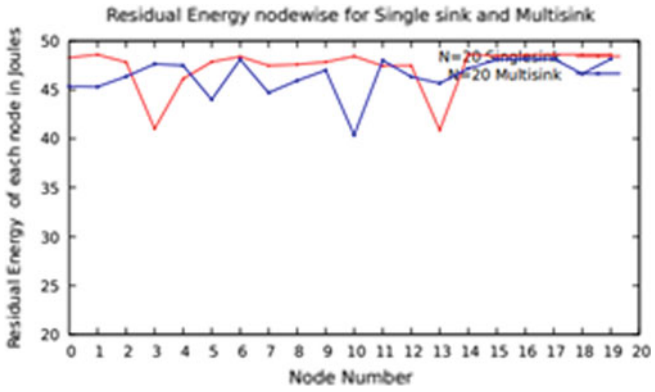


Fig. 21 Residual energy for 20 nodes

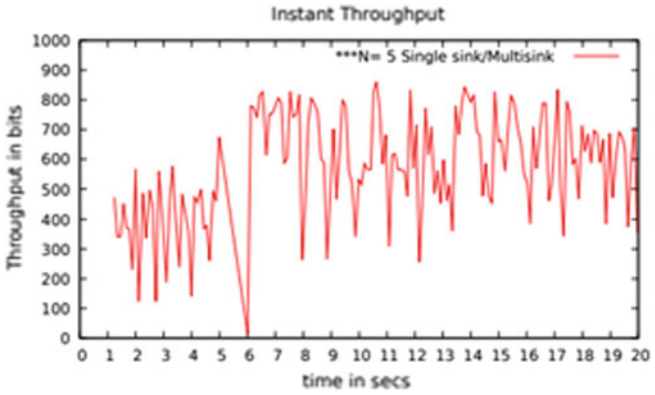


Fig. 22 Instant throughput for 5 nodes

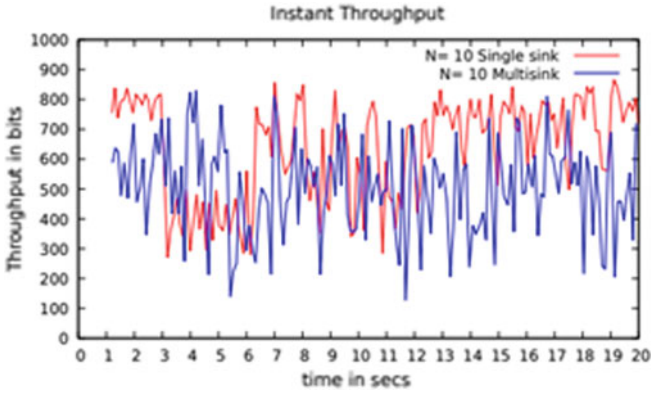


Fig. 23 Instant throughput for 10 nodes

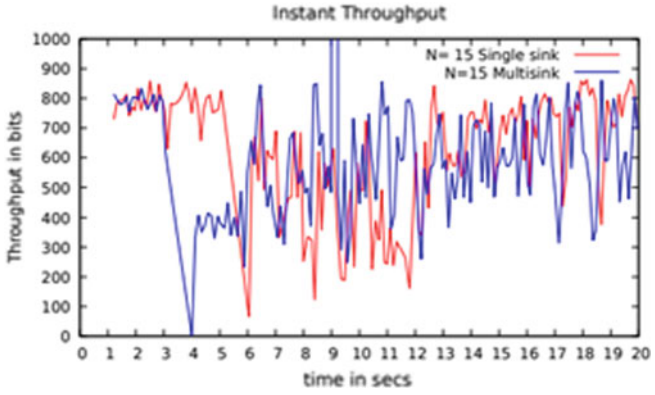


Fig. 24 Instant throughput for 15 nodes

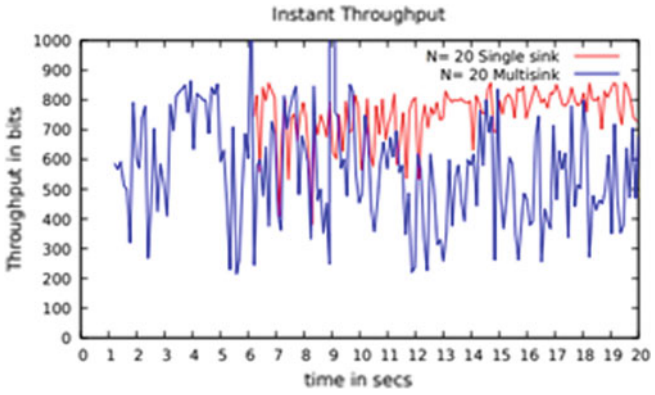


Fig. 25 Instant throughput for 20 nodes

5 Conclusion

In this paper, we have proposed multi-sink routing protocol for UWSNs. The experiments are conducted by simulation to evaluate performance metrics of the proposed protocol. For simulation as stated in scenario-1, only one sink is used on surface of water. As mentioned in scenario-2, the number of sinks is increased such that for each layer there is one sink. Comparison of performance metrics obtained in both scenarios shows that the proposed protocol improves % PDR and throughput for network of size 5, 10, 15 and 20. Energy tax is improved for network of size 10 and 20. The work carried out to compare performance metrics of static single sink architecture with static multi-sink architecture assumes that all nodes are stationary. As an extension to our research work, we evaluate the performance of the proposed protocol by giving mobility to nodes. Our next step is to study the effect on performance metrics of mobile sink architecture.

Reference

1. Xie, P., Cui, J.H., Lao, L.: VBF: vector-based forwarding protocol for underwater sensor networks. In: IFIP Networking 2006, Portugal, pp. 1216–1221 (2006)
2. Nicolaou, N., See, A., Xie, P., Cui, J., Maggiorini, D.: Improving the robustness of location–based routing for underwater sensor networks. In: Proceedings of IEEE OCEANS 2007 (2007)
3. Yifeng, J., Lin, S.: NIR: UWSN routing protocol based on node neighbour information. In: Proceedings of International Conference on Future Information Technology and Management Engineering, Changzhou, 2010, pp. 219–222
4. Wahid, A., Kim, D.: Connectivity-based routing protocol for underwater wireless sensor networks. In: Proceedings of ICTC 2012, pp. 589–590 (2012)
5. Ayaz, M., Abdullah, A., Low Tang, J.: Temporary cluster based routing for underwater wireless sensor networks. In: Proceedings of International Symposium in Information Technology (ITSim), 2010, pp. 1009–1014 (2010)
6. Jornet, J.M., Stojanovic, M., Zorzi, M.: Focused beam routing protocol for underwater acoustic networks. In: Proceedings of the Third ACM International Workshop on Underwater Networks, Series WuWNet'08, pp. 75–82. ACM, New York, NY, USA (2008)
7. Shin, D., Hwang, D., Kim, D.: DFR: an efficient directional flooding-based routing protocol in underwater sensor networks. In: OCEANS 2008, Quebec City, QC, 2008, pp. 1–7 (2008)
8. Yan, H., Shi, Z., Cui, J.H.: DBR: depth-based routing for underwater sensor networks. In: Proceedings of the IFIP Networking'08 Conference, Singapore, pp. 72–86 (2008)
9. Maryam Pouryazdanpanah K., Anjomshoa, M., Ahmad Salehi, S., Afroozeh, A., Marjan Moshfegh, G.: DS-VBF dual sink vector-based routing protocol for underwater wireless sensor network. In: Control and System Graduate Research Colloquium (ICSGRC), Shah Alam, 2014, pp. 227–232 (2014)
10. Ayaz, M., Abdullah, A.: Hop-by-hop dynamic addressing based (h2-dab) routing protocol for underwater wireless sensor networks. In: In: 2009 International Conference on Information and Multimedia Technology, Jeju Island, 2009, pp. 436–441(2009)

11. Wahid, A., Lee, S., Jeong, H.J., Kim, D.: EEDBR: Energy-efficient depth-based routing protocol for underwater wireless sensor networks. In: Advanced Computer Science and Information Technology. AST 2011 Communications in Computer and Information Science, vol. 195, pp. pp. 223–234. Springer, Berlin, Heidelberg (2011)
12. Wahid, A., Lee, S., Kim, D., et al.: MRP: a localization-free multi-layered routing protocol for underwater wireless sensor networks. Wireless Pers Commun **77**(2014), 2997–3012 (2014)

An Empirical Analysis of Fake Tweet Detection Using Statistical and Deep Learning Approaches

Vaishali Vaibhav Hirlekar and Arun Kumar

Abstract In today's era, fake news has become a rising threat. Fake news could be defined as news which involves deliberate lies or flaws, spread via traditional news media. The spread and propagation of such fake news poses significant threats to the individual as well as the national security. Moreover, the authenticity of social media article is never been verified frequently. A variety of approaches have been proposed so far with respect to fake news recognition, to avoid the victims of deceptive social media. This paper presented a data-driven perspective using natural language processing method, machine learning and deep learning approaches to detect fake tweets. We have analyzed the performance of various statistical classifier algorithms like logistics regression, random forest and Naïve Bayes with different performance features like count vectorizer and tf-idf. Also, we analyzed the performance of long short-term memory (LSTM) model, convolutional neural network and hybrid approach to improve the overall accuracy.

Keywords Fake news · Natural language processing · Feature extraction · Machine learning · Deep learning

1 Introduction

Fake news is nothing but the content generated to deceive the people. As we know, the Internet is a major information source, but now, it has become a Web of deception. Online social media are becoming popular communication networks to spread deceived information. There are numerous reasons why fake news is created; reason could be financial, political or related to business. The advancement of social networking has made all users instantaneous, no editing, no quality checking, and

V. V. Hirlekar (✉)
Sir Padmapat Singhania University, Udaipur, Rajasthan, India
e-mail: vaishali.hirlekar@spsu.ac.in

A. Kumar
Ir Padmapat Singhania University, Udaipur, Rajasthan, India
e-mail: arun.kumar@spsu.ac.in

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_95

obviously with no liability. The facts are getting presented without authority, and it itself becomes a certificate of truthfulness for millions of users who watch them on their site. This intentional disinformation is used to distract a person from true reality. Unfortunately, online social networks like Facebook, Twitter, Whatsapp and Instagram are being used to spread misinformation, deceit, abuse and fraud by people. Researches have also shown that the key sources of misinformation have always been concerns such as health, politics, finance and technology trends.

In manual fake news detection, general approach is to identify whether the news is from an actual or an alleged source, the relevance of the news matter and headline. Credibility of the author, publishing date of that article is also checked. Most importantly, the weightage is given to the news if the same information found on a variety of authentic sources. However, challenges are different in developing an automated system for fake news detection on any online social media. As online social media enables propagation of news in real time. This news is in text format so cleaning and translating textual data to machine readable format is an important task. There are various approaches used to simplify this procedure by means of machine learning and natural language processing (NLPs) techniques. However, due to the semantic nature of the content, automated approaches are limited in accuracy, and manual intervention is often needed in this process. Another most important thing is, online social networks produce vast volumes of data so the process becomes computationally costly to execute in real time.

2 Background and Motivation

In time of life-threatening pandemic of COVID, many would not have considered fake news to be a big threat. Nevertheless, this has become business as usual all over the world. There is a surge of disinformation around the globe about the pandemic that is emerging. But with varying severity, virtually all nations were affected by fake news about the virus' roots and its resulting proliferation and risks.

At present, the planet has confronted two viruses simultaneously: One is real and the other is artificial, but still dangerous. Fake news and disinformation have created many challenges in the fight against the pandemic at all levels of government.

2.1 Distorted News then and Now

If we look into history, distorted news was not a problem in society just few years ago. There was common concern about such news from a long time back. The scientific group began to concentrate its attention on the issue of fake information in the beginning of the decade which has started in the form of review spam, Web page spam, rumor detection and hidden paid posters. However, the term fake news coined by the 2016 US Election primarily. Among other counterfeit news that appeared

during the Trump campaign, "Pope Francis reported endorsement of Donald Trump for president of the US" was the most popular one. During that time, even Facebook itself was steadily profoundly affected because it makes the dissemination of false news coming from unchecked Web sites.

Along with this, we can observe that many of the companies/publisher are using Clickbait's techniques rather than phishing's techniques just to raise their advertising profits. Clickbait is something which is tricking people into consuming your content in such a way that they make you believe that it will be better than what it really is. And such fake information may have a detrimental effect on both safeties of individuals as well as national security too.

2.2 Challenges in Fake Tweet Detection

- The relation between news topics and credibility of the news creator infers the outcome of fake tweet.
- Checking the credibility of the content can improve the efficiency of fake tweet detection.
- Tweets are basically short text and text analysis is traditionally performed with natural language techniques. NLP tools alone may not be able to deal with tweet properly.
- Models must rapidly process the tweets as most of the time, news are time-sensitive.
- Many researchers have recommended mechanisms for identifying fake news from crowd channels in social networks. The expertise of the crowd has not been used effectively so far in alliances to make decisions.
- Many of the fake tweet datasets we see online are based on incidents like the 2016 US elections. This is an incredibly particular case, and if used solely, it may lead to significant bias in the model. Diversification of the domain and time periods is necessary
- Developments in deep learning can potentially benefit fake tweet research. However, the challenge is deep learning techniques often requires massive training data and time for model training.

3 A Review of Some Relevant Cases

Bing Liu et al. have discussed in the opinion spam and analysis about opinion spam which is basically a review spam. As we know, any potential customers use reviews to find feedback from the existing users before buying a product. Also, reviews are often used by the manufacturer to detect quality issues in their product and/or to gain the information about their competitors' product [1]. Due to the fact that no check is done on the comments/reviews written, it may happen that fake reviews can also get

submitted which in turn results in review spam. Here, feature identification concept has been used by the researcher. This feature identification is done based on, deliberately misleading readers, brands reviews and other unrelated reviews containing random texts. Author used SVM, Naïve Bayes and linear regression model and found accuracy of 72%, 75% and 78%, respectively. In order to avoid overfitting, the reviewer omitted all product price and product rating functionality for this study.

Michael et al. discussed about an online monitoring system which was capable of performing Twitter stream trend identification [2]. The framework detects real-time evolving trend on Twitter and offers substantive commentary that synthesizes a comprehensive overview of each topic. Twitter Monitor used context algorithms for extraction such as PCA, SVD for experimentation. Finally, the user communicates with the system to rate and to apply their own descriptions on the system according to the shown patterns. Twitter Monitor handles bursty keywords for pattern detection. However, there is no spam traffic detection mechanism analyzed in this paper. Carlos et al. [3] assumed that trend detection system exists (like Twitter Monitor) and emphasize on the marking of those trends or occurrences and proposed a model in the paper titled information credibility on Twitter that checks the credibility of news on Twitter and classifies them based on features extracted. This model has used retweeting behavior. The author has used 10,000 tweets dataset and proposed a set of features based on propagation, user-level, topic level and message text for experiment and used j48 decision tree with accuracy of 78.7%. However, the impact of context that is displayed in Twitter like followers count not included in the analysis.

Vahed Qazvinian et al. had proposed a framework for finding tweets that could be rumors [4]. For rumor identification, author has used different features based on tweet, network pattern and content and achieved 84% accuracy. Different attributes such as lexical trends, speaking trends, unigrams and di-grams features, re-tweet count, hashtags and URLs have been considered here to improve the overall accuracy of the model. Author has used manually annotated dataset of 10,000 tweets from five different controversial topics.

Cheng Chen et al. have proposed a methodology to classify a special community of Internet users, known as hidden paid posters or known as the Internet water army in China [5]. The Internet provides a different kind of online work openings with paying posters. They are paid for publishing comments or posts for secret reasons in numerous online forums and blogs. Author used a classification scheme to identify any user with the usage of the features as a possible paying poster or regular user. The outcomes of the five response characteristics are combined and used, i.e., average interval of posing time, days on which behavior happens, successful reports and degree of similarity. Used a support vector machine (SVM) with 88.79% accuracy. In this analysis, the researchers gathered news stories and related comments from two popular Chinese news pages surrounding this special social event: Sina.com and Sohu.com. However, this limited dataset may lead to biased output.

Krishna Kumar et al. [6] proposed disinformation identification model by cognitive science in online social networks. In this article, re-tweet activity decides the reliability of sources. In this model, the general acceptability score of the tweet was determined using the PageRank algorithm in the re-tweet graph [6]. However, this

approach considers only the re-tweet feature and PageRank algorithm used in this technique is prone to manipulation. Zhiwei et al. [7] proposed a three-layer hierarchical reputation network, composed of posts, sub-events and event data sets, from Sina Weibo, the leading Chinese micro-blog [7]. Researcher used support vector machine algorithm and achieved 88 percent accuracy with dataset 1 and 85% with dataset 2. In order to verify the feasibility of the proposed model, only two datasets on the micro-blog are obtained on the same subject: One with random bogus news a year and real news at the same time; the other with both false and factual news linked to the same subject.

Arushi Gupta et al. proposed a technique to classify an account as "spammer" or "non-spammer" [8]. Researcher has made use of databases composed of 1064 Twitter accounts and 62 user and tweet-related attributes. Three distinct algorithms for research—the Naive Bayes, clustering and decision trees have been used, but better precision achieved in clustering which was 87.9%. As a future work, author suggested the integrated solution can be further enhanced by retaining the high precision of the decision trees method to the identification of spammer accounts.

Inggrid Yanuar et al. proposed a study of hoax news detection in Indonesian language using Naïve Bayes classifier [9]. This study uses its own dataset on 250 pages of hoax and true news stories. Three reviewers have done a manual classification of the dataset. Researcher used tf-idf and Naïve Bayes classifier and achieved accuracy of 78.6%. Pyae and Nwed [10] proposed a framework for satire detection using emotional features [10]. Researcher has used news article (676 real-226 satire), Amazon product review (1000 satire-1000 serious) and tweets corpus about 20,000–30,000 tweets for each non-satire and satire distribution. The approach has used bigram and trigram BoW, tf-idf for text classification used SVM and ensemble bagging classifier and achieved accuracy of 66%. Researcher concluded that BoW worked better with long text and tf-idf works better in short text, i.e., with tweets.

4 Methodology Used for Fake Tweet Detection

The general approach used for fake tweet detection using basic classifier and neural network is shown in Fig. 1. In the initial phase of this approach, it would gather the data, execute the various text operations for pre-processing, such as text cleaning, tokenizing, stop word handling and punctuation, to eliminate the noise in the data. All of these pre-processing steps aim to reduce the size of the vocabulary without eliminating any relevant information. The smaller the vocabulary, the lower the complexity of the memory, and will use more robust the parameters for the estimated words. In the next phase, it would extract the characteristics and use various classifier and neural network models to achieve greater accuracy. The flow of the entire process will be as follows.

- Perform data cleaning which is used to remove noise in the data.

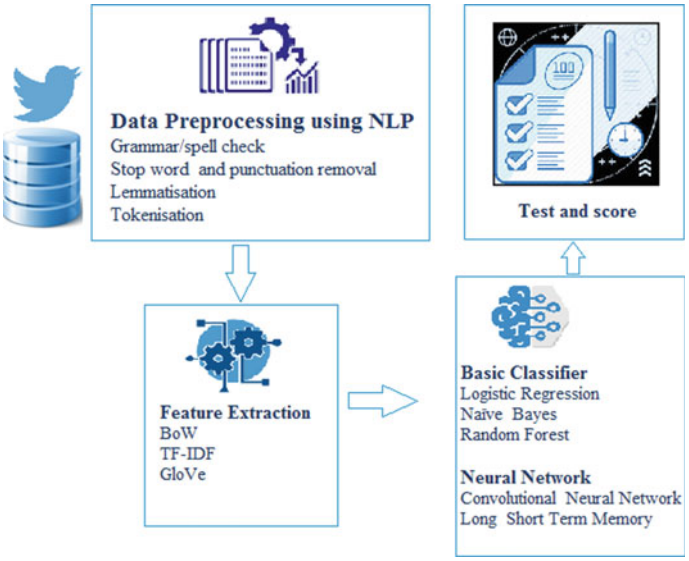


Fig. 1 General framework for fake tweet detection with basic classifiers and neural network approach

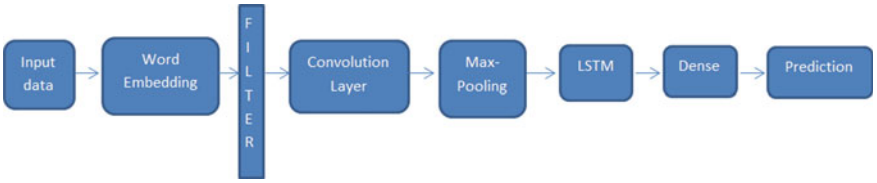


Fig. 2 CNN-LSTM model

- Text in the tweet is tokenized into tokens. It separates tokens based on white space and punctuation. Convert the tokens into lowercase. Remove punctuation from each token.
- Filter out remaining tokens that are not alphabetic.
- Filter out tokens that are stop words.
- Perform lemmatization to consider the context of the word and convert the word to its meaningful base form. To analyze and model the data further, it has to be converted into feature vector. Techniques include bag of words model (BoW), tf-idf and Glove model.
- Check the accuracy with basic classifier models such Naïve Bayes, logistic regression and random forest.
- Also check the accuracy with deep learning model like long short-term memory and convolution neural network.

4.1 Using Deep Learning Network Approach

Deep learning networks are recent development based on tensor variant to solve the discriminations through learning on scalable generative components. Recurrent neural network (RNN) is used to generate output based on the prediction of input data. In order to train RNN model, pre-segmentation such as tokenization of the training data need to be done along with data cleaning [11]. RNN captures the sequential information present in the input data which improves the results.

In convolutional neural network, first layers embed words into low-dimensional vectors. The next layer performs convolutions over the embedded word vectors using multiple filter sizes. For example, sliding over 3, 4 or 5 words at a time. Next, we can max pool the result of the convolutional layer into a long feature vector, add dropout regularization and classify the result using a softmax layer.

The CNN-LSTM architecture involves using convolutional neural network (CNN) layers for feature extraction on input data combined with LSTMs to support sequence prediction. This architecture consists of two sub-models: the CNN model for feature extraction and the LSTM model for interpreting the features across time steps (Fig. 2).

In the CNN-LSTM model, first layer is a convolutional layer. A convolution is a filter that can extract features from the data. Next, max pooling layer iterates over the tensors and takes the highest value. The next layer is a long short-term memory (LSTM) layer which is composed of a cell, an input gate, an output gate and a forget gate. The cell remembers value over arbitrary time intervals and the three gates regulate the flow of information into and out of the cell. The last layer before the prediction is made is a fully connected layer. This is a regular neural network layer where all nodes are connected with each other.

5 Dataset Analysis

We have used dataset from Kaggle (https://www.kaggle.com/smid80/coronavirus-covid19-tweets-early-april) related to coronavirus (COVID-19) tweets. This dataset includes Twitter-related variables: the text of some tweets and the tweeting identities, the hashtags that were used and the account locations. We have performed the analysis on number of fake and real tweets done during first 15 days of April and it shows that compared to real tweets, the number of fake tweets is much higher. Figure 3 shows the COVID tweet analysis done during first 15 days of April.

6 Results and Discussion

Table 1 shows the experimental results with statistical classifiers on COVID-19 tweet dataset. Count vectorizer and tf-idf have been used with logistic regression, random

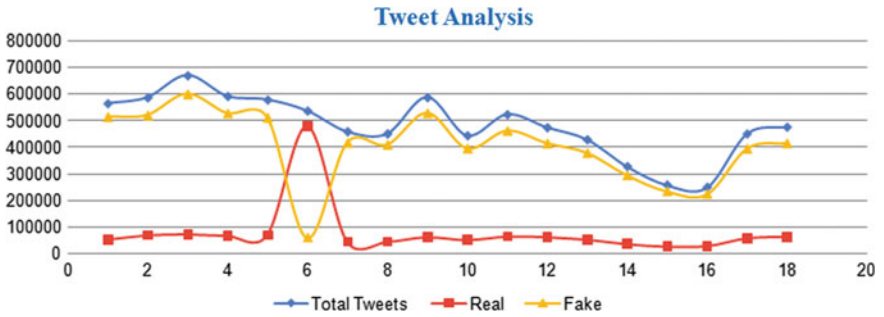


Fig. 3 COVID tweet analysis done during first 15 days of April

Table 1 Experimental results with statistical classifier

Algorithm	Precision	Recall	Accuracy
Logistic regression using count vectorizer	0.92	0.92	0.924
Logistic regression with TF-IDF	0.93	0.93	0.925
Random forest with count vectorizer	0.92	0.93	0.926
Random forest with tf_idf	0.92	0.93	0.928
Naïve Bayes using count vectorizer	0.92	0.92	0.922
Naïve Bayes using tf-idf	0.93	0.92	0.923

forest and Naïve Bayes classifiers with maximum feature of 1000. The accuracy achieved by the classifiers is around 92%. However, random forest with tf-idf has outperformed with 92.8% accuracy among all the statistical classifier. In Table 2, bidirectional LSTM has used with one hot encoding method. Glove vectors have been used with different neural network approaches such as convolutional neural network and CNN-LSTM model. However, a hybrid combination of CNN with LSTM has outperformed here with 94.2% accuracy. This experiment has been carried out on 313,036 tweets from the dataset with total 14 parameters.

Table 2 Experimental results with neural network

Algorithm	Precision	Recall	Accuracy
Bidirectional LSTM	0.89	0.90	0.896
Simple CNN model	0.92	0.92	0.929
Modified CNN model	0.93	0.93	0.934
CNN-LSTM model	0.94	0.94	0.942

6.1 Bidirectional LSTM with One Hot Encoding

A bidirectional LSTM consists of a forward and a backward recurrent neural network and final prediction is made combining the results of both the networks at any given time t. One hot encoding length of the word vector is equal to the length of the vocabulary, and each observation is represented by a matrix with rows equal to the length of vocabulary and columns equal to the length of observation, with a value of 1 where the word of vocabulary is present in the observation and a value of zero where it is not. In this model, we have used embedding layers with voc_size = 5000, embedding vector feature = 40 and input length = 20 as parameters. Bi-LSTM layer is used followed by dense layer with sigmoid activation and achieved accuracy of 0.896.

6.2 Simple CNN Model

In this experiment, we have used simple CNN method with three convolution layers (filter size = 5) and applied ReLu activation function to the matrix. Then performed pooling operation to reduce the dimensionality using max pooling layer. Flattened the output and feed into a fully connected layer and finally used softmax activation function to classify the output. Using this model, correct predictions are 145,509 and a total number of test examples are 156,518 with the 0.9297 accuracy.

6.3 Modified CNN Model

In this experiment, first layers embed words into low-dimensional vectors. The next layer performs convolutions over the embedded word vectors using multiple filter sizes, sliding over 3, 4 or 5 words at a time. Next, we max pool the result of the convolutional layer into a long feature vector and classified the result using a softmax layer. The convolutional layers followed by max pooling. Remember that we use filters of different sizes. Because each convolution produces tensors of different shapes, we need to iterate through them, create a layer for each of them and then merge the results into one big feature vector. Using this model, correct predictions are 146,317. Total number of test samples is 156,518 and achieved 0.9348 accuracy with this model. Size of train, validation and test is as follows: 250,428 156,518 156,518 (Figs. 4 and 5).

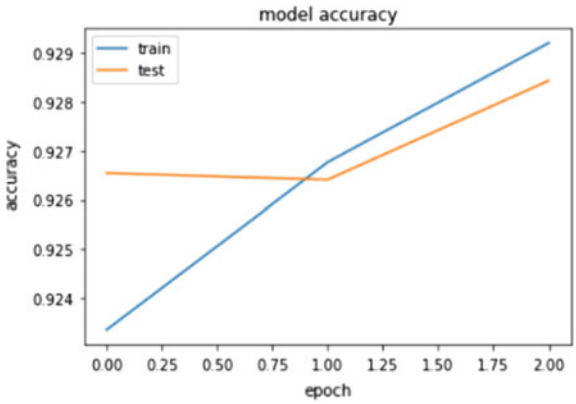


Fig. 4 Simple CNN model accuracy

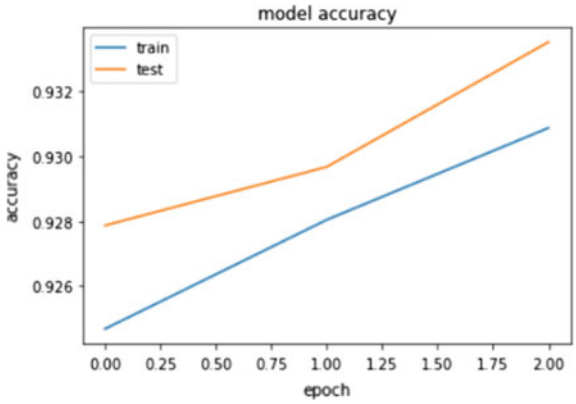


Fig. 5 CNN model 2 accuracy

6.4 Hybrid Approach (CNN with LSTM)

In this experiment, hybrid architecture comprises of two convolution layers followed by max pooling layer. In the first convolution layer, we have used 32 filters of kernel_size 5, whereas 64 filter used of kernel_size 3 in the next convolution layer and the ReLu activation function. LSTM layer is followed by batch normalization. The last layer before the prediction is made is a fully connected layer.

Size of train, validation, test: 250,428 31,304 31,304 (Figs. 6 and 7).

7 Conclusion

This paper presented a methodology for fake tweet detection using different word embedding method such as bag of word and tf-idf models with statistical classifiers.

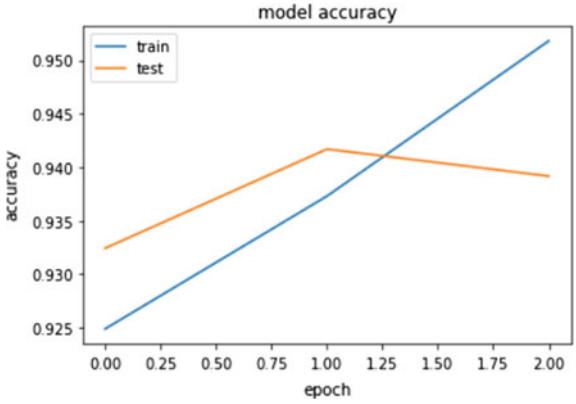


Fig. 6 CNN with LSTM model accuracy

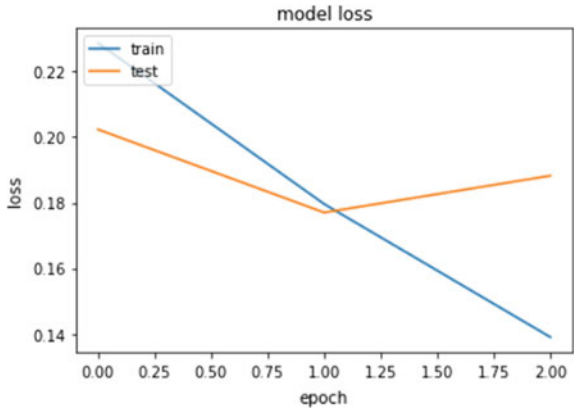


Fig. 7 CNN with LSTM model loss

It carries out pre-processing steps involving cleaning of data, removing the stop word and tokenizing sentences using NLTK toolkit. In the presented work, logistic regression, random forest and Naïve Bayes achieved 92.5%, 92.8% and 92.3% accuracy, respectively. We have also evaluated the model with neural network such as Bi-LSTM with one hot representation method, convolutional neural network and CNN-LSTM using Glove vector and achieved 89.6%, 93.4%, 94.2%, respectively. The experimental results show that CNN- LSTM model has outperformed in the given case. In future, we would evaluate the credibility of the tweet by checking metadata such as account statistics, media content, text and linguistic characteristics.

References

1. Jindal, N., Liu, B.: Opinion spam and analysis. In: Proceedings of ACM International Conference on Web Search and Data Mining, New York, 11–12 February 11–12

2. Michael, M., Nick, K.: TwitterMonitor: trend detection over the twitter stream. In: Proceedings of ACM International Conference on Management of Data, USA, 6–11 June 6–11, p. 1155
3. Carlos, C., Marcelo, M., Barbara, P.: Information credibility on twitter. In: Proceedings of International World Wide Web Conference Committee (IW3C2), Hyderabad. 28 March –1 April, p. 675 (2011)
4. Vahed, Q., Emily, R., Dragomir, R.R., Qiaozhu, M.: Rumor has it: identifying misinformation in microblogs. In: Proceedings of Conference on Empirical Methods in Natural Language Processing, Scotland, 27–31 July, p. 1589 (2011)
5. Cheng, C., Kui, W., Venkatesh, S., Xudong, Z.: Battling the internet water army: detection of hidden paid posters. In: Proceedings of IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, Canada, 25–28 August, p. 116 (2013)
6. Krishna Kumar, K.P., Geethakumari, G.: Detecting misinformation in online social networks using cognitive psychology. Springer Open J. Hum.-Centric Comput. Inf. Sci. **4**(1), 14
7. Zhiwei, J., Juan, C., Yu-Gang, J., Yongdong Z.: News Credibility evaluation on microblog with a hierarchical propagation model 2014. In: Proceedings of IEEE International Conference on Data Mining, China, 14–17 December, p. 230 (2014)
8. Arushi, G., Rishabh, K.: Improving spam detection in online social networks. In: Proceedings of International Conference on Cognitive Computing and Information Processing (CCIP), India, 3–4 March (2015)
9. Inggrid, Y.R.P., Rosa, A.A., Faisal, R., Study of hoax news detection using naïve Bayes classifier in Indonesian language. In: Proceedings of 11th International Conference on Information & Communication Technology and System (ICTS), Indonesia, 31 October, p. 73 (2017)
10. Pyae, P.T., Nwe, N.: Implementation of emotional features on satire detection. In: Proceedings of IEEE/ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD), Japan, 26–28 June, p. 149 (2017)
11. Smys, S., Chen, J.I.Z., Shakya, S.: Survey on neural network architectures with deep learning. J. Soft Comput. Parad. (JSCP) **2**(03), 186–194 (2020)

Cybersecurity During COVID-19

Sheeba Uruj Ahmad, Sunidhi Kashyap, Sujala D. Shetty, and Neeru Sood

Abstract The world today has become fully digitalized. The increase of digitalization has provided a huge platform for the cybercriminals to launch their attacks. With the outbreak of coronavirus (COVID-19) followed by lockdowns, digitalization has been the only way to stay connected virtually. With the growing technology every day, the attacks become more and more sophisticated as the attackers quickly learn and invent new ways to find their prey. Every sector whether public or private has been affected due to breaches in cybersecurity especially during the COVID-19. In this paper, we have tried to analyze various types of cyberattacks carried out during the pandemic with its impact and preventive measures to be taken for protection.

Keywords Cybersecurity · Coronavirus · COVID-19 · Digitalization

1 Introduction

COVID-19 has offered the best time for cybercriminals to attack. Identified in Wuhan, Hubei in China on December 9, 2019, the coronavirus was classified as pandemic on March 11, 2020 [1]. The outbreak of this deadly virus brought hysteria and fear in the whole world. Government of various countries declared lockdowns which brought devastating effects on various sectors including healthcare, education, energy and even households. This led to unemployment, shutdown of many companies and

S. U. Ahmad (✉) · S. Kashyap · S. D. Shetty · N. Sood
Department of Computer Science, Birla Institute of Technology & Science, Pilani, India
e-mail: sheeba@dubai.bits-pilani.ac.in

Dubai International Academic City, Dubai, UAE

S. Kashyap
e-mail: f20180191@dubai.bits-pilani.ac.in

S. D. Shetty
e-mail: sujala@dubai.bits-pilani.ac.in

N. Sood
e-mail: sood@dubai.bits-pilani.ac.in

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_96

businesses, poverty and a rapid increase of technology dependence, even for the people who had no basic computing skills. Not only for work, but also for the very basic daily activities like news, communication, entertainment and even shopping, the general public had to rely on technology to get their tasks done virtually.

The fear of pandemic raised curiosity of humans to study and get as much information as possible from any source provided and also to find its solution. It is this human curiosity which has offered a pool of opportunities to the cyberattackers to take complete advantage of the situation and play with the human psyche. From malicious URLs to spreading of misinformation, the cyber world has seen an immense influx of attacks. This time has served as the best time for the cyberattackers to plan their attacks and evade. Cyberattack seemed to be an easy way to earn for those who were unemployed or had shortage of money. In security systems, 'human' is the most essential key. Cyberattackers prey on general public having little or no knowledge of cybersecurity.

Social engineering attacks are at its peak as more and more people want to gain information to cure the disease but their lack of knowledge about the security systems only makes them a prey to the attackers with no valuable information in hand. A number of websites offering masks, sanitizers and gloves are nothing but bogus where the customers pay the ransom for the products but these are never delivered at their doorstep [1]. Phishing attacks and ransomware attacks have increased tremendously during this time period, and a lot of important and sensitive data have already been compromised. Not only the general public, but the hackers even targeted famous personalities like Barack Obama and Bill Gates by hacking their Twitter accounts and earned around $117,000 as bitcoins. Surprisingly, the hacker was just a seventeen-year-old boy from Florida who was arrested later, but his attack was completely online where he could succeed just by social engineering [2].

As of today, no sector is safe from cyberattacks. Businesses, educational sectors, healthcare sectors, financial sectors, energy sectors and even the general household have been affected due to some attack or the other on their systems. Data is money today, so stealing this immense amount of data from various sectors by applying numerous techniques can prove to be hazardous for the economy of every sector [3].

2 Literature Review

The outbreak of coronavirus has brought about a number of changes everywhere. It has forced public to follow a very different lifestyle as compared to what they have been living. This includes social distancing, wearing masks and using sanitizers all the time. Socializing is highly discouraged, and hence, any events held whether big or small are all cancelled due to this. In short, the era of COVID-19 has been forcing people to isolation as much as they can and maintaining more cleanliness than ever before. So to maintain isolation, working virtually has been the only option to get the work done.

Advancements in technology have made our lives very easy but at the same time has offered a huge platform for the digital evaders to study the same technology but use it in a destructive way. Cybersecurity has gained a lot of importance in recent years as organizations are getting more inclined towards digitalization and automation. The cyber world has seen numerous attacks of different types meant solely to gain unauthorized access and make money. The various types of cyberattacks are listed below:

- **Malware attacks**: Malware is nothing but a piece of malicious software designed to steal, modify or delete data on the systems. Usually, it enters the system by clicking a malicious link sent via email or available on fake websites. It has various forms like spyware, trojans, scareware and so on whose main objective is to enter the system and disrupt its activities. To a user, it looks very genuine but if observed carefully, they can be easily recognized. During the COVID era, these malwares have been spreading through embedded and interactive corona maps and websites [1, 3].
- **Ransomware attacks**: Ransomware, as the name suggests, is designed to gain ransom from the victimized user. It also spreads by clicking on malicious links, email attachments and even through compromised credentials of employees. Once it enters the system, the ransomware locks all the files having sensitive and critical data on the system and asks the user for a ransom amount in order to decrypt the data. These kind of attacks have increased on a large scale during this pandemic and affected almost all sectors including healthcare and educational institutions [3].
- **Phishing attacks**: Phishing attacks aim at stealing user's credentials, his bank details or any other such confidential information by disguising as a legitimate organization. This is a type of social engineering attack where the user is fooled by the attacker in an attempt to gain useful information like coronavirus or tax refunds. Phishing attacks have risen massively during these few months of corona attack by scam emails and fake websites [1, 4].
- **DDoS attacks**: DDoS attacks or Distributed Denial-of-Service attacks have also risen and have mainly affected the healthcare division and government sector. These attacks usually slow down the system responses or crash them by flooding the servers with huge numbers of unnecessary or bogus requests, hence causing delays in responding or not replying at all to legitimate requests. This can prove fatal in organizations where immediate availability of data plays an important role [3].
- **Malicious Domains**: Since the outbreak of coronavirus, a massive number of fake and malicious domains have been registered. A slight error in spelling such spelling 'coronovirus' instead of 'coronavirus' is enough to alter a legitimate website into a malicious site. Such websites are created every day for the purpose of scamming, phishing, spreading malware and so on. Fake COVID maps [1, 3].
- **Spamming attacks**: Spamming has been one of the oldest techniques to fool general users where the spammer disguises himself as someone from a legitimate organization or company so that the victim reveals his payment details or opens

some malicious link sent as attachment. Numerous cases of spamming have been reported during this pandemic where the spammers pretend to a member of WHO and ask for funds or send payment invoices. Not only emails, but spamming through WhatsApp, telephone calls, faxes and websites have seen an alarming increase recently [3].

- **Malicious Social Media Messaging**: Social media messaging has become really popular in the last decade. Whatsapp, Facebook, Twitter, Instagram and many such platforms are available today to help people stay connected with their loved ones and also to maintain professional ties virtually. Apart from offering so many advantages, it has also offered a lot of opportunity in the cyber world to design and attack the general public using these platforms. Massive influx of scams and phishing has been observed on these platforms. Alluring the users to get free subscriptions for apps like Netflix and capture their payment details is one of the examples observed as the world has become more dependent on technology [3].
- **Spreading Misinformation**: Misinformation is one of the biggest enemies today where the world seeks right knowledge about the pandemic and its cure. But unfortunately fake news articles are so common these days that it has become really difficult to make out the difference between what is right and what is wrong. Numerous fake websites created require the user to fill in their personal data to get some useful information which actually does not even exist. Emails with attachment like 'Safety Measures' is one of the common ways today to attract users to click on malicious attachments and links. Spreading misinformation is not only harmful for the general people but also affects the reputation of the country and its government which is why many a times interference of government was required to bring a stop to such spreads [1].

3 Analysis

Cyber threats and cybercrimes have always been in picture ever since the world started to move toward digitalization. But the pandemic due to COVID-19 has caused a tremendous increase in the number of cybercrimes. A huge number of individuals lost their jobs as a result of various restrictions put by the government to save economy and also for security purposes. This has encouraged more number of people to become cybercriminals so as to earn a living and survive. Also, in some countries like Nigeria, there are issues of power supply and Internet connectivity because of which the work sector has to depend on free Wi-Fi provided in public places. This can prove fatal for confidentiality [5].

About 4000 domains have been registered in 2020 solely for the purpose of cybercrimes [4]. The recent reports from FBI state that there are around 4000 complaints in the cyber division every day [2]. The damage costs during this pandemic might double due to the phishing attacks, ransomware attacks, insecure remote access to corporate networks, exposure of login credentials and access of confidential data to family

members and visitors by remote workers [4]. According to founder of 'Monstercloud,' the ransomware attacks have increased by 800% during the pandemic. Furthermore, these attacks are now converted to '*doxware*' wherein if the victim refuses to pay the ransom amount, then they are threatened about their data being leaked and their customers being notified about the same [2].

As per the reports by Cybersecurity Ventures, the cybercrimes will cost around $6 trillion by 2021 which will be the greatest transfer of wealth ever [4]. Interpol is also seeing an alarming increase in the number of cyberattacks especially aimed at big corporations, government organizations and critical infrastructures. Microsoft alone has reports about 20,000–30,000 phishing attacks and social engineering attacks a day only in USA. Corona research centers have been a major target for the hackers. In July 2020, the cyber officials of UK, USA and Canada had issues a warning about a hacker team from Russia who tried to hack and steal the vaccine [2].

Cyberattacks are more successful than ever due to compromised emotional states of humans at this time. A large number of malicious apps were launched for mobiles to take advantage of this opportunity. 'COVIDLock' an app that claimed to track COVID cases was actually a ransomware which locks the user's phone once installed and to unlock they are required to pay a ransom amount of 100 bitcoins within 48 h. If the victim failed to do so, then they are threatened about their data being deleted and leaked online on social media. There are also some Android apps which offer selling masks and safety kits but in reality they are SMSTrojans which once installed, collect contact list of victim's phone and send out SMS to spread itself [3]. Such apps look very legitimate to the user by their name, for example '*Corona Safety Mask*' and once installed they ask for user's permission to read contacts and send SMS messages. At this point, user needs to understand the red flag and immediately uninstall the app. Otherwise, it spreads itself by sending messages to all the contacts [6]. Apart from creating malicious apps, there are also certain webpages created for the purpose of donating to aid researchers. This idea takes advantage of the goodness of people and lures them to pay ransom amount in the form of cryptocurrency. In February 2020, about $2 million dollars were stolen in USA using this scam. The attackers ask only for bitcoins as it is impossible to trace where the money finally goes [1].

The COVID-19 pandemic has seen a huge influx of cyber threats and attacks in different forms whether involving social engineering or some sophisticated techniques. Here are a few of the very recent attacks listed below:

- Twitter attack of famous celebrities, companies and politicians, like Apple, Bill Gates, Kayne West and Barack Obama, where general public was conned to send bitcoins.
- Numerous malicious filenames have been created to trick individuals and steal data. Some of them like 'EESkiri-COVID-19′ may be a key-logger disguised as a COVID-19 help site [4].
- New Zealand Stock Exchange (NZX) has been hacked multiple times using DDoS attacks from overseas. This resulted in website crashing and also halting its operations. Such attacks have also affected many other companies including PayPal,

Yes Bank India and Brain Tree where the victims had to pay a hefty amount in bitcoins.

- There has also been a considerable increase in hacking of Linux systems. Several examples include 'Lemon Duck'—a cryptomining malware, 'Lucifer'—a hybrid DDoS botnet, 'Fritzfrog'—a botnet campaign, 'Drovorub'—a malware to exfiltrate sensitive data and many more.
- 'Zeppelin Ransomware' was a threat for health sectors which remained undetected by anti-virus apps. This is due to the use of a new Trojan downloader hidden in garbage text of VB scripts. Other such ransomwares that affected the health sectors include 'REvil', 'Maze' and 'Netwalker' that stole sensitive data.
- A security flaw in Bluetooth technology gave rise to a new attack called 'BLUR-tooth' which can overwrite authentication keys to grant unauthorized access. Other Bluetooth attacks include 'Spectra,' 'Bluetooth Impersonation Attacks (BIAS)' and Bluetooth Reconnection Flaw.
- The education sector has also been affected by cyberattacks during this pandemic. A huge number of 'Zoombombing' cases came up where the intruders could join in any session as guest and disrupt online classes. Many students and teachers have been fooled by installing a bogus version of Zoom, thus allowing malicious software to enter their systems. The Newcastle University was targeted by the 'DoppelPaymer' ransomware gang [7].

Figure 1 shows the distribution of the key COVID-19 inflicted cyber threats and the increase in percentage of the mentioned cyberattacks. The number of malicious domains increased by 22%, whereas ransomware increased by 36%. A steep increase of 59% was seen in phishing/fraud. Cyberthreats inflicted from fake news had a 14% increase.

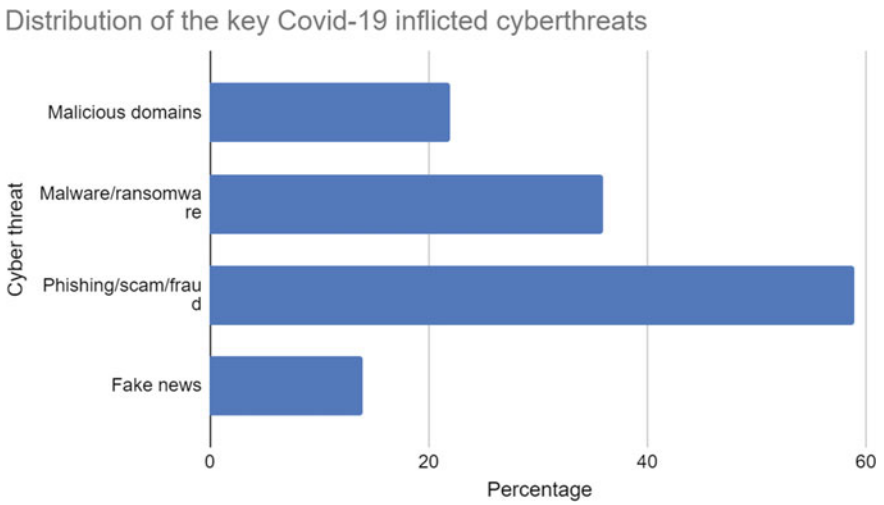


Fig. 1 Graph showing distribution of key cyberthreats [8]

Out of all the cyberthreats, ransomware continues to top the list. These attacks mainly target the government sectors, educational sectors and healthcare sectors. According to McAfee, ransomware grew 56% in the past four quarters. The most recent ransomware attack was reported by Microsoft for Android users where the latest variant of ransomware could simply block access to the device by displaying a screen which shows a ransom note providing instructions to pay the ransom. This malware, AndroidOS/MalLocker.B, uses Android features and incorporates a machine learning module which can bypass the anti-virus software too [9].

Healthcare sector has been most affected by ransomware attacks [8]. Payment of the ransom is usually done in the form of bitcoins as it is the best way to not trace where the money finally ends up. The diagram below depicts the same [10] (Fig. 2).

There are various ways in which the cyber criminals launch their attacks on their victims. Figure 3 depicts the percentage of different causes leading to a cyberattack. 85% of the people posting puppy pictures were scammers, 43% of data breaches are from cloud-based web applications. Cyberattacks starting with an email are 75% and external factors contribute to 70%. Stolen or weak credentials and phishing contribute to 37% and 25%, respectively. Ransomware is found in 27% of malware incidents—up from 24% in 2019. Human error accounts for 22% and organized crime gangs account for 55% of attacks. 80% of hacking breaches involve brute force or stolen credentials [11].

Apart from healthcare, education and government sectors, there have been other sectors which have suffered major financial losses due to the cybercrimes. Stock markets have fallen drastically in the past 30 years. Crude oil prices have also gone

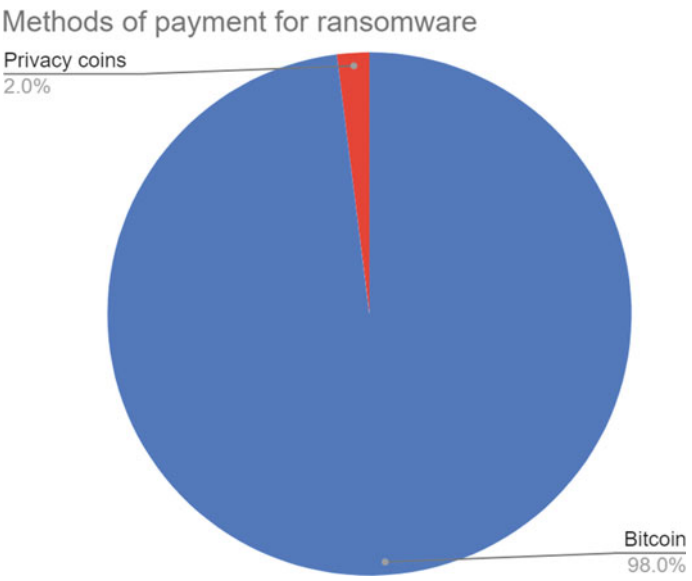


Fig. 2 Diagram showing payment methods for ransomware [10]

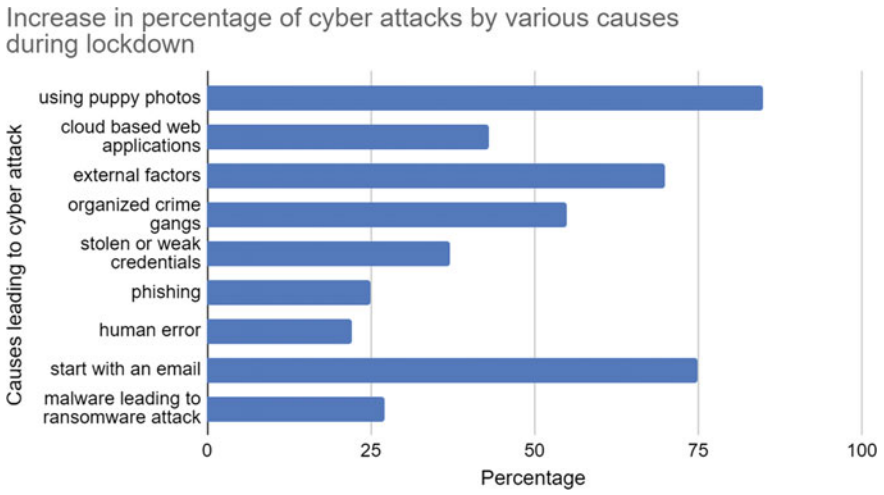


Fig. 3 Graph showing percentage increase on various causes of cyberattacks [11]

to the lowest since 1991. Fintech users are vulnerable to social engineering where hackers use their tricks to show that they are legit and gain personal information.

The digital world has advanced tremendously within a few years and so have the attacks. As the organizations today are shifting to cloud for efficiency and productivity, attacking and stealing information from cloud has become a major target for the attackers. With the advent of Internet of Things (IoT), cloud has found its way even through the households. The latest technologies involving artificial intelligence and machine learning have created a new horizon for the cybercriminals to attack their victims in more sophisticated ways. Latest attacks based on the newest technologies are as discussed below:

- **Cloud Jacking**: As the world shifts more toward cloud computing, attacking cloud will be the next target for the cyberattackers to victimize the users and steal or modify their sensitive and critical information. The cloud jacking may be carried out by injecting malicious code in the cloud either by SQL injection or by cross site scripting. Considering the security aspects of cloud, when an organization shifts to public cloud, the cloud provider has the responsibility to keep the infrastructure secure only, while majority of the responsibilities like data protection, access controls, configuring and monitoring the system, and patching rests on the shoulders of the customer himself [12]. Protection of APIs (Application Programming Interfaces) will be the most challenging part of protecting the cloud as it is only through the APIs that any third party can get access to the cloud [13].
- **Attacks on IoT**: The widely emerging Internet of Things will offer an even larger surface for cybercriminals to evade privacy and integrity of user's data. Botnet draining, draining off CPU and memory through malwares, less stability of devices are some of the security issues related to IoT. Attacks on IoT can specifically prove

to be harmful for the operations which are time-sensitive as any delays in response can result in heavy losses, be it health sector or financial sector [12, 13].

- **5G to Wi-Fi Vulnerabilities**: These days operations of IoT are shifting to 5G for better speed and response time. 5G offers better speed and efficiency by carrying more data and calls on Wi-Fi thus saving the bandwidth. Compared to 4G, 5G is hundred times faster [14]. As 5G is software based, it uses digital and distributed approach for routing. This means that it would pass through multiple data entry points and hence security of these checkpoints is mandatory to avoid any kind of security breaches [14]. 5G is widely available in most of the public places, and hence, a lot of security vulnerabilities in it can be explored by the hackers [12, 13].
- **Deepfakes**: As technology is advancing day by day, attacks and forgery are also getting more and more advanced every day. An example of this is the usage of 'Deepfakes', a resultant of advancements in artificial intelligence and machine learning. Deepfakes involve creating videos out of images and thus can prove really harmful in destroying person's image completely. Such synthetic identities can prove to be really helpful in performing more sophisticated phishing scams and making billions of money out of it [12, 15].
- **AI Fuzzing**: Numerous applications today are based on artificial intelligence and machine learning. AI fuzzing is one the techniques used to detect system vulnerabilities. On the other hand, it may also be used by cyberattackers to create and spread zero-day attacks wherein the organizations will be unaware of the new attack and even if detected, will not be able to find its solution in a day [15].
- **Machine Learning Poisoning**: This attack is based on poisoning the training set used by machine learning to develop its models [15]. The purpose is to inject bad data in the training set so as that machine learns incorrect classifications and biases. Such attacks can be carried out by transfer learning, data manipulation, data injection or logic corruption. To avoid this, careful monitoring of the training data should be done [16].

4 Management and Prevention of Cyberattacks

The advancements in the world of technology have brought new challenges to its security. The newer the technology, the newer the cyberattacks. The dependence on technology has increased further more during COVID-19 due to lockdowns, isolations and restricted travels. The considerable huge increase in cyberattacks is seen during the pandemic, namely because there has been a massive increase of employees working from home, children and college students attending school online, as well as the persistent anxiety and emotions faced by people during this difficult time. Inaccurate information is worse than no information at all as it causes a lot of mental effort to correct it later and is just a waste of time. The panic of the pandemic had led to panic buying of excessive products like grocery items and medicines. People started rushing to all the available grocery stores to pile up the daily essential items

cereals, dairy products, frozen items and other household products. This led to empty stores with hardly any items left for the other half of the public. Excessive tests for COVID also put a lot of pressure on medical industries. Swab test and blood test are the two common ways to test these days. Testing kits available for COVID can be really helpful in finding out the infected individuals quickly to help control the pandemic but when millions have to be tested, then it is the medical sector that has to speed up its production.

As this era gave way to work from home, a lot of non-skilled employees were also forced into technical skills. Moreover, productivity of large corporations was affected as their VPNs could not handle the load to access network. Massive strain was put on economy as the businesses suffered and many even closed down. There were major financial losses as stock markets crashed and due to cancellation of flights, travel sector suffered huge losses.

The pandemic affected all the sectors tremendously and cyberattacks further added on to the losses. That is why learning the best ways to protect you and your company's data are so critical nowadays and working with cybersecurity professionals is more important than ever.

The right preventive measures taken at the right time can really help improve the situation and save lots of money and thefts. A few of these are listed below: [12, 15]

- Automation of patching and vulnerability management so as to detect and correct any approaching cyberattack.
- 'Think before you click' mentality needs to be encouraged. Be careful while clicking on any fake URLs or email attachments. Always check for any spelling errors before visiting any website.
- Use only approved devices and update them as well as the router software regularly. Use of VPNs should be encouraged.
- Always backup all the systems to ensure fast data recovery in case of any cyberattacks.
- Ensure a complete security suite for any devices to be used outside the premises of an organization. These may include strong passwords, multi-factor authentication, malware protection, firewall settings, DNS filtering and disk encryption.
- Have an incident response plan ready in case of any emergency situation and keep updating them to account for workforce changes in a distributed network.
- Educate the employees about data protection and security protocols.
- For all confidential files, apply end-to-end encryption. Sensitive emails may also be encrypted with passwords.
- System hardening which involves deleting and uninstalling any unimportant apps so as to reduce the surface of attacks. Also, it may include deleting any default usernames and passwords which are automatically created on downloading some apps.
- A small misconfiguration in cloud can lead to numerous errors and can prove harmful for its security. Hence, it is of utmost importance the cloud is configured correctly.

- For Linux users, FBI and NSA have advised to update Linux kernel version to 3.7 or later and activate UEFI Secure Boot to allow kernel signing enforcement.
- For Bluetooth, it is advisable to not share any sensitive data over it and always turn it off when not in use. Also, keep updating the Bluetooth-enabled devices to patch any exploitable flaws.

5 Limitations

This paper covers the research analysis collected from various sources and research work done in this field. However, a practical implementation of the analytical work has not been done. For a further continuation of this research work, some of the cyberattacks may be performed on various devices for testing. Further vulnerabilities could be explored in digital devices, and any steps to overcome them can be designed.

6 Conclusion

The considerable huge increase in cyberattacks is seen during the pandemic, namely because there has been a massive increase of employees working from home, children and college students attending school online, as well as the persistent anxiety and emotions faced by people during this difficult time. These are the challenging times which require attention of individuals and awareness among them about the importance of cybersecurity and its possible consequences if not taken seriously. As the technology will progress, the attacks will become more and more sophisticated. Hence, apart from the many advantages and ease provided by the digital world, one must also take special care of keeping himself secured. The right education and awareness can really save a lot of destruction possible due to security breaches.

References:

1. Mouton, F., De Coning A.: COVID-19: impact on the cyber security threat landscape. In: Project: Social Engineering: Defining the field from Both an Attack and Defence Perspective (2020)
2. Top Cyber Security Experts Report: 4,000 Cyber Attacks a Day Since COVID-19 Pandemic. https://www.prnewswire.com/news-releases/top-cyber-security-experts-report-4-000-cyber-attacks-a-day-since-covid-19-pandemic-301110157.html. Last accessed 19 Oct 2020
3. Khan, N.A., Brohi, S.N., Zaman, N.: Ten deadly cyber security threats amid COVID-19 pandemic (2020). https://doi.org/10.36227/techrxiv.12278792
4. Ahmad T.: Corona virus (COVID-19) pandemic and work from home: challenges of cybercrimes and cybersecurity. SSRN Electron. J. (2020). https://doi.org/10.2139/ssrn.3568830

5. COVID-19's Impact on Cyber Security. https://www2.deloitte.com/content/dam/Deloitte/ng/Documents/risk/ng-COVID-19-Impact-on-Cybersecurity-24032020.pdf. Last accessed 19 Oct 2020
6. New Android App Offers Coronavirus Safety Mask But Delivers SMS Trojan. https://www.zscaler.com/blogs/security-research/new-android-app-offers-coronavirus-safety-mask-delivers-sms-trojan. Last accessed 08 Nov 2020
7. Cyber Risks in Education Sector. https://www.itgovernance.co.uk/blog/why-has-there-been-increase-in-cyber-risks-for-the-education-sector. Last accessed 19 Oct 2020
8. INTERPOL report shows alarming rate of cyberattacks during COVID-19. https://www.interpol.int/en/News-and-Events/News/2020/INTERPOL-report-shows-alarming-rate-of-cyberattacks-during-COVID-19. Last accessed 19 Oct 2020
9. Microsoft Alerts Android Users of the New Ransomware. https://www.livemint.com/technology/tech-news/android-users-alerted-of-new-ransomware-that-could-render-smartphone-useless-11602419891977.html. Last accessed 19 Oct 2020
10. 2018–2020 Ransomware Statistics and Facts. https://www.comparitech.com/antivirus/ransomware-statistics/. Last accessed 19 Oct 2020
11. The 2020 Cybersecurity Stats You Need to Know. https://www.fintechnews.org/the-2020-cybersecurity-stats-you-need-to-know/. Last accessed 19 Oct 2020
12. Top 10 Cyber Security Threats in 2020. https://www.kaseya.com/blog/2020/04/15/top-10-cybersecurity-threats-in-2020/. Last accessed 19 Oct 2020
13. Impact of Cyber Threats on 5G. https://www.gsma.com/aboutus/workinggroups/are-we-prepared-to-deal-with-the-impact-of-cyber-threats-on-5g. Last accessed 19 Oct 2020
14. 5G Cyber Security: Threats and Solutions. https://americansecuritytoday.com/5g-cyber-security-threats-and-possible-solutions/. Last accessed 19 Oct 2020
15. 5 Cyber Security Threats to be aware of in 2020. https://www.computer.org/publications/tech-news/trends/5-cybersecurity-threats-to-be-aware-of-in-2020. Last accessed 19 Oct 2020
16. Machine Learning Poisoning in 500 Words or Less. https://www.thinkautomation.com/eli5/machine-learning-poisoning-in-500-words-or-less/. Last accessed 19 Oct 2020

QoS Enhanced Optimum Vertical Handoff Decision Algorithm for Heterogeneous Wireless Networks

Soumya Peddi, Siddarama Patil, and Jayashree Agarkhed

Abstract Wild enhancement of remote network means to provide the steady constant affiliation through acquire to assorted advance as well as to have relationship through the finest organization which give the superlative quality of service (QoS). Vertical handover come keen on portrait when consumer moves in dissimilar organization. Plotting canny vertical handover thought is a mainly vital test pro heterogeneous wireless network (HWNs). In this manuscript, we plan a QoS upgraded idyllic vertical handoff (O-VHO) choice computation pro HWNs. The planned calculation comprise of two stages. In primary stage, accumulate all essential credit to recognize the require for handoff plus afterward start it. We consider the numerous credit for progress the consumer quality dimensions, pro instance, statistics broadcast, handoff immobility, received signal strength (RSS),consumer portability, power exploitation, network cost, consumer inclination, as well as QoS dimensions (handoff deferral, throughput, plus parcel misfortune proportion). At to tip we plan multi-useful Intuitionist feathery TOPSIS (MIF-TOPSIS) base dynamic computation to figure the consumer accurate organization region. At last, the planned O-VHO computation applies to assorted HWNs, pro instance, WLANs, MANETs plus VANETs. The recreation outcome shows the planned O-VHO choice computations perform influential than existing VHO calculation as far as QoS dimensions.

Keywords Optimum vertical handoff (O-VHO) · Decision-making algorithm · Received signal strength (RSS) · Multi-functional intuitionistic fuzzy TOPSIS (MIF-TOPSIS) · Heterogeneous wireless network (HWNs)

S. Peddi (✉) · J. Agarkhed
Department of Computer Science and Engineering, Poojya Doddappa Appa College of Engineering, Kalburgi, India

S. Patil
Dept of E &CE, Poojya Doddappa Appa College of Engineering, Kalburgi, India

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_97

1 Introduction

Nowadays through the headway of remote innovation, a little remote organization through assorted access advance can coincide all the whilst as well as resolve attempt to fulfill consumers necessitate plus requisite. In a heterogeneous wireless networks (HWNs), assorted network be synchronized so as to provide the consumer finest administration [1]. Cell phone outfitted through assorted interface fatal proceeds amongst the accessible option incorporate IEEE 802.11 wireless local area network (WLAN) [2], IEEE 802.16 worldwide interoperability for microwave access (WiMAX) [3], satellite framework [4] plus Bluetooth [5] present in heterogeneous environment. Be that as it might, assorted issue, pro instance, affirmation, quality of service (QoS), consumer portability, approval plus so on must be consider through plan as well as enhancement of HWNs [6, 7]. It has diverse access innovation, covering plus inclusion, as well as organization intend, convention pro transport, steering plus versatility executive. Likewise unique administrator offers diverse assist request from consumers, pro instance, voice, video, sight plus sound, text as well as so on in market. As a outcome of these variety, when the consumer move there is a necessitate to relinquish the correspondence channel starting through one organization then onto next via belief about its highlight as well as furthermore the consumer necessities. The consistent flexibility (consumer improvement) of HWN is one of the motive pro perpendicular handoff concerns [8–11]. The channel handover among two unique networks encompass been finished in vertical handoff.

2 Related Works

Çeken et al. [11] encompass planned a fluffy rationale base handoff choice computation pro remote heterogeneous organization. The information velocity, RSSI, plus adaptable rapidity limits consider as contributions of fluffy base structure so as to choose handoff investiture compute plus choose the finest up-and-comer passageway roughly a keen portable fatal. Likewise, rather than routine calculation plus this tactic takes encircling obstruction power, which is alluded to as impedance rate, as another contribution to choice series.

Fu et al. [12] encompass introduce a rapid as well as safe handoff confirmation plot reliant on certification permit pro WiMAX plus WiFi HWNs. Portable place shows its linking endorsement permit bent via recently visit base station/passage to intent one at whatever tip a handoff happen, moreover later the flexible consumer plus aim base station/passageway preserve finish the ordinary legalization plus concludes their mutual summit key lacking communicate through the corroboration, appreciation, plus bookkeeping worker, which effectively lessen the handoff confirmation interruption.

Jaraizsimon et al. [13] encompass build up a hereditary computation pro take care of advancement concern, which is to place the finest mix of load pro the QoS limits

of HWNs pro a portable fatal. The pact exposed permit the portable fatal to decide the finest organization to set up association in a perpendicular handoff measure, as indicate via consumers inclination. The planned computation is attempt to progress the common presentation of a versatile fatal overseeing productively the.

QoS asset of originate remote organization along its way, underlining key functionalities, pro instance, safety level, broadcast capacity, interruption, etc., pro the assorted administration mention to network.

Bin et al. [14] encompass planned Markov base computation plus it is considered the position vary of accessible organization resolve persuade QoS of vehicle terminal correspondence administration. Markov compute is utilized to anticipate the alter of remote organization position after choice via means of progress likelihood. At to tip, the load of assess limit resolve be dictate via fluffy rationale tactic. The all out livelihoods of every remote organization, counting handoff option salary, handoff execution wages plus correspondence administration earning after handoff, the idyllic organization to handoff resolve be chosen.

3 System Model

Network communiqué is one of quickest mounting areas in worldwide circumstances. The varied sort of consumers have stretched speedily throughout the mainly recent decade. The administrator be unfailingly invested in fraction of amount of energy to persuade the consumer necessities. In the HWNs, a consumer ought to encompass the option to shift preliminary through one organization then onto subsequently via performing vertical handoff (VHO). During VHO, it is indispensable to painstakingly alter the broadcast capability obligation as well as redistribution to provide most noteworthy QoS to current consumers. The structure replica of planned computation comprise of BS/AP, assorted kind of organization consumers (pro instance PCs/workstations, portable, plus vehicle). Incorporation of WLAN, MANET as well as VANET network has superfluous favorable circumstance concerning portability, inclusion territory as well as statistics relocate capacity. The organization replica through the consumer conveyance of planned work are appearing in Fig. 1.

4 Multiple Attributes Gathering from Network

Bandwidth is a measurement of available or consumed data communication resources expressed in bits per second. It is a good indicator of traffic conditions in the access network and is especially important for delay-sensitive applications.

Handoff latency is the duration from handover initiation to handover completion which equal to amount of time from when the user is disconnected from the existing BS and switching to new BS and the user receives the first packet from the new BS.

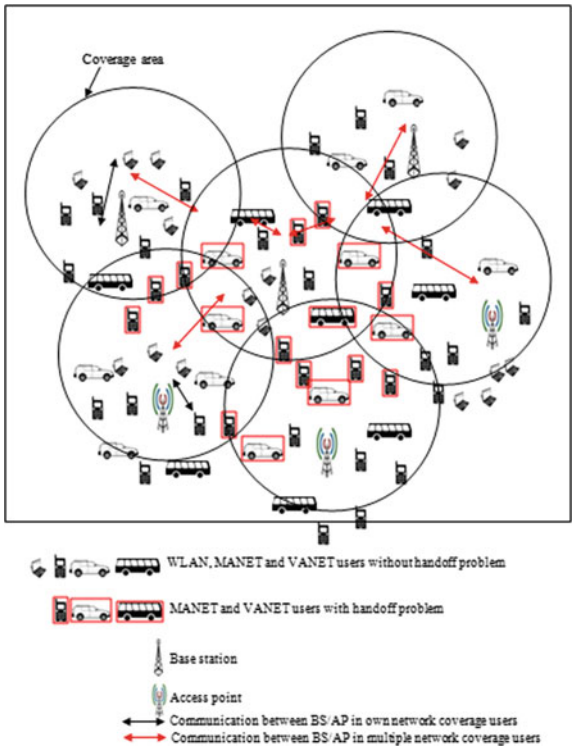


Fig. 1 User deployment of proposed work

Received signal strength handover decision is a handover initiation phase in homogeneous environment. Received signal strength (RSS) is the traditional handover decision criteria in almost all existing horizontal handover algorithms. RSS is also an important decision criterion in the VHO algorithms. RSS is the most widely used criterion because it is easy to measure and is directly related to the service quality. There is a close relationship between the RSS readings and the distance from the mobile terminal to its point of attachment.

5 Proposed O-VHO Decision-Making Algorithm

5.1 Multi-functional Intuitionistic Fuzzy TOPSIS

The MCDM circumstances preserve be uttered via a matrix call decision matrix (DM) as well as it is comprised of n numeral of multi-functional faith attribute A_1, $A_2 \ldots A_n$ as well as m numeral of probable Elucidation S, $S_2 \ldots S_m$ pro load computation.

$$\mathrm{DM} = \begin{array}{c} \\ S_1 \\ S_2 \\ \vdots \\ \vdots \\ S_m \end{array} \begin{array}{c} \begin{array}{cccc} A_1 & A_2 & \cdots\cdots & A_n \end{array} \\ \begin{bmatrix} i_{11} & i_{12} & \cdots\cdots & i_{in} \\ i_{21} & i_{22} & \cdots\cdots & i_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ i_{m1} & i_{m2} & \cdots\cdots & i_{mn} \end{bmatrix} \end{array} \tag{1}$$

In load computation, the set of load W_S is engaged to the m realistic solution plus the set of significance weight W_t is functional to computation of initial choice producer.

$$W_s = \{W_{s1},\ W_{s2}, \ldots, W_{sl}\} \tag{2}$$

$$W_t = \{W_{t1},\ W_{t2}, \ldots, W_{tl}\} \tag{3}$$

The choice structure (11) obliges evenness to modify it keen on a dimensionless lattice. Separating the network mechanism in each line (criteria) via square base of summative off the square of the mechanism of to column yield a consistent choice grid through identical worth mechanism of r_{xy}:

$$r_{xy} = \frac{i_{xy}}{\sqrt{\sum_{j=1}^{n} i_{xy}^2}} \text{ where } x = 1, \ldots, m \text{ and } y = 1, \ldots, n \tag{4}$$

The pact of entropy behavior pro each compute is then used to calculate the entropy weights. So as to outline entropy weights through which to alter the choice network pro a TOPSIS computation a stage of disparity is gotten via subtract entropy commencing one:

$$d_x = 1 - E(C_x) \quad \text{where} \quad x = 1, 2, \ldots, m \tag{5}$$

The level of peculiarity communicates the trait discrimination force amongst the appraisal of each standard. The more prominent the generally assessment of di the more noteworthy to replica is in extrication among the n choice, plus further noteworthy load it is neutrally relegate in computation. The entropy weight W pro each measures x is then unwavering to framework the pact of entropy load (W_e) (Fig. 2)

$$W_x = \frac{d_x}{\sum_{x=1}^{m} d_x},\ x = 1, 2, \ldots, m \tag{6}$$

$$W_e = (W_1, W_2, \ldots, W_x, W_m) \tag{7}$$

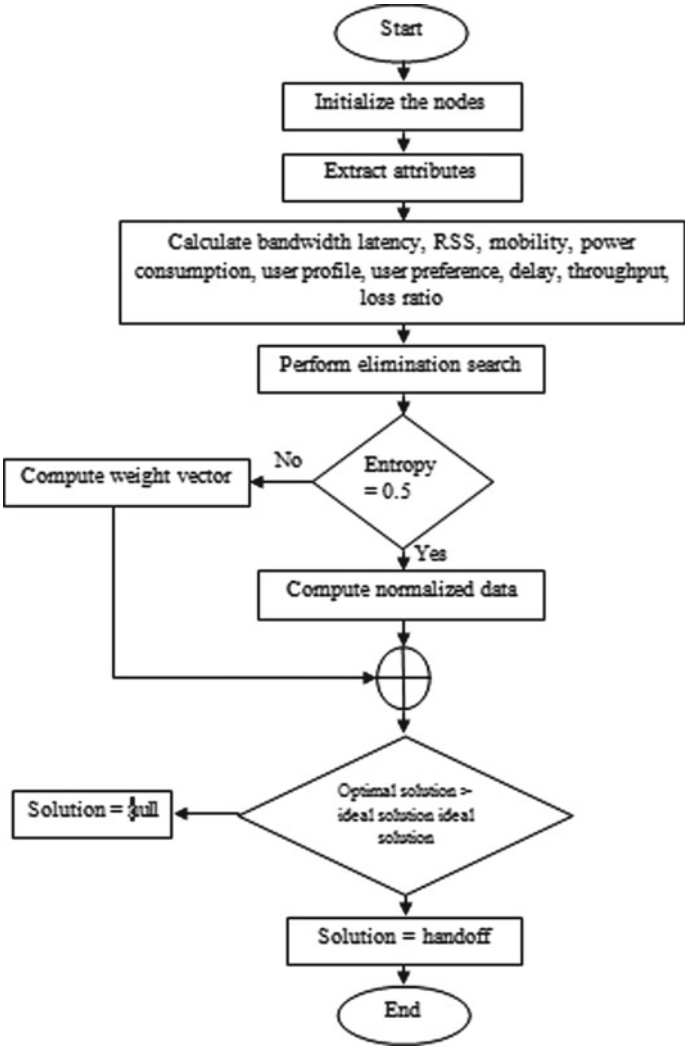


Fig. 2 Work flow of O-VHO decision making algorithm

6 Simulation Result

The planned O-VHO vibrant computation is reenacted to utilize network assessment scheme NS-2 apparatus. We believe a massive province dimension as 1000 × 1000 square through the elevated numeral of consumers through the BS/AP. The recreation period of planned structure is set as 100 s. In meticulous, we exploit the broadcast capability be 22 MHz pro all channel as well as the converse power pro each hub is 1 W. We further guess to the way adversity instance is 4 as well as alter of commotion

Table 1 Testing scenario

Scenario	Number of user	User mobility (ms)	Simulation time (s)
1	50–250	5	100
2	100	1–5	100
3	100	5	Pause time: 20, 40, 60, 80, and 100

is 10^{-10} W. The broadcast power pro a consumer is 17 dBm as well as for base station is 20 dBm. The consumers plus base station exploit the accompany 802.11n corporeal setting in 5 GHz band, MAC supervision information component accumulation plan as well as whole dimension of 5 edges pro each MAC supervision information component entirety through 1500 bytes load information pro each edge. The three distinct test circumstances use to scrutinize the exhibition of planned MORA plan as well as it is contrast plus the current MSA-VHO dynamic technique. The test circumstances classify in Table 1.

The appearance of the planned O-VHO dynamic computation assesses through assorted QoS dimensions, pro instance, handover interruption, dormancy, throughput, adversity proportion, numeral of handoff, plus usual statistics relocate capability (See area 4.1). Numeral of handover is vital to shrink the extent of handover as persistent handover would cause the depletion of asset of organization. A handover is view as added overhead when a surrender back to initial point of correlation is requisite within certain instance span, as well as such handover ought to be restricted.

6.1 Varying Number of Users

In this part, assembly of QoS dimensions examines via altering numeral of consumers through rigid portability as 5 ms as well as imitation instance as 100 s. The quantities of consumers be distorted via the 50, 100, 150, 200, as well as 250.

The handoff deferral of planned O-VHO computation is little disparity through MSA-VHO tactic which is appeared in Fig. 3a.

The handoff dormancy of planned O-VHO computation is little disparity through MSA-VHO method is appeared in Fig. 3b.

7 Conclusion

In this paper, QoS enhanced optimum vertical handoff decision algorithm (O-VHO) proposed for HWNs. The proposed algorithm consists of two phases such as attributes gathering and decision aggregated phase. First, we gather multiple attributes to identify the need for the handoff, then we have proposed multi-functional intuitionistic fuzzy TOPSIS (MIF-TOPSIS)-based decision-making algorithm to compute

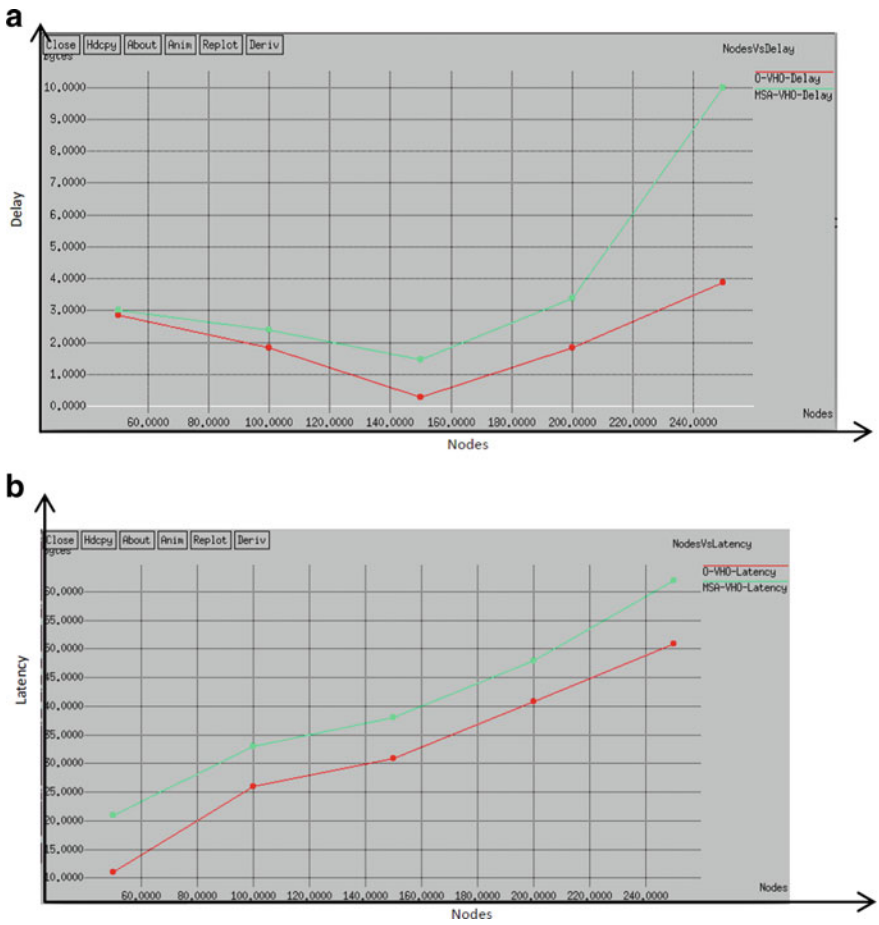


Fig. 3 **a** Nodes versus delay, **b** nodes versus latency

the moving user's exact location. For the testing purpose, the proposed O-VHO algorithm is applied to different HWNs with the combination of WLANs, MANETs, and VANETs. From the simulation result, the proposed O-VHO decision algorithm performs very efficient than existing MSA-VHO method by the quality metrics such as handoff delay, latency, throughput, packet loss ratio, number of handoff, and average bandwidth.

References

1. Tragos, E., Tsiropoulos, G., Karetsos, G., Kyriazakos, S.: Admission control for QoS support in heterogeneous 4G wireless networks. IEEE Netw. **22**(3), 30–37 (2008)

2. Deng, D.-J., Yen, H.-C.: Quality-of-service provisioning system for multimedia transmission in IEEE 802.11 wireless LANs. IEEE J. Sel. Areas Commun. **23**(6), 1240–1252 (2005)
3. Wang, F., Ghosh, A., Sankaran, C., Fleming, P., Hsieh, F., Benes, S.: Mobile WiMAX systems: performance and evolution. IEEE Commun. Mag. **46**(10), 41–49 (2008)
4. Beakley, G.: Overview of commercial satellite communications. IEEE Trans. Aerosp. Electron. Syst. **20**(4), 455–464 (1984)
5. Shepherd, R.: Bluetooth wireless technology in the home. Electron. Commun. Eng. J. **13**(5), 195–203 (2001)
6. Niyato, D., Hossain, E.: Competitive pricing in heterogeneous wireless access networks: issues and approaches. IEEE Netw. **22**(6), 4–11 (2008)
7. Heath, R., Debbah, M., Larsson, E., Kim, D., Viswanathan, H., Guvenc, İ: Introduction to the issue on signal processing in heterogeneous networks for future broadband wireless systems. IEEE J. Sel. Top. Signal Process. **6**(3), 213–215 (2012)
8. Lim, J., Hong, D.: Mobility and handover management for heterogeneous networks in LTE-advanced. Wirel. Pers. Commun. **72**(4), 2901–2912 (2013)
9. Hasib, A., Fapojuwo, A.: Mobility model for heterogeneous wireless networks and its application in common radio resource management. IET Commun. **2**(9), 1186 (2008)
10. Lahby, M., Essouiri, A.: A novel modeling approach for vertical handover based on dynamic k-partite graph in heterogeneous networks. Digit. Commun. Netw. **5**(4), 203–316 (2019)
11. Çeken, C., Yarkan, S., Arslan, H.: Interference aware vertical handoff decision algorithm for quality of service support in wireless heterogeneous networks. Comput. Netw. **54**(5), 726–740 (2010)
12. Fu, A., Zhang, G., Zhu, Z., Zhang, Y.: Fast and secure handover authentication scheme based on ticket for WiMAX and WiFi heterogeneous networks. Wirel. Pers. Commun. **79**(2), 1277–1299 (2014)
13. Jaraiz-Simon, M., Gomez-Pulido, J., Vega-Rodriguez, M.: Embedded intelligence for fast QoS-based vertical handoff in heterogeneous wireless access networks. Pervasive Mob. Comput. **19**, 141–155 (2015)
14. Bin, M., Hong, D., Xianzhong, X., Xiaofeng, L.: An optimized vertical handoff algorithm based on Markov process in vehicle heterogeneous network. China Commun. **12**(4), 106–116 (2015)

Plant Disease Detection and Classificatiion Using Convolutional Neural Network

S. Muthurajkumar, S. Safiyyah Sanovar, G. Sridhar, and G. Priyanga

Abstract Agriculture has become even more important in recent years than the years back where plants were used only to feed animals and humans. Nowadays, to improve the livelihood of humans, plants are used to generate electricity and other forms of energy. For these reasons, plants should be taken proper care to get at most benefits from them. Even food security is a major threat in recent years, so there is a need to take proper care of plants. Therefore, plant diseases have to be prevented and reduced. And there are various plant diseases that cause huge harm to society and economy and also lead to huge ecological losses. To avoid such losses, the plant disease should be detected accurately and timely. The most not unusual plant illnesses are the sicknesses brought about by means of fungi which appear as spots on the leaves. The spots affect the chlorophyll in the leaves of the plant; it becomes difficult for plants to carry out photosynthesis by means of which they prepare their food; and this affects the growth of plants and thus the yield to a great extent. And the leaves are fully covered with spots in case of severe infection and results in withering of the plant. Plant disease can be detected and categorized using convolutional neural community and the lessons having maximum possibility is chosen. The algorithm is estimated to give up to about 97% accurate results.

Keywords Plant disease · Deep learning · Convolution neural networks · D-CNN

1 Introduction

Agriculture is the major source of income in India. Farmers grow many varieties of crops but the production of crops decreases as the plants is infected by different types of diseases on their stem, leaf, and fruits. The leaf diseases are caused mainly by fungi, viruses, and bacteria. To overcome the leaf sicknesses, the sicknesses in leaves are categorized primarily based on diseased leaf sorts. This may be carried

S. Muthurajkumar (✉) · S. Safiyyah Sanovar · G. Sridhar · G. Priyanga
Department of Computer Technology, Anna University, MIT Campus, Chrompet, Chennai 600044, India

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_98

out the usage of neural networks, and through this, the loss in production can be minimized by taking the appropriate steps. To achieve this, first off, the farmer clicks an photograph of the leaf of the infected plant and uploads it on the server; the photograph is processed, and the functions are extracted from the photograph. The important features are alone identified, and based on only those images, the image is classified using neural networks algorithms and the leaf disease is identified. The plant disease detection and classification system are important because plant disease has become a dilemma and has an adverse effect on crops, thereby causing huge reduction in the quality as well as quantity of yield. The plant disease and pests reason loss to farmers and are a risk to meals protection. In the latest years, the spread of pests and plant sicknesses has dramatically elevated. The motive at the back of this is weather alternate, alternate, globalization, and reduced resilience in manufacturing structures because of many years of agricultural intensification. The sudden outbreaks and increase in spread of plant disease cause huge loss of crops which threatens the livelihood of poor farmers and nutrition and food security of many people at a time. From the beginning of agriculture, farmers have been practicing methods for reducing the various plant diseases. By understanding the host and pathogen interactions, the measures for controlling plant diseases were found out in early twenty-first century. These help in managing of new plant diseases by finding some common principles for controlling plant disease. For those who are into agriculture without any prior expertise about plant disorder and controlling measures, an automatic system for detection of plant disease could be very beneficial and has many advantages; it video display units big fields of crops and detects the plant illnesses robotically by figuring out the symptoms at the leaves of the plant. The system can be implemented using various techniques like image processing and classification. For the detection of plant disease, many algorithms have been studied out system implements detection using neural networks. Neural network is an artificial intelligence gadget that imitates the functionality of the human mind. Bernard Widrow of Stanford University turned into the first to develop neural networks within the early fifties. In many fields in computing especially in voice recognition systems, image recognition systems, robotics, medicine, etc., neural networks are used.

The system implements convolutional neural network to categorize the plant disorder. In deep learning , a category of deep neural networks is known as convolutional neural network (CNN or ConvNet). Regularized variations of multilayer perceptrons are convolutional neural community. For visual imagery, convolutional neural network is commonly applied. Completely connected organizations are alluded to as multilayer perceptrons, where every neuron of one layer is associated with all the neurons in the following layer. They are liable to overfit information because of the fully connectedness of the network. For regularization, to the loss characteristic a few form of value dimension of weights is brought. A different approach for regularization is done through CNN, and they use the hierarchical pattern in data, and using smaller patterns, they assemble complex patterns. The system can be implemented using a customized architecture which is VGG architecture having 16 –19 layers. In VGG architecture, the convolutional layers are stacked on top of each other. Training of VGG model is challenging when it comes to convergence of deeper network.

Hence, the smaller network of VGG is trained with less weights to make the training easier. Then the smaller networks are converged and used to initialize larger deep networks. Deep convolutional neural networks have improved performances. Here a system based on deep CNN using AlexNet framework is developed. AlexNet is much faster than VGG and provides more accuracy. The system is proposed to classify the leaves of apple, blueberry, cherry, corn, grape, orange, peach, pepper, potato, raspberry, soybean, strawberry, tomato plants and can classify 37 classes of diseases.

2 Related Work

Zhang et al. [1] stepped forward the maize leaf illnesses identification accuracy and the wide variety of community parameters were decreased, based on deep learning proposed leaf disease recognition. It proposed a device with stepped forward accuracy of leaf ailment and reduced the convergence iterations the use of the stepped forward strategies Cifar10 and GoogleNet which improved the model training and recognition effectively.

Jiang et al. [2] proposed a deep mastering technique the usage of progressed convolutional neural networks for detection of diseases of apple plant in real time. The apple plant sicknesses labeled was alternaria leaf spot, brown spot, mosaic, grey spot and rust using deep CNN model implemented through GoogleNet inception structure and rainbow concatenation.

Sardogan et al. [3] It proposed learning vector quantization (LVQ) algorithm and a convolutional neural network version for detection and category of tomato leaf sickness. In this model, the filters were applied to three channels based on RGB components.

De Luna et al. [4] used a motor-controlled picture catching box made to catch four sides of each tomato plant which was utilized for the location and acknowledgment of tomato plant leaf infections. A particular variety of tomato which is Diamante Max was utilized as the guinea pig.

Pardede et al. [5] used convolutional auto-encoder and proposed an unsupervised feature learning algorithm for the detection of leaf diseases. The author produced discriminative capabilities and manner to give enter to SVM-based totally classifiers which used the output of car encoder for automatic detection of leaf illnesses.

Howlader et al. [6] recognized guava leaf sicknesses through the proper D-CNN and proposed a mode which classified major diseases of guava leaf such as whitefly, rust, and algal leaf spot. They used stochastic gradient descent(SGD) optimization algorithm with some hyperparameter.

Kosamkar et al. [7] proposed a device that did preprocessing, feature extraction of the snapshots of the plant leaves from plant village dataset and accompanied by way of type of the plant disorder through convolutional neural community and endorsed insecticides the use of Tensor glide technology.

Sabrol et al. [8] using classification tree, segmentation approach, and CNN algorithms

3 Proposed Work

3.1 System Overview

The idea of the system is to provide information about the disease the plant is been affected, to detect, and classify the plant disease. The farmer takes image of the leaf of the inflamed plant and uploads it to the server through webpage. The processing of picture is carried out, and the features are extracted from the photograph. Then class of the image is found using convolutional neural network based totally at the extracted features and the magnificence with the best possibility is selected and the disease is detected.

3.2 System Methodology

The following system work drift explains the ordered technique for detection and classification of plant disorder. Firstly, it acquires the image of the plant leaf which is unprocessed. Then to remove distortion and noise in the input image, few pre-processing techniques are applied. Then from the image, the features are extracted for non-redundant and greater informative enter for training the model. Finally, the type of the plant sickness is diagnosed using convolutional neural networks (Fig. 1).

3.3 Image Acquisition

The motion of retrieving an photo from any source is defined as image acquisition, broadly speaking from a hardware-based totally supply. Without no image as the input, the further processing cannot be done so image acquisition is the first step in workflow sequence. The acquired image is completely not processed.

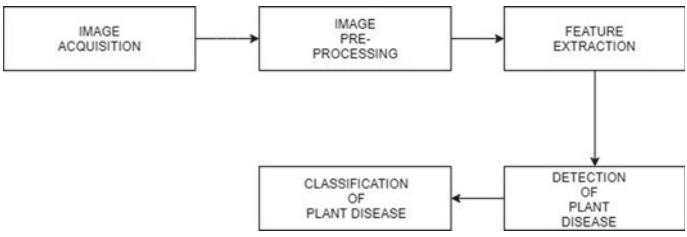


Fig. 1 System workflow

3.4 Image Preprocessing

Image processing is very important to get more informative and non-redundant data to be given as input. It helps to extract many features, and discrimination ability is increased. All the enter photographs in the dataset using which the model is trained need to be of the same format and size. Therefore, all of the snapshots are resized to 256 X 256 pixels.

3.5 Convolutional Layer

The CONV layer parameters comprise of an arrangement of learnable channels. Each channel is little spatially (along width and height), however, reaches out through the full profundity of the information volume. For instance, an ordinary channel on a first layer of a CNN may have measure $5 \times 5 \times 3$ (i.e., 5 pixels width and stature, and 3 since pictures have profundity 3, the shading channels RGB). Amid the forward pass, we slide each channel over the width and stature of the information volume and figure speck items between the passages of the channel and the contribution at any position. As we slide the channel over the width and tallness of the information volume, we will deliver a two-dimensional actuation outline that gives the reactions of that channel at each spatial position. The system will learn channels that enact when they see some kind of visual element, for example, an edge of some introduction or a fix of some shading on the primary layer, or in the end whole example on higher layers of the system. Every neuron in a CNN isn't totally associated with the neurons in the previous layer. It is associated with a gathering of neurons, and this is known as the open field. Responsive field relies upon the extent of the channel. The yield delivered by the layer relies upon three qualities:

Depth: The intensity of the output volume relies upon on the range of filters used. If we use twelve filters, then the intensity of the output produced might be 12.

Stride: Stride is the cost with which we slide the clear out throughout the input values. At the point when the step is 1, then we move the channels each pixel in turn. At the point when the step is 2, then the channels ricochet 2 pixels; all at once as we slide them around. This will deliver more modest yield volumes spatially.

Size of Zero Padding: It will allow us to govern the spatial period of the output volumes. We can compute the spatial length of the output volume as a characteristic of the enter extent length (W), the receptive subject size of the convolutional layer neurons (F), the stride with which they may be applied (S), and the amount of zero padding used (P) on the border using the equation:

$$\text{Output size} = (wf + 2p)/s + 1 \tag{1}$$

3.6 Pooling Layer

Intermittently, we embed a layer in the middle of progressive Conv layers in a ConvNet design and this layer is called pooling layer. Its ability is to progressively decrease the spatial size of the depiction to diminish the amount of boundaries and computation in the framework, and in this manner to similarly control overfitting. The pooling layer works unreservedly on every significance cut of the data and resizes it spatially, using the MAX task (we moreover use typical pooling). The most generally perceived shape is a pooling layer with channels of size 2×2 associated with a stroll of 2 down-models every significance cut in the commitment by 2 along both width and tallness, discarding 75% of incitations. Every MAX undertaking would this circumstance be taking the greatest multiple numbers.

3.7 Fully Connected Layer

Neurons in a totally connected layer have complete connections to all activations within the previous layer, as seen in everyday neural networks. This layer takes the values from the sooner layers and study weights and bias values to eventually classify the given enter into considered one of many enter classes (Table 1).

Table 1 Structural parameters of proposed CNN model

CNN parameter	Filter size/stride	Output size
Conv 1	$54 \times 54 \times 96$	4
Pool 1	$27 \times 27 \times 96$	2
Conv 2	$17 \times 17 \times 256$	1
Pool 2	$8 \times 8 \times 8 \times 256$	2
Conv 3	$6 \times 6 \times 384$	1
Conv 4	$4 \times 4 \times 384$	1
Conv 5	$2 \times 2 \times 256$	1
Pool 3	$1 \times 1 \times 256$	2
Fully connected	–	4096
Fully connected	–	4096
Fully connected	–	1000
Softmax	–	38

4 Experimental Evaluation

4.1 Dataset Description

A proper dataset is needed for evaluation of the performance of the model. An openly accessible dataset new plant disease dataset is utilized here which comprises of 87,000 RGB pictures of both sick and sound leaves and ordered into 37 distinct classes. The dataset is separated into proportion of 80:20 as preparing and approval set (Fig. 2).

4.2 Execution Environment and Procedure

The experiment is conducted on a 64 bits operating system with Windows10, and system has i5 core along with 16GB RAM and 8GB GPU. The training of the model is done using GPU as it gives high speed of computation. The model is implemented using TensorFlow framework and a neural network library keras 2.2.0.

Step 1: Data preprocessing: all the dataset pics are resized to a format of 256x256 pixels .

Step 2: Dataset is divided into a ratio of 80:20 as training and validation set.

Step 3: Data augmentation: to the training set, the augmentation procedure is carried out for rotation, resizing of photograph, and adding some random noise to pictures for overfitting not to occur.

Step 4: Feature extraction: In the early layers of convolutional neural network architecture, the features are extracted through convolutional operations.

Step 5: Training the model: AlexNet framework is used. After developing the architecture, the model is trained using training set features.

Step 6 Evaluation: Accuracy of model would be evaluated with the help of test set.

Step 7: Tuning: If effects aren't exceptional tune the model by changing the parameters of structure consisting of kernel length, nodes in remaining are absolutely related layer.

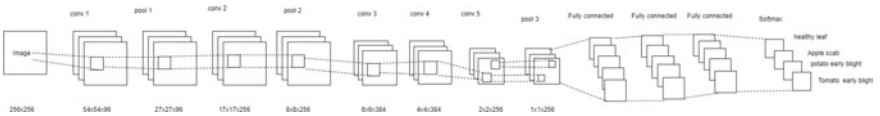


Fig. 2 Layers of CNN

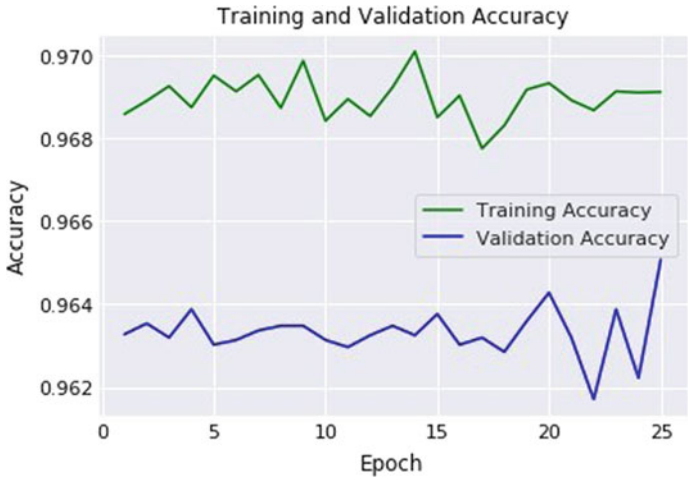


Fig. 3 Accuracy of training and validation

Step 8: Store the weights: Final version which has trained save it in model name configuration record in order that it may be used for brand new information.

Step 9: Extract the features and evaluate with trained model

4.3 Experimental Result

We have implemented the plant disease detection and classification using CNN system. The graphs for training/testing loss and accuracy Vs epoch is drafted. We have sampled the inputs for 25 epochs.

In Fig. 3, the training and validation accuracy graph are plotted and performance of the version is evaluated.
In Fig. 4, the training and validation loss graph are plotted for figuring out whether the model is overfitting/underfitting.

The plant sickness is detected, and the form of disease is recognized (Fig. 5).

5 Conclusion

Thus, in conclusion, the detection of plant sickness is carried out the usage of CNN architecture based on deep learning. The system used AlexNet architecture instead of VGG architecture to reduce the error rates. This implementation has scaled up the accuracy, and the described system gives 96.77 accuracy. In this model, the plant leaf can be utilized to hit upon the plant issue. At the point when different components of

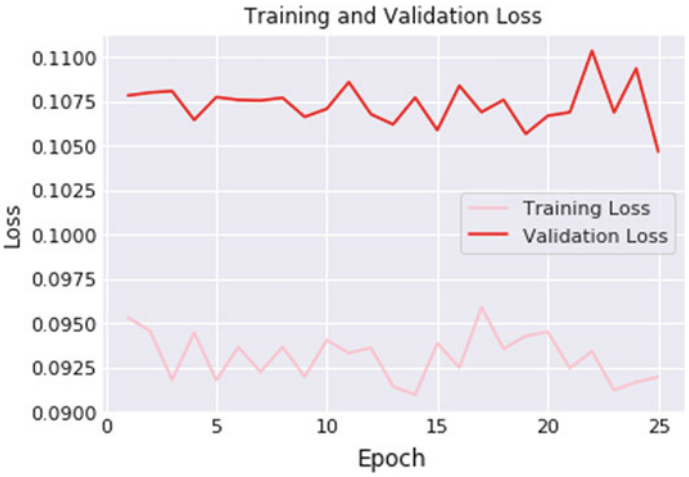
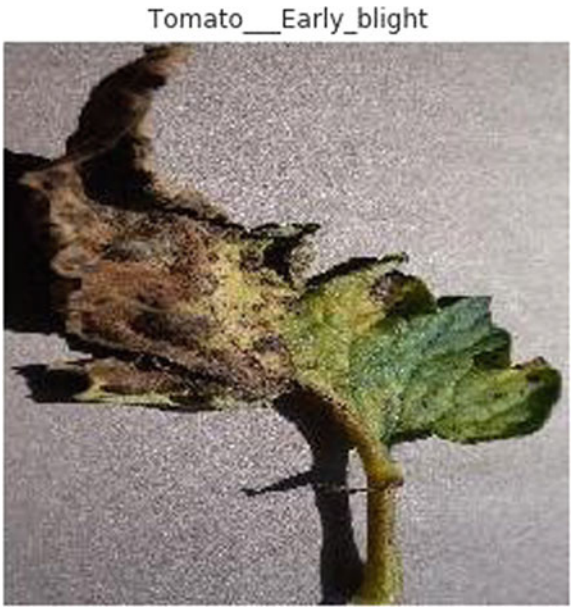


Fig. 4 Loss of training and validation

Fig. 5 Disease identification



the plant alongside stems, roots and branches are considered to find plant infirmity and the recognition precision may be extended. Future compositions comprises of plant sickness location and order through thinking about different components of the plant.

References

1. Zhang, X., Qiao, Y., Meng, F., Fan, C., Zhang, M.: Identification of Miaze leaf diseases using improved deep convolutional neural networks. IEEE Access **6**, 30370–30377
2. Jiang, P., Chen, Y., Liu, B., He, D., Liang, C.: Real-time detection of apple leaf diseases using deep learning approach based on improved convolutional neural networks. IEEE Access **7**, 59069–59080 (2019)
3. Sardogan, M., Tuncer, A., Ozen, Y.: Plant leaf disease detection and classification based on CNN with LVQ algorithm. In: 2018 3rd International Conference on Computer Science and Engineering (UBMK), Sarajevo. pp. 382–385 (2018)
4. De Luna, R.G., Dadios, E.P., Bandala, A.A.: Automated image capturing system for deep learning-based tomato plant leaf disease Detection and recogition. In: TENCON 2018—2018 IEEE Region 10 Conference (2018)
5. Pardede, H.F., Suryawati, E., Sustika, R., Zilvan, V.: Unsupevised convolutional autoencoder-based feature learning for automatic detection of plant diseases. In: 2018 International Conference on Computer, Control, Informatics and Its Applications (IC3INA) (2018)
6. Howlader, M.R., Habiba, U., Faisal, R.H., Rahman, M.M.: Automatic recognition of guava leaf diseases using deep convolutional neural network. In: 2019 International Conference on Electrical, Computer and Communication Engineering (ECCE), Cox'sBazar, Bangladesh, pp. 1–5 (2019)
7. Kosamkar, P.K., Kulkarni, V.Y., Mantri, K., Rudrawar, S., Salmpuria, S., Gadekar, N.: Leaf disease detection and recommendation of pesticides using convolution neural network. In: 2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA) (2018)
8. Sabrol, H., Satish, K.: Tomato plant disease classification in digital images using classification tree. In: 2016 International Conference on Communication and Signal Processing (ICCSP), Melmaruvathur, pp. 1242–1246 (2016)
9. Singh, V., Varsha, Misra, A.K.: 'Detection of unhealthy region of plant leaves using image processing and genetic algorithm. In: 2015 International Conference on Advances in Computer Engineering and Applications, Ghaziabad, pp. 1028–1032 (2015)
10. Nachtigall, L.G., Araujo, R.M., Nachtigall, G R.: Classification of apple tree disorders using convolutional neural networks. In: 2016 IEEE 28th International Conference on Tools with Artificial Intelligence (ICTAI) (2016)

Recognition of Kannada Handwritten Words from Answer Scripts Using Machine Learning Approaches

M. Ravikumar and S. Sampathkumar

Abstract In this paper, handwritten word recognition system for answer scripts written in Kannada language is proposed, which comprises a new concept for handwriting word recognition architecture based on CNN model and SVM classifier. Paper provides a detailed concept and implementation of Kannada word recognition. Real-time dataset is used for training the model. The proposed system has achieved the accuracy of 90% for Kannada handwritten words.

Keywords Answer scripts · Kannada · Pre-processing · Segmentation · OCR

1 Introduction

Kannada is a Dravidian language spoken predominantly by the people of Karnataka in the southwestern region of India [1]. Handwritten word recognition has been an evolving field of research from past few decades' recognition of word in printed text evolving which has now dawned upon the recognition of words in handwritten documents [2]. Human beings are perennial optimizers; hence, they are on the lookout of technologies for making their task easier [3]. Handwriting is harder to recognize for humans as well as computers. The nature of writing styles of persons' changes depending on font styles, cursive, touching characters. To segment these types of handwritten words is strenuous task. Before going to build any efficient OCR, some of the stages have to be processed—they are word segmentation and character segmentation. Handwritten word recognition provides solution for many applications like postal automation, bank cheque reading, automatic data entry, reading customer filled forms and answer script evaluation, etc. There are many difficulties in dealing with word recognition in case of vary in both size and shape, slope factor of writing. Any handwriting OCR system consists of the following processes:

1. Image acquisition

M. Ravikumar (✉) · S. Sampathkumar
DOS in Computer Science, Kuvempu University, Shivamogga, Karnataka, India
e-mail: ravi2142@yahoo.co.in

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_99

2. Pre-processing
3. Feature extraction
4. Classification.

Recognition of words from answer scripts is very challenging task. To build robust and efficient Kannada OCR, regional language contains ottaksharas present below the characters. A great portion of answer script however still takes place in physical manual evaluation, i.e., answer evaluation is still not easy for teachers to evaluate. The overwhelming volume of answer scripts data to store. As a result, Kannada handwritten recognition is a time-consuming process.

Some of the major difficulties while word recognition of handwritten text are:

1. There can be variation in shapes and font styles of different writers.
2. Connected or overlapped nature of handwriting.
3. Kannada language contains ottakshara as it is overlapped with other characters.
4. Some characters have similar contours.
5. Kannada scripts contain semi-cursive characters which are difficult to segment.

To get a desirable result in a word recognition phase, a robust and efficient pre-processing and segmentation are needed. Only proper segmentation will improve the performance of OCR. Without segmentation, we cannot directly build an efficient word recognition OCR.

The paper is organized in the following sections: Sect. 2 contains the review of related work. Section 3 gives brief idea of proposed method used. Section 4 contains results and discussion, and finally Sect. 5 gives the conclusion of the proposed work.

2 Related Work

In this review paper [4], Kannada district name recognition is focused using classifier. The database is created by 60 different writers with district names of Karnataka. Euclidean distance and DTW algorithm are used to recognize the word images. Preprocessing stage includes binarization, noise removal, skew detection and correction, thinning. The recognition accuracy provides considerable results.

Zone-wise horizontal and vertical profile-based method is proposed [5], and dataset consists of 750 images of Kannada words. Preprocessing and segmentation are done for the purpose of word recognition neural network used for training the network. Generated features are used to train the model. Thinning, bounding box generation and feature extraction are done to create the zones. The system works in two stages—training and testing. Experiments give recognition accuracy of 97.1%.

Locality-preserving projections method is proposed [6], system design contains data acquisition, preprocessing, feature extraction classification and recognition. The experiment was conducted on white sheet, and images are scanned. For the purpose of proposed algorithm, very small network of five nodes is used. Algorithm provides very well-efficient transmission of data. The accuracy gained is 80%.

Convolutional neural network model is proposed [7], and proposed system works for two stages—training and testing. In the first step, input document is converted to colored image to gray scale. In the step 2, denoising is processed to remove noise. In the step 3, preprocessing is done to extract features and to configure CNN model, web application is developed. Efficacy for char 74k is 99% and for own created dataset is 96%.

Online recognition of Kannada characters is proposed [8], recognition of words for top, middle and bottom strokes. Data is collected by different age groups of Karnataka. Genius mouse pen was used to collect dataset. These are validated, segmented and features that were extracted. Character segmentation is done by individually dividing into top strokes, middle stroke and bottom stroke. For recognition, KNN and SVM pattern are used. The maximum yielding rate of accuracy is 92.5% for KNN and 94.35% for SVM.

HMM-based character recognition method is proposed [9], the stage involved to develop a model is preprocessing to enhance and render the suitable data. Segmentation to separate individual character feature extraction to store only required feature for recognition. Training is carried out by Baum–Welch re-estimation procedure. The model is trained for 100 training samples. Accuracy of Kannada characters is 76%.

Grid-based approach method is proposed [10] for 28 district names of Karnataka state. Each written from 40 different people is collected for experimentation. The proposed method segregates the input word into four grids. PCA is applied for better representation. For classification distance measure technique is applied and obtained result is 68%.

Template matching method for recognition on correlation analysis is proposed [11]. In the step1, the dataset is collected by test images of many writers and passed for preprocessing stage to process the image. In the step 2, images are segmented into individual characters using simple segmentation method. The segmented images correlated with the already stored images. The maximum related images are shown in the editable format.

Hybrid features for recognition of handwritten Kannada characters are proposed [12], and 3600 samples are collected. K-nearest neighbor classifier is used to classify characters. Proposed techniques use local and global features. Features are extracted for each individual image. Overall accuracy gained is 87.33%.

From the literature, some more good works on recognition of word are available [1, 13–16].

3 Proposed Method

In this section, we discuss the proposed method for handwritten word recognition, and the block diagram of the proposed method is given in Fig. 1. Scanned handwritten answer scripts are collected from different schools. In the step 2, preprocessing is done to get rid of unwanted noise like ink smudges, blur, and it is necessary to process

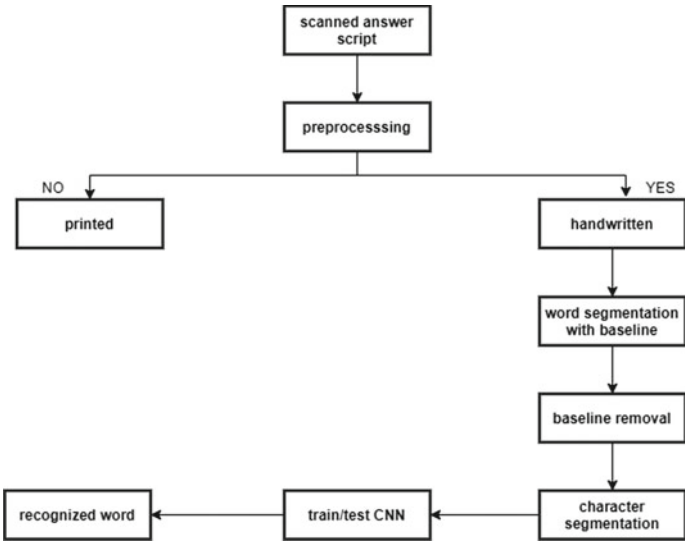


Fig. 1 Various stages of proposed method

the answer script, necessary data. It helps in saving of memory required to store the answer script data. In the step 3, we have two stages that is whether we want to process printed or handwritten text. Our main focus is on written answer so we select only text for further stage. In the step 4, handwritten answer is extracted from answer script which is of one-word answer by using horizontal line detection method. In the step 5, removal of baseline is applied, so that word is sent to next stage. In the step 6, segmentation of characters are done by using bounding box and contour method and segregated characters are stored. In the step 7, stored individual characters are separated for training and testing phases. In the step 8, by using CNN and SVM, comparison is done by passing the handwritten word with trained characters. Finally in the last step, word recognition method is chosen for handwritten word recognition which provides considerable results.

In the next section, experimentation is discussed.

4 Results and Discussion

For the purpose of experimentation, we have created our own dataset such a way that initially we have formulated questions from 8th and 9th standard Kannada textbooks and ten questions are opted, i.e., in the form of printed form collected from government and private schools consists of 200 answer scripts of Kannada handwritten one-word answer, all the answers are written by different students. To write the answers, baseline is provided for the students to write the answers. In this proposed work

Table 1 Recognition accuracy of handwritten Kannada words

S. No.	Words	Number of scripts	Recognized words	Recognition rate (%)	Average
1	Bharatha	200	185	92.5	90.09%
2	Kannada	160	150	93.45	
3	Pampa	150	142	94.66	
4	Muvathu	150	125	83.33	
5	Shivamoga	150	128	85.33	
6	Puducherry	150	148	98.6	
7	Batriya	150	122	98.60	
8	Thota	150	122	81.30	
9	Kalara	160	150	93.45	
10	Saheba	150	125	83.33	

to prove the efficacy of the proposed method, recognition of Kannada handwritten word with different handwritten styles written by more number of students is very big task for teachers to evaluate and to recognize for OCR. Our proposed work aims at word recognition for answer scripts which help evaluation tool in the future. Most of the words are presented with ottaksharas with one or two touching characters. To segment character individually, contour and bounding box method is used. The classifiers SVM and CNN are used for word recognition. The obtained word recognition accuracy of proposed method is 90%. However, the dataset contains characters with touching with one another which is very hard to classify and recognize the word. The main drawback of the proposed method is whenever more than one ottakshara is present below the character, the accuracy will be reduced and it can be observed from Table 1, i.e., character with ottakshara with varied handwritten style and the answer script containing character without ottakshara gives good results because there is no touching character with baseline.

The results of the aforementioned word recognition techniques and results are displayed (Fig. 2).

By observing from Table 1, we conclude that the word bacteria has least percentage, i.e., 81% as it contains ottakshara with more than one and Puducherry text has high rate of percentage as it doesn't contain any ottakshara, and the overall accuracy is 90% (Fig. 3).

5 Conclusion

We have proposed a method to extract words from answer scripts and recognition of Kannada answers. The proposed works are based on different approaches like bounding box, straight line detection, feature extraction. SVM and CNN are used to train/test recognize handwritten word. Experimentation is carried out on our own

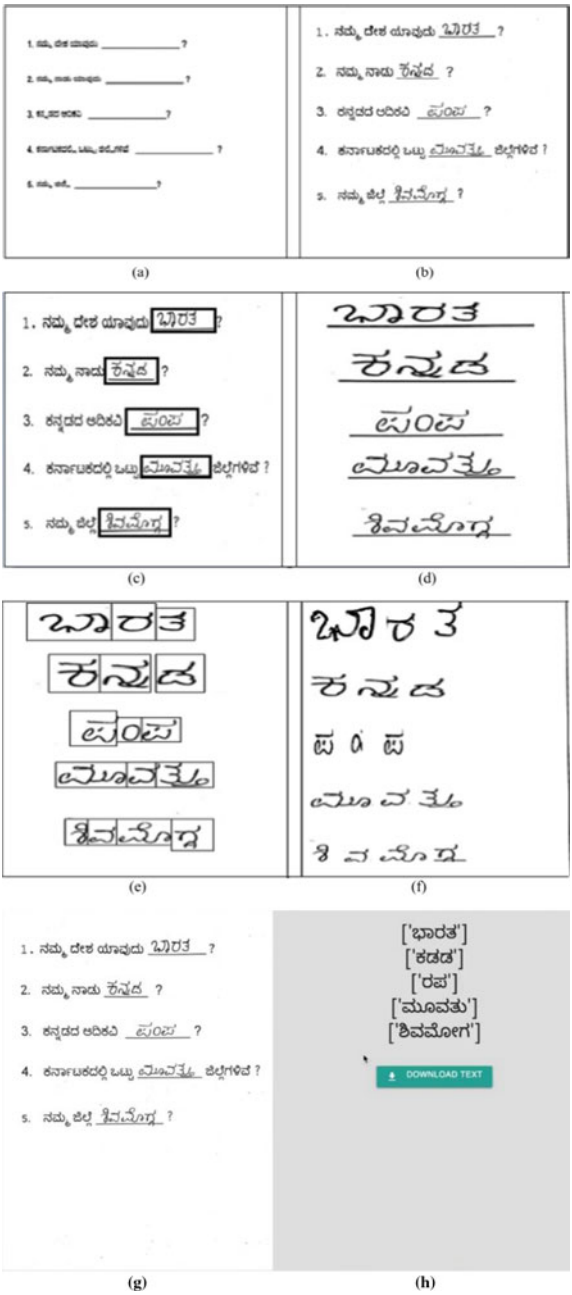


Fig. 2 Representation of word recognition stages

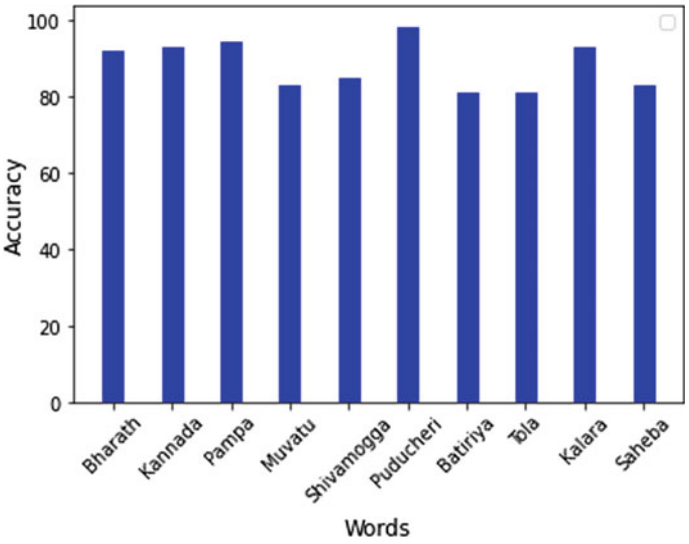


Fig. 3 Representation of word recognition accuracy

dataset containing 200 one-word answer scripts and proposed approach yields better results for word recognition. The overall accuracy of 90.09% is achieved. Hence, we conclude that the words with ottaksharas reduce the accuracy and words without ottaksharas give maximum accuracy.

Reference

1. https://en.wikipedia.org/wiki/Kannada
2. Sureka, T., Swetha, K.S.N., Arora, I., Dr. Mamatha, H.R.: Word recognition techniques for Kannada handwritten documents. In: IEEE, 10th ICCCNT 2019 6–8 July 2019. IIT, Kanpur, Kanpur, India
3. Shakunthala, B.S., Naveen, N.: Unconstrained handwritten Kannada documents leading to line and word segmentation. Int. J. Eng. Res. Technol. (2018)
4. Naveena, C., Manjunath Aradhya, V.N.: Handwritten character segmentation for Kannada scripts. In: World Congress on Information and Communication Technologies, pp. 144–149 (2012)
5. Mr. Chethan Kumar G.S., Dr. Kiran, Y.C.: Kannada handwritten character segmentation using curved line and recognition using support vector machine. Int. J. Combined Res. Dev. (IJCRD) **5**, 641–645 (2016)
6. Mamatha, H.R., Srikantamurthy, K.: Morphological operations and projection profiles based segmentation of handwritten Kannada document. Int. J. Appl. Inf. Syst. **4**(5) (2012). ISSN 2249-0868
7. Azhaguvarthani, G., Ramani, B.: Handwritten text recognition system for English. Int. J. Pure Appl. Math. **117**(16) (2017). ISSN 257-263

8. Sah, R.K., Indira, K.: Online Kannada character recognition using SVM classifier. In: Proceedings of International Conference on Computational Intelligence and Computing Research (2017)
9. Sushma, A., Veena, G.S.: Kannada handwritten word conversion to electronic textual format using HMM model. In: Proceedings of International Conference on Computational Systems and Information Systems for Sustainable Solutions (2016)
10. Patel, M.S., Reddy, S.L.: An impact of grid based approach in offline handwritten Kannada word recognition. IEEE (2014)
11. Aravinda, C.V., Prakash, H.N.: Template matching method for Kannada handwritten recognition based on correlation analysis. IEEE (2014)
12. Pasha, S., Padma, M.C.: Recognition of handwritten Kannada Characters using hybrid Features. In: Proceedings of IET
13. Sen, S., Prabhu, S.V., Jerold, S., Pradeep, J.S., Dr. Choudhary, S.: Comparative study and implementation of supervised and unsupervised models for recognizing handwritten Kannada characters. In: 3rd IEEE Conference on Recent Trends in Electronics, Information International & Communication Technology (RTEICT-2018), pp. 774–778 (2018)
14. Muller, A.C., Guido, S.: Introduction to Machine Learning with Python—A Guide for Data Scientists. O'ReillyMedia, Inc., Sebastopol, CA (2016).
15. Goodfellow, I., Bengio, Y., Courville, A.: Deep Learning. MIT Press (2017)
16. Klep, D., Van Gerven Sanne Schoenmakers, M.: Data augmentation of a handwritten character dataset for a convolutional neural network and integration into a Bayesian linear framework (2016)

Cognitive Internet of Things: Exploring New Horizons for Elderly Care by Training Intelligent Devices

Vandita Srivastava

Abstract Currently, most elder care is provided fully by a caregiver, nurse, family, or friend who remains with the older person in order to do so. An increasing number of nuclear families means that a person cannot always remain with the elder. To combat this problem, they are usually sent to old age homes. Some, however, choose to live in their own homes. This means that lots of older people are living alone with little to no supervision. This is not ideal as when they face a problem, there is sometimes no way to inform someone about it. In this paper, we argue about the need for such a system which monitors the elderly and informs a responsible adult if and when a problem arises so that proper care can be taken even when a person is not present with the elder throughout the day. I first present the urgent need for such a system, inspired by the day-to-day struggles of a typical elder. I then provide a comprehensive understanding of CIoT and machine learning. I look into the various issues faced by them and how to properly combat them by using well-trained intelligent devices formed by cognitive internet of things. CIoT provides us with multiple options and opportunities to use objects in our everyday life in an innovative way. Various services are provided by many companies which can be used to implement an intelligent system to drastically change the way we take care of our elderly.

Keywords Cognitive IoT · Machine learning · Intelligent devices · CIoT platforms

1 Introduction

Joint families are becoming increasingly rare with nuclear families on the rise. This means that more and more elderly people are having to live on their own. According to the 2011 census, India has a population of 104 million elderly people, comprising of 51 million males and 53 million females [1] and this number is only increasing. Of this, more than 15 million [2] have been recorded to be living alone. They have

V. Srivastava (✉)
Centre for Converging Technologies, University of Rajasthan, Jaipur, India

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_100

to face the challenges in their daily life alone and if any problem arises, they have no way of helping themselves or contacting a responsible adult.

We need to work towards improving the lives of our elderly. The new and emerging field of cognitive internet of things provides us with multiple opportunities to do just this. CIoT makes use of the internet of things combined with intelligent thought process to make cognitive decisions.

We can make use of sensors and actuators to create a system that monitors the activities of the elderly person.

Consider a scenario where a grandma is living alone in her house. She has a daily set routine of waking up at 5 am everyday then going to the kitchen to make herself a pot of tea at 5:30 am after which she goes and sits in the temple room at 5:45 am till 7 am. She has her whole day planned put and barely ever deviates from this set pattern. Now one day, she wakes up as usual, goes to make tea, and due to some unfortunate circumstance, collapses on the kitchen floor. She has no way of letting anyone know about this until and unless someone comes to her by chance on their own. She has basically been rendered helpless.

Now consider another scenario in which when she collapses, a sensor detects that something is not quite right. It detects that her heart rate is elevated, or that it is now 6 am and she has still not made it to the temple room. It can deduce that this is abnormal behavior and can send a message to the concerned responsible adult with the help of an actuator. This way, help can arrive for her in a much faster and efficient manner.

The later scenario makes use of the CIoT to make an intelligent decision based on its knowledge gained with the help of machine learning. That scenario is obviously much more preferable.

Much of the technology-based research done till now in elderly care is focused on IoT. Not much has been done in it by using CIoT.

The rest of the paper is organized as follows. Section 2 discusses and aims to understand IoT, CIoT, machine learning, and how they all are connected and must work together. Section 3 looks in depth into the need of this technology and the devices that can be used to achieve it. Section 4 talks about the various services provided by multiple organizations for machine learning and IoT and the process by which this system will work using those services. Section 5 lists out the applications of this idea. Section 6 talks about challenges which comes with this technology. Lastly Sect. 7 contains the concluding remarks.

2 IOT, CIoT, and Machine Learning

2.1 Internet of Things

The history of Internet goes back to the development of communication between two computers through a computer network in the late 1960s. Since then, the evolution of

the Internet has passed three main phases: Internet of Computers, Internet of People (mainly via social networking), and Internet of Things (including computers, people, and any other physical/virtual objects) [3].

It allows us to communicate with each other because the computers are capable of sending and receiving information in the form of signals, over the network.

Using IoT we are now able to communicate with our day-to-day devices. It allows the integration of multiple otherwise mundane devices to the computer-based system. The expression "Internet of Things" (IoT), coined back in 1999 by Kevin Ashton, the British technology pioneer [4]. Internet of things term represents a general concept for the ability of network devices to sense and collect data from around the world, and then share that data across the internet where it can be processed and utilized for various interesting processes [5]. This means that we can use IoT to establish communication between devices which does not come in the category of computers, and then make them perform various tasks. For instance, IoT is being used as security system, pet care system, driverless cars, etc.

Most research in IoT, however, has been focused on tasks which does not require intelligent decision-making. This is done with the use of Cognitive Internet of Things.

2.2 Cognitive IoT

Cognitive IoT can be considered as the very future of IoT.

CIoT revolves around the concept of taking IoT to the next level by attempting to add "thinking". Cognition refers mainly to understanding, reasoning, and learning. In cognitive computing we try and integrate these into the computer system. Understanding, for computers, means that it takes large amount of heterogeneous data and finds meaning from it. Reasoning relates to the ability of the computer to derive answer to some problem based on the surrounding conditions and information. Learning refers to the ability of the computer to automatically infer new knowledge from data. The aim of CIoT is to combine all these things so that our everyday devices are able to make decisions on their own.

In CIoT, objects interact with each other and follow a context-aware perception–action cycle, while using the methodology of understanding-by-building to learn from both the physical environment and social networks. They then store the learned knowledge in databases, and adapt themselves to changes or uncertainties via resource-efficient decision-making mechanisms [6].

A cognitive management framework for IoT is proposed, in which dynamically changing real-world objects are represented in a virtualized environment, and where cognition and proximity are used to select the most relevant objects for the purpose of an application in an intelligent and autonomic way [7]. This combats the issues of IoT which prevent it from responding to an ever-changing environment.

We can apply Cognitive IoT by using the sensors that are already ubiquitous in our everyday life to compute more things and work together. This leap can make our *things* understand and interact with us in our language(s) [8].

In the cognitive process, the multi-domain cooperation can increase network capacity and the machine learning can enhance the intelligence for future.

This is a newly emerging field and lots of research is going on about its implementation and uses. It has a lot of potential and will be a great step forward in the betterment of the standard of living.

This capability of decision-making is given to the computer with the help of machine learning.

2.3 Machine Learning

Machine learning is a tool to make smart computers. It is the process of making an intelligent device that is capable of taking decisions, making predictions, and much more.

Machine learning is programing computers to optimize a performance criterion using example data or past experience [9].

The process of machine learning basically has two steps

1. Training, i.e., concept generation
2. Testing.

Training can be done in a supervised or unsupervised manner, depending on the need (Fig. 1).

The need for machine learning is now more than ever due to the flood of available data because of the internet. This "big data" is nearly impossible to process by humans due to limited time and resources. Machine learning allows for a much easier way to deal with it, especially due to increased computational power, new algorithms, and support from multiple industries.

Examples of ML applications include-

1. Learning association: This means that the computer learns to form association.
2. Classification: The computer leans to divide given data into classes based on some attributes.
3. Pattern recognition: The computer learns to look and find particular patterns in the data.

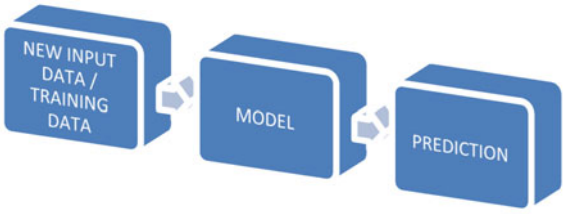


Fig. 1 Pipeline for machine learning model

Pattern recognition can be used in medical learning. It can be applied to CIoT to make devices capable of understanding a set pattern and responding if there is any sort of deviation in the pattern.

2.4 Need for Machine Learning in IoT

Machine learning is a crucial part of Cognitive Internet of Things.

It allows us to do a number of things, such as

1. Develop systems that automatically adapt & customize themselves as per user's requirements.
2. Discover new patterns from large databases, specifically with the unstructured data.
3. Ability to mimic humans and replace certain monotonous tasks which require some intelligence.
4. Develop systems that are too difficult because they require specific skills/look at big data.

By using CIoT with machine learning we can design specific systems and network of devices which all work together to perform any specialized task. The system will use machine learning to think for its own. It will develop a cognitive thought process and do things which till now require human intelligence. This is especially useful in the medical field where it can be used for diagnoses purposes and as a security system, or as an alert system.

We can apply these principles in elderly care to make a specialized network of devices geared toward helping the elders.

3 CIoT and Elderly Care

According to a recent survey done in some major states of India, 4.8% people above 60 years live alone. 9.8% of them are living with one other older person, generally their spouse. 15.0% people live in homes with no one below 60 years of age [2].

All of them deserve a better standard of living. We want to make their lives easier by using CIoT to make an eldercare system.

3.1 Intelligent Devices

The sensors and actuators used in the system are intelligent agents or devices which have been trained to collect information from their environment and respond according to it.

The term IoT is mainly used for devices that wouldn't usually be generally expected to have an internet connection, and that can communicate with the network independently of human action [10].

For CIoT we need to expand this definition further to mean those devices that have the capability to think and learn as well as having normal computing abilities [11].

Intelligent devices are different from normal computer-coded programs that are abundantly used everywhere, and are currently used in IoT. Intelligent devices are able to learn and respond. Normal computers can only do what has been coded in them. A program is written as lines of code by a programmer. Its functionality is limited for only those scenarios which have been coded into it. They cannot conceivably think of every possible scenario and provide the action needed to be performed with it.

Intelligent devices don't need this kind of hard coding of every single possible case to be able to properly respond to it. With the help of machine learning, they develop a cognitive thought process which can deduce the right action on its own when faced with unknown data and new scenarios. It uses past knowledge and experience to automatically conclude the correct course of action.

These intelligent devices can be connected to form a network of such tools to create CIoT system.

For our elder care CIoT system, we need the devices in our network to be able to communicate with each other as well as with humans and work for, and in harmony with, the elder.

Wearables are a type of intelligent device used in IoT. There are many existing wearables that can be modified to become CIoT based intelligent devices. Smart bands, specially fitness bands have a good future in eldercare system.

Fitness bands are already quite advanced and we can train them to become CIoT and then use them for eldercare. Many also already offer the functionality of being connected to another device, most commonly a smart phone. We can expand on this and make it connected with other sensors placed in each room as well as the smart phone, which can act as the messaging system and thus as an actuator. It sends a message to the responsible adult when abnormal behavior is detected.

We require another intelligent sensor to be placed in each room to monitor the times the elder enter and exists the room. It will be connected to the band in such a way that whenever they are in the same room, they communicate with each other and the time is recorded. For this, an intelligent Bluetooth device can be used.

4 Process by Which This System Will Be Implemented

4.1 Platform

IoT platform requires [12].

1. Sensors or controllers.

2. A gateway device to aggregate and transmit data back and forth to the data network.
3. A communications network to send/receive data.
4. Software for analyzing and translating data.
5. The end application service, which creates much of the value.

A sensor is a device that receives and responds to a signal. Once a sensor detects one or more of these signals (an input), it converts it into an analog or digital representation of the input signal [13].

An actuator is a part or element of machines that can make the machine responsible for moving or controlling a mechanism or system [14].

A gateway device is a gate to the network. All data coming in and out of the network must flow through it. It is considered to be on the "edge" of the network.

A communication network is established between all the devices. Many different wireless communication technologies and protocols can be used to connect the smart device such as Internet Protocol Version 6 (IPv6) [15].

Software for analyzing and translating data is the cloud and the machine learning model we will use to make cognitive designs based on all the data collected by the sensors.

The end application service is the final physical product created to cater to our needs.

Platforms allow us to manage the system by providing all the required resources to connect devices, acquire data, process the data, etc.

Here is a list of 5 promising platforms that are being considered for our CIoT application [14]:

1. Amazon: In 2014, Amazon acquired 2lemetry, an IoT startup with a strong M2M focus, which emerged as AWS IoT service.
2. IBM: Enterprise IoT is the cornerstone of IBM Watson platform.
3. PTC: PTC acquired ThingWorx in 2013 to bring enterprise IoT capabilities to its platform.
4. GE: GE's Predix platform targets industrial IoT deployments.
5. Microsoft: Microsoft has built Azure IoT Suite as a turnkey solution that customers can deploy.

From all these, probably the best platform to use for CIoT would be IBM Watson.

4.2 Building the Cognitive IoT System

Cognitive IoT employs the use of machine learning and has 2 basic requirements:

1. Algorithms-these learn the behavior required to be done by our system
2. Storage-big data will be used to implement our system which requires immense storage capabilities.

The storage and data processing capabilities are usually provided in cloud.

In CIoT, the machine learning should be done using edge computing architecture. This means that the computing happens at the edge gateway near the sensors and actuators. This is important because of two main reasons- latency and transfer cost.

When computation is done on the cloud, it takes time for the data to be transferred to and from the cloud. This time maybe just a few seconds but even that can be damaging in a practical situation. For instance, if a grandma falls down, we need an immediate response. We cannot tolerate any amount of latency. Additionally, since the amount of data being generated by our sensors is very high, the time to transfer it to the cloud completely will also become too much. This can be avoided by edge computing.

4.3 Edge and Cloud Computing

The first step is to connect the edge gateway to the cloud. The de facto standard for this connection is using MQTT and an MQTT message broker. MQTT is a machine-to-machine (M2M)/"Internet of Things" connectivity protocol [15].

Publish/subscribe models facilitate deployments of *n–n* connection scenarios. Sending data from the gateway to the queue is not sufficient, so an ETL (Extract, Transform, Load) component is needed. To complete the architecture, cognitive API services can be called in real-time or during batch processing [16].

To further optimize the model, the help of various complex machine learning algorithm is taken. The best algorithms for this CIoT include those made for time-series forecasting, anomaly detection, and optimization.

For our purposes, the anomaly detection algorithm will be best.

The IBM Watson platform which includes the IBM Cloud and IBM Watson Studio can be used to make the CIoT system for elder care.

4.4 Working and Implementation

The system will be installed in the home of the elderly. They will be given the band to wear on their wrist and Bluetooth sensors will be put in each room of the house. The band will continuously transmit health-related data of the elderly to the cloud storage. When they will enter any room, the band and Bluetooth device will start interacting and communicating with each other and the time of entry and exit will be noted. The system is already trained and has already learned all the normal conditions in which the elderly person lives, with the help of machine learning. When any sort of abnormality takes place, the system detects it with the help of the anomaly detection algorithm. It then quickly processes the data and makes a cognitive decision. It sends a signal to the actuator to send an alert message to a responsible adult. The responsible adult will then take the appropriate action.

The system will have a list of responsible adults and emergency contacts from which it will cognitively decide which adult to send the message to. This will depend on the severity of the problem and the location of the adult-how far from home the adult is right now, and so, how soon he might be able to respond. Additionally, if the first responder is not giving any response due to some reason, the system will direct the message to the second, and so forth.

Suppose the grandma is sitting at home and she starts feeling feverish. Her temperature is a little up. Now, this is not a high alert situation but it is still noteworthy. So the system sends a message to her son, who reads and acknowledges the message. The son then takes the required action.

If the grandma has a sudden heart attack, this is a very high alert situation and needs instant medical attention. The system will then immediately first contact a hospital and then contact the respondents.

5 Applications

We have mainly focused the application of our product on elderly people who live alone. In addition, this system can also be used with elderly people who live with their families but their families have to go to their jobs, leaving them in the house for the majority of the day. Further, even for elders who have someone with them in the house, this system can be implemented. This system monitors the health status of the elder and that is always a useful information to collect and get notified about.

Apart from these, this system can be used for people of all ages if needed.

6 Further Challenges

It is not a trivial task to detect every action of an elderly person, since a lot of complex actions and diverse surroundings need to be taken into account. This is made even more challenging by adding both indoor as well as outdoor activities. For the purpose of this work, only the home environment is considered, like “Ability to cook food,” “Housekeeping,” “Laundry,” which can produce a lot of information about the physical and cognitive status of the elderly” [17].

We need to further expand the scope of the eldercare system so that it covers not only their home but all the places outside it as well.

7 Concluding Remarks

Our elders will benefit greatly from the use of CIoT. The world is moving forward fast and technology is advancing at a rapid pace. So, it is only fitting that we take our elders with us on this journey into a new and better future.

CIoT is still in its early stages of development. We need to make new algorithms and find new ways of implementing them.

Intelligent devices are the future in every field, and they are especially beneficial in the medical field as they can be used to identify patterns and abnormalities. They should be exploited to their full potential for the benefit of everyone.

This venture is just the start in modern healthcare and eldercare system. With advancement in machine learning and CIoT, new and better features are sure to be added in this field.

References

1. Senior Citizens—status in India Homepage. https://vikaspedia.in/social-welfare/senior-citizens-welfare/senior-citizens-status-in-india. Accessed 05 Oct 2018
2. Nagarajan, R.: 15 million elderly Indians live all alone: census. 1 OCT 2014 Homepage, https://timesofindia.indiatimes.com/articleshow/43948392.cms?utm_source=contentofinterest&utm_medium=text&utm_campaign=cppst. Accessed 05 Oct 2018
3. Wu, Q., et al.: Cognitive internet of things: a new paradigm beyond connection. IEEE Internet of Things J. **1**(2), 129–143 (2014). https://doi.org/10.1109/JIOT.2014.2311513
4. Kramp, T., van Kranenburg, R., Lange, S.: Introduction to the Internet of Things. In: Bassi, A., et al. (eds.) Enabling Things to Talk. Springer, Berlin, Heidelberg (2013). https://doi.org/10.1007/978-3-642-40403-0_1
5. Pramanik, P.K.D., Pal, S., Choudhury, P.: Beyond automation: the cognitive IoT. Artificial intelligence brings sense to the internet of things. In: Sangaiah, A., Thangavelu, A., Meenakshi Sundaram, V. (eds.) Cognitive Computing for Big Data Systems Over IoT. Lecture Notes on Data Engineering and Communications Technologies, vol. 14. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-70688-7_1
6. Vlacheas, P., et al.: Enabling smart cities through a cognitive management framework for the internet of things. IEEE Commun. Mag. **51**(6), 102–111 (2013). https://doi.org/10.1109/MCOM.2013.6525602
7. Matthews, S.: What is cognitive IoT? Homepage, https://www.ibmbigdatahub.com/blog/what-cognitive-iot. Accessed 06 Oct 2018
8. Zhang, M., Zhao, H., Zheng, R., Wu, Q., Wei, W.: Cognitive internet of things: concepts and application example. Int. J. Comput. Sci. Issues **9**, 151–158 (2012)
9. Alpaydin, E.: Introduction to machine learning. In: Al-Turjman, F. (ed.) Cognitive Sensors and Iot Architecture, Deployment and Data Delivery, 2nd edn. The MIT Press (2010)
10. What is the IoT? Everything you need to know about the Internet of Things right now Homepage, https://www.zdnet.com/article/what-is-the-internet-of-things-everything-you-need-to-know-about-the-iot-right-now/. Accessed 03 Feb 2020
11. Sarker, D., Subramanain, S.: Cognitive IoT incorporating intelligence in building smart environment. In: IOP Conference Series: Materials Science and Engineering, vol. 263, pp. 042012 (2017). https://doi.org/10.1088/1757-899X/263/4/042012
12. IoT Plateform: What They Are and How to Select One Homepage. 03 Aug 2016. https://www.link-labs.com/blog/what-is-an-iot-platform, Accessed 07 Oct 2018

13. Al-Sarawi, S., Anbar, M., Alieyan, K., Alzubaidi, M.: Internet of Things (IoT) communication protocols: review. In: 2017 8th International Conference on Information Technology (ICIT), Amman, 2017, pp. 685–690. https://doi.org/10.1109/ICITECH.2017.8079928
14. MSV, J.: 6 Enterprise IoT platforms to consider for your digital transformation strategy homepage. https://www.forbes.com/sites/janakirammsv/2016/11/19/6-enterprise-iot-platforms-to-consider-for-your-digital-transformation-strategy/#76ac70901898. Accessed 07 Oct 2018
15. Mqtt. https://mqtt.org. Accessed 07 Oct 2018
16. https://developer.ibm.com/tutorials/iot-cognitive-iot-app-machine-learning/
17. Mulero, R., Almeida, A., Azkune, G., Abril, P., Arredondo, M.T., Paramo, M., Patrono, L., Rametta, P., Sergi, I. An IoT-aware approach for elderly-friendly cities. In: IEEE Access, vol. PP. pp. 1–1 (2018). https://doi.org/10.1109/ACCESS.2018.2800161

Blockchain-Sidechain Based Data Storage for Reimaging Electronic Health Record via Optimized Interplanetary File System

R. Deepa and Meenakshi S. Arya

Abstract Modern healthcare systems are extremely complicated and expensive. This can be minimized through better monitoring of health records, using Blockchain technology. The first implementation of Blockchain was done to provide distributed records of money-related transactions which were not reliant on unified authorities or financial institutions and seem to provide the way for trend in the traditional healthcare industry benefitting by its unique features such as data isolation and transparency. The system proposes a Sidechain-Blockchain Technology to store and manage patient data and authentication of the same. In addition to that, the main blockchain can transmit or share the patient information record with other smart systems. The system is implemented using optimized Interplanetary File System in order to provide faster and safe access of information between the main Blockchain and side blockchain using Ethereum Dapp. The proposed model uses different development environments to build a highly effective EHR that is not controlled by a single point of authority and uses a decentralized system for storage. It's worth noting that the implementation of Blockchain-sidechain on top of ethereum results in patient-centered access to the data and improves scalability.

Keywords Blockchain · Ethereum · EHR · Healthcare · Side-chain · IPFS

R. Deepa · M. S. Arya (✉)
Department of Computer Science and Engineering, SRM Institute of Engineering and Technology, Vadapalani, Chennai, India

R. Deepa
e-mail: deepar@srmist.edu.in

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_101

1 Introduction

An Electronic Health Record (EHR) is a digital version of a patient's medical records. It was introduced earlier in hospitals to do away with organizational problems associated with paper records. However, it underwent several changes from being self-managed system to a web-based/cloud-based system. These changes further introduced risks to the data and the system itself such as data integrity issues, privacy issues, single point of failure, and so on. These disadvantages called for a modern, secure solution that maintains Electronic Medical Records (EMRs) in the healthcare departments using Blockchain technology. Blockchain eliminates data integrity and privacy issues as the information stored on a Blockchain are immutable and encrypted. It further solves the single point of failure being a decentralized platform. All these features of Blockchain thus make it the perfect technology to implement EHR systems. In this paper, a system with Sidechain storage has been introduced where the actual data is stored offline and the hash of the data is stored in the main Blockchain. This way, the proposed distributed storage ensures safety against manipulation and confidentiality of the data. In the current times, the hospitals collect the patient's data, process it and then store it in the system in order to improve information safety, security, healthcare quality, and emergency response abilities [1, 2]. For effective treatment and care of the patients, secure data sharing is very important as patients tend to switch healthcare providers during the course of their treatment. It is thus imperative for the service providers to be updated with the latest information about the health condition of the patient [3]. By establishing a Blockchain at each level of stakeholders, the method altogether manages the provider's system authentication process and stores the data for all the smart systems in separate chain.

The contribution of this paper is two-fold. The paper proposes an enhanced sidechain structure for stakeholder's authentication in the peer-peer network. Rather than storing and retrieving the whole main Blockchain after every block generation practice as conventional sidechain technology, the proposed structure uses a consensus algorithm by saving a reference mainchain block to demonstrate the presence of the patient's data. The proposed structure has lesser storage requirements and is more effective while searching for the target patent's information. Additionally, with an aim to provide maximum efficiency and improve the performance of the EHS, an optimized IPFS protocol based on dynamic analysis of the trust value of the target system has been proposed.

The paper is organized as follows. Section 2 briefly introduces the available literature on the proposed topic, theoretical details of the Blockchain technology. An overview of the Sidechain-Blockchain technology and its dependencies is discussed in Sect. 3. System Design and Architecture is presented in Sect. 4. Section 5, describes the implementation of the proposed Blockchain-Sidechain-based authentication scheme with an optimized IPFS protocol. The results are discussed in Sect. 6, the results are presented and concluding remarks and future scope is provided in Sect. 7.

2 Literature Survey

Blockchain innovation empowers a distributed network background with no requirement for a middle entity. Transactions are safe as well as dependable due to the utilization of cryptographic standards. Satoshi Nakamoto invented this technology for his famous digital currency or crypto-currency work, that is, bitcoin. Though the main use of this technology was to address bitcoin's double spending issue however soon this innovative technology found its use in various other applications [4]. The blocks (building blocks of a blockchain) are connected with other blocks like a chain. The bocks are continuous and each block is used for storing transactions [5]. For information distribution and ownership sharing of a piece of distributed information, a decentralized approach is used. The main protection mechanism for transaction batches being carried by a Blockchain is hashing, and these batches are handled by P2P networks. Blockchain has the certain advantages of being secure, anonymous, and has data integrity with no intermediatory intervention [6] which proves it to be a right method to store patient's medical records, as the top priority in healthcare industry is the security and anonymity of patient's medical data. Blockchain technology as a feasible solution to healthcare systems has been justified by many researchers.

Lately, Blockchain innovation has turned out to be popular, trendy and has percolated in various areas [7], generally because of the prevalence of digital forms of money. One prominent area where in Blockchain innovation has gigantic potential is healthcare as these systems demand patient-centric techniques to deal with Electronic healthcare systems with an aim to increase the interoperability and precision of the EHRs.

The broad application areas of use of Blockchain in healthcare are Electronic Medical Records, Tracking and Tracing Medical data, Artificial Intelligence for the analysis of complex medical data, Secure and Guaranteed Payments, Medical Research.

The most promising and prominent implementation of the Blockchain framework is Ethereum and IPFS. Ethereum was presented in the paper [8], the paper also discussed some drawbacks of the scripting language of Bitcoin. Modified GHOST (Greedy Heaviest Observed Subtree) protocol is the basis of the consensus within the Ethereum network [9], it is designed to fix the network's problem of stale blocks. The stale blocks will arise as a result of more computing power at the disposal of one group of miners as compared to others, i.e. blocks from one pool contribute more to the network thereby resulting in the problem of centralization which is against the whole idea of a Blockchain. GHOST protocol integrates such old blocks into the longest chain calculations. Ethereum is based on the modification of the GHOST where uncles up to seven generations are included [10].

In [11] the authors explain the use case of Ethereum Blockchain in creating a decentralized application executing a smart contract. It focuses on the elimination of TTP (Trusted Third Party). One of the main components of the applications is the smart contracts hosted on top of the Blockchain, the local Ethereum client, and a web app. The GUI for the local Ethereum client is provided by the web app, which in turn communicates with the smart contract on the Ethereum Blockchain.

The initial intention of designing Sidechain was to facilitate bitcoin and other cryptocurrencies to exchange money between different networks [12]. In [13] the authors used a sidechain for permitting the important data to transfer from one chain to others more securely at very less time thereby reducing the overhead of the main blockchain, which enables the system to increase both agility and autonomy of using multiple networks. In order to share data among main blockchain and sidechain-blockchains, IPFS protocol was applied by authors in [14].

One of the applications of this technology is being worked on by the True field team. The practice of medical record-keeping reduces healthcare cost and also assures quality health care. The True field platform aims at creating a patient-centric approach to storing of medical information. They plan to achieve this by creating unique digital medical IDs for each user [15, 16].

3 Blockchain Technology Background

3.1 Blockchain Architecture

A Blockchain is a distributed, shared, and immutable digital database that retains track of transactions that take place across the network. A block is a sequence of transactions. It contains the transactions chronographically. A Blockchain is made of a series of blocks. The blocks are chained together as each block contains the hash of the previous block(except the first block), thus making a series [17]. The first block in a Blockchain is also known as the genesis block. A "node" is also stated as the shareholder of the Blockchain. Every node in the network has two keys, a public key and a private key. The private key is used for decrypting the messages and it permits a node to read that message. The public key is used for encryption of messages sent to a node. Some nodes also act as the miners of the Blockchain ecosystem. The miners are referred to as special nodes that use an algorithm to validate the transactions submitted by other nodes and make a block that will be a candidate for the next block to be added in the Blockchain (Fig. 1).

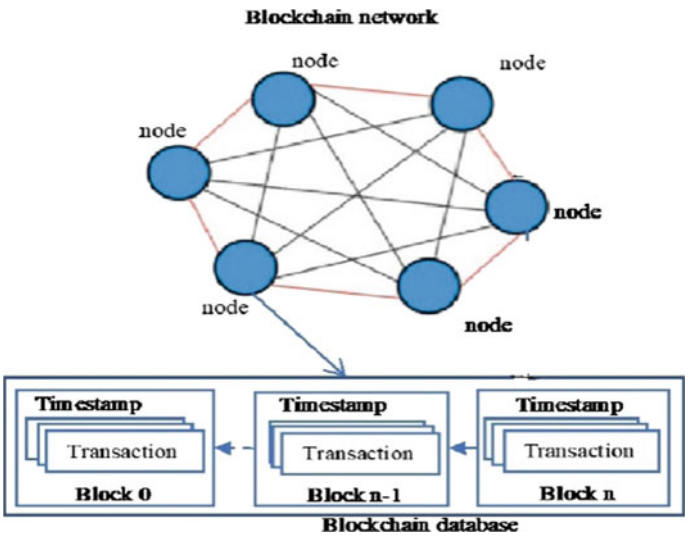


Fig. 1 Typical Blockchain architecture

3.2 IPFS

IPFS or Interplanetary File System is a protocol that uses P2P network used to transfer data storage from one node to another node. One of the basic strategies of IPFS is secure data storage, which has become right option for securing more sensitive and crucial data. The users can share large file of any format on the network through IPFS gateway [3]. IPFS protects stored data from any kind of alteration. When the users try to share the data on the network, IPFS will generate Hash value and Node ID (cryptographic hash) to the corresponding data to avoid redundancy. Node ID/CID (cryptographic identifier) can be shared in public network to provide access to the data to another nodes present in the network. The hash value generated can be stored on the DAPP to reduce the comprehensive computational tasks over the Blockchain. IPFS network is constructed on top of a data structure called as IPFS object which holds data and its link in it. Data can be either structures or unstructured binary data and the data get linked in an array. Work flow of IPFS protocol can be described as follows: Redundant data's are not allowed to exist on the IPFS network. In node ID the network stores data and its linked information in the form of index value.

3.3 Smart Contracts

A smart contract is a computer program designed to automatically facilitate, check or execute a contract's negotiation or execution. Smart contracts allow the performance of credible transactions without third parties [18]. These transactions are traceable and unalterable. They permit trusted transactions and agreements to be carried out among disparate, anonymous parties without the need for a central authority, legal system, or external enforcement mechanism (Fig. 2).

4 System Design and Architecture

The systems proposed earlier show Blockchain implemented EHR that concentrates on access-control and data sharing. The given systems do not really concentrate on an efficient storage, storing the data in the main Blockchain network. Therefore, a system with a side-chain storage is proposed where the actual data is stored offline and the hash of the data is stored in the main-Blockchain. IPFS is used for this purpose of side-chain storage and we take advantage of its scaling features. For the implementation of the proposed system Ethereum is used.

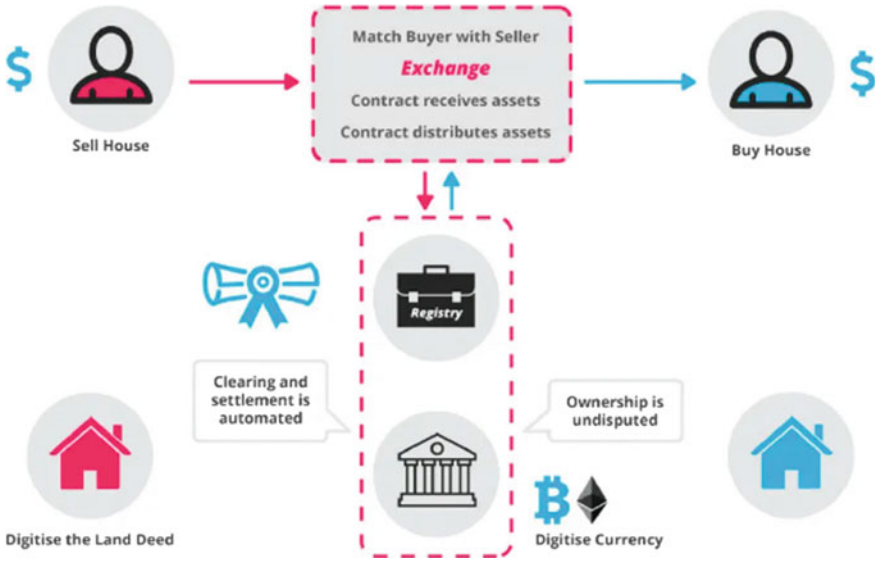


Fig. 2 How smart contracts work

4.1 System Design

There are three entities or components within the proposed structure or scheme. The conceptual system consists of participants who may be patients, clinicians, admins, and lab technicians. There are different access rights for different participants namely—Read, Write, Modify, and Revoke.

4.1.1 Network

This consists of an interconnection of peers/nodes that enables the transfer of data without any central authority playing the middle man. This ensures complete privacy of data. This creates a distributed system where no one controls the network. All the nodes in the network are treated equally and have equal powers."

4.1.2 Transactions

The system has the following functionalities:

Create Medical Record (By patient/Clinician): This adds a medical record to the existing patient records. For a clinician to be able to create records, the need the access permission of the patient beforehand. Only those clinicians with access permission

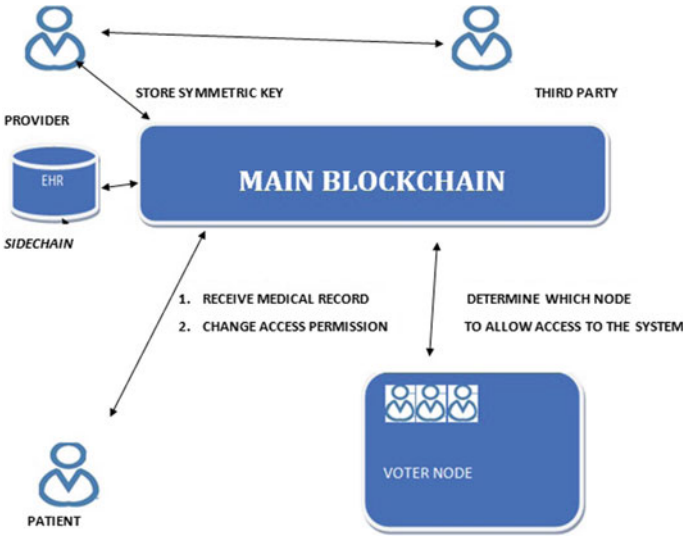


Fig. 3 Proposed sidechain based EHR model

can create medical entries. A patient, by default has the access right to add medical data to their profile.

The specific medical records of the patient are stored along with the hash value and CID generated by IPFS, which includes the uploaded file containing the patient's test reports or other medical information.

Grant Access (To Clinician/Lab): This functionality allows patient to grant access to the above-mentioned records and information to the clinician/Lab.

Revoke Access (To Clinician/Lab): This would let the patient revoke the access of clinician/Labs to add or view data after a certain point or after their treatment is over.

View Medical Record (By patient/Clinician): This operation Is used both by doctors and patients. This function also requires access permission of the patient. Only after the patient grants access to view, the participants can access those data. The patient has access rights to view their own data. This data is stored in a decentralized way in IPFS and accessed through the hash of the data (Fig. 3).

5 System Implementation

The implementation of the business logic of the system is done through smart contracts. It consists of three smart contracts:

Migration contract: This is present in all Ethereal DApps that lets us migrate the contracts to the Blockchain via the truffle framework.

Record contract: This defines the access rules and has information regarding which users have what access rights over the record.

Patient contract: This consists of the logic behind the operations that the patient has access to. It also initializes the patient address.

5.1 Proposed Algorithm

Algorithm 1 Grant Access Permission to Authentic Third Party By Patient At Blockchain Mainchain

```
Create new record for patient registration
patient.sol
pragma solidity ^0.5.0;
contract patient {
address patientAddress; mapping (address=>bool) canCreate;
constructor()
public {patientAddress = msg.sender;}
function
giveWritePermission(addresswriter)
public{
require(msg.sender = = patientAddress,
"You don't have rights!!");
require(canCreate[writer] = = false,
"You alredy have write permission!!");
canCreate[writer] = true;}
Patient revokes access from third party
function
revokeWritePermission(address writer)
public{
require(msg.sender = = patientAddress,
"You don't have rights!!");
require(canCreate[writer] = = true,
"No need to revoke permission!!");
canCreate[writer] = false;}
function canCreateRecords(address writer)
public view returns(bool){
return canCreate[writer];}
end of algorithm
```

CID represents the list of authorized Clinicians and PID represents the list of authorized patients; msg.sender in Ethereum using Solidity language which denotes the Ethereum address of the user/sender of the request/message.

6 Results and Discussion

Access control is achieved where in only those clinicians who are granted access to a patient's record can access its contents. When a fraudulent access is attempted by someone without access permission, records are not displayed and an appropriate error message is shown. Thus, data security is achieved. The record fetching is also faster while using IPFS than a centralized storage. Thus, in addition to providing off-chain storage, IPFS also makes storage and retrieval of records faster and more efficient. The usage of IPFS for storage also avoids single point of failure being a decentralized storage solution (Figs. 4, 5, 6 and 7).

With the built-in smart contract compilation, linking, deployment, and binary management the transaction for new records can be created, viewed and even the patient can change the access grant based on the provider request. It also provides scriptable deployment and migrations framework. It gives network management for deploying public and private networks. It provides interactive console for direct contract communication. It instantly rebuilds assets during development. We can also use an external script runner that executes scripts within the premises. we have used Ganache. Ganache is a personal blockchain that runs on the local machine and can be used to deploy contracts, run applications, and test them. It is available both as desktop application and commandline tool. By default, it provides 10 accounts with 100 ethers each that can be used for transactions. We can also see block history, events, logs.

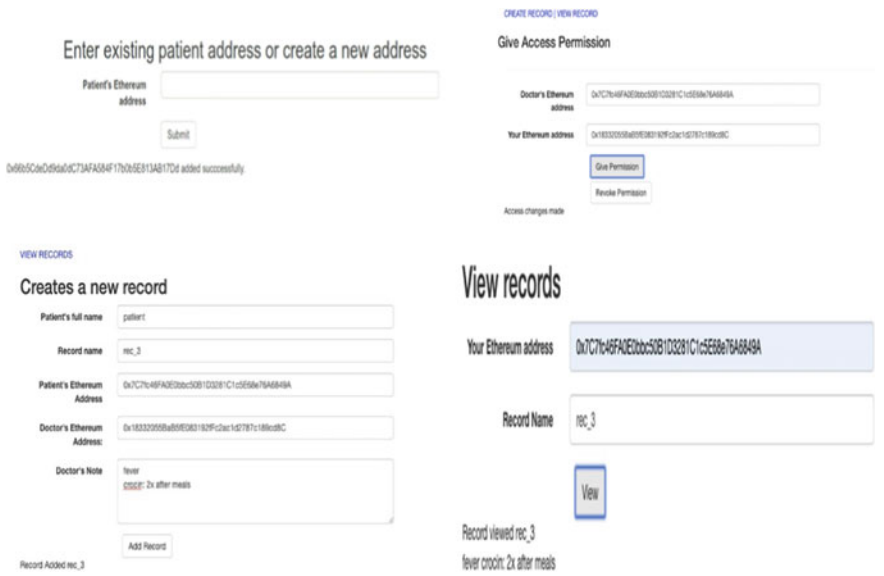


Fig. 4 Creating new record and providing access grand to the providers using Sidechain Blockchain via IPFS

Fig. 5 Sample transaction receipt for adding new file/medical record

Fig. 6 Migration and deployment of smart contracts in the network

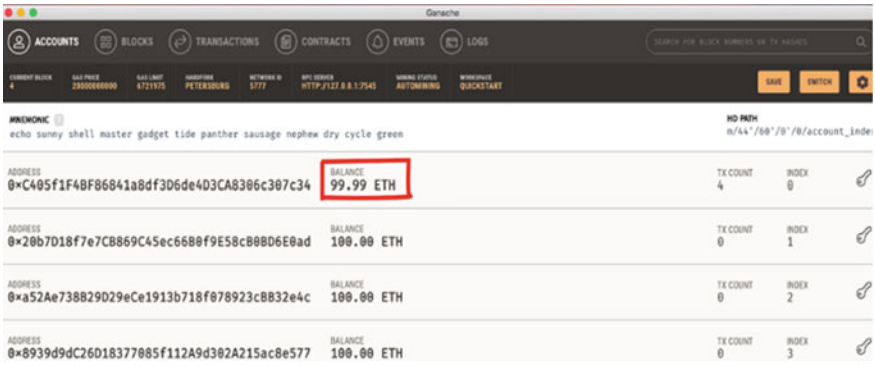


Fig. 7 Deploying the contracts on Ganache

6.1 Discussion

In addition to the proposed EHR system and the potential of Blockchain in EHR, there are a few challenges that could be carefully addressed and rectified.

6.1.1 Medical Data Upload Privacy

While uploading medical data, the Blockchain faces few critical privacy issues. Even though the Blockchain platform have more restriction on entities to keep track of their transaction records, intrusive miners can gather personal information such as user location or usage pattern information during the entire process. In this case novel strategies which combine uploading optimization and user data protection will be very useful. Especially, reinforcement learning (RL) uploading techniques have just emerged as a promising solution to solve critical problems in healthcare.

6.1.2 Scalability

In modern healthcare scenarios, there are many stakeholders who want to participate in collaborative electronic healthcare systems, i.e. mutual healthcare systems between multiple hospitals which need to be highly scalable and manageable IoT architecture that ensures dependency and robust interconnection of multiple IoT devices within the Blockchain network. Besides, different healthcare providers can have different roles based on their functionality, such as doctor, clinicians, physicians, nurses,

Supplychain, insurance companies. Therefore, there is a requirement to extend the present EHR sharing system, aiming to adapt to scalable electronic healthcare systems while promising high quality of user experience and e-health sharing reliability.

7 Conclusion and Future Scope

As discussed earlier, there are various benefits of using Blockchain technology to create an application that stores Electronic Health records. A modern, secure solution is introduced that maintains EMRs in the healthcare departments using Sidechain-Blockchain technology.

This application currently has very few use cases which satisfies just the most necessary functionalities of an EHR management system. The application can further be extended to store all types of medical documents other than prescriptions like X-rays, blood reports etc. Furthermore, the application can be extended to provide templates to fill in the details directly and then generate and upload reports directly.

This can save the overhead created from the file uploads. One of the features that can be added to such a kind of medical application is the payments. While approving the medical data, i.e., expecting some kind of medical report from the provider, that patient can also send the respective amount of ether as a payment, which if not sent, the data won't be added to the patient's profile. To implement these kinds of features, the functions in the smart contract can be made payable by using the 'payable' keyword that can accept payments. For this purpose, we can either create our own cryptocurrency for the application or use ether directly. Both the approaches will have different effects. Creating a cryptocurrency for the application alone will bring market value of the application into the picture.

References

1. Guo, R., Shi, H., Zhao, Q., Zheng, D.: Secure attribute-based signature scheme with multiple authorities for blockchain in electronic health records systems. IEEE Access **6**, 11676–11686 (2018)
2. Uddin, M.A., Stranieri, A., Gondal, I., Balasubramanian, V.: Continuous patient monitoring with a patient centric agent: A block architecture. IEEE Access **6**, 32700–32726 (2018)
3. Ramani, V., Kumar, T., Bracken, A., Liyanage, M., Ylianttila, M.: Secure and efficient data accessibility in blockchain based healthcare systems. In: 2018 IEEE Global Communications Conference (GLOBECOM), pp. 206–212. IEEE (2018)
4. Nakamoto, S.: Bitcoin: A peer-to-peer electronic cash system. Manubot (2019)
5. Sompolinsky, Y., Zohar, A.: Secure high-rate transaction processing in Bitcoin. In: Böhme, R., Okamoto, T. (eds.) LNCS Financial Cryptography and Data Security, pp. 507–527. Springer, Berlin, Heidelberg (2015). Retrieved 20 Oct 2020, from https://doi.org/10.1007/978-3-662-47854-7_32
6. Ocheja, P., Flanagan, B., Ueda, H., Ogata, H.: Managing lifelong learning records through blockchain. Res. Practice Technol. Enhanced Learn. **14**(1) (2019). https://doi.org/10.1186/s41039-019-0097-0
7. Zhao, H., Bai, P., Peng, Y., Xu, R.: Efficient key management scheme for health blockchain. CAAI Trans. Intell. Technol. **3**(2), 114–118 (2018)
8. Bogner, A., Chanson, M., Meeuw, A.: A Decentralised Sharing App running a Smart Contract on the Ethereum Blockchain, pp. 177–178 (2016). ACM. Retrieved 20 Oct 2020, from https://doi.org/10.1145/2991561.2998465
9. Sompolinsky, Y., Zohar, A.: Secure high-rate transaction processing in Bitcoin. In: International Conference on Financial Cryptography and Data Security, pp. 507–527. Springer, Berlin (2015)
10. Buterin, V.: Ethereum White Paper A Next Generation Smart Contract and Decentralized Application Platform [Epaper] (2013). Retrieved 20 Oct 2020, from https://cryptorating.eu/whitepapers/Ethereum/Ethereum_white_paper.pdf
11. Li, M., Tang, H., Hussein, A.R., Wang, X.: A Sidechain-based decentralized authentication scheme via optimized two-way peg protocol for smart community. IEEE Open J. Commun. Soc. **1**, 282–292 (2020)
12. Parizi, R.M., Homayoun, S., Yazdinejad, A., Dehghantanha, A., Choo, K.K.R.: Integrating privacy enhancing techniques into blockchains using sidechains. In: 2019 IEEE Canadian Conference of Electrical and Computer Engineering (CCECE), pp. 1–4. IEEE (2019)
13. Kuo, T.T., Ohno-Machado, L.: Modelchain: decentralized privacy-preserving healthcare predictive modeling framework on private blockchain networks (2018). arXiv preprint arXiv:1802.01746

14. Kolekar, S., More, R., Bachal, S., Yenkikar, A.: Review Paper on Untwist Blockchain: A Data Handling Process of Blockchain Systems—IEEE Conference Publication (2020). Ieeexplore.ieee.org. Retrieved 20 Oct 2020, from https://ieeexplore.ieee.org/document/8533868.
15. Zheng, Z., Xie, S., Dai, H., Chen, X., Wang, H.: An overview of blockchain technology: architecture, consensus, and future trends. In: 2017 IEEE International Congress on Big Data (BigData congress), pp. 557–564. IEEE (2017)
16. Liu, X., Chen, R., Chen, Y.W., Yuan, S.M.: Off-chain data fetching architecture for ethereum smart contract. In: 2018 International Conference on Cloud Computing, Big Data and Blockchain (ICCBB), pp. 1–4. IEEE (2018)
17. Zhang, X., Poslad, S.: Blockchain support for flexible queries with granular access control to electronic medical records (EMR). In: 2018 IEEE International conference on communications (ICC), pp. 1–6. IEEE (2018)
18. Sato, T., Himura, Y.: Smart-contract based system operations for permissioned blockchain. In: 2018 9th IFIP International Conference on New Technologies, Mobility and Security (NTMS), pp. 1–6. IEEE, Paris (2018). Retrieved 20 Oct 2020, from https://doi.org/10.1109/NTMS.2018.8328745

Computational Fluid Dynamic Analysis of Flow Patterns Around Rhizophora Mangrove Roots

Sini Rahuman, Mohamed Ismail, and Shyla Manavalan Varghese

Abstract This study aims to investigate the wind and water flow patterns around stilt (Prop) roots of Rhizophora mangrove species. ANSYS Computational Fluid Dynamics (CFD) Software is used to Simulate and Analyze the wind and water flow around the roots. During cyclones, typhoons or hurricanes, the velocities of wind and water flow are not constant. The change in speed of Wind and water flow under these circumstances is simulated using the step function inlet velocity model. Velocity and pressure at various locations are obtained to analyze the flow patterns. The simulation results show that the stilt roots of Rhizophora mangrove tree steadily reduces the velocity of wind and wave flow. This study can be used to improve the efficiency of break water models and thereby protect the coast from the hazards caused by natural calamities.

Keywords Rhizophora species · Mangroves · Water flow · Wind flow · Jet flow · Turbulent · Stagnation · CFD · Stilt roots · Incompressible · Unsteady · Step function

1 Introduction

All over the world, one of the most natural phenomena that occurs is erosion which has threatened human activities in the coastal areas creating hazards to human livelihoods. Mangroves have been acclaimed in recent years as main stream defense against coastal erosion. It is a well-known reality that vegetation protects the coast, to a certain

S. Rahuman (✉)
Sathyabama Institute of Science and Technology, Chennai, India
e-mail: sini.rahuman@polytechnic.bh

M. Ismail
Department of Mathematics, Sathyabama Institute of Science and Technology, Chennai, India

S. M. Varghese
Department of Mathematics, Houston Baptist University, Houston, TX 77074, USA

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_102

degree, from the anger of waves, storm surges, and tsunamis caused by extraordinary climate occasions.

Rhizophora species are the common species of Mangroves widely identified in Asia. They have up root structures (prop or stilt roots) which contribute to higher drag coefficient. These prop roots show significant resistance to the stream of water. *Rhizophora apiculate* and *Rhizophora mucronata* are very effective in minimizing the impact of tsunamis because based on field survey [1]. This is further reinforced by another research, which assumes that the Mangrove width density, its versatility, and complexity of uproot systems are a response to the wind and wave strength reduction process [2]. In short-lived atmospheric fluctuations such as hurricanes or storms, the storm rises due to abnormally high levels of sea water in coastal areas [3]. Mangroves decrease the forward velocity leading to a decrease in the flooding region [4]. Mangrove zone would shield the area behind them from floods from a Category 5 typhoon with a fast forward speed of 11.2 m/s [5]. The Mangrove roots produce barriers that create complex 2D flow with eddies, jets, and vegetation-scale turbulence [6, 7]. An examination performed on CFD simulation of unsteady turbulent two-dimensional water flow with constant inlet velocity around the Mangrove species focused on the flow patterns within the root area and analyzed the velocity dissipation process [8, 9].

In order to enhance coastal protection systems, it is important to extend the existing knowledge of Mangrove roots in realistic weather conditions. This research, therefore, emphasizes on the flow of wind and water through the roots of the Mangroves by gathering knowledge from the root structure of the *Rhizophora mangroves*. This detail is changed to data suitable for CFD simulation [10] and examination.

2 Methods

2.1 Study Approach

The CFD simulation and analysis of unsteady flow of air (Incompressible) and water around the roots of *Rhizophora mangrove* species are performed using k-ε turbulence model. In Fig. 1, the primary trunk (represented by the big circle T) is the root that grows along the same line with main tree trunk. The circles S_1 to S_7 indicate the primary prop roots with dimensions 9 cm. 113 stilt roots are considered. The diameters of the stilt roots are within the range of 2–5 cm. 'Figure 2' shows the computational model of the geometry with roots plotted in 3 m by 3 m lattice along with far-field open boundary (14 m $\times$ 6 m). A mesh is generated with 609,138 number of cells based on the Grid Independent Test. The boundary is defined and the mesh is exported for CFD Analysis using the CFD software Ansys Fluent 18.1. The results are validated with the help of Grid Independence Test, Residual plots for convergence, and the results available in the Literature Survey.

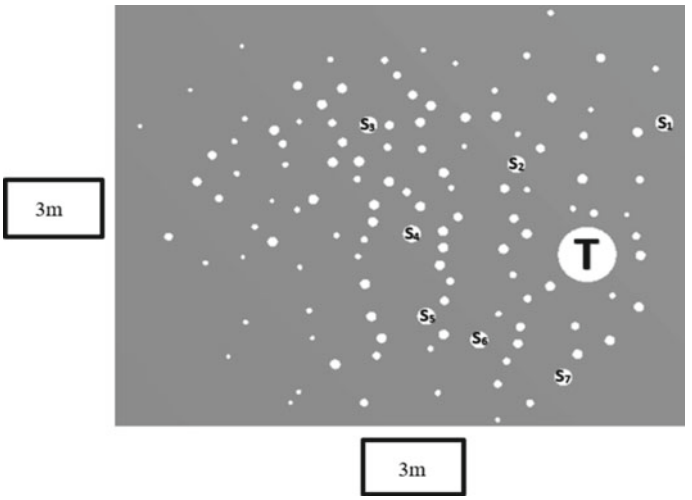


Fig. 1 The roots of Rhizophora species—computational model geometry

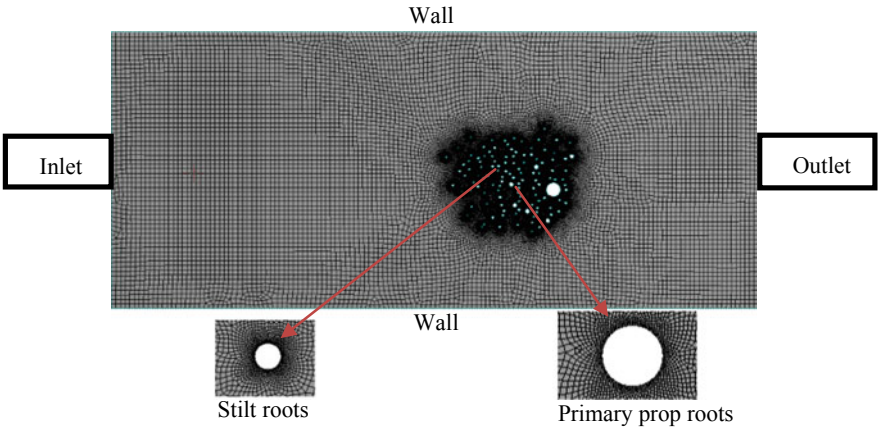


Fig. 2 Computation model geometry generated in Ansys design modeler

2.2 Governing Differential Equation

Navier-Stokes system of equations governing incompressible viscous fluid flow are as follows:

Continuity Equation:

$$\nabla \cdot \vec{q} = 0$$

Momentum Equation:

$$\rho\left(\frac{\partial \vec{q}}{\partial t} + \vec{q} \cdot \nabla \vec{q}\right) = -\nabla p + \mu \nabla^2 \vec{q} \tag{1}$$

2.3 Boundary Condition

Case 1 is considered as incompressible unsteady air flow with the inlet velocity modeled as step function.

$$\begin{aligned} \text{Inlet velocity} &= 5 \text{ m/s}, \ 0 \leq t \leq 3 \\ &= 10 \text{ m/s}, \ 3 \leq t \leq 6 \\ &= 15 \text{ m/s}, \ 6 \leq t \leq 9 \end{aligned}$$

$\mu = 1.7894\text{e}{-}05$ kg/m s (coefficient of viscosity of air); $\rho = 1.225$ kg/m^3 (density of air).

Case 2 as Unsteady water flow with the inlet velocity modeled as step function.

$$\begin{aligned} \text{Inlet velocity} &= 5 \text{ m/s}, \ 0 \leq t \leq 3 \\ &= 10 \text{ m/s}, \ 3 \leq t \leq 6 \\ &= 15 \text{ m/s}, \ 6 \leq t \leq 9 \end{aligned}$$

$\mu = 0.001003$ kg/m s (coefficient of viscosity of water); $\rho = 998.2$ kg/m^3(density of water).

The inlet velocity is modeled as step function in this paper, because, in reality, the speed of the wind and water will change with respect to time. Incompressible Viscous flow is considered in both Case 1 and Case 2 as Mach number is less than 0.3. The Reynolds number for this flow in Case 1 and Case 2 is larger than 5.0×10^5, so turbulence model is used to simulate the flow. k-ε turbulence model is considered and Finite Volume Method is used to solve this problem. The body forces are eliminated. Outlet Pressure is 0 Pa for Cases 1 and 2.

3 Analysis and Result

Case 1 as unsteady air flow(Incompressible) with the inlet velocity modeled as step function.

See Figs. 3, 4, 5 and 6.

Case 2 as unsteady water flow with the inlet velocity is modeled as step function.

See Figs. 7, 8, 9 and 10.

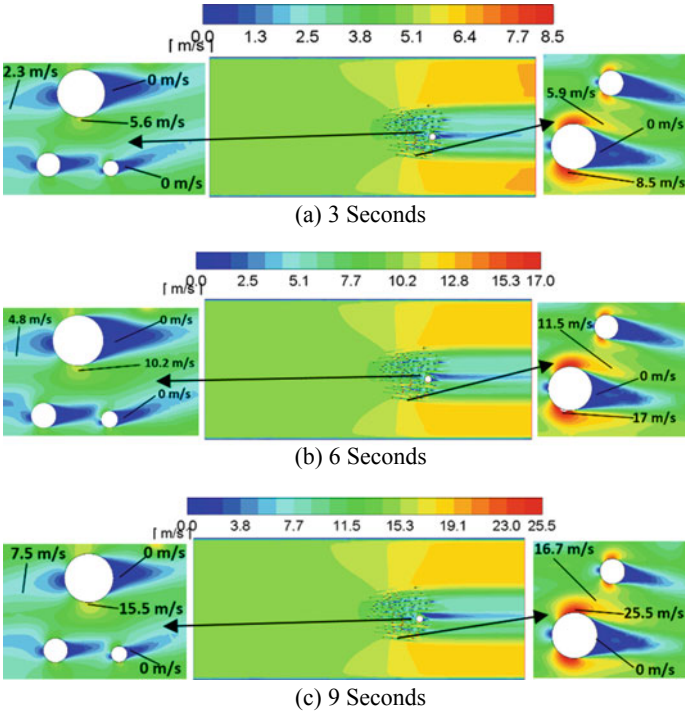


Fig. 3 Velocity profile of incompressible unsteady air flow in case 1

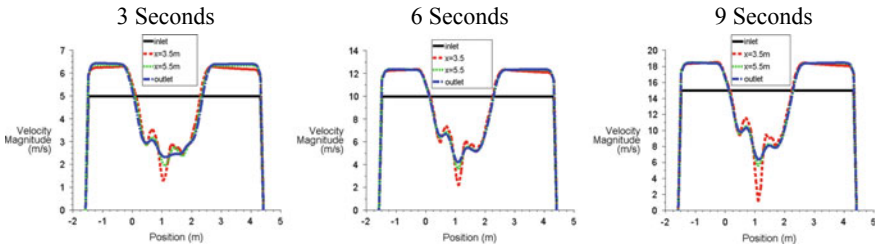


Fig. 4 Velocity dissipation of incompressible unsteady air flow by *Rhizophora mangrove* stilt roots at $x = 3.5$ m and $x = 5.5$ m in case 1

The velocity of the flow in Case 1, after 3, 6 and 9 s indicated in the left side enlarged images of 'Fig. 3a–c' shows that the velocity is reduced to 0 m/s (represented by the dark blue color) as the fluid flows toward the downstream side of the roots near the primary trunk. The speed of the wind from 5 m/s, 10 m/s, and 15 m/s reduces to 2.3 m/s, 4.8 m/s, and 7.5 m/s respectively, closer to the root indicated by the light blue color. The velocities on the perpendicular side of the root at different time intervals are shown as 5.6, 10.2, and 15.5 m/s (represented by light green color). The right side

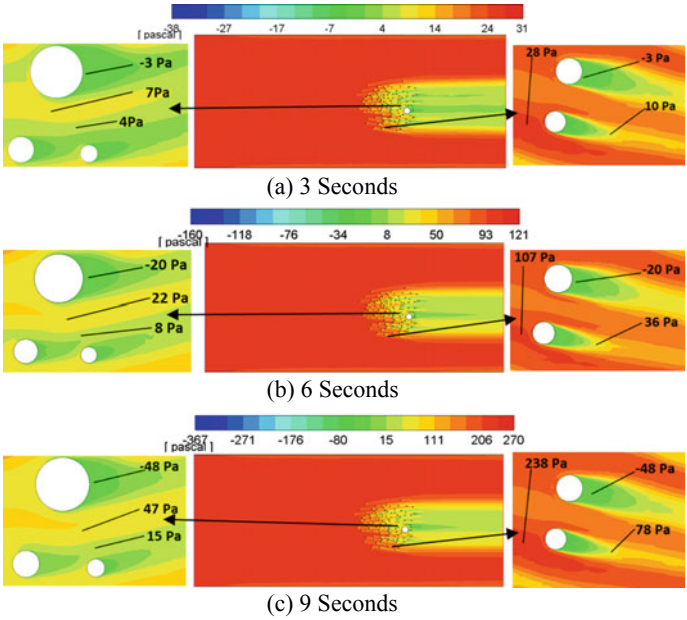


Fig. 5 Pressure distribution of incompressible unsteady air flow in case 1

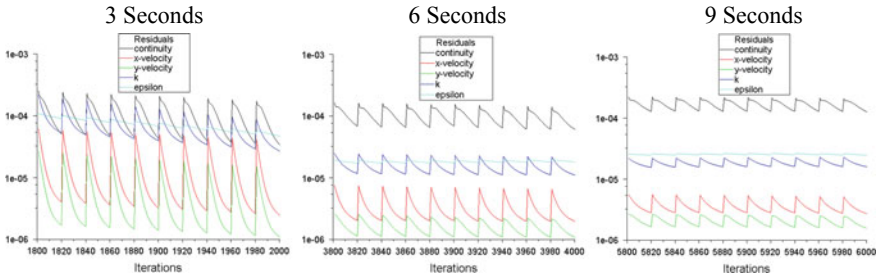


Fig. 6 Convergence plot of incompressible unsteady air flow in case 1

enlarged images in 'Fig. 3a–c', are the roots presented away from the primary trunk. The velocity is very high in the perpendicular side of these stilt roots, which are indicated by red color (velocities 8.5, 17, and 25.5 m/s) and yellow color (velocities 5.9, 11.5, and 16.7 m/s) after 3 s, 6 s and 9 s respectively, in Case 1.

Similarly, the left side enlarged image in 'Fig. 7a–c' shows that, after 3, 6, and 9 s, the velocity is reduced to 0 m/s as the wind flows towards the downstream side of the root presented near the primary trunk which is represented by the dark blue color. The speed of the wind from 5, 10, and 15 m/s reduces to 2.1, 4.5, and 7 m/s closer to the root indicated by the light blue color. The perpendicular side of the stilt root velocities at different time intervals are shown as 4.6, 10.3, and 14.5 m/s, are

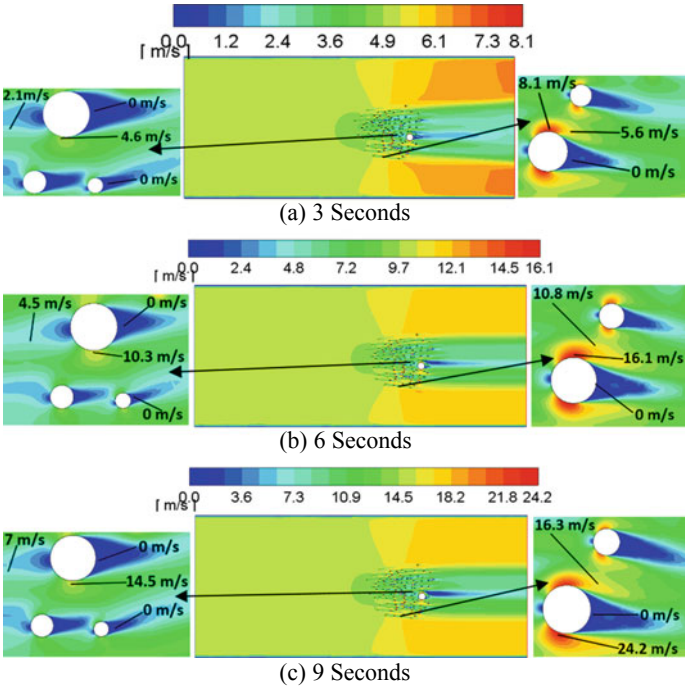


Fig. 7 Velocity profile of unsteady water flow in case 2

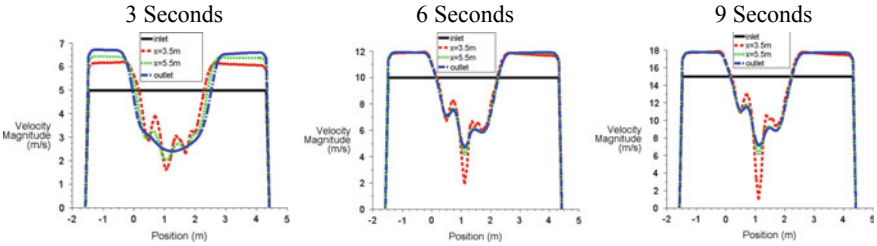


Fig. 8 Velocity dissipation of unsteady water flow by *Rhizophora mangrove* stilt roots at $x = 3.5$ m and $x = 5.5$ m in case 2

represented by light green color. The right side enlarged images in 'Fig. 7a–c' are the roots presented away from the primary trunk. The velocity is very high as shown in the perpendicular side of these roots, which is indicated by red color represent the velocities as 8.1, 16.1, and 24.2 m/s and yellow color as 5.6, 10.8, and 16.3 m/s after 3, 6, and 9 s in Case 2. Velocity dissipation analysis is done at $x = 3.5$ m and $x = 5.5$ m in Cases 1 and 2 as indicated in 'Figs. 4 and 8'. A clear observation indicate that magnitudes of the velocity reduce by more than 70%, when the flow passes through the *Rhizophora mangrove* roots.

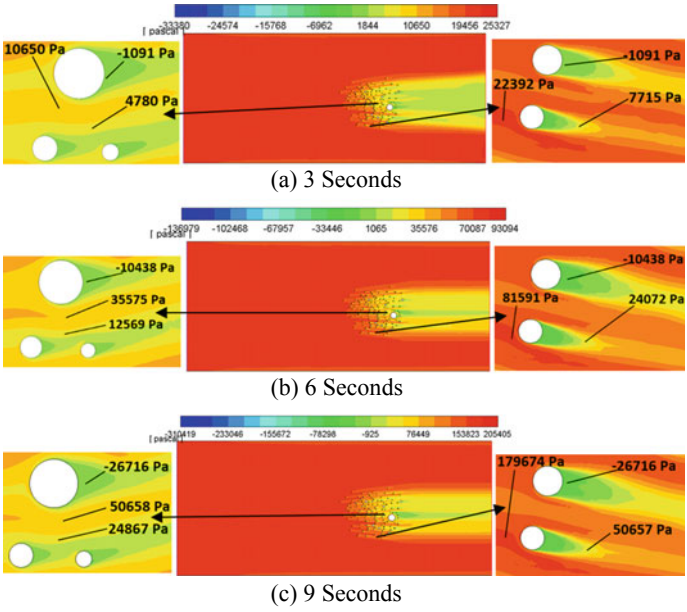


Fig. 9 Pressure distribution of unsteady water flow in case 2

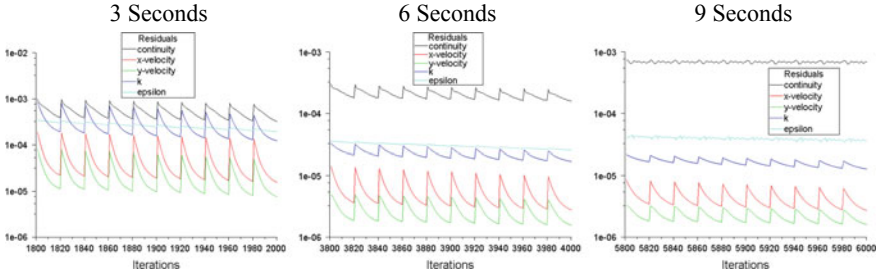


Fig. 10 Convergence plot of unsteady water flow in case 2

The pressure distribution of the air flow in Case 1, after 3, 6, and 9 s indicated in the left side enlarged images in 'Fig. 5a–c' shows that, in the close proximity the pressure near the primary trunk roots are −3, −20 and −48 Pa (green contour), 4, 8, and 15 Pa (light green color) and 7, 22, and 47 Pa (yellow contour). The right side enlarged images in 'Fig. 5a–c', which are the pressure distribution of the roots presented away from the primary trunk, explain that in close proximity the pressure near the stilt roots are −3, −20, and −48 Pa (green contour), 10, 36, and 78 Pa (yellow color), and 28, 107, and 238 Pa (red color).

Similarly, the pressure distribution of the water flow in Case 2, after 3, 6, and 9 s indicated as the left side enlarged images in 'Fig. 9a–c' shows that, in close

proximity the pressure near the primary trunk roots are −1091, −10,438 and −26,716 Pa (green contour), 4780, 12,569 and 24,867 Pa (yellow color) and 10,650, 35,575 and 50,658 Pa (light orange color). The right side enlarged images in 'Fig. 9a–c', which are the pressure distribution of the roots presented away from the primary trunk, after 3, 6 and 9 s, demonstrated that in close proximity the pressure near the stilt roots are −1091, −10,438 and −26,716 Pa (green contour), 7715, 24,072 and 50,657 Pa (yellow color), and 22,392, 81,591 and 179,674 Pa (red color).

The convergence of the solution is observed and correct to 3 decimal places for incompressible unsteady air flow in Case 1 and water flow in Case 2 as displayed in 'Figs. 6 and 10'.

4 Conclusion

From all simulation results, it can be visualized with the help of different contours that the speed of wind and water energy steadily decreases from red to blue after the air and water flows through the Rhizophora stilt Mangrove roots. The magnitudes of the velocity reduce by more than 70% when the flow passes through the *Rhizophora mangrove* roots. The geometry of root systems of these Mangrove trees aid hold the soil together and thereby defend the coastal regions against erosion and extreme weather conditions Mangroves reduces coastal erosion by reducing the speed of storms and waves that create coastline erosion. This study proves that the roots of *Rhizophora mangroves* play a remarkable role in dissipating the speed of wind and water flow and protect the coast from erosion and save the area near the coast from hazards caused by natural calamities. This study is therefore an evidence to show that CFD is the best tool to investigate problems in actual environment with minimum cost.

Limitations And Future Scope

The investigation can further be undertaken by considering three-dimensional flow of heavy wind and water flow around the mangrove roots to fix the issues in a more realistic situation.

References

1. Tanaka, N., Sasaki, Y., Mowjood, M., Jinadasa, K., Homchuen, S.: Coastal vegetation structures and their functions in tsunami protection: experience of the recent Indian Ocean tsunami. Landscape Ecol. Eng. **3**(1), 33–45 (2007)
2. Kathireshan, K., Rajendran, N.: Coastal mangroves forests mitigated tsunami. Estuar. Coast. Shelf Sci. **65**(3), 601–606 (2005)
3. McIvor, A.L., Möller, I., Spencer, T., Spalding, M.: Reduction of wind and swell waves by mangroves. In: Natural Coastal Protection Series (2012)

4. Liu, H., Zhang, K., Li, Y., Xie, L.: Numerical study of sensitivity of mangroves in reducing storm surge and flooding to hurricane characteristics in Southern Florida. Cont. Shelf Res. **64**, 51–65 (2013)
5. Zhang, K.Q., Liu, Y., Hongzhou, X., Jian, S., Rhome, J., Smith, T.J.: The role of mangroves in attenuating storm surges. Estuar. Coast. Shelf Sci. **102**, 11–23 (2012)
6. Furukawa, K., Wolanski, E., Mueller, H.: Currents and sediment transport in mangrove forests. Estuar. Coast. Shelf Sci. **44**(3), 301–310 (1997)
7. Burger, B.: Wave attenuation in mangrove forests: numerical modeling of wave attenuation by implementation of a physical description of vegetation SWAN. Master thesis, Delft University of Technology, Netherlands (2005)
8. Zamin, M., Jusoh, M., Aziz, N.A., Inayatullah, O.: Computational fluid dynamics simulation of flow velocities dissipation by mangrove roots structure. ARPN J. Eng. Appl. Sci. (2016)
9. Aziz, N.A., Inayatullah, O., Jusoh, M., Bin, M.Z.: The mechanism of mangrove tree in wave energy propagation. Adv. Mater. Res. **614**, 568–572 (2012)
10. Shyla, M.V., Naidu, K.B., Vasanth Kumar, G.: Identification of sensor position using CFD Simulation, pp 302–307. ISBN 978-93-5107-261-4 (2014)

ICT Systems to Empower a City's Infrastructure in the Case of Kochi

Divya Susanna Ebin, Meghana K. Raj, and B. S. Sheethal

Abstract The entire process of structuring a city and making it smart and innovative depends on the method of improving the efficiency and quality of life of people. This process involves collection of data, assessment of information, and using that information appropriately. As the city's population is increasing the demand for services and facilities is also increasing and this can be addressed through the use of ICT. ICT is a three-letter word which helps to improve the quality of life in the city in a short period of time. In the recent years, its use has increased enormously all over the world and it has played a prime role during the pandemic situation and has kept life moving without any obstacles. Hence good management of ICT is very important for a country's development. This paper studies the use of ICT in various field to improve the quality of life of people and also provides proposals which bridges the gap of development.

Keywords ICT · Infrastructure · Transport · Urban development · Education

1 Introduction

As per United Nation's Department of Economic and Social Affairs, the urban population has surpassed 50% in 2008 from 30% in 1950. It has also predicted that the population will increase to 64% in developing countries and 86% in developed countries by 2050. Due to increase in the population the demand for various facilities such as infrastructure, social services, modern technologies has also increased. A developing country like India suffers more from this issue compared to developed countries hence there is a need to address this situation immediately. ICT which is information and technology is one of the recent technologies used to solve the problems of inadequate facilities. It stresses on the importance of communication, computer application and other necessary soft wares. The impact of ICT on society has increased in last few years, due to which many problems faced by the cities such

D. S. Ebin (✉) · M. K. Raj · B. S. Sheethal
MS Ramaiah Institute of Technology, Mathikere, Bangalore 560054, India

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_103

as congestion of traffic, shortage of infrastructure, etc. are solved. ICT makes the city smart and innovative and it is directly proportional to city's development and prosperity.

2 Aim, Objective, and Purpose of the Project

2.1 Aim

The paper aims at providing some ICT solutions to improve the living conditions of a fast-growing city. ICT is considered as a tool to improve the city's structure so as to improve quality of life and generate maximum efficiency on public environment and new opportunities in the city.

2.2 Objective

Living Conditions. To analyze and understand the impact of ICT on improving the living conditions of the city.

Development of City. To incorporate ICT as a bridge for development of the city.

2.3 Purpose of the Project

The purpose of the project is to incorporate ICT in the context of Kochi. The city of Kochi has the largest railway station which consists of the city's CBD and a historically relevant institutional zone. The precinct is defined by the major arterial roads and caters to 15,000–20,000 commuters daily with huge infrastructure development works in the offset. With this development, the city is seeing urban problems such as traffic congestion, vehicular-pedestrian conflict, sewage, improper waste management, parking, etc. So an attempt has been made to mitigate these issues by using ICT in the city level services and providing a better quality of life.

3 Theoretical Background

ICT or "Information and communication technology" is a means to improve quality of life, efficiency of urban operation services, and competitiveness with respect to

economic, social and environmental aspects to make a city smart, innovative, and sustainable.

3.1 Scope of City Level Services Through ICT

Infrastructure. Street lighting, Parking, Education, Sustainability, Water metering, consumption and leakage, Waste Management, Air quality, CO_2 emission [1].

Smart Waste. With increasing demand for consumer goods, wastage has increased exponentially. Hence it is very important to manage the waste with smart solution such as Implementing waste tracking systems; to monitor and control the movement of different kinds of waste, sorting of waste without the operator coming into contact with it [1]. Leveraging technology to collect and share data from source to transportation to disposal of waste and connecting various smart waste management systems with local waste management service providers.

Smart Physical Safety and Security. The existing security technology such as video surveillance, video analytics, and biometrics will remain the main focus of a city's security [1]. Method to manage information flow and analyze the data will be the main areas of concern in the next generation of security.

Smart Education. In the long run, education may be the most important smart city service for adults as well as for children. The role of Schools and Universities is, therefore, a key element to consider in the design of smart education solutions.

Scope of Quality of Life Dimension ICT. Living conditions of the residing city, Basic needs that influence the "quality of life" like food, shelter, water, safety, and security, health, jobs, education.

4 Existing City Structure

Kochi is a port city on the south west coast of India and is part of the district of Ernakulum in the State of Kerala [2]. Kochi is often called Ernakulum which refers to the mainland part of the city [2]. Kochi is considered as the commercial and industrial capital of Kerala. In the past, the city has traded with foreign countries like China, Dutch, the British, and the Portuguese. In the present, port and IT parks evolving a Software hub add to the character of the city. There are some other important characteristics of the city such as the Economic capital of Kerala, most urbanized district (47.56%), Green Open Space—1% of corporation area, 1/4th of the city area covered with water bodies.

4.1 Population Growth

The total population of the city is 1.3 million [3] and the population growth is more in the north-eastern part of the city and less in the central part. It is analyzed that in the older local bodies of the city, it is witnessing a slow down while the newly added local bodies like Thrikkakara is experiencing a higher rate which is due to an increase in the commercial activities in the main city area as well as the growth in residential areas in the periphery of the city.

4.2 Existing Land Use and Urban Pattern

Kochi Corporation shows that a major part of the land other than the water bodies and agricultural patches are already developed. The primary land use is residential, while commercial activity is seen concentrated in the city core. The observed trends of the land-use conversion are change in land use are from residential to commercial or mixed use. Ribbon development of commercial activities is seen along the highways. As the city grew, because of the availability of developable land the city extended towards the east, away from the backwaters. The new growth corridors in the city are determined for the implementation of better movement network through road. Better access and new transport nodes have facilitated the new commercial nodes (Fig. 1).

5 Existing City Structure

Located in the South of Kochi, in the Ernakulam District of Kerala is the Ernakulam South Railway Station and its precincts. It has major arterial roads M.G Road along the western edge, Kaloor Kadavantra Road on the eastern edge and the north and south are outlined by Banerjee Road and Sahodaran Ayappan Road. It consists of the city's CBD with many office and commercial spaces, an institutional zone that houses historically relevant government schools and it is currently the focal point of

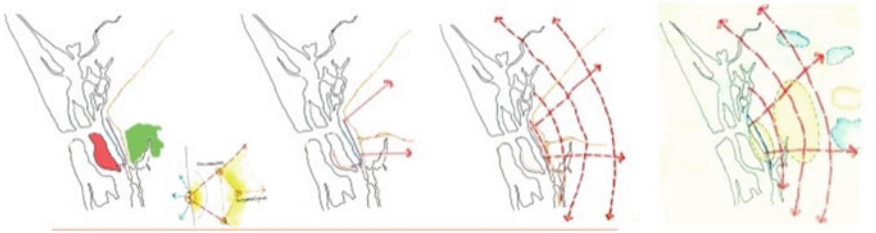


Fig. 1 Urban pattern of the city. *Source* Author generated

Fig. 2 Context precinct level. *Source* Author Generated

infrastructure development attracting entrepreneur investments. This precinct can be termed as a terminal of various modes of transit including the existing railway station, the KSRTC bus stand, and the proposed water transit along the Thevara-Perandoor canal running through the precinct. The main focus of the study being the Ernakulam South Railway Station, the study area taken is about 2 km around this central core. The Rail acts as a barrier dividing the city into two parts one being predominantly commercial in character and the other side being predominantly residential in nature (Fig. 2).

5.1 History and Evolution

The history of Ernakulam dates back to 2.5–4 million years. [3] Initially, the lowlands of Ernakulam were covered by forests, later the transgression of sea up to midland submerged the forests. The present development marks the formation of seaport-airport road, lately the A.L Jacob railway over bridge was opened along with the other infrastructure development and simultaneously there was increase in intensity of residential, commercial activities (Fig. 3).

Fig. 3 Evolution of the site area. *Source* Author Generated

5.2 Built Use

Land Use pattern of a city is essentially a reflection of economic and social form and structure. There is a very organic growth on the western side of the railway station while the eastern side of railway shows rigidity of planning in large parcels of land. The western side of the railway line is witnessing a slow transformation from residential use to commercial (Fig. 4).

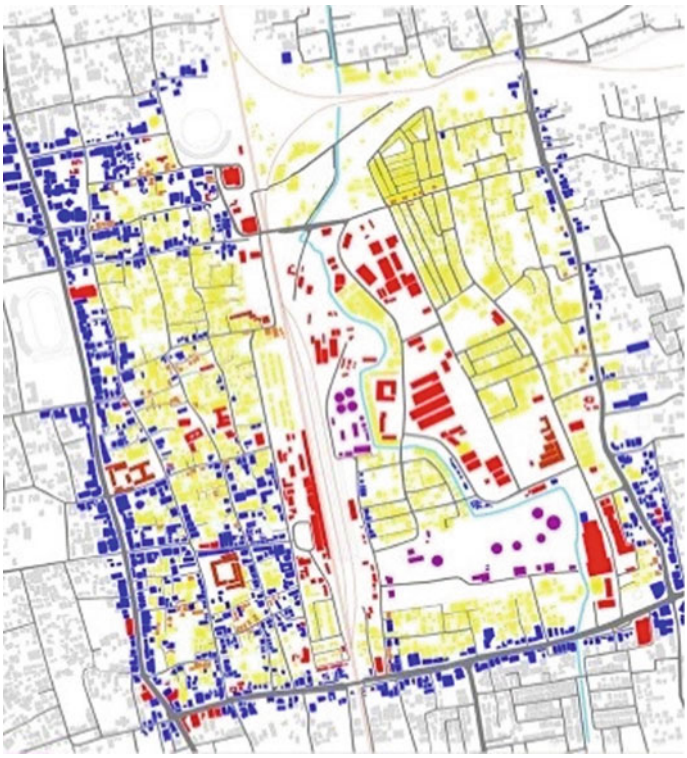


Fig. 4 Built use map. *Source* Author Generated

Fig. 5 Road network map. *Source* Author Generated

5.3 Transportation

The total road network of Kochi city is 614 km, road density of 1.03 km/1000 population and 6.47 km/km^2 of surface area. The share of arterial roads in the total road network in the city was only 2.75%, sub arterial was around 9%, local streets formed the major share 64% of the road network in the city. Due to the lack of connected pedestrian pathways, there are a lot of conflict points in most of the major nodes in the site area (Fig. 5).

5.4 Bus Routes, Bus Stopes, and Parking

Most of the bus routes and bus stops are seen to be in the peripheral sides of the site area catering to the western side of the railway and to M.G Road. Lack of continuous roads and lack of public transport systems to the eastern side of the railway makes it again an underutilized area. The private bus routes do not cater to the internal areas of the precinct which mainly consists of the residential area (Fig. 6).

Fig. 6 Bus routes and bus stop map. *Source* Author Generated

5.5 Drainage and Sewage System

Drainage systems provided in the site area lack continuity and improper management of the same leads to waste accumulation. The sewage waste from more than 60% of the residences are drained into the storm water drain which again leads to the canal. The spaces under the flyover are used for dumping wastes due to lack of waste collection (Fig. 7).

Overall 5% of the entire district has sewage connections as per City Sanitation Plan of Kochi which covers the central part and along marine drive of the district. The waste management in Kochi is mainly through septic tank and other localized means. Due to high water table, the septic tank and two pit latrines do not function properly and hence there is high pollution of water and soil. To overcome this situation there is a need to expand the network of sewage connection. However, the idea of expansion also faces many challenges, as the streets are narrow in the area (Fig. 8).

5.6 Education

ICT is a mode that will support, enhance and improve the delivery of information for student learning. In the recent years, we have witnessed that education is also

Fig. 7 Drainage system map. *Source* Author Generated

dependent on ICT and online interactions help to reduce the time and make it more organized to assess the reports, assignment, etc. It is of great importance in the pandemic situation, as it keeps the learning of students going without wasting time. But there is a certain limitation in this area, the rural and the poor suffer a lot because of poor connection or lack of technology. Hence there is a need to expand the ICT throughout the country and address the situation in a smart way as education is a main tool for the development of the country.

6 Issues Identified

The major issues identified and which is considered for developing strategies are,

- East west connectivity hampered by railway line.
- No designated parking area leading to street parking which causes traffic congestion and reduction in capacity of the road.
- Defunct Railway Station parking area.
- Underutilized Eastern entry to the station due to poor accessibility and lack of signage.
- Lack of facilities for lateral and cross movements of pedestrians in major pedestrian intensive locations in the area.

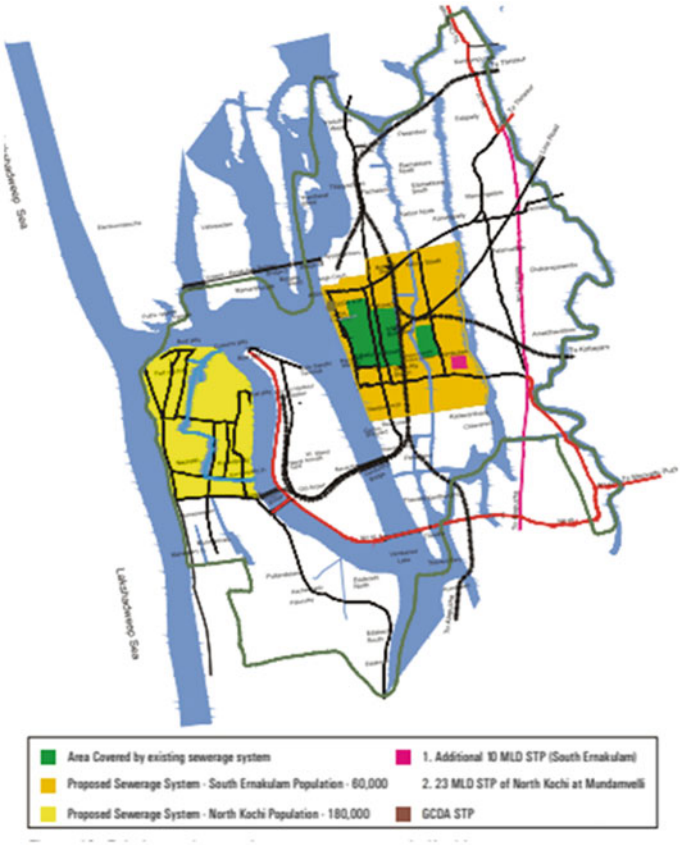


Fig. 8 Sewage system map. *Source* City Sanitation Plan

- Lack of a continuous system of pedestrian walkways.
- Footpaths in various part encroached by street hawkers.
- Lack of sufficient bus routes into the central residential regions.
- Lack of street lighting in the major and the interior roads makes the place unsafe to travel after dusk.
- The sewage waste from more than 60% of the residences are drained into the storm water drain which leads to the canal.
- Waste accumulation in Urban voids due to lack of collection (Fig. 9).

7 Proposals

1. Sensors to help create free movement of the vehicles to decongest the junctions by regulating the signal cycle.

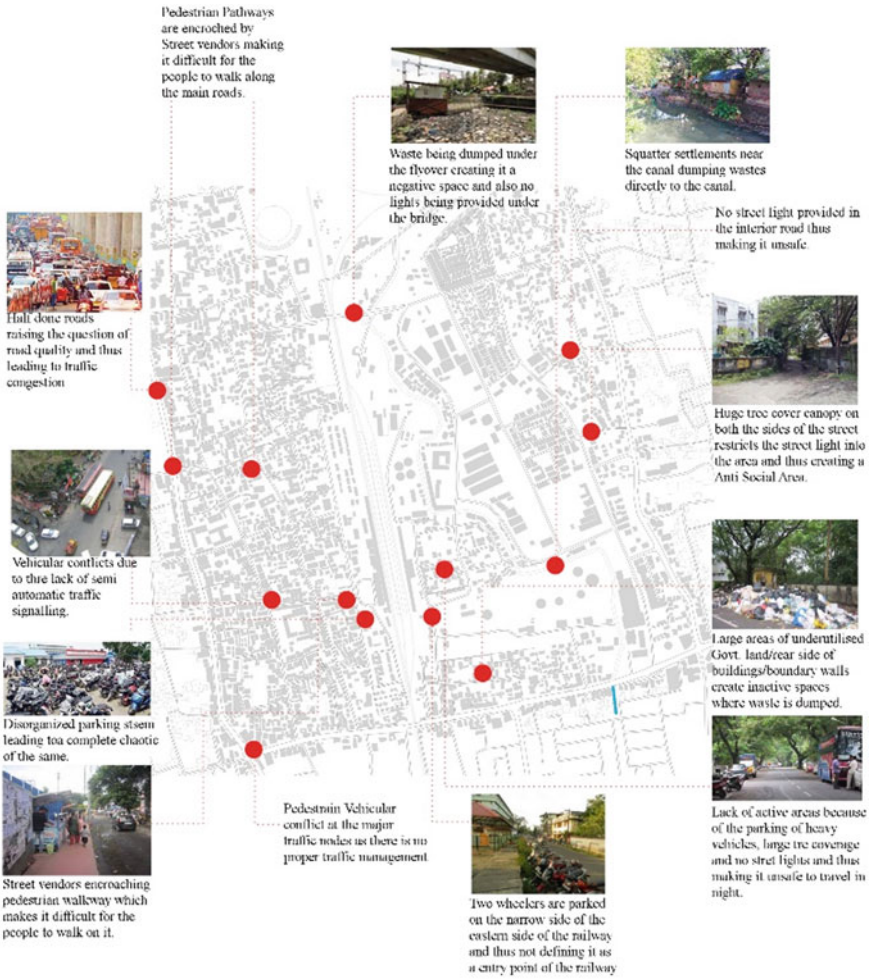


Fig. 9 Issues mapping. *Source* Source Generated

2. Speed sensors to be installed at various streets to regulate the city's traffic and fines to be charged for over speeding.
3. Footpaths to be provided and the same to be used to generate energy which can be used for street lighting and dynamic pedestrian crossing system to be installed at the junctions to avoid pedestrian-vehicular conflict.
4. Vendors should be relocated and provided with web based services for selling their goods and also to collect payment, so as to avoid physical interaction.
5. Space under the flyover can be converted into community activity centres with digital display, CCTV etc., which can be used for entertainment purpose during various occasions. This will also deter unwanted activities and dumping of waste.

6. Implementing waste tracking systems will help to monitor and control the sorting and movement of different kinds of waste without the operator coming into contact with it.
7. Source reduction, proper segregation of waste to be done at various points in the area and development of appropriate ways to reuse, recycle the waste.
8. Leveraging technology to collect and share data from source to transportation to disposal and connecting various smart waste management systems with local waste management service providers.
9. To use Sensors to monitor the discharge of grey and black water to avoid mixing of the same at appropriate locations. Sewage disposal to the lake to be treated and the lake to be rejuvenated.
10. Automated multi-level parking to solve the parking issue near the Metro Stations.
11. For safety the existing security technology such as video surveillance, video analytics, and biometrics will remain the main focus. Along with it the use of computer controlled signages at all necessary points to guide people and also indicate the timings and routes of the public transport system.
12. Smart education solution to be designed to bridge the lacunae in the Education system.

8 Conclusion

Taking into consideration the characteristics of a developing country like India the implementation process is delayed due to lack or delay in funding and other obstructions. The proposal given in this research paper will be implemented under long term planning period which will take 5–15 years. For a growing city like Kochi, ICT plays a prime role in structuring the city and improving the quality of life. It helps to solve various major problems of the city such as traffic, sewage, drainage, connectivity, etc., and in turn helps to improve the city.

References

1. TechniCity.: https://www.Thecitiesoftomorrow.Com/Re-Thinking/ChallengestheChallengesOfASustainableCity (2020). [online] Available at: https://technicity.osu.edu/technology-across-the-globe-www-thecitiesoftomorrow-com-re-thinking-challenges-the-challenges-of-a-sustainable-city. Accessed 19 Nov 2020
2. Lsgkerala.gov.in.: CDP Kochi—Executive Summary (2020). [online] Available at: https://lsgkerala.gov.in/htm/CDPKochi/Executive%20Summary.htm. Accessed 19 Nov 2020
3. Acyutamēnōn, S.: The Cochin state manual. MLibrary, Ann Arbor, Mich (1911)

Analysis of Cloud-Based Intrusion Detection System

S. Sobin Soniya and S. Maria Celestin Vigila

Abstract Information technology companies start deploying cloud computing as their backbone of business throughout the world day by day. Because cloud computing offers flexible, user-friendly, and pay-peruse services to the companies. In order to be successful in the platform, the companies have to face privacy and security issues of the cloud network's nature. Because the cloud network is open and well-connected architecture, it is unsafe to the user's data. Intrusion detection system is very helpful in detecting intruders on the cloud network. This paper furnishes an introduction of possible intrusive activities on the cloud network. Also, this paper analyzes some of the cloud-based intrusion detection systems. The analysis is marked with respect to the parameters such as type of systems, the technique used by the intrusion detection system, merits and demerits of the techniques. This analysis also helps to conclude which of the techniques can be employable in the cloud network by pointing the limitations of each method. For eliminating security challenges of the cloud system, the intrusion detection system should use multiple detection methods.

Keywords Machine learning techniques · Security · Cloud computing · Support vector machine · Deer hunting optimization algorithm · Intrusion detection system

1 Introduction

Securing the network means preventing or protecting the network from unauthorized intruders. This requires permitting only the entitled person to approach, enter or modify the information in a network. It is necessary to build a strong intrusion detection system (IDS). Network security system prevents unauthorized access to the network [1].

S. Sobin Soniya (✉)
Department of CSE, Noorul Islam Center for Higher Education, Kumaracoil, Kanyakumari 629180, India

S. Maria Celestin Vigila
Department of IT, Noorul Islam Center for Higher Education, Kumaracoil, Kanyakumari 629180, India

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_104

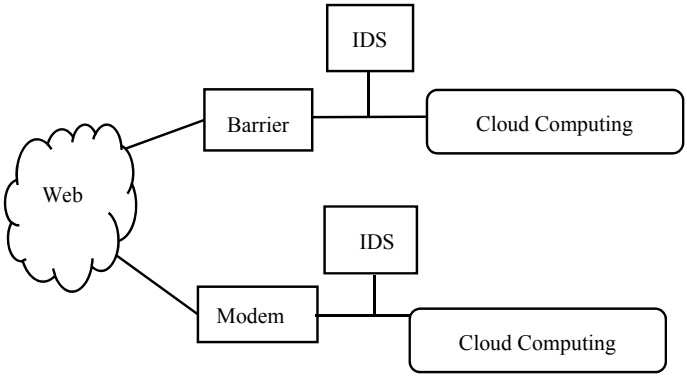


Fig. 1 Model of intrusion detection system

1.1 Intrusion Detection System

As technology develops, people attempt to take information from computer networks that are unauthorized to them. Intrusion detection system is an art of capturing or monitoring the third parties' interest. This is like an immune system to the network which monitors every activity within the network to identify the possible and probable attacks. This is mostly in the form of a software application that captures the network activities and also checks for any attempt to violate the protocol of the network. It also reports the same to the base station. Fig. 1. shows the operation of an intrusion detection system.

1.2 Cloud Computing

This is an infrastructure for Web-based computer and information technology assets. This helps many industries and other Web users to use the assets like operating systems, software applications and complex hardware equipment which are in an economical way through the Internet [2]. The role of cloud computing is practically unlimited.

1.3 Intrusions on Cloud Computing

Cloud computing offers services that are open for everyone. Even though the services are open for everyone, the cloud network should not have redundant information in order to secure the network from intrusions. Cloud computing offers services like cloud-based software services [3], ready-made infrastructure [4], platform services

[5] and expert services [6].Cloud computing uses the same technology as the same operating systems and system applications that are used in common computers and networks. Cloud also offers new technologies which are susceptible to the new form of attacks.

Cloud computing is a new platform where a user can store information and work on the information through the cloud itself. [7]. The novel techniques like the Internet of things (IoT) and mobile computing require more instructions to come up with a decision. In order to use big data technology in the cloud network, the system should be capable of handling more instructions to improve production. These grants access to multiple connections over the cloud which acts like a shared virtual network infrastructure. The nature of cloud computing is known as the wealth like being large-scale and dynamic. Also, it is known for the handling of time variation and ownership change. Because of these properties, the security concerns are problematic when compared to traditional networks.

The cloud environment is manufactured with the technique called virtualization of networks, which means the handling of flow between virtual machines unseen and unimaginable by the modern intrusion detection systems. This makes intrusion detection challenging on the cloud network [8]. The IDS should secure the network by detecting the attacks from the entry as well as based on the patterns of traffic in the network. The need for an automatically evolving system makes the development of such an IDS a highly challenging job. The following paper explains cloud-based intrusion detection and the challenges of cloud-based intrusion detection techniques, the proposed approach and discussion on the conclusion.

2 Related Works

In this module, different instruction detection systems are surveyed. This module also cascades different intrusion detection techniques and the effectiveness of their securing capability from the new type of attacks. A newly proposed technique is also analyzed in this section.

2.1 Analysis of Cloud-Based Intrusion Detection Systems (CIDS)

CIDS [9] techniques are based on three detection techniques: signature-based, anomaly-based and hybrid detection. Figure 2 represents the classification of cloud-based intrusion detection methods.

Signature-based detection techniques: The purpose of this technique is to secure the network by the memory of previous attacks. In these methods, the previous attacks are taught to the system to secure the network from future repeated attacks. This works

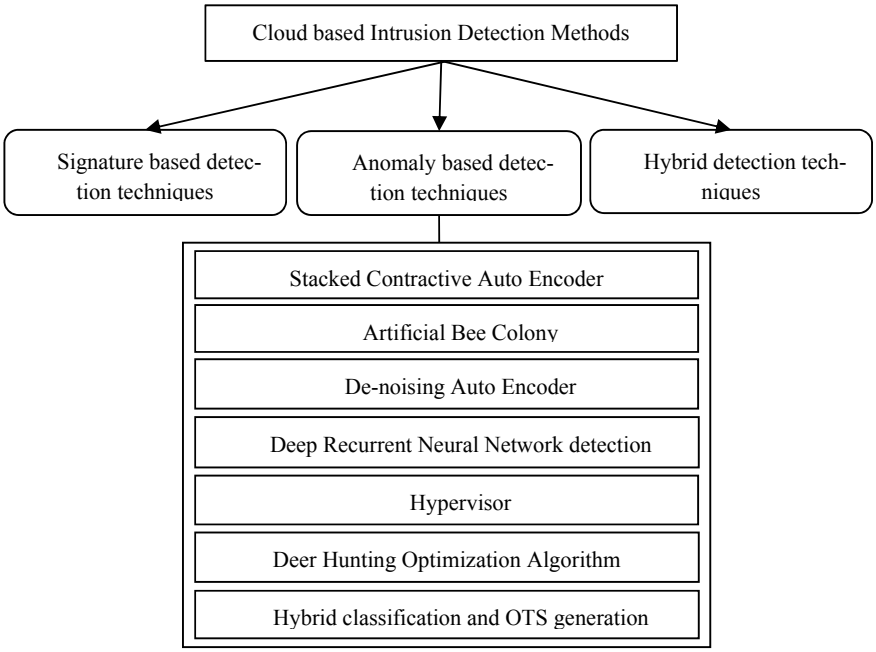


Fig. 2 Classification of cloud-based intrusion detection methods

like the human immune system. This is similar to the working principle of antivirus software in computers. The security system is injected with the information about the already happened attacks, and the system is programmed with the procedure to stop the attack if happens again.

Anomaly-based detection techniques: This detection technique detects the anomaly by the nature of the traffic in the network. It classifies the traffic by normal and abnormal as compared with the base station. It classifies according to the bandwidth used for traffic, ports and protocol used by the traffic. When malicious traffic is detected, it alerts the administrator. The unknown attacks can also be detected by this technique.

Hybrid-based detection techniques: The hybrid detection technique uses the combination of both methods. So this technique will help to detect both previously known attacks as well as unknown attacks. The hybrid method will definitely improve the detectability of the intrusions in the network. The hybrid detection techniques in cloud computing were used by Aldribi et al. [10] and Modi et al. [11] to increase the efficiency of intrusion detection system (Table 1).

Table 1 Overview of the existing cloud-based IDS

Authors	Techniques used	Merits	Demerits
Wang et al. [12], (2020)	Stacked contractive auto-encoder SVM + (SCAE) intrusion detection method	Achieved better detection performance	The SVM classifier cannot be effectively recognized some of the new attacks existed in the test data. It failed to lower the controller's issue
Hajimirzaei and Navimipour [13], (2019)	Combining multilayer perceptron (MLP), fuzzy clustering algorithms and artificial bee colony (ABC)	Reduced wrongly classified occurrences, root mean squared error and mean absolute error	Failed to combine meta-heuristic methods genetic algorithm
Abusitta et al. [14], (2019)	Denoising auto-encoder	Capable of making decisions about suspicious activities without feedback from the IDSs	The performance was not evaluated with other deep learning methods
Aldribi et al. [10], (2020)	The cloud IDS is based on hypervisor	Achieved improvisation of anomaly detection and cloud feature extraction	Failed to capture the hypervisor's whole power
Wang et al. [15], (2019)	Improved dynamic immune algorithm (IDIA)	Detected the malicious data samples and detect homogeneous ones effectively	Potential risks remained may lead to complicated issues to this full-scale network. It was difficult to avail of some components and services from vendors to inject our program in the server
Balamurugan and Saravanan [16], (2019)	Normalized K-means recurrent neural network (NK-RNN)	Produced lower false alarms. NK-RNN was a good classifier for detecting intruders for higher frequent and lower frequent attacks. Also, it detects probe attacks	The method was not implemented in real-time applications
Patil et al. [7], (2019)	Hypervisor-level distributed network security (HLDNS)	Able to detect known and new attacks. Reduced overall computation cost	Failed to detect network attacks like viruses, SQL Injection, worms, etc

(continued)

Table 1 (continued)

Authors	Techniques used	Merits	Demerits
Achbarou et al. [17], (2018)	Distributed IDS (DIDS)	Build the system more efficient and secure. Detected a significant number of attacks in a relatively short time	It was not realistic in real-time applications
Maria Celestin Vigila and Soniya [18], (2018)	Rough set-based future selection using SVM	Successfully eliminates the hurdles in the connection data used in the network and resulting in higher detection accuracy	Standalone SVM is not performing well compared to combinational detection methods
Ajay et al. [19], (2016)	Decision tree split (DTS) algorithm	It overcomes the issues of split value and feature selection in the construction of a decision tree	It does not give uniform weightage to the domain value
Bramma et al. [20], (2019)	Deer hunting optimization algorithm (DHOA)	It uses a new approach like each time the hunters upgrade their location till they reach the deer. So, the unexpected attacks were able to be detected	The overall performance of this method lacks as it misses out on the regular detectable attacks

3 Limitations of This Work

Cloud computing was not the only platform used by the referenced works. Some works were implemented on fog computing or edge computing platforms. So, the performance and efficiency derived from those platforms may not be completely accurate with the cloud computing platform. Also, the dataset used by the researchers was not the same. So, the results mentioned in the results are subject to vary when the dataset is changed.

4 Proposed Methodology

This paper focuses on using different machine learning techniques as a possible combination to effectively detect intrusions. The combined machine learning techniques are expected with more efficient results. The primary objective of this research will be designed to develop an intrusion detection method in a cloud environment using a newly proposed method called feedback deer hunting optimization (FDHO) algorithm. The proposed intrusion detection process will be carried out by deep

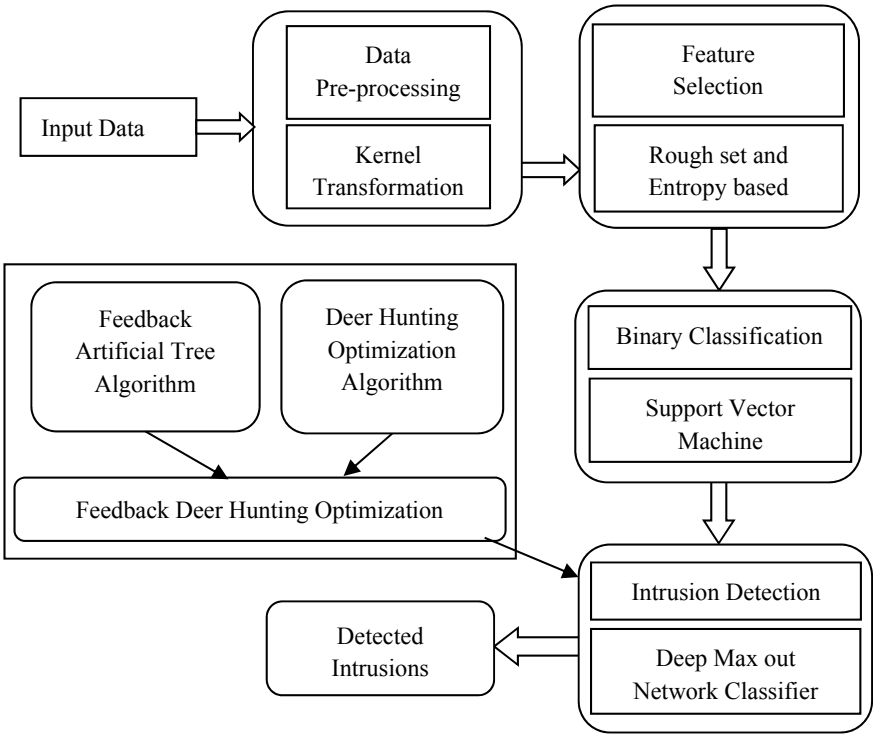


Fig. 3 Block diagram of the proposed approach

maxout network [DMN] classifier. Figure 3 represents the working principle of the proposed approach.

The procedure of the proposed IDS model will be carried out by the following steps: In the beginning step, the data will be preprocessed to reduce the noise and external artifacts using kernel transformation. Once the data was preprocessed, the feature selection process will be carried out to obtain the selected feature using rough set and entropy-based method. The selected features will be allowed to perform binary classification such that the features will be classified by support vector machines. However, the classified results will be used to detect attacks using deep maxout network classifiers whereas the training will be on the basis of feedback deer hunting optimization algorithm [FDHOA].

The proposed FDHO algorithm is derived by the combination of the feedback artificial tree algorithm [FAT] [21] and the DHOA. However, the proposed method will be implemented in the MATLAB tool, and the performance will be evaluated using metrics such as accuracy, sensitivity and specificity. The dataset used for the implementation process is the knowledge discovery dataset [KDD] cup 99 [22] and the BOT-IOT dataset [23].

4.1 Metrics

The tested performance results of IDS are evaluated by using metrics such as accuracy, sensitivity and specificity. The basic attributes [18] are as follows:

- False Positive: Triggering positive without actual attack.
- False Negative: Triggering negative for an actual attack.
- True Positive: Triggering correctly as an attack.
- True Negative: Triggering correctly as non-attack.
- Sensitivity: It is used to determine the propagation of actual positive cases that got predicted correctly.
- Specificity: It is used to determine the propagation of actual negative cases that got predicted correctly.
- Accuracy: The ratio between true negative and correctly predicted observation.

5 Conclusion

The intrusion detection systems need to be improved to be capable of being protective of the increase in intrusions. In this paper, different types IDS were studied and analyzed. The existing methods have many merits and demerits. Some of the systems were incapable of adapting to the environment and some were very complex to be integrated into the network. Some other techniques were not accurate in detecting the intrusions. So, there is an opportunity for researching new methods or combinations of techniques to increase the performance of the IDS in the cloud. The upcoming researches will also depend on improving the current IDS techniques.

References

1. Sobin Soniya, S., Maria Celestin Vigila, S.: Intrusion detection system classification and techniques. In: IEEE International Conference on Circuit, Power and Computing Technologies, pp. 147–152 (2016)
2. Milani, B.A., Navimipour, N.J.: A comprehensive review of the data replication techniques in the cloud environments: major trends and future directions. J. Netw. Comput. Appl. **64**, 229–238 (2016)
3. Sucahyo, G., Rotinsulu, Y.: Software as a service quality factors evaluation using analytic hierarchy process. Int. J. Bus. Inf. Syst. **24**(1), 51–68 (2017)
4. Tao, C., Gao, J.: On building a cloud-based mobile testing infrastructure service system. J. Syst. Softw. **124**, 39–55 (2017)
5. Bassiliades, N., Symeonidis, M.: A semantic recommendation algorithm for the PaaSport platform-as-a-service marketplace. Expert Syst. Appl. **67**, 203–227 (2017)
6. Navimipour, N.J., Rahmani, A.M., Navin, A.H., Hosseinzadeh, M.: Expert cloud: a cloud-based framework to share the knowledge and skills of human resources. Comput. Hum. Behav. **46**, 57–74 (2015)

7. Patil, R., Dudeja, H., Modi, C.: Designing an efficient security framework for detecting intrusions in virtual network of cloud computing. Comput. Secur. **85**, 402–422 (2019)
8. Diaz-Verdejo, J., Garcia-Teodora, P., Macia-Fernandez, G.: Anomaly based network intrusion detection: techniques, system and challenges. Comput. Secur. **28**, 18–28 (2009)
9. Lee, J.H., Park, M.W., Eom, J.H., Chung, T.M.: Multi-level intrusion detection system and log management in cloud computing. ICACT 552–555
10. Aldribi, A., Traoré, I., Moa, B., Nwamuo, O.: Hypervisor-based cloud intrusion detection through online multivariate statistical change tracking. Comput. Secur. **88**, 101646 (2020)
11. Modi, C.N., Patel, D.R., Patel, A., Muttukrishnan, R.: Bayesian classifier and snort based network intrusion detection system in cloud computing. In: Third International Conference on Computing, Communication and Networking Technologies (2012)
12. Wang, W., Du, X., Shan, D., Qin, R., Wang, N.: Cloud intrusion detection method based on stacked contractive auto-encoder and support vector machine. IEEE Trans. Cloud Comput. (2020)
13. Hajimirzaei, B., Navimipour, N.J.: Intrusion detection for cloud computing using neural networks and artificial bee colony optimization algorithm. ICT Exp. **5**(1), 56–59 (2019)
14. Abusitta, A., Bellaiche, M., Dagenais, M.: A deep learning approach for proactive multi-cloud cooperative intrusion detection system. Futur. Gener. Comput. Syst. **98**, 308–318 (2019)
15. Wang, W., Ren, L., Chen, L., Ding, Y.: Intrusion detection and security calculation in industrial cloud storage based on an improved dynamic immune algorithm. Inf. Sci. **501**, 543–557 (2019)
16. Balamurugan, V., Saravanan, R.: Enhanced intrusion detection and prevention system on cloud environment using hybrid classification and OTS generation. Clust. Comput. **22**(6), 13027–13039 (2019)
17. Achbarou, O., El Kiram, M.A., Bourkoukou, O., Elbouanani, S.: A new distributed intrusion detection system based on multi-agent system for cloud environment. Int. J. Commun. Netw. Inf. Secur. **10**(3), 526 (2018)
18. Sobin Soniya, S., Celestin Vigila, M.: Improving the efficiency of classifiers for detecting intrusion using rough set theory. J. Adv. Softw. Eng. Test. **1**(1), 1–14
19. Ajay Guleia, A., Raj, K., Devi, S.: Decision tree based algorithm for intrusion detection. Int. J. Adv. Netw. Appl. **7**(04), 2828–28344 (2016)
20. Brammya, G., Ninu Preetha, N.S., Ramya, R., Rajakumar, B.R. Binu, D.: Deer hunting optimization algorithm: a new nature-inspired meta-heuristic paradigm. Comput. J. (2019)
21. Li, Q., He, Z.C., Li, E.: The feedback artificial tree (FAT) algorithm. Soft Comput. 1–28 (2020)
22. Aggarwala, P., Sharmab, S.K.: Analysis of KDD dataset attributes—class wise for intrusion detection. Procedia Comput. Sci. **57**, 842–851 (2015)
23. The BoT-IoT Dataset. https://www.unsw.adfa.edu.au/unsw-canberra-cyber/cybersecurity/ADFA-NB15-Datasets/bot_iot.php. Accessed on Oct 2020

Analysis and Prediction of Cotton Yield with Fertilizer Recommendation Using Gradient Boost Algorithm

B. M. Sagar, N. K. Cauvery, Prashant Abbi, N. Vismita, B. Pranava, and Pranav A. Bhat

Abstract The majority of every nation's economy depends on agriculture. However, unpredictable changes in climatic conditions have affected the cultivation of crops. There is a declining interest for individuals to remain in this field since farmers are not encouraged to use technology for processes. This happens primarily because there is a lack of involvement of information technology in the farming sector. This obstacle can be overcome by applying several techniques of programming to help farmers estimate the yield of their produce and motivate them to remain in this field. Crop yield prediction includes anticipating yield of the crop from available historical information like climatic, acreage parameters and soil parameters. Machine learning is one technique that can be used to predict the crop yield. Using parameters that are associated with the environment, agriculture and farming, machine learning models can predict the crop yield with good accuracy. The research previously done in this sector has proven to be instrumental in forming the basis of this model. The findings have shown that parameters such as nutrient levels, acreage and historical data have a greater impact on the overall yield in addition to the climatic conditions of the place being observed. In this work, a model has been proposed which helps cultivators predict the yield of the crop even before cultivating directly onto the agricultural lands. The yield prediction based on location, acreage and fertilizer data for several

B. M. Sagar (✉) · N. K. Cauvery · P. Abbi · N. Vismita · B. Pranava · P. A. Bhat
Department of ISE, RV College of Engineering®, Bengaluru, India
e-mail: sagarbm@rvce.edu.in

N. K. Cauvery
e-mail: cauverynk@rvce.edu.in

P. Abbi
e-mail: prashantabbi.is19@rvce.edu.in

N. Vismita
e-mail: vismithan.is17@rvce.edu.in

B. Pranava
e-mail: pranavab.is17@rvce.edu.in

P. A. Bhat
e-mail: pranavabhat.is17@rvce.edu.in

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4_105

regions in the United States has been experimented using various algorithms. The parameters tested and the features used introduce a unique insight into the relationship between nitrogen, phosphorous and potash (N-P-K) used (acreage covered and quantity) and the lint yield of cotton. A point to note here is that the model has taken into account the historical observations (1975–2017) in order to recommend fertilizers. This has been made possible using the multi-variate regression algorithm. Additionally, the yield prediction results have shown that the gradient boosting regressor proved to be the best performer, returning a standalone accuracy score of 90.73% and 80.72% on being validated using the fivefold cross validation method. The proposed model aims to design, develop and implement the training model by using different input data. The machine will be able to learn the features and extract the crop yield from the data by using data mining and data science techniques.

Keywords Gradient boosting regressor · Crop yield · Machine learning · Data mining

1 Introduction

Agriculture is the foundation of every country's economy. Throughout history, agriculture has been considered as the fundamental and the foremost cultural practice around the world. In a country like the United States of America, which has a prominent demand of food because of the population, it is required to incorporate information technology by applying new techniques in various areas of agricultural process. Crop yield prediction is one of them.

Nowadays, people do not have the awareness about the right time and correct place for cultivation of the crops. There is a lack of knowledge in the current generation about agricultural processes, cultivation methods and crop growing as compared to their ancestors. However, incorporating technology as a part of these programmes would greatly help the agriculturists.

Various approaches are implemented in order to increase and improve the yield of the crop. Leveraging the use of data mining is one such concept which has proved to be helpful for anticipating the crop production. Basically, data mining is the way of analysing various perspectives in the domain and summarizing the useful information to convert it into result-oriented methods of delivering solutions in the area.

The procedure involves finding the relation between the various features in an extremely large collection of the data obtained through datasets. Every single feature considered is analysed to predict the required output. In mining, mathematical models are applied in order to convert the raw data into processed information which are later retrieved using information retrieval techniques. Extracting the relation between the parameters applied to the algorithm and the final outcome is of importance. Relevance of every feature used and their percentage of contribution in predicting the final outcome are also considered to be prominent.

Based on this solution, the concepts of machine learning can be implemented to predict the crop yield based on the factors affecting the crop production. Prediction of annual crop yield will significantly contribute to the economic development and would add to the farmers profit as well.

It is of significant importance that work has to be carried out to inculcate technological methodologies in the domain since a major portion of the country's population rely on agriculture as their livelihood and are contributing to the economic factors of the nation.

2 Literature Survey

Once the data is processed, analysing the category of the problem is critical. It is identified to be one of classification, clustering or a regression problem. The work involves the problem statement to be under classification category. For a given problem and dataset, it is important to decide on supervised learning, unsupervised learning and reinforcement learning algorithms. Extensive research has been carried out on prediction of crop yield, and a few of them considered to be relevant are discussed here.

Priya [1] demonstrates that the random forest worked efficiently on large databases with higher accuracy. The random forest algorithm primarily consists of three essential parameters, namely ntree, mtry and nodesize. The authors selected rainfall (mm), kharif and rabi season of each year of every district, maximum temperature (degree Celsius) and crop production (Tonnes) as the key features for observation.

Manjula et al. [2] propose millet crop yield prediction by taking high-dimensional datasets. The authors have found that by using random forest classifier they have found an accuracy of 99.74%. The parameters that were considered included soil, min temp, max temp, humidity, rainfall, etc.

Bhanumati et al. [3] highlight the analysis of various attributes like the location, pH value from which the alkalinity of the soil is determined. Along with it, percentage of nutrients like nitrogen (N), phosphorous (P) and potassium (K) prove to be a defining factor in predicting the yield of the crop.

Based on the above readings, the following points were noted and are now considered in the proposed model:

- Features to be used for successful prediction involve area under cultivation, harvested area, soil type, input of fertilizers (preferably nitrogen, phosphorous and potash content).
- Models previously implemented with a high accuracy score are predominantly linear regression, random forests classification and KNN.
- The dataset should include consistent entries with a few outliers for a good result. If possible, inclusion of historical data aids in the recommendation of fertilizers since the yield can vary from state to state.

- Metrics to be included after the completion of the model implementation include RMSE, R^2 and adjusted R^2 values.

3 Proposed Methodology

The proposed methodology contains two phases: training phase and the testing phase. In the training phase, the data was collected and preprocessed from the **US Department of Agriculture** [4].

The conflicting data in this step was removed from the dataset. The rows which did not have consistent data (NA/INF) were replaced with the mean of that particular feature of the corresponding state. The preprocessed data was operated upon the train dataset using the gradient boosting regressor (for predicting the lint yield) and multi-variate regression (for fertilizer recommendation subject to historical data). This training set contained 80% of the complete data, while 20% was part of the test dataset (to determine the accuracy of the model). The training phase ends with generating the rules by quantifying the relationship of the output—lint yield, with the given input features.

In the testing phase, the yield values are predicted based on the generated rules. The existing actual values are compared to the generated predicted values (for the 20% of the data) to determine the accuracy of the model. Key metrics such as standard and adjusted R squared values are used to determine the correctness of the model (Fig. 1).

Data Collection—Obtaining the Dataset

The dataset used is a combined data frame from six different datasets taken from a consolidated report available at the **US Department of Agriculture** [4] for 14 states in the United States of America, from 1975 to 2017.

The final dataset contained the following features:

- **State**—Name of the 14 states in USA
- **Year**—Data from 1964 to 2017 (historical occurrences and records)
- **Name of Fertilizer (%)**—Acreage receiving the particular fertilizer
- **Name of Fertilizer (Pounds/Acre)**—Amount of the fertilizer being given to that acreage
- **Area Planted (Acres)**—As mentioned
- **Harvested Area (Acres)**—As mentioned
- **Lint Yield (Acres)**—Yield of the cotton crop.

Data Preprocessing and Feature Selection

As part of the preprocessing stage, the missing values were replaced with the mean values. Further to increase accuracy, the values were later replaced with the mean values of that feature, corresponding to the state, since states were showing varied

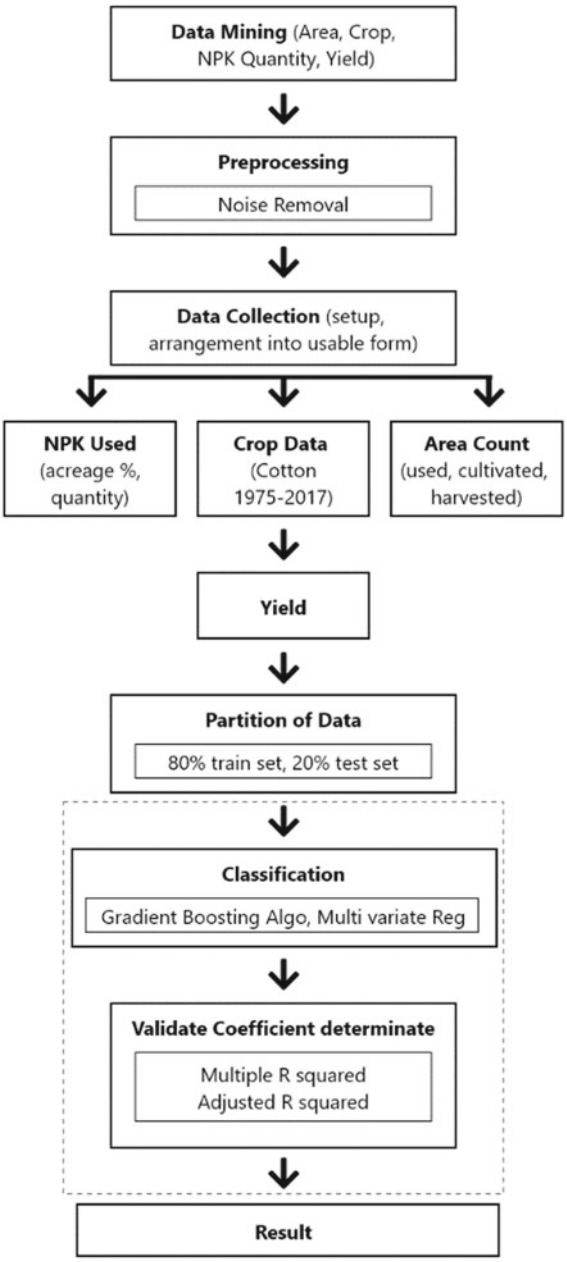


Fig. 1 Methodology diagram

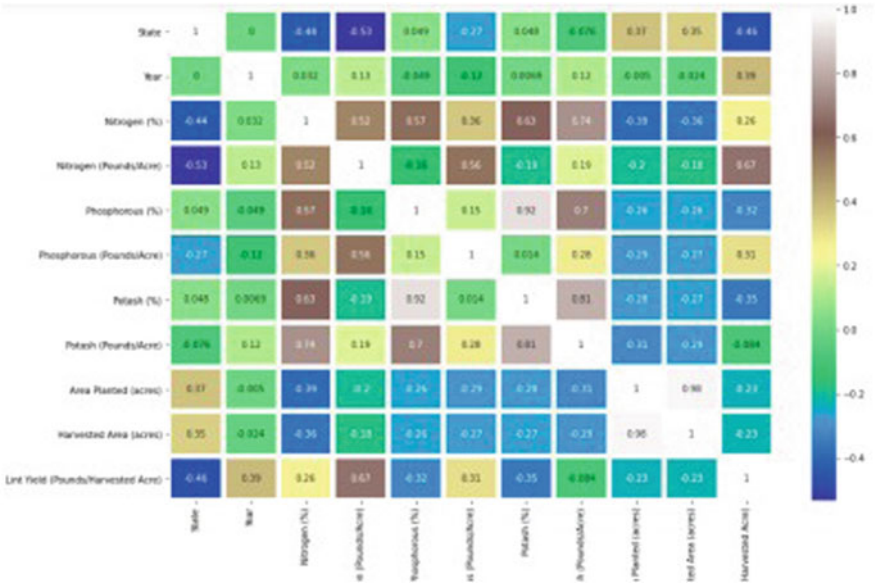


Fig. 2 Correlation matrix of features

values that were not in correlation with each other. This prevented overfitting of the model since it was optimized for 14 different states.

Unique values of States were noted, mapped and labelled to corresponding integral values, to get a dataset that was capable to be trained on the regression model used.

Pearson's correlation was used to check for redundant features. Fortunately, all the features showed a value greater than |0.2| which shows a medium–high relation between the features and the target variable (lint yield) (Fig. 2).

A correlation matrix provides a visualization of the correlation between features in the dataset by analysing the coefficients. Each cell in the table shows the correlation between two features/variables. A correlation matrix is used to summarize data, as an input into a more advanced analysis and as a diagnostic for advanced analyses. Here is a correlation matrix of the features considered relevant in the dataset.

Algorithms used in this proposed model

Algorithms used are linear regression, linear regression with L2 regularization, polynomial regression, random forest regression, gradient boosted regression and nearest-neighbour regression.

These models mentioned as above are inclined to over fitting which can be taken care of by carefully tuning the hyper-parameters. Using fivefold cross validation, comparison is done for performance and found that using mostly default settings, the random forest regressor worked out the best, trailed by nearest-neighbour regression and L2 linear regression with polynomial features.

Table 1 Accuracy report

S.No.	Metric	Score
1	R^2 value	90.73%
2	Adjusted R^2 value	90.60%
3	Root mean squared error (TRAIN)	44.51
4	Root mean squared error (TEST)	96.53
5	Mean absolute error	71.42
6	Mean squared error	9318.90

This affirmed the previous idea that absolute linear models are not appropriate for this data/feature space. Interestingly, gradient boosted regression—a recent favourite among many machine learning competitions—performed effectively. But this algorithm has a sizable number of hyper-parameters. Some tuning of the parameters brought the performance up to levels that surpassed the random forest regressor.

Tuning the hyper-parameters of the random forest regressor could not accomplish the same performance as with the gradient boosted regressor algorithm.

The following two algorithms were identified for the model that are as follows:

Gradient Boosting Regression: The intuition behind gradient boosting algorithm is to leverage the pattern in residuals and strengthen a weak prediction model, until the present residuals become randomly distributed. Once the residuals do not have any pattern that could be modelled, the process is terminated to prevent any case of overfitting. Algorithmically, this means minimizing the loss function, such that test loss reaches its minima. This was used to predict the lint yield of the cotton crop.

Multi-variate Regression: This algorithm is used to derive and define the linear relation between more than one independent input variable, referred to as predictor, and more than one dependent output variable, known as response. This was used to recommend fertilizers for a given state, depending on area inputs.

Accuracy Report

Apart from the above metrics, the model was validated using the fivefold cross validation method to rule out cases of over fitting or under fitting of the model. It returned an array with accuracy score of **80.72%** that portrayed the correctness of the model (Table 1).

A similar approach was used for checking the accuracy of the fertilizer recommendation system (using multi-variate regression). Here, the R2 value returned for the testing process was **81.11** and **93.18%** for train (Fig. 3).

4 Results

This figure depicts the output of the total lint yield for the top ten states in the USA since 1975. California tops the table with the maximum total yield showing that

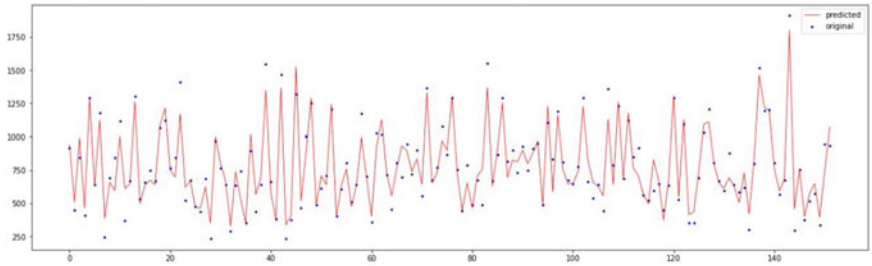


Fig. 3 Visualization of the relation between original and predicted values

in spite of a relatively smaller area that is planted and harvested, the state has still managed an above average yield (Fig. 4).

This indicates that there is something that the state is executing differently (Fig. 5).

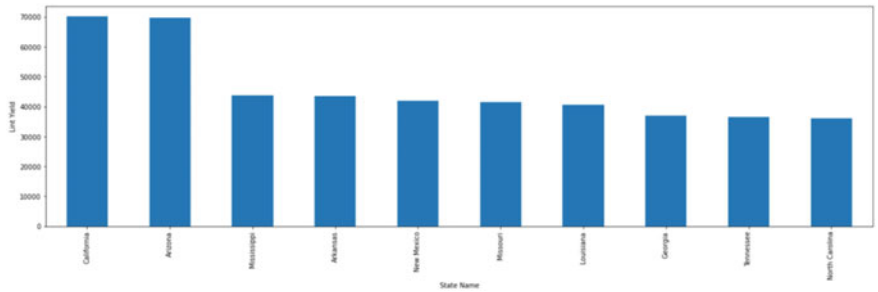


Fig. 4 Lint yield versus state

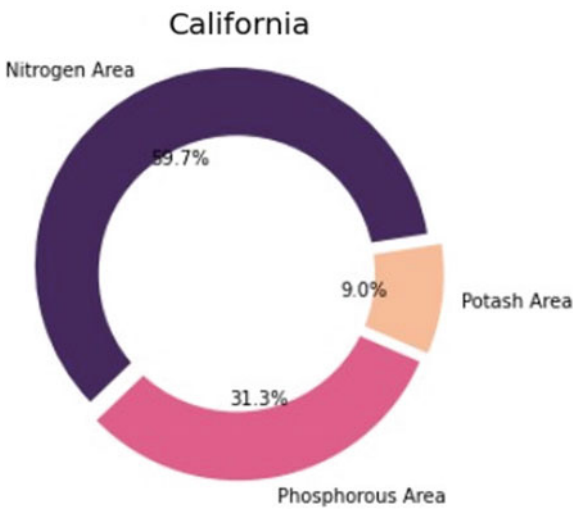


Fig. 5 NPK acreage in California

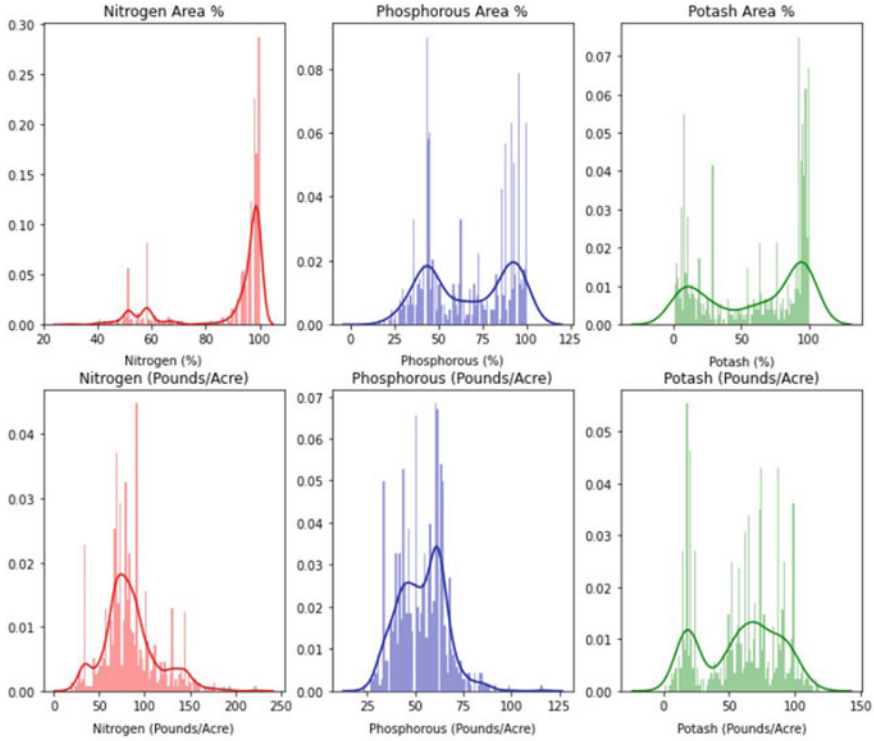


Fig. 6 Relation of NPK with lint yield of cotton

On conducting further analysis, it was observed that California has been implementing a strategy involving the greater use of **nitrogen and phosphorous** as compared to potash. Hence, over the years, they have been able to maintain a consistently increasing yield, in spite of having lesser area for plantation and harvesting. Correctly, this fact has also been verified by science, since nitrogen is the most absorbed mineral from the soil and is an essential factor in determining the yield of a crop (Fig. 6).

On visualizing all the data concerned to NPK acreage use and its quantity, it has been established that nitrogen and phosphorous have contributed towards the yield of the cotton crop in the USA, while the net yield does not entirely depend on the area planted and area harvested.

5 Conclusion

The above discussed results section clearly shows that gradient boosted decision tree regressor is best when compared to other regression technique. The tuned model

achieved an R^2 value of ~0.91. Switching to this algorithm might have helped, but (possibly) at the cost of performance. A more careful feature engineering can possibly offset this effect or by using certain ensemble techniques. Additionally, it was hard to overcome overfitting completely. Getting more data would have been an obvious solution.

However, the findings from this paper can be used by researchers in this field in the future to predict the crop yield efficiently by leveraging the use of historical data, NPK ratings and acreage covered, along with the climate parameters in different regions of the world to help farmers gain profit by estimating their yield and accordingly managing their resources to maximize the effects of their efforts.

References

1. Priya, P., Muthaiah, U., Balamurugan, M.: Predicting yield of the crop using machine learning algorithms
2. Manjula Josephine, B., Ruth Ramya, K., Rama Rao, K.V.S.N., Kuchibhotla, S., Venkata Bala Kishore, P., Rahamathulla, S.: Crop yield prediction using machine learning. Int. J. Sci. Technol. Res. **9**(02) (2020). ISSN 2277-8616
3. Bhanumathi, S., Vineeth, M., Rohit, N.: Crop yield prediction and efficient use of fertilizers. In: 2019 International Conference on Communication and Signal Processing (ICCSP)
4. Fertilizer Use and Price, Economic Research Service. US Department of Agriculture. https://www.ers.usda.gov/data-products/fertilizer-use-and-price.aspx

Author Index

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*, Lecture Notes in Networks and Systems 191,
https://doi.org/10.1007/978-981-16-0739-4