Sridaran Rajagopal · Kalpesh Popat ·
Divyakant Meva · Sunil Bajeja ·
Pankaj Mudholkar (Eds.)

Communications in Computer and Information Science 2428

Artificial Intelligence Based Smart and Secured Applications

Third International Conference, ASCIS 2024
Rajkot, India, October 16–18, 2024
Revised Selected Papers, Part V

Part 5

Springer

Communications in Computer and Information Science 2428

Series Editors

Gang Li, *School of Information Technology, Deakin University, Burwood, VIC, Australia*

Joaquim Filipe, *Polytechnic Institute of Setúbal, Setúbal, Portugal*

Zhiwei Xu, *Chinese Academy of Sciences, Beijing, China*

Rationale
The CCIS series is devoted to the publication of proceedings of computer science conferences. Its aim is to efficiently disseminate original research results in informatics in printed and electronic form. While the focus is on publication of peer-reviewed full papers presenting mature work, inclusion of reviewed short papers reporting on work in progress is welcome, too. Besides globally relevant meetings with internationally representative program committees guaranteeing a strict peer-reviewing and paper selection process, conferences run by societies or of high regional or national relevance are also considered for publication.

Topics
The topical scope of CCIS spans the entire spectrum of informatics ranging from foundational topics in the theory of computing to information and communications science and technology and a broad variety of interdisciplinary application fields.

Information for Volume Editors and Authors
Publication in CCIS is free of charge. No royalties are paid, however, we offer registered conference participants temporary free access to the online version of the conference proceedings on SpringerLink (http://link.springer.com) by means of an http referrer from the conference website and/or a number of complimentary printed copies, as specified in the official acceptance email of the event.

CCIS proceedings can be published in time for distribution at conferences or as post-proceedings, and delivered in the form of printed books and/or electronically as USBs and/or e-content licenses for accessing proceedings at SpringerLink. Furthermore, CCIS proceedings are included in the CCIS electronic book series hosted in the SpringerLink digital library at http://link.springer.com/bookseries/7899. Conferences publishing in CCIS are allowed to use Online Conference Service (OCS) for managing the whole proceedings lifecycle (from submission and reviewing to preparing for publication) free of charge.

Publication process
The language of publication is exclusively English. Authors publishing in CCIS have to sign the Springer CCIS copyright transfer form, however, they are free to use their material published in CCIS for substantially changed, more elaborate subsequent publications elsewhere. For the preparation of the camera-ready papers/files, authors have to strictly adhere to the Springer CCIS Authors' Instructions and are strongly encouraged to use the CCIS LaTeX style files or templates.

Abstracting/Indexing
CCIS is abstracted/indexed in DBLP, Google Scholar, EI-Compendex, Mathematical Reviews, SCImago, Scopus. CCIS volumes are also submitted for the inclusion in ISI Proceedings.

How to start
To start the evaluation of your proposal for inclusion in the CCIS series, please send an e-mail to ccis@springer.com.

Sridaran Rajagopal · Kalpesh Popat ·
Divyakant Meva · Sunil Bajeja ·
Pankaj Mudholkar
Editors

Artificial Intelligence Based Smart and Secured Applications

Third International Conference, ASCIS 2024
Rajkot, India, October 16–18, 2024
Revised Selected Papers, Part V

Springer

Editors
Sridaran Rajagopal
Marwadi University
Rajkot, Gujarat, India

Kalpesh Popat
Marwadi University
Rajkot, Gujarat, India

Divyakant Meva
Marwadi University
Rajkot, Gujarat, India

Sunil Bajeja
Marwadi University
Rajkot, Gujarat, India

Pankaj Mudholkar
Marwadi University
Rajkot, Gujarat, India

ISSN 1865-0929 ISSN 1865-0937 (electronic)
Communications in Computer and Information Science
ISBN 978-3-031-86301-1 ISBN 978-3-031-86302-8 (eBook)
https://doi.org/10.1007/978-3-031-86302-8

© The Editor(s) (if applicable) and The Author(s), under exclusive license
to Springer Nature Switzerland AG 2025

This work is subject to copyright. All rights are solely and exclusively licensed by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

If disposing of this product, please recycle the paper.

Preface

"A Little Progress Each Day, Adds up to a Big Result"

Our 3rd International Conference on Advancements in Smart Computing & Information Security (ASCIS 2024) invited original research papers under five tracks, namely, **AI & ML, Smart Computing, Cyber Security, Networking and Cloud Computing and Computer Applications for Sustainability.** We received 667 submissions and 181 were recommended for publication in Springer's CCIS series. The papers will appear as 6 volumes (Volume Numbers ranging 2424–2429). We are very much thankful to the editorial board of **Springer CCIS** for their continuous support.

The submitted papers were involved in an open review process, involving three reviewers for each paper. We are thankful to the Technical Program Committee (TPC) members who were very supportive in the review process.

Several sponsors who contributed generously towards the successful running of ASCIS 2024, including **Science and Engineering Research Board (SERB), K7 International and D- Link Academy**. These collaborations enriched the conference and extended its outreach.

ASCIS 2024 was conducted during 16–18 October, 2024, and featured insightful keynote addresses from national and international experts in their respective fields. The line-up included **Prabir Kumar Biswas** from IIT Kharagpur, Cybersecurity expert **Ram Kumar G, Arnav Bhavsar** from IIT Mandi, **Murali Raman** from Asia Pacific University, and **Aravind Benegal** from NLB Services.

The keynotes addressed several recent technological advancements whereas the hands-on workshops provided practical exposure to the students and participants. The PhD forum featuring the research works of research scholars and recent PhD graduates added value to the conference and acted as a platform for knowledge sharing.

These volumes titled **"Artificial Intelligence based Smart & Secured Applications"** are the consolidation of all such solutions employing technologies of AI, ML, Smart Computing and, Cybersecurity in every significant field seeking research solutions.

October 2024 Sridaran Rajagopal

Organization

General Chair

R. Sridaran	Marwadi University, India

Program Committee Chairs

R. Sridaran	Marwadi University, India
Kalpesh Popat	Marwadi University, India
Divyakant Meva	Marwadi University, India
Sunil Bajeja	Marwadi University, India
Pankaj Mudholkar	Marwadi University, India

Steering Committee

Ketanbhai Marwadi	Marwadi University, India
Jitubhai Chandarana	Marwadi University, India
P. Bhanu Prasad	Future Technology
R. B. Jadeja	Marwadi University, India
Sanjeet Singh	Marwadi University, India
Prabir Kumar Biswas	IIT Kharagpur, India
Anil Kumar Sao	IIT Bhilai, India
Arnav Bhavsar	IIT Mandi, India
Naresh Jadeja	Marwadi University, India
Shamala Subramanyam	Universiti Putra Malaysia, Malaysia
Padmavathi	Avinashilingam Institute for Home Science and Higher Education for Women, India
T. Devi	Bharathiar University, India
K. Iyakutti	SRM Institute of Science and Technology, India
Ram Kumar G.	Global Automotive Company, India

Technical Program Committee Chairs

AI and ML Track

Hitesh Chhinkaniwala	Adani University, India
Ahlad Kumar	National Forensic Sciences University, India
Mahendra Kumar Gourisaria	Kalinga Institute of Industrial Technology, India

Cyber Security Track

Ruchira Naskar	Indian Institute of Technology, Kharagpur, India
Shravani Shahapure	Deloitte Touche Tohmatsu, India
Ritika Vivek Ladha	Adani University, India

Smart Computing Track

Manish Khare	Dhirubhai Ambani Institute of Information and Communication Technology, India
Dinesh Kumar	Bennett University, India
Rebakah Geddam	Nirma University, India

Networking and Cloud Computing Track

Rachna Jain	Bhagwan Parshuram Institute of Technology, India
Santosh Kumar Majhi	Guru Ghasidas Vishwavidyalaya (Central University), India
Nishant Doshi	Pandit Deendayal Energy University, Gandhinagar, India

Computer Applications for Sustainability Track

Manas Ranjan Pradhan	Skyline University College, UAE
J. Ramkum	Krishna Arts and Science College, India

Additional Reviewers

Arun Raj Lakshminarayanan
Dhanamma S. Jagli
Tushar H. Jaware
Kamal Sutaria
Mahalakshmi G.
Nebojsa Bacanin
Rahul Ashok Patil
Satvik Khara
Vijay D. Katkar
Vipin Sharma
Amarinder Kaur
Anita Gautam Khandizod
Bhoopendra Singh
Delecta Jenifer Rajendren
Harishchander Anandaram
Karuna Nidhi Pandagre
Amrita Kumari
Dipti H. Domadiya
Prashant Sen
Rajasekaran Selvaraju
Saran Raj Sowrirajan
Harshal Anil Salunkhe
Sumathy S.
Hemant Ingale
Jamberi K.
Jayashree Nair
Jyoti Kharade
Keerti Jain
Kumaresh N.
Mythili Shanmugam
Nikita Bhatt
Parvathaneni Naga Srinivasu
Praveen P. V. Sundar
Ravirajsinh S. Vaghela
Richa Adlakha
Ritesh Patel
Sharanyaa S.
Srinivasulu Raju S.
Swetta Kukreja
Vinothina V.
Ajay Kushwaha
Anbumani K.
Ashwin B. Makwana
Chandresh K. Kumbharana
Divyakant T. Meva
Raji C. G.
Anita M. Indu
Kailash Patidar
Kannadhasan Suriyan
Hardik K. Molia
Jayashri Pattar
Jonnadula Narasimharao
Kaviyarasi R.
Ramakrishna Garine (Ram)
Nageswari D.
Prema C.
Rajender Kumar
Ram Ratan
S. Amuth Sachin H. Gajjar
Sreenaga Chidambara Vinjamuri Dattu
Vipul Shah
A. Yovan Felix
Abhay Singh Bhadauria
Aditi Sharma
Ajay Kumar Kushwaha
Ajita Deshmukh
Akash Saxena
Amita Sharma
Ankit Subhash Didwania
Anubhav Kumar Prasad
Arun Adiththan
Asha V.
Ashokkumar Baldevbhai Prajapati
Avinash Sharma
Bhagvati Parekh
Brijesh R. Jajal
Dafni J. Rose
Dhivya K.
Dipti Chauhan
Divya R.
Divyanshu Chandra
Ashwin R. Dobariya
Prabhat Sharma
Rohit Kanauzia
Abhinav Tomar
Akshara Dave

Ashish Saini
Avnip Deora
Dimple M. Thakar
Gaurav Kumar Ameta
Jayant R. Nandwalkar
Kapil Joshi
Lokesh P. Gagnani
Nisha Khurana
Samir B. Patel
Sri Hari Nallamala
Sunil L. Bajeja
T. Sathish Kumar T.
Parwinder Kaur
Rajesh Bansode
Saraswathi S.
T. Buvaneswari Muruganantham
Vijayalakshmi P. S.
Galiveeti Poornima
Hari Kumar Palani
Hemraj S. Lamkuche
Jafar Ali Ibrahim Syed Masood
Jay A. Dave
Jebakumar Immanuel D.
Juliet Rozario
Kajal S. Patel
Karthikeyan Eswaramurthy
Kavitha Ganesh
Krishnakumar A.
Lata Suresh
Latchoumy P.
M. R. Ramesh
Madhu Shukla
Manisha Rawat
Maruthamuthu R.
Maulika Patel
Megha P. Mudholkar
Mohamed Iqbal M.
Mohit Tiwari
Muralidharan R.
Narayan Joshi
Neeru Sharma
Neha Sharma
Nidhi M. Patel
Nilu Singh
Pankaj K. Mudholkar
Parth Gautam
Pranav Verma
Praveen Kumar
Priyanka Sharma
Purnendu Bikash Acharjee
Radha B.
Rajesh Yadav
Rakesh Kumar Yadav
Renjith V. Ravi
Richa Ashok Modiyani
Rohit Goyal
Rupali Atul Mahajan
S. S. Priscila
Saifullah Khalid
Sandeep Mathur
Saswati Mahapatra
Senthil Kumar R.
Shamala K. Subramaniam
Shanti Verma
Shweta Sharma
Srinivasa Rao Pallapu
Subramanian Balambigai
Sudipta Hazra
Sunil Gautam
Suresh B. Kumar
Tanmay Kasbe
Thirumurugan Shanmugam
Uma N. Dulhare
Umesh S. Pinjarkar
Vaibhav C. Gandhi
Vandani Verma
Vipin Sharma
Vishvjit K. Thakar
Yogesh Ghodasara
Chandran C. P.
Raghu N.
Jaimin N. Undavia
Kavipriya P. P.
Mohamed Mosbah
Padma Selvaraj
Remegius Praveen Sahayaraj
Shadab Siddiqui
Vineet Kumar Singh
Abhishek Sharma
Anilkumar Suthar

Bharanidharan G.
Chandra J.
Disha J. Shah
John T. Abraham
P. Rizwan Ahmed
B. Surendiran
Dushyantsinh B. Rathod
Priya K.
Rushikumar R. Raval
Ajitha V.
Kajal Patel
Gourav Gupta
J. Grace Hannah
Jasminder Kaur Sandhu
Jinal H. Tailor
Karthik B.
Kruti K. Sutaria
K. P. Malarkodi
Naresh Kumar
Nirav Bhatt
Pradip Mathuradas Jawandhiya
Rajeswari S.
Revathy G.
Ripal D. Ranpara
Senthilkumar Meyyappan
Shikha Maheshwari
Sweeti Sah
Vijay Karnatak
A. Maheswary
Anamika Rana
Anjali Diwan
Biswaranjan Mishra
Chintan Patel
Himanshu K. Maniar
T. S. Murugesh
Darshankumar C. Dalwadi
Anitha K.
Priya Chandran
Jalpesh Vasa
Jeevitha R. R.
Jose M. Molina
Lilly Florence M.
Mohan Subramani
Naresh Kshetri
Qixia Zhang
Rajeswari M.
Roshan Anant Gangurde
Safvan Vahora
Vijaya Kittu Manda
Yugendra Devidas Chincholkar
A. P. Nirmala
Abhilasha Vyas
Ahmed Banimustafa
Ajay M. Patel
Ajitha I.
Akoramurthy Balasubramaniam
Angeline R.
Ankur N. Shah
Anup Palsokar
Aruna Animish Pavate
Ashish Kumar
Ashwin Raiyani
Balraj Verma
Boopathi Raja G.
Chandra Mohan
Darshita S. Pathak
Dipak Ramoliya
Divya Didwania
Anant G. Kulkarni
Chintan B. Thacker
Deepak Kumar Verma
Rashmi Soni
Saraswathi S.
Abhishek Sharma Padmanabhan
Ankit J. Faldu
Asmita Manna
Daxa Vekariya
Disha H. Parekh
Jay Kumar Jain
Jaypalsinh A. Gohil
Krupa S. Mehta
Nafees Akhter Farooqui
Premkumar Borugadda
Rupesh Kumar Jindal
Shadab Siddiqui
Sudhanshu Maurya
Sunil Gupta
Vatsal H. Shah
Rajan Patel
Ramesh T. Prajapati

Sreejith Vignesh B. P.
Veena Soni
G. Charles Babu
Gaurav Agrawal
Harish Padmanaban
Himanshu Rai
Jane K. Nithya
Jaydeep R. Ramani
Juhi Singh
Jyotirmoy Pathak
Kamal Saluja
Karthikeyan R.
Kedir Lemma Arega
Kumuthini C.
Lataben J. Gadhavi
Lipsa Das
M. Vinoth Kumar
Mahalakshmi G.
Manoj Kr. Mishra
Mastan Vali Shaik
Meet Patel
Minal S. Shukla
Mohammed Wajid Khan
Monther Ali Tarawneh
Nagaraju Kilari
Neha Parashar
Neha Ripal Soni
Nilesh Vijay Sabnis
Pagalla Bhavani Shankar
Paresh V. Virparia
Pragadeswaran S.
Prashant P. Pittalia
Preethi S. R.
Priyanka Suyal
Pushparaj Pal
Rajib Biswas
Ramkumar Jaganathan
Rinkoo Bhatia
Ruchika Pharswan
Rutvi R. Shah
Sadhana Singh
Samriti Mahajan
Sanjeevkumar D. Angadi
Savitha Jaganathan
Shafi Pathan
Shantha Mary Joshitta
Shilpa Mehta
Sivanesan Rajangam
Subhadip Sarkar
Sudha R.
Suhasini Vijaykumar
Sunil Saxena
Surya Kameswari Uduga
T. D. Sudhakar
Tarannum Bloch
Tripti Tiwari
Umang Thakkar
Vaibhav A. Gandhi
Vallidevi Krishnamurthy
Vikas Tripathi
Vipul A. Shah
Yogendra Kumar
Chetan Dudhagara

Our Sponsors

Sr. No.	Name of the Sponsor	Logo
1	SERB India	
2	D-Link	
3	Digital India	
4	K7 International	
5	Samatrix.io	
6	Marwadi Technologies	
7	Stelcore	

Abstracts of Keynotes

Transforming Industry 4.0 with Generative AI and Cloud Computing

Gundarapu Pavan

Safran DIFIT, France

Industry 4.0, characterized by the convergence of automation, IoT, and advanced analytics, is reshaping industries globally. As organizations strive for greater efficiency, innovation, and adaptability, the synergy between cloud computing and generative AI is emerging as a transformative force. This keynote will explore the vital role of cloud computing as the backbone of Industry 4.0, providing scalable infrastructure and real-time data accessibility. It will delve into what generative AI is, how it acts as a catalyst for innovation, and how it accelerates the development of intelligent systems, from predictive maintenance to automated design and decision-making. By examining the synergies between these two technologies, the keynote will illustrate their potential to revolutionize industrial operations and unlock new business value.

Attendees will gain insights into the tangible benefits and ROI of integrating cloud and AI solutions. The keynote will also address the challenges organizations face—ranging from data security to infrastructure optimization—and offer strategic considerations for successful implementation. Looking ahead, we'll explore future trends in Industry 4.0, where cloud-enabled generative AI could redefine the industrial landscape, driving sustainable growth and continuous innovation.

The session will conclude with a call to action, encouraging students and industry leaders to embrace these technologies to remain competitive in a rapidly evolving digital economy.

Brain Decoding with Deep Learning

Arnav Bhavsar

IIT, Mandi, India

The talk will focus on visual brain decoding via EEG using deep learning techniques. Brain decoding, also known as Perceptual Brain Decoding (PBD), involves using brain responses to different stimuli to identify the original stimulus or its characteristics.

Specifically, we target an external visual stimulus, which involves object images on a computer screen. The goal is to reconstruct the class of the image from the EEG data captured with observing the image. The decoding involves either classifying reconstructing the type of image shown during the visual task.

For reconstructing class-specific images, we will discuss the utilization of generative networks such as GANs, using embeddings from a pre-trained EEG classifier network as input. Additionally, we will also cover an end-to-end generative network (NeuroGAN) for image synthesis, where the generator produces class-specific embeddings and images from EEG signals.

Advanced Insights into Deep Fake Recognition

Deepak Kumar Jain

Dalian University of Technology, China

The rapid advancement of artificial intelligence and machine learning has led to the proliferation of deep fake technology, enabling the creation of highly realistic, AI-generated synthetic media. While deep fakes offer exciting possibilities in fields like entertainment and virtual reality, they also pose significant threats to public trust, cyber security, and political stability. This presentation, *Advanced Insights into Deep Fake Recognition*, explores the complex techniques used to produce deep fakes and highlights the growing challenges in detecting them. We delve into the mechanics of deep fake creation, particularly the role of Generative Adversarial Networks (GANs), and discuss the real-world implications of this technology's misuse, from identity theft to political disinformation.

The presentation also examines cutting-edge detection methods, including AI-driven facial and audio analysis, and emerging technologies like block chain and quantum computing, which are being employed to verify the authenticity of digital content. Furthermore, we explore current applications of deep fake detection tools across social media platforms, law enforcement, and the media, while considering the ethical and legal issues surrounding the use and regulation of deep fake technology. By providing a comprehensive overview of the challenges and future trends in deep fake recognition, this presentation aims to equip audiences with the knowledge to navigate this evolving digital landscape.

Security Models and Strategies of Cloud Computing in Business and Public Administration

Zdzislaw Polkowski

WSG University, Poland

Globalization and the development of information and communication technologies (ICT) have caused significant changes in business and public administration in the European Union. Currently, the digital economy is driven by modern ICT systems that offer employers and employees new tools for effective operation. Cloud Computing systems deserve special attention, as they have an impact on business, public administration and the everyday life of citizens. Cloud Computing creates many benefits not only for users but also has a positive impact on the environment, which is in line with the & quot; European Green Deal & quot; strategy implemented in Europe. However, the security of Cloud Computing and many previously unseen threats are always in the focus of both cloud users and providers of this modern solution. Threats, such as zero-day exploits, advanced persistent threats (APTs), and insider attacks are some popular threats that strained the cloud services in business and public administration. These threats create a large barrier to widespread use and in many cases delay the further development of these systems. The regulations constantly introduced and updated in Europe take into account the development of secure Cloud Computing, sensibly regulating the ways of processing not only personal data, but also data processing in general. It is worth mentioning that research and analysis have been carried out in real companies and public administration institutions to build robust and secure cloud platforms that can tackle modern-day cyber threats. The analysis highlights that the best outcomes in cloud computing implementations are driven by the collaboration of interdisciplinary teams, specifically those composed of IT specialists, legal experts, and managers. This synergy ensures a holistic approach, balancing technical efficiency with legal compliance and strategic oversight. Special attention should be given to security strategies, such as robust data encryption, identity and access management, regular auditing, and compliance management to mitigate the negative consequences of security breaches in cloud environments. These advancements will be particularly beneficial for a wide audience, including IT professionals, researchers, students, and business leaders, all of whom play a role in implementing and managing cloud solutions. Furthermore, future research could explore how these interdisciplinary teams tackle emerging challenges, such as evolving regulatory landscapes, data sovereignty, and the integration of advanced technologies like AI and machine learning in cloud infrastructure. Such investigations would contribute to the ongoing development of best practices and position cloud computing as a more secure and sustainable solution in a rapidly changing digital ecosystem.

Keywords: Cloud Computing · IT security · Business · Public administration · European Green Deal

Role of Computing and Technology in Sustainable Development Goals

Murali Raman

Asia Pacific University of Malaysia

We are living in a highly Volatile, Uncertain, Complex and Ambiguous (VUCA) world, which threaten the long-term sustainability of global business, economy and natural living. The Sustainable Development Goals (SDGs) outlined by the United Nations serve as a global roadmap for achieving a more equitable and sustainable future. As the world grapples with pressing challenges such as climate change, poverty, and inequality, computing and technology have emerged as powerful tools for addressing these complex issues. This keynote explores the pivotal role of computing and technology in advancing the SDGs, highlighting specific applications and their potential impact.

Computing and technology can contribute to the SDGs in various ways. For instance, in the realm of climate change, advanced analytics can be used to optimize energy consumption and identify sustainable energy sources. Artificial intelligence can enhance the efficiency of renewable energy systems and predict natural disasters, enabling early warning systems and disaster preparedness. Additionally, computing-powered solutions can facilitate sustainable agriculture by optimizing crop yields, reducing water consumption, and minimizing the use of harmful chemicals.

In the context of poverty and inequality, computing and technology can empower marginalized communities through digital inclusion and education. Affordable and accessible internet connectivity can bridge the digital divide, providing opportunities for education, healthcare, and economic development. Furthermore, online platforms can facilitate access to financial services, enabling individuals and small businesses to escape poverty.

Computing and technology can play a crucial role in promoting good health and well-being. Telemedicine and remote healthcare services can improve access to quality healthcare, especially in rural and underserved areas. Additionally, data analytics can be used to identify disease outbreaks and track their spread, enabling effective public health interventions.

In summary, computing and technology offer immense potential for addressing the SDGs. By harnessing the power of these tools, we can create a more sustainable, equitable, and prosperous world for all. This keynote uses case examples from different parts of the world, showcasing how SDGs can be addressed with proper utilisation of computing and technology. However, realizing this potential requires collaboration among governments, businesses, and academia to develop innovative solutions and ensure equitable access to technology.

Cybersecurity in the Age of 4.0 IR

Srinath Doss

Botho University, Botswana

In the era of Industry 4.0, where cyber-physical systems, the Internet of Things (IoT), and artificial intelligence drive unprecedented connectivity and automation, cyber security has never been more critical. As industries leverage smart technologies to enhance productivity, efficiency, and innovation, the digital landscape expands, bringing both opportunities and vulnerabilities.

The convergence of operational technology with information technology in smart factories, autonomous vehicles, and interconnected supply chains creates a complex and dynamic environment. A single security breach can disrupt production, compromise sensitive data, and even endanger lives. Cyber attacks are evolving in sophistication, targeting the very backbone of our digital infrastructure.

Cybersecurity in the age of Industry 4.0 demands a proactive and multi-layered approach. It requires the integration of advanced threat detection systems, robust encryption protocols, and comprehensive risk management strategies. Organizations must foster a culture of cybersecurity awareness, ensuring that employees at all levels understand and adhere to security best practices.

Collaboration across sectors is essential. Governments, private enterprises, and academic institutions must work together to develop innovative solutions, share threat intelligence, and establish global cybersecurity standards. Continuous investment in research and development is crucial to staying ahead of cybercriminals and safeguarding our digital future. In this digital revolution, cybersecurity is not just a technical necessity but a fundamental pillar of trust and resilience. By embracing a holistic and forward-thinking approach, we can protect our assets, secure our data, and pave the way for a safer, more interconnected world.

Keywords: Industry 4.0 · Cyber security · Internet of things · Cyber Attack · Risk Management

Trust in the Realm of Generative AI – Opportunities and Threats

Arvind Benegal

Chief Risk Officer & Sr. Advisory Consultant, USA

Trust is of paramount importance in Information Technology systems and data confidentiality, integrity and availability. AI and Generative AI have reshaped the technology landscape. A balanced approach to innovation is needed that incorporates both Trust and Risk Management. The threats represent the clear and present risks and challenges posed by emerging technologies, while the opportunities speak to the benefits brought by these innovations. This highlights the dual nature of technological advancement. Wherever there is an opportunity, there is also a significant threat.

The opportunities include the efficiency gains of automation of repetitive tasks and enhanced decision-making. Personalization addresses tailored experiences for users in healthcare, finance, and customer service. Innovation allows for creation of new products, services, and business models. Democratization of technologies provides access, challenging elite layers, and Just in Time Knowledge suggests just enough to solve problems.

The risks include Bias and Fairness since AI systems may perpetuate or amplify existing biases. Security vulnerabilities pose increased attack surfaces and potential misuse of AI capabilities. Data Privacy presents risks of unauthorized data usage and breaches, and Prompt Injections increase exposure to sophisticated AI attacks.

The Audit Focus Areas should consider Data Integrity. It must be ensured that data used by AI systems is accurate, complete, and secure, and ethical AI establishes frameworks to prevent bias and ensure fairness. Accountability and transparency are achieved by creating mechanisms for auditing AI decisions and actions. Institutional guidance provides adherence to IT governance frameworks such as COBIT to manage AI risks. Robust control measures are implemented to enhance trust. For instance, ISO/IEC 42001 is an international standard that provides a governance framework for implementing and continually improving artificial intelligence management systems.

Trust is built by by incorporating strategies for IT Auditors. Continuous Monitoring regularly reviews AI systems and processes for compliance and performance. Collaboration with stakeholders by engaging with developers, data scientists, and business leaders aligns AI initiatives with ethical standards. Education and Awareness train employees on the risks and best practices related to AI. And finally, a Risk-Based Approach is adopted by prioritizing audits and controls based on the risk levels associated with different technologies.

Machine Learning and Data Analytics at Scale - Rise of Unified AI and Data Platforms

Kumar Gangwani

BDB.ai, India

The last 10+ years have seen a breakneck pace of innovation in areas of Data Analytics, Machine Learning and AI, something that remained unseen in many decades prior. The availability of inexpensive storage, affordable compute, innovations in parallel computing frameworks, global availability of open-source tools and libraries, automation of data engineering, algorithmic innovations and availability of specialized hardware like GPUs and TPUs have ushered in the modern Data and AI era.

The open-source ecosystem has democratized access to cutting-edge tools and technologies to the larger technology community. For instance, frameworks like TensorFlow and PyTorch allow developers and researchers to build sophisticated models quickly and efficiently.

Yet, most organizations, despite good intentions, sizeable investments in technology and large technology teams, struggle to scale after successfully operationalizing a few initial analytics and ML use cases. In fact, the rate of failures of corporate data science and machine learning teams is so high that the notion of the data and AI-driven enterprise often seems a myth.

Despite the availability of a plethora of tools for data engineering and machine learning, most organizations fail to build and scale with these effectively. Part of this has to do with the overpromise around capability of enterprise tools sold to them, which sometimes are good at a particular solution like ingestion or discovery only; this results in organizations purchasing a number of tools that don't integrate well with each other or scale for increasing data volumes and diverse data demands. But that's not the only reason. Fragmented, inconsistent and incomplete data is often a factor responsible for failure of Machine Learning projects. Integrating ML workflows into production environments, handling versioning, scaling models and automating monitoring – a discipline being referred to as ML Ops – remain challenging for many businesses; and add to that rising complexity of algorithms itself. Most technical resources have specialization in perhaps one or two areas, but a very limited view of the entire data chain, and struggle to solve platform scale issues.

The whole industry today longs for an integrated approach that allows operation of analytics and ML at scale. The past few years have seen the rise of a few unified platforms that integrate the entire data infrastructure and plethora of data processes – from ingestion (batch, micro batch, real-time), data cleansing, enrichment and preparation, storage into data lake, with right security and governance, notebook/tools to experiment with and deploy Data Science and ML models, enable ML automation, use of AI services and self-service dashboards and visualization – all under one roof. These unified platforms

greatly simplify and accelerate the time to commercialize analytics use cases, often 4–5 times faster than commonly used tools, and the end-to-end automation of all data processes greatly reduces resource effort by a large margin, providing for sizeable cost savings. This integrated capability offered by unified data and AI platforms is the future of enterprise analytics and Machine Learning.

Implementing Cyber Security Best Practices to Protect Your Crown Jewels

Ram Kumar G.

Global Automotive Company, India

In a hyper-connected digital world, cyber threats are ever growing and evolving, putting confidential data supporting IT infrastructure at risk. Cyber crimes are common today and no organization is immune to cyber attacks. The cyber crime industry has grown into a multi-billion dollar illegal business and the profile of the attacker has changed over the years – from script kiddies to well-organized hacker groups supported by nation states.

In this prevailing scenario, it is important for organizations to be cognizant of the cyber threats and proactively put in place cyber security measures by way of security controls, processes backed up by skilled talent. A strong understanding of the multi-faceted domains of cyber security is essential to appreciate the range and depth of the field. This exposure leads to exploring the security best practices that need to be implemented in organizations as per their risk profile to safeguard their critical IT infrastructure and confidential business information including intellectual property in what has come to be characterized as "crown jewels". The security controls are defined by the results of a risk assessment exercise, situational awareness of ongoing cyber attack campaigns enabled by inputs from cyber threat intelligence, obligations to customers and regulatory requirements for cyber security and data privacy.

Keywords: Cyber security · Best practices · Crown jewels · Cyber attacks · Data protection · Critical assets · Cyber threats · Security risks

Smart Spatial Computing and Its Application

P. Bhanu Prasad

Future Technology, France

Spatial computing is an emerging technology that blends digital and physical worlds, allowing users to interact with smart devices in more seamless and immersive ways. It leverages a range of technologies, including augmented reality (AR), extended reality (XR), Artificial intelligence, Internet of things (IoT), haptic feedback, computer vision, sensor fusion, geospatial data analysis, and spatial mapping. Spatial computing integrates various technologies to interpret the physical world and merge it with digital information, effectively combining both realms. These experiences can range from fully immersive environments to the addition of digital elements in the real world, creating interactive experiences that enhance our understanding of spatial relationships. At its core, spatial computing empowers users to visualize, manipulate, and analyze complex spatial data in real time with ease.

Advancements in hardware and software enable developers to create rich, immersive, interactive environments that captivate users. This evolution not only improves the overall user experience but also opens up new possibilities. Despite the significant potential of spatial computing, challenges persist. Concerns regarding data privacy, security, and ethical considerations need to be addressed before these technologies can be widely deployed. Additionally, the lack of standardization and interoperability among different spatial computing platforms presents further obstacles to widespread adoption.

This immersive, natural way of interacting can enrich our lives in numerous ways, enhancing hobbies like entertainment, gaming, and sports, as well as educational experiences and professional activities. The applications of spatial computing are extensive and diverse, impacting vital sectors that shape daily life, including navigation, astronomy, virtual home design, learning through gaming, skill development, healthcare and fitness, virtual shopping, industrial settings, remote assistance, architecture and design, the stock market, and logistics and supply chain management.

Spatial computing signifies a paradigm shift in how we interact with various environments, unlocking a vast array of new applications. Here are some key applications of spatial computing that can help those interested in pursuing a career in this field get started.

Keywords: Spatial computing · Extended reality (XR) · Industrial Internet of Things (IIoT) · Sensor fusion, Digital Twins · Artificial intelligence, Haptic

Enhancing Recommender Systems with Extensions of Fuzzy Sets for Considering Stakeholders' Uncertainty in Software Requirements Prioritization

Vassilis C. Gerogiannis

University of Thessaly, Larissa, Greece

In large, complex and distributed software development projects, it is often not feasible to implement all of the requirements in a single software release because of budget, time and staff constraints. Software requirements prioritization is an important activity in software requirements engineering which aims to evaluate and prioritize the candidate software features to be implemented in the next release of a software system. Requirements prioritization is often applied iteratively, according to various criteria, by the multiple involved project stakeholders who may have different roles, needs and knowledge. In large-scale software development projects, it is often practically impossible to ensure that all stakeholders have complete knowledge on all candidate requirements. Thus, stakeholders may show some degree of uncertainty and hesitation, as it is difficult to precisely evaluate each requirement according to each prioritization criterion. This talk will present results of our ongoing research work on combining Recommender Systems and soft computing techniques based on extensions of Fuzzy Sets to address problems of requirements prioritization in the context of large and complex software development/maintenance projects. Recommender Systems can be a useful solution to information overload situations when stakeholders have to evaluate a large number of alternatives. Fuzzy sets extensions, such as Intuitionistic Fuzzy Sets, can deal with stakeholders' uncertainty and hesitation regarding the prioritization criteria importance and requirements ratings. The results from applying our methods are promising since they demonstrate that they can be effectively used to support multiple stakeholders to prioritize large requirements sets.

Keywords: Software Requirements Prioritization · Recommender Systems · Intuitionistic Fuzzy Sets

Generative Adversarial Network (GAN): An Efficient Tool for Data Generation and Data Completion

Prabir Kumar Biswas

IIT, Kharagpur, India

Generative Adversarial Network (GAN) is a relatively recent concept in deep generative learning wherein two neural networks are pitted against each other in a zero-sum non-cooperative game to match a non-stationary distribution to an intractable stationary distribution. In this lecture we shall talk about how the generative modelling capability of GAN can be exploited to design efficient frameworks for image/video inpainting. Semantic inpainting refers to realistically filling up large holes in an image or video frame. Our approach is to first train a generative model to map a latent noise distribution to a natural image manifold and, during inference time, search for the best-matching noise vector to reconstruct the signal. The primary drawback of this approach is its inference time iterative optimization and lack of photo-realism at higher resolution. In this talk both of the abovementioned shortcomings are addressed. This is made possible with a nearest neighbour search-based initialization (instead of random initialization) of the core iterative optimization involved in the framework. The concept is extended for videos by temporal reuse of solution vectors. Significant speedups of about 4.5-5x on images and 80x on videos is achieved. Simultaneously, the method achieves better spatial and temporal reconstruction qualities.

Next, we will pose the following question and try to get the answer - 'Do we at all need an iterative inference framework?' - to answer this, a data-driven parametric network is trained to directly predict a matching prior for a given masked image. This converts an iterative paradigm to a single feed-forward inference pipeline with around $800\times$ speedup. Finally, recent advancements in high-resolution GAN training are leveraged to scale an inpainting network for higher resolution.

Artificial Intelligence and Machine Learning: From Early Symbolic Systems to More Complex System

Salisu Mamman Abdulrahman

Kano University of Science and Technology, Nigeria

Artificial Intelligence (AI) and Machine Learning (ML) have evolved significantly over the years, moving from simple symbolic systems to more complex and sophisticated systems capable of learning and adapting. Advances in AI and ML systems have revolutionized various sectors, changing how people interact with technology. For four decades (mid-1950s until the mid-1990s), Symbolic AI was the dominant paradigm. These algorithms work by processing symbols that represent objects or concepts in the real world and their relationships. In Symbolic AI the main approach is to use logic-based programming where rules and axioms are used to perform inference and deduction. High expectations surrounding AI's potential to solve complex problems and mimic human intelligence were met with disappointing results, leading to what is known as the AI Winter. Despite the challenges faced by researchers during the AI Winter, the field experienced a revival in the 1990s and early 2000s. AI Researchers shifted their focus from Symbolic AI and rule-based systems to statistical and probabilistic approaches that are more promising. The emergence of machine learning algorithms, such as Support Vector Machines and Bayesian networks, enabled AI systems to demonstrate tangible successes in various applications. Machine learning systems have evolved dramatically over the years, progressing from basic models capable of simple tasks to complex neural networks that drive cutting-edge AI applications. Throughout this evolution, AI and ML have transitioned from rule-based systems to learning-based approaches, greatly expanding their capabilities and applications. In this keynote address, we will present an overview of their journey from inception to recent advances.

Keywords: Artificial Intelligence · Machine Learning · Symbolic System · Complex System

Attention Mechanism for Abnormal Behavior Detection from Video Surveillance

Vinothina V.

Kristu Jayanti College (Autonomous), Bengaluru, India

In todays fast-paced environment, detecting abnormal behaviour in real time is increasingly challenging, yet video monitoring remains vital for public safety. Traditional surveillance systems still rely heavily on manual observation, which is not only time-consuming but also prone to errors, especially when managing large volumes of footage. While machine learning (ML) algorithms have been widely used for behaviour detection, they often struggle with the complexity of unstructured video data. These algorithms typically require vast amounts of labelled data and can falter when dealing with imbalanced datasets, reducing their accuracy in identifying rare, yet critical, behaviours. Moreover, ML techniques tend to have slower learning rates, limiting their effectiveness in real-time scenarios and large-scale data applications. In contrast, deep learning models, especially convolutional neural networks (CNNs), have proven more effective by automatically learning feature representations directly from raw video data. However, CNNs still face challenges when it comes to capturing the temporal dynamics of behaviour across video frames. To address these challenges, the proposed method integrates a hybrid model that combines the strengths of CNNs and transformers with a context-based attention mechanism.

Modelled after the human brain's ability to focus on the most pertinent aspects of a scene, attention mechanisms integrated into deep learning frameworks have significantly improved both the accuracy and efficiency of abnormal behaviour detection. When combined with CNNs, these attention mechanisms emphasize key spatio-temporal features. Spatial attention allows the model to detect critical body movements and gestures, while temporal attention tracks changes in behaviour over time, identifying sudden or repetitive actions. This approach allows for real-time tracking of individual actions within the video frame, enabling timely alerts and interventions when abnormal behaviours are detected. When evaluated against existing methods, the proposed solution demonstrated superior performance, effectively overcoming the limitations associated with traditional machine learning approaches in behaviour recognition.

Keywords: Attention Mechanism · Abnormal Behaviour Detection · Video Surveillance · Deep Learning · Spatial Attention · Temporal Attention · Contextual Information

Zero Trust and Its Impact

Chander Shekhar Sharma

AUNZ - Cybersecurity Unit, Australia

The Zero Trust security model represents a fundamental shift in how organizations approach cybersecurity. Unlike traditional security frameworks that rely on a strong perimeter to protect internal networks, Zero Trust operates on the principle of "never trust, always verify." This approach assumes that threats can exist both inside and outside the network, and therefore, every access request—whether from a remote employee or an internal user—must be authenticated and authorized. By focusing on the identity of users and devices rather than the network location, Zero Trust transforms the way organizations secure their resources. At the heart of Zero Trust is the concept of identity as the new security perimeter. This means that organizations must implement robust identity verification mechanisms to ensure that only authorized individuals can access sensitive information. Multi-factor authentication (MFA) is a critical component of this strategy, requiring users to provide multiple forms of identification before gaining access. Additionally, risk-based adaptive authentication can further enhance security by adjusting access levels based on various factors, such as the user's location, device, and behaviour patterns. This continuous verification process helps organizations mitigate the risk of unauthorized access and data breaches.

The benefits of adopting a Zero Trust identity strategy are substantial. For instance, organizations can significantly reduce the likelihood of data breaches by ensuring that every access request is scrutinized. This not only protects sensitive information but also fosters a culture of security awareness among employees. Furthermore, Zero Trust can improve the user experience by streamlining access to resources while maintaining stringent security measures. For example, remote workers can securely access company applications from any location without compromising security, thereby enabling greater flexibility and productivity.

Use cases for Zero Trust identity are increasingly relevant in today's digital landscape. Organizations that have adopted cloud services or remote work policies can leverage Zero Trust principles to secure their environments effectively. For instance, a financial institution might implement Zero Trust to protect customer data by requiring MFA for all transactions and continuously monitoring user behaviour for anomalies. Similarly, a healthcare provider could use Zero Trust to ensure that only authorized personnel can access patient records, thereby complying with regulations and safeguarding sensitive information. By humanizing the approach to cybersecurity, Zero Trust not only addresses technical challenges but also aligns with the evolving needs of organizations and their employees in a rapidly changing digital world.

Empowering Academic Advising Through Smart Computing and Predictive Analysis

Ahmed Al-Brashdi

UTAS, Oman

The role of academic advising has evolved from static and transactional to proactive support, emphasizing innovative approaches to address the variations of students' performance and needs. By leveraging historical data on student performance, course complexity, and former evaluation, predictive analytics can identify students at risk of academic difficulty. An early warning system with comprehensive student data and predictive analysis based on historical data enables timely interventions, such as academic coaching or tutoring, to improve student retention and success. Moreover, using predictive analytics and AI will optimize resource allocation by balancing advisor workloads, and optimizing course scheduling. This data-driven approach contributes to cost-efficiency and improved student satisfaction. This approach can personalize the academic experience for each individual student by recommending tailored academic plans based on individual student strengths, goals, and performance patterns without overloading advisors. This approach fosters student engagement and accelerates progress towards degree completion. In UTAS – Oman, an in-house system enables advisors and students to effectively and efficiently engage with personalized study plans and course scheduling with emphasis on on-probation students. It shows improvements in student satisfaction and resource allocation.

Keywords: Academic advising · Predictive analysis · Smart computing · AI

Best Practices of Research Ethics

Vijay Singh Rathore

Shree K. Karni Universe College, Jaipur

Research ethics constitute a cornerstone of scientific inquiry, ensuring the credibility, integrity, and societal relevance of scholarly endeavours. This study delineates the fundamental principles governing ethical research practices, emphasizing adherence to honesty, objectivity, and transparency in all stages of research, from data collection to dissemination. Key ethical mandates include the avoidance of fabrication, falsification, and plagiarism, as well as the active protection of confidentiality, intellectual property rights, and the dignity of research participants.

The role of supervisors is highlighted as pivotal in fostering an ethical research environment, ensuring adherence to legal and institutional norms, and mentoring researchers in responsible practices. Equally, researchers are tasked with diligent participation in their projects, rigorous compliance with safety protocols, and acknowledgment of collaborative contributions.

Ethical challenges, such as authorship disputes, data ownership, and conflicts of interest, are addressed through structured protocols and mutual respect among collaborators. Furthermore, responsible dissemination of findings underscores the importance of accuracy and societal impact, avoiding premature or misleading claims. This framework underscores the collective responsibility of the research community to uphold ethical standards, thereby advancing scientific progress while maintaining public trust.

Contents – Part V

Smart Computing

Survey on Security Models Employed in Blockchain Based Applications 3
Lochana Mahajan, Sridaran Rajagopal, and Rujuta Shah

Comparative Analysis of Improved K-Means Clustering for Human Freedom Index 26
F. Mohamed Ilyas, S. Silvia Priscila, K. Sheela, J. Vimala Roselin, K. V. Sona, and R. Prema

Advancements and Innovations in Mixed Reality: A Comprehensive Review 40
Jitesh Solanki and Kalpesh Popat

Enhancing Sleep Disorder Prediction: CNN with IQR and Lasso Preprocessing 57
V. Nagarajan and C. Meenakshi

A Survey on Various Sentiment Analysis Applications Employing Opinion Mining 69
Meghna Bhatt, Sridaran Rajagopal, and Kinjal Raval

Mitigating the Technological Challenges of Regenerative Farming with an Integrated Framework Merging IoT, Ontologies, and GIS 95
Tshepiso L. Mokgetse, Rajagopal Sridaran, and Hlomani Hlomani

Leveraging Blockchain Technology: Decentralized Finance (DeFi) in Health Insurance 107
D. Susana and V. Srividya

Analysis of Security with Countermeasures in IoT Based Applications – A Survey 121
V. Sini and Sridaran Rajgopal

Efficient Train Traffic Prediction Using Support Vector Machines with Advanced Preprocessing Techniques 132
C. Radhika and D. Kerana Hanirex

A Comparative Analysis of Nerve Vessels Enhancement and Segmentation Techniques Through Edge Detection on Retinal Fundus Images 145
N. Durga, D. KeranaHanirex, and A. Muthukumaravel

Solidarity Linking Women and Technology: An Analysis 161
Anuradha Acharya and Kenil Thakkar

Detection and Classification of Epileptic Seizures Using Machine Learning and Deep Learning: A Systematic Review and Challenges 183
Shirish Mohan Dubey, Kamlesh Gupta, and Shikha Agrawal

Application of Smart Computing in Steel Structural Health Monitoring: Sensor Based Damage Detection for Smart Infrastructures 195
Husain Rangwala and Tarak Vora

Application of Smart Computing Systems for Smart Cities and Urban Infrastructure: Framework, Data Management, and Smart Monitoring Attributes 212
Ankur Bhogayata, Amit Thoriya, Tarak Vora, and Divyakant Meva

Decentralized EHR Exchange in Healthcare: Enhancing Privacy and Security with Blockchain and Cryptographic Techniques 235
Aqeel A. Yaseen, Kalyani Patel, Abdulla J. Y. Aldarwish, and Ali A. Yassin

A Proposed Model to Transliterate Digitalized Braille Text into Gujarati: An Aid for Visually Challenged People 247
Nikisha Jariwala and Kalpesh Popat

Financial Data-Driven Classification of Manufacturing Companies: A Case Study of the Malaysian Industry 261
Rathimala Kannan and Kannan Ramakrishnan

Smart Choices: How Computer Applications Shape Consumer Decisions on Household IoT Appliances 276
R. Chitra, M. Sathish, and Jone Stervin

A Study on Augmented Reality's Impact on Online Shopping in India: Enhancing Consumer Experience and Behaviour 289
N. Vijaykumar, T. M. Charuhasini, and Arul Rajan

Utilizing Digital Twins for Predictive Modelling in Manufacturing Cyber-Physical Systems 309
B. Uma Maheshwari, R. Sujatha, V. Deekshitha, and Viswanath Ananth

Multi-sensor Data Fusion for Accurate Prediction of Suspicious Activities in Video 322
R. Radhika and A. Muthukumaravel

The Emergence of the Social Internet of Vehicles (SIoV): A Comprehensive Analysis of Architecture, Technologies, and Applications 335
Tarandeep Kaur and Pankaj Deep Kaur

Network and Cloud Computing

Innovative Perceptions: A Survey on Named Data Networks Through an Inventive Lens 351
A. Shanmuga Priya and K. Vaitheki

Optimizing Resource Utilization and Performance in Multi-cloud Infrastructures Through Predictive AI Analytics 365
Arun Pandiyan Perumal

VANET Security Through Blockchain Based Location Verification: Innovation and Challenges 375
S. Sandosh, B. Thriam Bakesvar, J. Subash Saajan, and R. Eniyavasanthan

Classification and Analysis of Online Business Models: A Framework for Practical Application 391
Roland Schmuck

A Comprehensive Analysis of Low-Energy Adaptive Clustering in Underwater Acoustic Networks 411
N. Karthika and D. Kavitha

Research Perspectives on Load Balancing Strategies in Serverless Computing 422
Yashwanth Balan and R. Arokia Paul Rajan

Real-Time Node's Power-Aware Kubernetes Scheduler in a Cloud Environment 433
Kiran Kumari and Dhananjay Dakhane

EAPTEEN: Energy-Aware Adaptable Routing for Continuous Monitoring in Event-Driven Wireless Sensor Networks the EAPTEEN Approach 444
D. Jeya Rani and Nagarajan Munusamy

Author Index 461

Smart Computing

Survey on Security Models Employed in Blockchain Based Applications

Lochana Mahajan[1,2(✉)], Sridaran Rajagopal[3], and Rujuta Shah[4]

[1] Fergusson College (Autonomous), Pune, Maharashtra, India
lochana.mahajan@gmail.com
[2] Marwadi University, Rajkot, Gujarat, India
[3] Faculty of Computer Applications, Marwadi University, Rajkot, Gujarat, India
[4] Chimanbhai Patel Institute of Computer Applications, Ahmedabad, Gujarat, India

Abstract. Blockchain technology has already gained significant importance in various fields involving medium to large scale applications. It has also gained research importance, in recent years, in order to make it application-friendly in terms of factors like security and de-centralization. Further research is also consistently going on towards Blockchain adoption, absorption, and business values. Out of these factors, security plays a significant part, so it is essential to understand the different threats affecting the Blockchain based applications to be considered while developing them. The proposed survey discussed here is based on various security risks revolving around Blockchain-based applications with the support of current state-of-the art research being carried out in different fields, such as IOT, Metaverse, Cloud based applications, and other relevant fields, it is believed that the proposed survey would greatly support the Blockchain application users and developers for the proper application of the technology.

Keywords: Blockchain · Security · Cryptocurrency · Smart Contract · IoT · Proof of Stake · Proof of work

1 Introduction

Blockchain, an emerging technology, is crucial in decentralized technologies as well as applications like computation of storage, security, communication, and transaction. It has the potential to revolutionize business activities and has been applied across various disciplines and industries. There exist many good definitions for Blockchain. A Blockchain consists of decentralized ledgers integrated into a network, providing anonymity, cryptographic data, consensus protocol, smart contract, and reputation mechanisms [44]. It offers benefits such as decentralization, intractability, strategic applications, security, and decision-making, [1] making it a valuable tool for businesses and management. In other words, asset tracking and transaction recording in a network of enterprises are made simpler by blockchain, a shared, immutable ledger [1, 44]. A house, automobile, cash, or land are examples of tangible assets. Intangible assets include intellectual property, patents, copyrights, and branding. Nearly everything of value may be registered and traded on a blockchain network, reducing expenses and risk for those involved.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
S. Rajagopal et al. (Eds.): ASCIS 2024, CCIS 2428, pp. 3–25, 2025.
https://doi.org/10.1007/978-3-031-86302-8_1

Blockchain technology has gained significant attention in academia, with scholars focusing on its mode of operation, technical characteristics, and possible challenges. Most research focuses on Blockchain-based service provision and technology application [1], reflecting the trend towards adoption, absorption, and business value of Blockchain technology. Blockchain uses an encrypted database and P2P network for data storage and communication, forming a special framework for value exchange.

The fundamental essentials of Blockchains are discussed in the subsequent sections below.

1.1 Significance of Blockchain Based Applications

Blockchain technology offers instantaneous, shareable, and transparent information storage for authorized network users, enhancing trust and efficiency in business operations. It allows for monitoring accounts, production, payments, and orders, ensuring all members have access to the same truth. This eliminates the need for external validations, duplicate record-keeping, and potential fraud and cyberattacks.

The introduction of IoT has increased transaction volume, causing a slowdown in organizations and a negative impact on the bottom line. Blockchain is a solution for increased productivity, security, and trust. It fosters confidence by providing a common record of truth, which in turn fuels other emerging technologies that boost productivity, openness, and trust. By implementing blockchain, organizations can improve their bottom line and reduce costs associated with external validations and duplicate record-keeping.

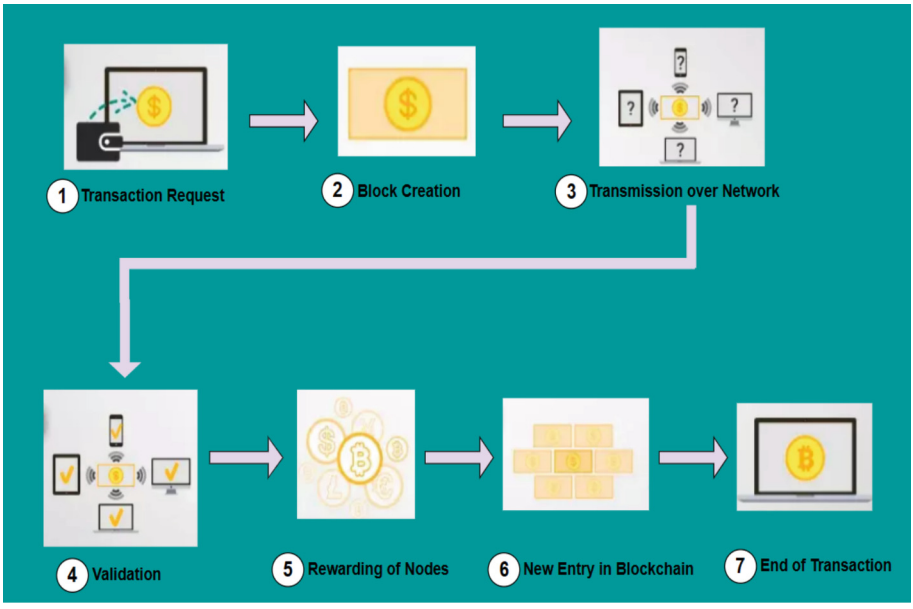


Fig. 1. Architecture of Blockchain Technology.

As seen in Fig. 1 above, the Blockchain architecture involves of several key components which includes nodes, Blockchain protocol, smart contracts, blocks, distributed ledger, peer-to-peer networks, and wallets. Nodes are hardware devices that connect to the Blockchain, maintaining a copy of the ledger and validating transactions. Consensus methods include Delegated Proof of Stake (DPoS), Proof of Work (PoW), Proof of Stake (PoS), and Practical Byzantine Fault Tolerance (PBFT). The Blockchain protocol outlines rules for maintaining the Blockchain, validating transactions, and reaching consensus. Smart contracts [6, 7] are self-executing contracts that work automatically when preset requirements are met. Transactions represent the transfer of resources or data, and decentralization reduces the need for a central authority. Cryptographic hash functions safeguard data integrity, and distributed ledgers reduce the risk of a single point of failure. Peer-to-peer networks allow direct communication between nodes without intermediaries.

Following are the key factors need to be considered while developing the applications using Blockchain:

i) **Scalability:** Blockchain technology is being researched for its scalability, privacy, and interoperability. Researchers are exploring techniques like sharding, layer-2 protocols, and improvements to consensus algorithms to increase transaction throughput and reduce latency.
ii) **Privacy and confidentiality**: are also being addressed through the development of privacy-preserving technologies like zero-knowledge proofs and confidential transactions.
iii) **Interoperability**: is crucial for building a cohesive and productive Blockchain ecosystem, enabling cross-chain transactions and interoperable decentralized apps.
iv) **Security**: is a key focus of Blockchain security research, focusing on identifying and thwarting potential risks such as weak consensus algorithms, 51% attacks, and smart contract vulnerabilities.
v) **Sustainability**: is another focus, with research focusing on more environmentally friendly consensus techniques to reduce the environmental impact of Blockchain mining. Proof-of-stake (PoS) and other energy-efficient consensus techniques are being investigated.
vi) **Decentralization and governance**: are also being explored, with discussions on on-chain and off-chain governance systems to ensure the longevity and adaptability of Blockchain networks.
vii) **Decentralized finance (DeFi) and tokenomics**: Tokenomics research focuses on the economic and incentive systems found in Blockchain networks, with the best way to create tokens and incentive systems for various DeFi applications still under investigation.
viii) **Regulatory and legal considerations**: are also being considered, with governments and regulatory agencies increasingly involved in Blockchain technology development.

The Blockchain technology has showed impressive application opportunities since its early development. Blockchain has been used in countless sectors, starting with the first cryptocurrency, and continuing with the present smart contract. The security of

Blockchain systems has not been thoroughly examined, despite a few research on privacy and security concerns pertaining to Blockchain technology applications. This work surveys real-world assaults against Blockchains by looking at well-known Blockchain systems, and we perform a systematic analysis on the threats to Blockchain security.

The remaining parts of this paper are organized as follows: Sect. 2 highlights some of the similar surveys available in the literature. Section 3 contains the proposed survey on threat models used in Blockchain based applications based on majority of key factors discussed above. Section 4 provides the conclusion and future directions of the on-going research work.

2 Related Work

Blockchain technology is gaining popularity for secure transactions due to its unique time stamp feature. However, it also poses security threats like hacking attacks. A survey [41] aims to highlight these threats and discuss available techniques to manage them. The paper explores the integration of blockchain with cloud technology [48], assessing access control, searchable encryption schemes, and cloud data center performance. It also discusses cloud threats, intrusion detection, and data breaches. The aim is to assist blockchain researchers and users. One survey reviewed the emerging Blockchain-as-a-Service (BaaS) [65] model, assesses access control and searchable encryption schemes, and examines cloud data center performance with blockchain support. Whereas, this paper discusses cloud threats [9] Intrusion detection and data breaches.

The metaverse and blockchain are gaining attention for their potential to revolutionize Building Information Modeling (BIM). Blockchain technology ensures fairness and security in resource transactions. The integration of BIM components, metaverse applications, and blockchain research, aiming to foster collaboration between academia and industry for further development and interdisciplinary research was discussed [66]. Another survey [67] explores the fusion of Blockchain and Artificial Intelligence in the metaverse, exploring components, digital currencies, AI applications, and blockchain-empowered technologies. Collaboration between academia and industries is needed for further research. Current paper focuses on virtual property and virtual harassment threats.

Cryptocurrencies like Bitcoin and Ethereum [63] use decentralized blockchain technology. However, blockchain is vulnerable to attacks from malicious users. This survey article provides a comprehensive overview of blockchain, including taxonomies, structure, and transactions in the cryptocurrency network. It also discusses consensus protocols, smart contracts, forks, and consensus generation techniques. The article also discusses key blockchain platforms like Hyperledger and multichain, emerging vulnerabilities, defensive methodologies, and future trends in blockchain.

The state-of-the-art in cryptocurrencies, including consensus mechanisms, applications, literature contributions [68], and ongoing projects. It also highlighted open challenges and suggests research directions in the domain.

The potential of blockchain for big data [69] which was discussed, focused on its decentralization and security. It discussed the integration of blockchain and big data, its services for secure data acquisition, storage, analytics, and privacy preservation, and its applications in smart cities, healthcare, transportation, and grids. It also presented representative blockchain-big data projects and discusses challenges and future directions

for further research in this promising area. A Survey introduced a blockchain-based enhanced cloud relational database (RDB) [70] for improved security in cloud computing. The system used self-verification mechanisms to detect and prevent erroneous manipulation.

Chang et. al., [45] discusses about decentralized IoT [64] platforms are being implemented due to Blockchain technology's inherent capabilities, enabling the safe and auditable exchange of data across connected smart devices in diverse, context-aware scenarios. Blockchain interoperability improves conventional trade, e-commerce, and transportation networks by enabling independent and secure real-time payment services.

The literature mainly discusses traditional challenges like scalability, energy consumption, and cost in developing Blockchain applications. While there is ongoing research on security, it is crucial to focus on specific threats faced by different applications. Decentralization is a major factor in Blockchain-based applications, and the increase in internet access users introduces additional threats. A complete survey is needed to address these challenges and present a more comprehensive approach to security in Blockchain applications.

3 Categorization of Threats to Blockchain Based Applications

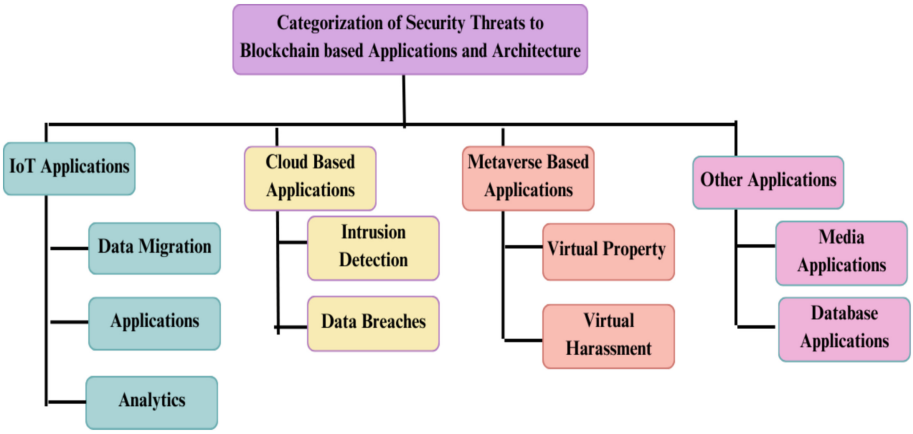


Fig. 2. Categorization of Security Threats to Blockchain Based Applications.

Blockchain is the fundamental technology driving this profound change in the internet's landscape, and WEB3 is a trending buzzword in the tech industry. Blockchain technology and cybersecurity typically complementing one another and are interconnected. Since Blockchain-based systems operate on a distributed design [6] rather than the conventional client-server architecture [6], they are, by nature more secured than traditional systems. Blockchains do, however, have some special attack routes and their own set of cybersecurity issues. Both the core Blockchain and the application levels are potential sources for these attack vectors [7]. The few security measures unique to

enterprise Blockchain solutions are management of identity and access, central administration, data confidentiality, safe correspondence in the operations, security of smart contracts, endorsement of the transaction and so on. The overall security approach of a commercial Blockchain solution uses both conventional security controls and controls specific to the technology.

The goal of this work is to examine some of the most important security holes in the fundamental Blockchain designs that could be exploited to compromise Blockchain-based applications.

The classification in the proposed survey is based on various security risks revolving around Blockchain-based applications with the present state-of-the art research being carried out in different fields, such as IOT, cloud, metaverse, and other relevant fields, are shown in Figure Fig. 2. These categories are further explored in the subsequent sections below:

3.1 IOT Applications

The IoT is a rapidly growing technology that integrates physical objects for data exchange and management. The Internet of Things (IoT) has revolutionized productivity and financial gains by enabling remote sensing and control of devices [43] across networks. However, security lapses [24] can occur due to the rapid growth of industrial data volume [23] needs data security. IoT [64] can enhance smart grids, allowing monitoring and control anytime, but its reliance on Internet-based protocols exposes it to ICT threats like DoS and domain-specific attacks. A safe, dependable, and easily accessible infrastructure is crucial for processing and storing [20, 21] computed data.

Blockchain technology offers numerous benefits for IoT, but its adoption is hindered by resource-constrained environments with limited computational, storage, and energy capabilities. Blockchain is not suitable for IoT due to its bandwidth overhead, delay, limited scalability, and high computational complexity. To address these issues, new technologies or a combination of existing ones are needed [19, 43]. Deep learning and federated learning have shown promising results [24] in detecting security breaches, but they still face security vulnerabilities [23] from hackers. This work proposes a Blockchain-enabled federated learning application model for IoT [64] and a data protection aggregation scheme, combining distributed random forest, Ada-Boost, and K-means clustering techniques for data protection. A lightweight blockchain-based IoT information sharing security framework combines transaction and data Blockchains for distributed storage and tamper-proofing, improving data registration efficiency, transaction efficiency, and privacy protection [22] through partial blind signature algorithms. A node cooperation dynamic game strategy is suggested to prevent malevolent local dominance behaviour. The framework's defence against assault, double-strand processing capability, and delay in verification are tested, proving safe, effective, and feasible.

The study proposes using an energy-restricted Unmanned Aerial Vehicle (UAV) [22] to collect sensing data from IoT devices in a wireless sensor network. The data is stored in the IoT devices, which may or may not communicate. Two novel data collection problems were formulated: a closed tour for the UAV, which maximizes the total volume of facts collected, and a unique data collecting system for collecting sensory data from multiple IoT devices simultaneously within the UAV's hovering range. The

study also presents NP-hard problems and heuristic and approximation techniques for data collection maximization.

The Internet of Things (IoT) offers numerous benefits [25], including data transmission, analysis, and real-time dashboards. Blockchain technology has potential in IoT [29], but faces design challenges such as regulation, energy consumption, standardization, scalability, and interoperability. This study examines the integration of Blockchain technology into 5G-IoT devices and identifies issues like regulation, energy consumption, standardization, scalability, and interoperability. Combining Blockchain technology with artificial intelligence [29] can resolve these problems [31]. A marketplace architecture and methodology are provided to facilitate selling streaming information, with Ethereum Smart Contracts used to evaluate trade-offs. IoT applications raise the standard of living but still face security and privacy concerns. Decentralized trust can be handled using Blockchain technology [32, 33].

3.1.1 Data Migration

Data migration involves the transfer of digital information to new locations, file formats, environments, storage systems, databases, datacenters, or applications. With the increasing amount of data produced by IoT sensors, effective data handling and processing platforms [25] are needed for modern and next-generation computer systems. These platforms must securely and seamlessly migrate large amounts of data between locations and offer preprocessing and analysis capabilities.

SSD technology is being integrated into multi-tiered storage systems to improve performance by enhancing random access memory input/output. Successful integration relies on automatic data transfer between SSD and HDD, considering unique deadlines, IO profiles, and workload characteristics for efficient migration. The challenge is to fully utilize data migration while meeting migration deadlines to optimize performance in SSD-enabled systems.

The increasing number of IoT sensors and devices is generating a significant amount of data, necessitating effective data handling and processing platforms. These platforms must be able to migrate large amounts of data securely and efficiently between locations and offer preprocessing and analysis capabilities. Several data pipeline designs have been proposed, but none have built-in features for data veracity, security, accuracy, and trustworthiness. An innovative data pipeline architecture called Coherent Coordination of Data Migration and Processing (CCoDaMiC) [25] was introduced to centralize data migration and processing, ensuring accuracy and security. However, the current architecture relies heavily on external cloud provider performance, potentially leading to significant processing lag.

A software framework has been developed to facilitate the dynamic distribution of the web of things [27] and live migration between computing nodes in an edge-cloud continuum. A monitoring system [27] is needed to monitor and rearrange workloads in accordance with real-time data streams and Quality of Service (QoS). Conventional cloud computing is widely used for storage, analysis, and processing IoT data, but its ability to manage high latency problems [27] in time-sensitive applications is limited. High computational, network, and communication latencies are essential for real-time applications in IoT. To address this issue, an end-user, cloud, and IoT latency-aware

model is needed. HDFS [28] was established to improve cloud computation and data access storage techniques, using hash values to optimize data access evidence storage settings. The enhanced fault tolerance rate, data processing efficiency, and file upload and download efficiency demonstrate the effectiveness of the proposed method for optimizing data [28] access and storage in cloud computing.

3.1.2 Analytics

IoT analytics is a method of analyzing data from various IoT devices to extract meaningful insights. It can be categorized into past, real-time, and predictive analytics. Context is used to extract meaningful information from IoT data, enabling lasting decisions. The Blockchain Protocol-based database [34], however, does not meet certain consistency models in conventional IoT environments. Simulation was used to investigate network quality and system dynamics' effects on consistency. A method to enable read/write latency and eventual consistency is being developed, but this model is still in the process.

The survey discusses the use of Blockchain technologies in IoT applications, focusing on their potential improvements and changes. It emphasizes the importance of streamlined transactions between nodes and users, as well as the need for privacy and security measures [35]. The study also highlights the need for secure system design that considers the unique features of IoT devices. The energy efficiency challenge of IoT analytics [36] infrastructures is highlighted, with the potential to reduce cloud computing resources by using existing computer resources at the network's edge. The study also discusses machine learning approaches for IoT-generated data analysis and highlights the current obstacles to improving these solutions. A case study is provided to demonstrate how the framework can be applied to real-world research and provide a paradigm for adaptive learning in IoT apps. The survey also discusses cybersecurity issues and the need for distributed systems [37] to distribute resources efficiently and enable scalability of linked IoT devices. The future of IoT-enabled human lives is achieved when edge devices converge, enabling concurrent improvements in machine intelligence technology in limited resource IoT systems.

3.2 Cloud Based Threats

Blockchain-based cloud applications provide with a lot of benefits, like improved security [49, 50], transparency, and decentralization [49, 50]. But there are also a number of risks and weaknesses that they face. Because cloud settings store, analyze, and transfer data over multiple servers and networks [50] that are managed by third-party providers, there are hazards to data privacy in cloud computing [51].

Although blockchain technology adds unique hazards that need to be carefully managed and mitigated, it also improves the security of cloud services [51]. Bugs and exploits are examples of smart contract vulnerabilities that might result in loss of currency, data breaches [50], or unlawful transactions. Once a flaw is deployed, it cannot be undone or changed unless the contract is redeployed. These bugs are permanent. Assaults using consensus mechanisms, such Sybil and 51% assaults, can be used to control the blockchain by preventing or double-spending transactions. Phishing and social engineering, supply chain attacks, malicious updates, scalability problems, quantum computing threats,

side-channel attacks, interoperability problems, denial of service (DoS), private key compromise, regulatory and compliance risks, and public ledger exposure are some of the risks associated with data privacy [50].

Smart contract audits, multi-signature wallets, data encryption, frequent upgrades, and employee training are examples of risk mitigation techniques. These safeguards assist in preventing insider threats, data availability and downtime, shared environment risks, hypervisor attacks, inadequate encryption, and API vulnerabilities, as well as unauthorized access, accidental exposure, data loss, corruption, or deletion, inadequate backups, inadequate data control, third-party management, lack of transparency, compliance and regulatory violations, and cross-border data transfers. Two main categories are mentioned below.

Blockchain technology adds unique dangers that need to be carefully managed and mitigated even while it improves the security of cloud services [49, 50]. Regular smart contract audits, the usage of multi-signature wallets, data encryption, frequent upgrades, and staff training are examples of mitigation techniques. Organizations may guarantee the security and integrity of their cloud infrastructure by mitigating these threats.

3.2.1 Intrusion Detection

Enhancing security and privacy requires addressing cloud security challenges. Although security measures have been implemented, they can yet be improved. Multiple IDS units are used in hierarchical intrusion detection techniques, which makes them vulnerable to tampering and necessitates additional security measures. Certain strategies are not applicable in distributed systems and are specialized to particular cloud environments [51]. Although they increase communication overhead, cooperative network intrusion detection systems still need to be validated against the most recent assaults.

Cloud computing also faces the difficulty of managing huge data [49, 51]. There is a need for ways to reduce data volume for intrusion detection since rapid data output can negatively impact system performance. Feature selection techniques are one such strategy that can increase system accuracy and lower false alarm rates. To decrease feature dimensionality, feature selection strategies have only been discussed in a small number of publications [51]. Intrusion detection, blockchain applications, and cloud systems are discussed in relation to cyberattacks. It covers blockchain-based trusting of intrusion detection, virtualization and containerization technologies, and cooperative anomaly detection systems. For cloud operations to be mitigated and restored, early identification is essential. The paper gives a summary of cloud architecture and uses deployment models to classify security events. Network Intrusion Detection Systems (NIDS) [51] in the cloud and cooperative NIDS for blockchain applications hosted on cloud infrastructure are also covered.

3.2.2 Data Breaches

Examining cloud computing security, dangers, and problems published between 2010 and 2020 is the goal of this systematic literature review (SLR). Data tampering, leaking, infiltration, and storage were the seven main security risks to cloud computing services that the assessment found [50]. There is still difficulty with data outsourcing

for cloud users and (Cloud Service Providers) CSPs. Blockchain was also mentioned by the SLR as a collaborating technology to allay security worries. The results point to further research into enhancing data availability, secrecy, and integrity as well as investigating blockchain as a collaborating technology to address these security issues. Integration issues include scalability, interoperability, data privacy [50], energy efficiency, and regulatory compliance arise when integrating blockchain with large-scale cloud systems. Concerns are raised by its high energy consumption, intricate platform and cloud provider connection, and intrinsic constraints in transaction processing speed. Another difficulty is ensuring data privacy [49], especially in public blockchains. Integration cannot be successful unless regulatory and compliance challenges are resolved [52]. In cloud computing environments, where data owners cannot control basic features like storage and access, data integrity is critical. The practical viability of blockchain technology is impeded by its constraints, such as limited throughput, high latency, and weak stability, which present potential answers to threats to data integrity [53]. In order to better understand the architecture of a safe by-design cloud federation platform for the public sector; a case study from the European SUNFISH project was examined. It outlines the requirements for data integrity in cloud computing systems and discusses unanswered research concerns and difficulties in finding answers. Additionally, a draft design of a blockchain-based database that works well in cloud computing settings was presented [53].

3.3 Metaverse Related Threats

The metaverse, which consists of all virtual worlds combined with augmented reality and the Internet, is a cooperative virtually shared space [56] that is produced by the convergence of virtually improved physical reality [56] and physically persistent virtual reality. It includes an extensive variety of virtual worlds [56] in which users can communicate with virtual items and one another in real time. The metaverse, which is thought to represent the next generation of the Internet, has garnered a lot of interest from academics and business. The metaverse can be conceptualized as a three-dimensional, fully immersive virtual environment [56] that users access and interact with via digital avatars using Augmented/Virtual Reality (AR/VR) [55, 56] devices. Although several Massively Multiplayer Online (MMO) games have early forms of the metaverse, the full-flesh metaverse is anticipated to be more sophisticated and made possible by a variety of cutting-edge technology [55]. One of the key technologies that could transform the metaverse into a democratic, decentralized virtual society with a separate economy and government is blockchain.

Key features of the metaverse are:

- Persistent Virtual Worlds: These are continuous, shared worlds that exist even when users are not actively involved.
- User Interaction: Users can interact with the environment and one other using avatars, which are digital representations of themselves.
- Interoperability: Different virtual worlds and platforms can interface with each other, allowing users to travel smoothly between them.
- Augmented Reality (AR): Enhancements to the real environment using digital information and media [55].

- Virtual Reality (VR): Fully immersive digital settings that replace the physical world [55, 56, 56].
- Social Interaction: Emphasis on social connection [55] and community building inside these virtual environments.
- Economies: Virtual economies where users can trade, buy, and sell digital assets and services.
- Content Creation: Users can produce and alter content within the metaverse, often contributing to its continuing development and evolution.

To maintain this virtual world's decentralization, security, and privacy, blockchain integration [55] with the metaverse is essential. Nevertheless, several problems with blockchain technology have also been found, suggesting that more work needs to be done before it can be used completely in the metaverse.

3.3.1 Virtual Property

Virtual property encompasses intangible assets [55] and resources found in digital environments, which includes social media platforms, online games, virtual worlds, and other digital ecosystems. These properties can encompass a variety of elements:

Digital Real Estate refers to the ownership of virtual land or spaces within online environments such as virtual worlds (e.g., Second Life) or multiplayer online games (e.g., Fortnite, Minecraft).
Virtual Items are in-game assets that players can acquire, trade, or sell. These assets include vehicles, weapons, or clothing.
Domain Names: Internet addresses that can be purchased, exchanged, and transferred.
Digital Goods: Items that are intangible, such as music files, e-books, and software licenses.
Exploring the world of cryptocurrency and NFTs: Delve into the realm of digital currencies such as Bitcoin and Ethereum, as well as the fascinating concept of non-fungible tokens (NFTs) that symbolize ownership of one-of-a-kind digital items or art.

Social media accounts encompass various digital assets such as profiles and followers, all of which contribute to one's online presence.

These properties have significant value as they can be bought, sold, and traded within their respective platforms and even in real-world markets. Virtual property ownership and transfer are typically regulated by platform terms of service and, in certain instances, by developing legal frameworks.

A crucial component of the metaverse's long-term growth is creation [56]. Furthermore, the metaverse's high degree of freedom and open environment strongly promote the growth of the creative activity [56] motivated by personal psychological demands or financial gain, which implies that a significant quantity of commodities created through creation [55] will manifest. Whatever the reason, the owner of the goods will not tolerate misuse or unauthorized duplication. This makes it vital to take precautions to protect it. As a result, certain features are realized, such as content protection, authentication, and tamper-resistance [55, 56], which discourages malevolent avatars from pilfering and unauthorizedly duplicating products. There have been numerous highly valuable virtual contents that have been issued and traded in the market. A Non-Fungible Token artwork

called The Merge, created by the digital artist [55], has recently been sold for a staggering a very big amount on Nifty Gateway. This Non-Fungible Token holds the record for being the most expensive ever sold, as a staggering number of collectors came together to acquire number of parts of it. Nevertheless, the evaluation of virtual assets remains unclear. Numerous Non-Fungible Token scams have emerged, artificially inflating the prices of utterly worthless Non-Fungible Tokens to astronomical levels. Specifically, companies involved in Non-Fungible Token projects often align themselves with established successes or artificially inflate demand by purchasing their own Non-Fungible Tokens at high prices using multiple accounts they create [55]. When considering the metaverse, certain conventional valuation methods like the Discounted Cash Flow model and the Market approach in the financial sector can be utilized. However, it is important to acknowledge that these methods do have their limitations. Although the Network Value to Transactions Ratio in the Market approach cannot anticipate a bubble in advance, the Discounted Cash Flow model is susceptible to errors due to the numerous assumptions required for cash flow projection. The effectiveness and precision of these methods have not been definitively proven, so the valuation of virtual assets/property in the metaverse remains an unresolved matter. Additional research is necessary to explore more appropriate solutions.

The recent emergence and advancement [55, 56] of diverse technologies has made it feasible to actualize the metaverse, which has garnered significant attention once more. It might have numerous positive effects on human society, including a decrease in discrimination, the eradication of individual distinctions, and socialization. But security and privacy are issues with everything, and the metaverse is no different. In contrast to existing virtual reality (VR) [55, 56] technologies, firstly the idea of the metaverse is examined and suggested a superior VR ecosystem [56]. Subsequently, we thoroughly examine and expound upon plausible security and privacy issues from four angles which involves goods, scenario, communication, and user information. It is suggested that to adequately address security and privacy concerns from a philosophical standpoint, it is necessary to utilize the new buckets effect.

Blockchain technology offers a viable method for managing the intricate provenance and ownership tracing of digital assets in the metaverse. This is achieved by storing proof of content/asset originality and involved operations on public ledgers. The majority of entities in the metaverse retain the historical records of activities on blockchain ledgers, guaranteeing their democratic, immutable, transparent, auditable, and non-repudiable nature. Furthermore, by encoding the ownership management logic into scripts that are executed on top of the blockchain, smart contracts provide an intelligent tracing solution. Blockchain technology has already been applied to the sharing of cloud storage, food, charging pile sharing, and ride sharing etc. Different ownership forms, such as collective and shared ownership [56], can exist in the metaverse in addition to private ownership, which presents additional difficulties in the ownership management of virtual objects and metaverse assets [55, 56]. There has been a growing interest in using Non-Fungible Tokens [55, 56] for ownership provenance [55, 56] and asset identification in current metaverse efforts. However, Non-Fungible Tokens are also susceptible to ransomware, phishing, inflation attacks, and cross-chain fraud [55, 56]. One instance is that malicious actors might mint the identical NFT on several blockchains at the same time.

3.3.2 Virtual Harassment

Virtual harassment is the term for the use of digital communication means to harass, threaten, or intimidate a person or group. It is also referred to as online harassment or cyber harassment. Numerous online venues, such as social media, email, chat apps, forums, and online games, may exhibit this behaviour. Cyberbullying can manifest in a variety of ways, including:

Cyberbullying is defined as persistently hostile activity aimed at an individual, usually consisting of insults, threats, or the dissemination of misleading information.

Doxxing: Trolling is the uninvited publication of personal information, often with malevolent intent. It involves posting controversial content to agitate others. Stalking is the persistent observation of someone's internet activity. Sexual harassment involves [57] inappropriate remarks or content. Hate speech is online speech that disparages or encourages violence against a person or group based on their characteristics. Impersonation involves taking on someone else's identity to mislead or hurt them [14, 57]. Spamming is the practice of sending repetitive, disruptive messages or content. Victims of virtual harassment can experience psychological, emotional, and sometimes physical effects. Digital platforms have guidelines and resources to report [57] and counteract harassment, and some regions have implemented legal actions to address and sanction such conduct.

When harassing others, attackers may pilfer or fake identities [54]. This can be especially harmful in the metaverse since it impacts not just the person but also their virtual possessions [54, 57] and reputation. It can be difficult to manage and moderate user-generated content in the metaverse, which raises concerns about negative behaviours like hate speech and disinformation [54]. Compromised or fake identities [54] can be used to launch various attacks [57], including Sybil attacks, where a single adversary controls multiple fake identities to manipulate the network. Sybil Attacks [57] can jeopardize the integrity of the wireless network application when a hostile actor assumes several false identities to exert disproportionate influence over the network. Virtual stalking [57] is the practice of repeatedly pursuing or reaching out to someone in the metaverse, creating an ongoing sense of harassment or surveillance. Stalking can be narrowed by giving users privacy settings [57] to regulate their interactions and visibility as well as by putting in place systems to report and ban harassers. Inadequate privacy protections [57] may leave users vulnerable to harassment and unsolicited encounters. Harassment [54, 57] can be avoided by providing users with detailed privacy options [57] that let them decide who can communicate with them, see their profile, or access their online areas. The problem may get worse if people are unaware of virtual harassment [57] and how to deal with it. Regular awareness campaigns [57] and the provision of tools for identifying and handling harassment can enable users to defend others as well as themselves. Harassment in real time [57] can have both short-term and long-term psychological effects. Putting in place real-time monitoring systems/tools [57] that employ AI to identify patterns of harassment and notify moderators can facilitate prompt user protection actions.

Blockchain-based metaverse applications need a multipronged strategy to address virtual harassment [54, 57], cyberbullying, and online harassment. This involves strong

governance frameworks, user empowerment, and technology solutions. Proactive measures [57] and blockchain technology's unique features, such as better security and decentralization, can create secure virtual environments. Cyberbullying and online harassment [57, 58] are potential issues due to anonymity and the risk of malware and ransomware attacks. Metaverse developers should implement robust security measures [58] like encryption and secure authentication, educate users on data protection, and establish policies to address these concerns. Future crime patterns and legal challenges [58] in the metaverse suggest that law enforcement must be alert to new crime patterns and update systems to ensure safe usage [58]. Strong security mechanisms, well-defined legal frameworks, and continuous research and development are essential components of a multi-faceted strategy [54].

3.4 Other Applications Related Threats

Blockchain technology, while providing a tamper-proof ledger of transactions, is not immune to cyberattacks and fraud. Hackers can manipulate vulnerabilities in blockchain infrastructure and have successfully exploited them over the years. Examples include code exploitation by the Decentralized Autonomous Organization (DAO). Blockchains are threatened through four primary methods: phishing, routing, Sybil, and 51% attacks [50, 57]. Phishing involves scamming users to obtain their credentials through fake emails. Routing attacks intercept data during real-time data transfers, allowing fraudsters to extract confidential data or currencies. Sybil attacks flood the network with false network identities, causing system crashes. 51% attacks involve miners gaining over 50% of a blockchain network's mining power, allowing them to control the ledger and manipulate it. Private blockchains are not vulnerable to 51% attacks. To ensure the security of your blockchain design and environment, consider using X-Force Red blockchain testing services.

3.4.1 Cryptocurrency

A digital currency, or cryptocurrency, is an alternative payment method made possible by encryption techniques. Because they employ encryption technology, cryptocurrencies can be used as a virtual accounting system in addition to a medium of exchange.

Cryptocurrency's popularity has declined since its peak, but the risks associated with it persist. Cybercriminals are increasingly targeting cryptocurrencies, with attacks on Bitcoins and other cryptocurrencies increasing rapidly. Crypto assaults [63] involve attempting to exploit blockchain networks, wallets, or transactions involving crypto assets. Crypto attacks aim to exploit victims, while retail investors and major players in the cryptocurrency space, such as exchanges and blockchain networks, can fall victim to cyberattacks. Users of cryp-tocurrencies are vigilant for emerging attack types. Blockchain technology, known for its confidentiality, security, and dependability, is seen as a revolutionary new data storage, transmission, and management system. It began with cryptocurrencies and expanded into the assets and credit sector before finding use in the information and communication domain. Blockchain allows users to collaborate in data computing, storage, authenticity verification, and database maintenance. The rapid development [60] of blockchain technology is gaining recognition in various sectors.

To deal with these attacks different consensus algorithms are surveyed and tested, but still the efficiency, safety and convenience, scalability, privacy leakage, selfish mining issues persists [61].

Blockchain technology requires the development of standards, clear definitions, and descriptions for its language. The public blockchain network lacks assurances on miners' identities, which could lead to potential dangers and attacks, including DoS [59, 62]. To address this, the authors propose improving consensus algorithms like PoW, PoS, DPoS, and PoA by introducing a simple identification procedure that distinguishes machined nodes from regular nodes. This will reduce DoS/DDoS [62] attacks. The focus is on consensus algorithms for managing smart contracts and cryptocurrency transactions, with research needed to determine their suitability for small devices and real-time transactions. Some proposed algorithms have a power of tolerance of 50% against adversaries, but in low-power cases, algorithms need strengthening to withstand future attacks [62]. The main causes of current blockchain systems' functionality and security issues are examined, along with new vulnerabilities [1] discovered in Ethereum and Bitcoin, such as vector 76, finney, and double spending attacks. Despite these advancements, blockchain security remains limited for dealing with bitcoin and ethernet networks [63].

3.4.2 Database Applications

Cybercriminals constantly develop new methods to breach web applications' security and access private information. Security professionals must improve vulnerabilities and strengthen their systems [38] to protect against these attacks. Smart cities, supply chain traceability, and government services using blockchain technology still face open issues. A survey on blockchain's applications in smart city [39] technology covers topics like smart citizen, supply chain management, smart healthcare, smart transportation, and smart grid.

There are several Blockchain enabled applications which are still vulnerable to attacks, some of them are discussed below.

Cryptocurrency popularity has decreased since its peak, but the risks remain. Cybercriminals are increasingly targeting cryptocurrencies, with attacks increasing rapidly. Crypto assaults [42] involve attempting to take advantage of blockchain networks, wallets, or transactions involving crypto assets. Healthcare sectors face cybersecurity issues, including data theft, ransomware attacks, and distributed denial of service (DDoS) attacks [42]. The most serious problem is the hacking of healthcare applications and APIs, which prioritize patient privacy, vital life-saving infrastructure, and data accuracy. These threats pose significant threats to the healthcare sector.

The real estate [56] industry faces challenges such as lack of transparency, inefficiencies, fraud, corruption, exorbitant expenses, and trust concerns. A conceptual framework is presented for smart applications for buyers/lenders to deal with real estates [56], highlighting six levels of Blockchain technology. Intellectual property rights grant the ability to restrict the use of an intellectual invention or forbid others from using it. With the advancement of digital technology and the Internet, multimedia [46, 47] and digital works have become more important, necessitating immediate attention to protect copyrights and information integrity. Digital watermarking technologies can help achieve digital copyright protection.

Online news sources (OSNs) [16] are a major basis for local and worldwide news, but they also have drawbacks, such as the propagation of fraudulent information (fake news). Identifying fake news on OSNs is a complex and unsolved problem due to the large volume of content generated and spread [4], as well as the complexity of online information. Digital watermarking [16] technologies can help achieve digital copyright protection. Despite the benefits of OSNs, identifying fake news [18] remains a complex and unsolved issue due to the complexity of online information [17].

Table 1. Appendix A: Consolidation of Blockchain Based Security Threats

Sr. No.	Threat Name	Classifier	Significance	Blockchain based solutions
1	Data Migration	IoT Based Threats	There are significant risks in data migration like data loss, semantic errors, data corruption, data security etc	There is a necessity to develop a vigorous data migration strategy
2	Analytics	IoT Based Threats	Certain challenges are there like data quality, data volume and velocity, scalability etc. needs to resolved	IoT data analytics revolutionizes business operations across industries, despite challenges and high implementation costs, making it economically viable to invest in
3	Intrusion Detection	Threats to Cloud Based Applications	Cloud IDS helps organizations identify cyber threats, providing security staff with critical alerts. Cloud intrusion prevention systems (IPS) thwart detected threats before they affect the cloud environment and corporate data storage	There are blockchain based solutions are in implementation phase to provide a log of suspicious activities, which will ensure data integrity and security
4	Data Breaches	Threats to Cloud Based Applications	Hackers continuously exploit network defences to gain unauthorized access to sensitive data, leading to severe consequences like financial loss, reputational damage, and legal liabilities	Blockchain technology enables organizations to maintain data integrity by creating cryptographic hashes and storing them securely

(*continued*)

Table 1. (*continued*)

Sr. No.	Threat Name	Classifier	Significance	Blockchain based solutions
5	Virtual Property	Threats to Metaverse Based Applications	Virtual assets, such as cryptocurrencies, NFTs, gaming tokens, and governance tokens, enable trading, transfer, and payment transactions	As payment transactions are involved security is crucially required. NFTs, or non-fungible tokens, are digital assets that individuals can own by using a unique blockchain code to represent the deed to a property in the metaverse
6	Virtual Harassment	Threats to Metaverse Based Applications	Virtual avatars and online personas offer anonymity, increasing the likelihood of improper conduct and shielding against immediate real-world consequences	Blockchain enables the metaverse to function as a public platform with a decentralized open-source ecosystem, facilitating digital commerce, application design, and user-friendly interfaces for decentralized wallets and exchanges
8	Cryptocurrency	Threats to other Applications	The system allows accountholders to securely manage their cryptocurrency movements by connecting their cryptoasset data with their financial institution and third parties	Blockchain technology, a digital ledger, records and confirms cryptocurrency trades, storing information about buying, selling, or exchanging digital assets without a central authority overseeing the market
9	Database Applications	Threats to other Applications	Database applications facilitate data consumption for end users and higher-level applications, enabling storage, retrieval, transaction processing, and machine learning calculations through UIs or APIs	A blockchain is a decentralized, distributed ledger replicated and synchronized across network computers, ensuring data security and parallel updates and validation by all participants

A consolidated view on the threats affecting the different applications under the categories like IoT, Cloud, Metaverse and other discussed so far has been tabulated under Appendix A.

4 Discussion

The Internet of Things (IoT) is revolutionizing the digital landscape by integrating physical objects for data exchange and management. Blockchain technology, particularly Ethereum Smart Contracts, can help address security, transparency, and decentralization in IoT environments. However, it also introduces unique risks and weaknesses, such as data privacy hazards. To address these issues, new technologies or a combination of existing ones are needed. Blockchain technology can be combined with artificial intelligence to improve data handling and processing platforms, enabling real-time data analysis and dashboards. The metaverse, a virtual space combining augmented reality and the Internet, is also transforming into a democratic, decentralized virtual society. However, challenges persist in efficiency, safety, convenience, scalability, privacy leakage, and selfish mining.

5 Conclusion & Future Work

The survey discussed in this paper is based on various security risks revolving around Blockchain-based applications belonging to fields, such as IOT, cyber security, intellectual property rights and so on. We have attempted to provide a classified view on the threats affecting Blockchain based applications with the support of state-of-the art literature. The consolidated view provided in the Appendix would also serve as a quick reference for the application developers. Right now, we are in the process of identifying more threats affecting various other Blockchain based applications. As a future work it has been planned to develop a mitigation model to deal with the application or data related threats affecting IoT applications (Table 1).

References

1. Barua, S., Talukder, D.: A Blockchain based decentralized video streaming platform with content protection system. In: 23rd International Conference on Computer and Information Technology (ICCIT), DHAKA, Bangladesh, pp. 1–6 (2020). https://doi.org/10.1109/ICCIT51783.2020.9392746
2. Sathish, S.K., Patankar, A.A., Khanna, H.: Aurum: a blockchain based decentralized video streaming platform. In: 2019 IEEE Wireless Communications and Networking Conference (WCNC), Marrakesh, Morocco, pp. 1–8 (2019).https://doi.org/10.1109/WCNC.2019.8886050
3. Zhao, H., Liu, Y., Wang, Y., Huang, Y.: Hiding data into blockchain-based digital video for security protection. In: 2020 3rd International Conference on Smart BlockChain (SmartBlock), Zhengzhou, China, pp. 23–28 (2020). https://doi.org/10.1109/SmartBlock52591.2020.00012

4. Evsutin, O., Meshcheryakov, Y.: The use of the Blockchain technology and digital watermarking to provide data authenticity on a mining enterprise. Sensors **20**(12), 3443 (2020). https://doi.org/10.3390/s20123443
5. Gimenez-Aguilar, M., Maria, J., de Fuentes, L., Gonzalez-Manzano, D.A.: Achieving cybersecurity in Blockchain-based systems: A survey. Future Gener. Comput. Syst. **124**, 91–118 (2021). https://doi.org/10.1016/j.future.2021.05.007
6. Jagpreet Kaur, K., Ramkumar, R.: The recent trends in cyber security: a review. J. King Saud Univ. Comput. Inf. Sci. **34**(8), 5766–5781 (2022). https://doi.org/10.1016/j.jksuci.2021.01.018
7. Wylde, V., et al.: Cybersecurity, data privacy and blockchain: a review. SN Comput. Sci. **3**(2), 1–12 (2022). https://doi.org/10.1007/s42979-022-01020-4
8. Jain, A.K., Sahoo, S.R., Kaubiyal, J.: Online social networks security and privacy: comprehensive review and analysis. Complex Intell. Syst. **7**, 2157–2177 (2021). https://doi.org/10.1007/s40747-021-00409-7
9. Butt, U.A., Amin, R., Mehmood, M., Aldabbas, H., Alharbi, M.T., Albaqami, N.: Cloud security threats and solutions: a survey. Wirel. Pers. Commun. **128**(1), 387–413 (2023). https://doi.org/10.1007/s11277-022-09960-z
10. Liu, L., Zhang, W., Han, C.: A survey for the application of blockchain technology in the media. Peer-to-Peer Netw. App. **14**(5), 3143–3165 (2021). https://doi.org/10.1007/s12083-021-01168-5
11. Kuss, D.J., Kristensen, A.M., Lopez-Fernandez, O.: Internet addictions outside of Europe: a systematic literature review. Comput. Human Behav. **115**, 106621 (2021). https://doi.org/10.1016/j.chb.2020.106621
12. Naslund, J.A., Bondre, A., Torous, J., Aschbrenner, K.A.: Social media and mental health: benefits, risks, and opportunities for research and practice. J. Technol. Behav. Sci. **5**(3), 245–257 (2020). https://doi.org/10.1007/s41347-020-00134-x
13. Poongodi, M., Nguyen, T.N., Hamdi, M., Cengiz, K.: Global cryptocurrency trend prediction using social media. Inf. Process. Manage. **58**(6), 102708 (2021). https://doi.org/10.1016/j.ipm.2021.102708
14. Wang, C., Zhu, H., Yang, B.: Composite behavioral modeling for identity theft detection in online social networks. IEEE Trans. Comput. Soc. Syst. **9**(2), 428–439 (2022). https://doi.org/10.1109/TCSS.2021.3092007
15. Senthil Raja, M., Arun Raj, L.: Detection of malicious profiles and protecting users in online social networks. Wirel. Pers. Commun. **127**, 107–124 (2021). https://doi.org/10.1007/s11277-021-08095-x
16. Bhattacharya, A., Bathla, R., Rana, A., Arora, G.: Application of machine learning techniques in detecting fake profiles on social media. In: 2021 9th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), Noida, India, pp. 1–8 (2021). https://doi.org/10.1109/ICRITO51393.2021.9596373
17. Jing, T.W., Murugesan, R.K.: Protecting data privacy and prevent fake news and deepfakes in social media via blockchain technology. In: Anbar, M., Abdullah, N., Manickam, S. (eds.) Advances in Cyber Security: Second International Conference, ACeS 2020, Penang, Malaysia, December 8-9, 2020, Revised Selected Papers, pp. 674–684. Springer Singapore, Singapore (2021). https://doi.org/10.1007/978-981-33-6835-4_44
18. Shammar, E.A., Zahary, A.T., Al-Shargabi, A.A.: A survey of IoT and blockchain integration: security perspective. IEEE Access **9**, 156114–156150 (2021). https://doi.org/10.1109/ACCESS.2021.3129697
19. Shammar, E.A., Zahary, A.T., Al-Shargabi, A.A.: A survey of IoT and blockchain integration: security perspective. IEEE Access **9**, 156114–156150 (2021). https://doi.org/10.1109/ACCESS.2021.3129697

20. Dwivedi, S.K., Amin, R., Vollala, S.: Blockchain based secured information sharing protocol in supply chain management system with key distribution mechanism. J. Inf. Secu. App. **54**, 102554 (2020). https://doi.org/10.1016/j.jisa.2020.102554
21. Ge, C., Liu, Z., Fang, L.: A blockchain based decentralized data security mechanism for the Internet of things. J. Parallel Distrib. Comput. **141**, 1–9 (2020). https://doi.org/10.1016/j.jpdc.2020.03.005
22. Li, Y., Liang, W., Xu, W., Jia, X.: Data collection of IoT devices using an energy-constrained UAV. In: IEEE International Parallel and Distributed Processing Symposium (IPDPS), New Orleans, LA, USA, pp. 644-653 (2020).https://doi.org/10.1109/IPDPS47924.2020.00072
23. Jia, B., Zhang, X., Liu, J., Zhang, Y., Huang, K., Liang, Y.: Blockchain-enabled federated learning data protection aggregation scheme with differential privacy and homomorphic encryption in IIoT. IEEE Trans. Indus. Inf. **18**(6), 4049–4058 (2022). https://doi.org/10.1109/TII.2021.3085960
24. Amanullah, M.A., et al.: Deep learning and big data technologies for IoT security. Comput. Commun. **151**, 495–517 (2020). https://doi.org/10.1016/j.comcom.2020.01.016
25. Dehury, C.K., Srirama, S.N., Chhetri, T.R.: CCoDaMiC: a framework for Coherent Coordination of Data Migration and Computation platforms. Future Gener. Comput. Syst. **109**, 1–16 (2020). https://doi.org/10.1016/j.future.2020.03.029
26. Aguzzi, C., Gigli, L., Sciullo, L., Trotta, A., Di Felice, M.: From cloud to edge: seamless software migration at the era of the web of things. IEEE Access **8**, 228118–228135 (2020). https://doi.org/10.1109/ACCESS.2020.3045632
27. Shukla, S., Hassan, M.F., Tran, D.C., Akbar, R., Paputungan, I.V., Khan, M.K.: Improving latency in Internet-of-Things and cloud computing for real-time data transmission: a systematic literature review (SLR). Cluster Comput. **26**, 1–24 (2021). https://doi.org/10.1007/s10586-021-03279-3
28. Wang, M., Zhang, Q.: Optimized data storage algorithm of IoT based on cloud computing in distributed system. Comput. Commun. **157**, 124–131 (2020). https://doi.org/10.1016/j.comcom.2020.04.023
29. Huo, R., et al.: A Comprehensive Survey on Blockchain in Industrial Internet of Things: Motivations, Research Progresses, and Future Challenges. IEEE Commun. Surv. Tutor. **24**(1), 88–122 (2022). https://doi.org/10.1109/COMST.2022.3141490
30. Haris, R.M., Al-Maadeed,, S.: Integrating blockchain technology in 5G enabled IoT: a review. In: IEEE International Conference on Informatics, IoT, and Enabling Technologies (ICIoT), Doha, Qatar, pp. 367–371 (2020). https://doi.org/10.1109/ICIoT48696.2020.9089600
31. Li, H., Pei, L., Liao, D., et al.: BDDT: use blockchain to facilitate IoT data transactions. Cluster Comput. **24**(1), 459–473 (2020). https://doi.org/10.1007/s10586-020-03119-w
32. Arshad, Q.A., Khan, W.Z., Azam, F., et al.: Blockchain-based decentralized trust management in IoT: systems, requirements and challenges. Complex Intell. Syst. **9**, 6155–6176 (2023). https://doi.org/10.1007/s40747-023-01058-8
33. Tseng, L., Yao, X., Otoum, S., et al.: Blockchain-based database in an IoT environment: challenges, opportunities, and analysis. Cluster Comput. **23**, 2151–2165 (2020). https://doi.org/10.1007/s10586-020-03138-7
34. Unal, D., Hammoudeh, M., Khan, M.A., Abuarqoub, A., Epiphaniou, G., Hamila, R.: Integration of federated machine learning and blockchain for the provision of secure big data analytics for Internet of Things. Comput. Secur. **109**, 102393 (2021) https://doi.org/10.1016/j.cose.2021.102393
35. Zschörnig, T., Wehlitz, R., Franczyk, B.: IoT analytics architectures: challenges, solution proposals and future research directions. In: Dalpiaz, F., Zdravkovic, J., Loucopoulos, P. (eds.) Research Challenges in Information Science: 14th International Conference, RCIS 2020, Limassol, Cyprus, September 23–25, 2020, Proceedings, pp. 76–92. Springer International Publishing, Cham (2020). https://doi.org/10.1007/978-3-030-50316-1_5

36. Adi, E., Anwar, A., Baig, Z., et al.: Machine learning and data analytics for the IoT. Neural Comput. App. **32**, 16205–16233 (2020). https://doi.org/10.1007/s00521-020-04874-y
37. Hakak, S., Khan, W.Z., Gilkar, G.A., Imran, M., Guizani, N.: Securing smart cities through blockchain technology: architecture, requirements, and challenges. IEEE Network **34**(1), 8–14 (2020). https://doi.org/10.1109/MNET.001.1900178
38. Shrimali, B., Patel, H.B.: Blockchain state-of-the-art: architecture, use cases, consensus, challenges and opportunities. J. King Saud Univ. Comput. Inf. Sci. **34**(9), 6793–6807 (2022). https://doi.org/10.1016/j.jksuci.2021.08.005
39. Makani, S., Pittala, R., Alsayed, E., et al.: A survey of blockchain applications in sustainable and smart cities. Cluster Comput. **25**, 3915–3936 (2022). https://doi.org/10.1007/s10586-022-03625-z
40. Ocheja, P., Agbo, F.J., Oyelere, S.S., Flanagan, B., Ogata, H.: Blockchain in education: a systematic review and practical case studies. IEEE Access **10**, 99525–99540 (2022). https://doi.org/10.1109/ACCESS.2022.3206791
41. Shah, R., Sridaran, R.: A study on security and privacy related issues in blockchain based applications. In: 6th International Conference on Computing for Sustainable Global Development (INDIACom), New Delhi, India, 2019, pp. 1240–1244 (2019)
42. Poongodi, M., Nguyen, T.N., Hamdi, M., Cengiz, K.: Global cryptocurrency trend prediction using social media. Inf. Process. Manage. **58**(6), 6793 (2021). https://doi.org/10.1016/j.ipm.2021.102708
43. Oughton, E.J., Jha, A.: Supportive 5G infrastructure policies are essential for universal 6g: assessment using an open-source techno-economic simulation model utilizing remote sensing. IEEE Access **9**, 101924–101945 (2021). https://doi.org/10.1109/ACCESS.2021.3097627
44. Gamage, H.T.M., Weerasinghe, H.D., Dias, N.G.J.: A Survey on blockchain technology concepts, applications, and issues. SN Comput. Sci. **1**(2), 1–15 (2020). https://doi.org/10.1007/s42979-020-00123-0
45. Chang, S.E., Chen, Y.: When blockchain meets supply chain: a systematic literature review on current development and potential applications. IEEE Access **8**, 62478–62494 (2020). https://doi.org/10.1109/ACCESS.2020.2983601
46. Patil, L.K., Popat, K.A.: Comparative analysis of several approaches of encoding audio files. In: Rajagopal, S., Faruki, P., Popat, K. (eds.) Advancements in Smart Computing and Information Security: First International Conference, ASCIS 2022, Rajkot, India, November 24–26, 2022, Revised Selected Papers, Part II, pp. 128–143. Springer Nature Switzerland, Cham (2022). https://doi.org/10.1007/978-3-031-23095-0_10
47. Patil, Lakhichand Khushal, Popat, Kalpesh A.: Design and Implementation of Multilayer Encryption for Audio File Security. In: Rajagopal, Sridaran, Popat, Kalpesh, Meva, Divyakant, Bajeja, Sunil (eds.) Advancements in Smart Computing and Information Security: Second International Conference, ASCIS 2023, Rajkot, India, December 7–9, 2023, Revised Selected Papers, Part III, pp. 179–191. Springer Nature Switzerland, Cham (2024). https://doi.org/10.1007/978-3-031-59100-6_14
48. Bharathi Murthy, C.V.N.U., Lawanya Shri, M., Kadry, S., Lim, S.: Blockchain based cloud computing: architecture and research challenges. IEEE Access **8**, 205190–205205 (2020). https://doi.org/10.1109/ACCESS.2020.3036812
49. Alkadi, O., Moustafa, N., Turnbull, B.: A review of intrusion detection and blockchain applications in the cloud: approaches, challenges and solutions. IEEE Access **8**, 104893–104917 (2020). https://doi.org/10.1109/ACCESS.2020.2999715
50. Alouffi, B., Hasnain, M., Alharbi, A., Alosaimi, W., Alyami, H., Ayaz, M.: A systematic literature review on cloud computing security: threats and mitigation strategies. IEEE Access **9**, 57792–57807 (2021). https://doi.org/10.1109/ACCESS.2021.3073203

51. Lata, S., Singh, D.: Intrusion detection system in cloud environment: Literature survey & future research directions. Int. J. Inf. Manage. Data Insights. **2**(2), 100134 (2022). https://doi.org/10.1016/j.jjimei.2022.100134
52. Murthy, C.V.N.U.B., Shri, M.L., Kadry, S., Lim, S.: Blockchain based cloud computing: architecture and research challenges. IEEE Access **8**, 205190–205205 (2020). https://doi.org/10.1109/ACCESS.2020.3036812
53. Wei, P.C., Yu, D.W., Zhao, S.K., Tyagi, S., Kumar, N.: Blockchain data-based cloud data integrity protection mechanism. Future Gener. Comput. Syst. **102**, 902–911 (2020). https://doi.org/10.1016/j.future.2019.09.028
54. Cheong, B.C.: Avatars in the metaverse: potential legal issues and remedies. Int. Cybersecur. Law Rev. **3**, 467–494 (2022). https://doi.org/10.1365/s43439-022-00056-9
55. Truong, V.T., Le, L., Niyato, D.: Blockchain meets metaverse and digital asset management: a comprehensive survey. IEEE Access **11**, 26258–26288 (2023). https://doi.org/10.1109/ACCESS.2023.3257029
56. Wang, Y., et al.: A survey on metaverse: fundamentals, security, and privacy. IEEE Commun. Surv. Tutor. **25**(1), 319–352 (2023). https://doi.org/10.1109/COMST.2022.3202047
57. Ali, M., Naeem, F., Kaddoum, G., Hossain, E.: Metaverse communications, networking, security, and applications: research issues, state-of-the-art, and future directions. IEEE Commun. Surv. Tutor. **26**(2), 1238–1278 (2024). https://doi.org/10.1109/COMST.2023.3347172
58. Bhardwaj, A., Kaushik, K.: Metaverse or metaworst with cybersecurity attacks. IT Professional **25**(3), 54–60 (2023). https://doi.org/10.1109/MITP.2023.3241445
59. Rao, S., Verma, A.K., Bhatia, T.: A review on social spam detection: Challenges, open issues, and future directions. Expert Syst. App. **186**, 115742 (2021). https://doi.org/10.1016/j.eswa.2021.115742
60. Dai, F., Shi, Y., Meng, N., Wei, L., Ye, Z.: From bitcoin to cybersecurity: a comparative study of blockchain application and security issues. In: 4th International Conference on Systems and Informatics (ICSAI), Hangzhou, China, pp. 975–979 (2017). https://doi.org/10.1109/ICSAI.2017.8248427
61. Li, W., He, M., Haiquan, S.: An overview of blockchain technology: applications, challenges and future trends. In: 2021 IEEE 11th International Conference on Electronics Information and Emergency Communication (ICEIEC)2021 IEEE 11th International Conference on Electronics Information and Emergency Communication (ICEIEC), Beijing, China, pp. 31–39 (2021). https://doi.org/10.1109/ICEIEC51955.2021.9463842
62. Alsunaidi, S.J., Alhaidari, F.A.: A survey of consensus algorithms for blockchain technology. In: 2019 International Conference on Computer and Information Sciences (ICCIS), Sakaka, Saudi Arabia, pp. 1–6 (2019). https://doi.org/10.1109/ICCISci.2019.8716424
63. Ghosh, A., Gupta, S., Dua, A., Kumar, N.: Security of cryptocurrencies in blockchain technology: state-of-art, challenges and future prospects. J. Netw. Comput. App. **163**, 102635 (2020). https://doi.org/10.1016/j.jnca.2020.102635
64. Lyimo, A.J., Kumar, K.A.: Efficient way in sharing of IoT data and uses of blockchain in auditing the stored data. Int. J. Internet Technol. Secur. Trans. **12**(1), 38 (2022). https://doi.org/10.1504/IJITST.2022.119669
65. Gai, K., Guo, J., Zhu, L., Yu, S.: Blockchain Meets Cloud Computing: A Survey. IEEE Commun. Surv. Tutor. **22**(3), 2009–2030 (2020). https://doi.org/10.1109/COMST.2020.2989392
66. Huang, H., Zeng, X., Zhao, L., Qiu, C., Wu, H., Fan, L.: Fusion of building information modeling and blockchain for metaverse: a survey. IEEE Open J. Comput. Soc. **3**, 195–207 (2022). https://doi.org/10.1109/OJCS.2022.3206494
67. Yang, Q., Zhao, Y., Huang, H., Xiong, Z., Kang, J., Zheng, Z.: Fusing blockchain and AI with metaverse: a survey. IEEE Open J. Comput. Soc. **3**, 122–136 (2022). https://doi.org/10.1109/OJCS.2022.3188249

68. Quamara, S., Singh, A.K.: A systematic survey on security concerns in cryptocurrencies: state-of-the-art and perspectives. Comput. Secur. **113**, 102548 (2022). https://doi.org/10.1016/j.cose.2021.102548
69. Deepa, N., et al.: A survey on blockchain for big data: Approaches, opportunities, and future directions. Future Gener. Comput. Syst. **131**, 209–226 (2022). https://doi.org/10.1016/j.future.2022.01.017
70. Awadallah, R., Samsudin, A.: Using blockchain in cloud computing to enhance relational database security. IEEE Access **9**, 137353–137366 (2021)

Comparative Analysis of Improved K-Means Clustering for Human Freedom Index

F. Mohamed Ilyas[1], S. Silvia Priscila[2(✉)], K. Sheela[3], J. Vimala Roselin[4], K. V. Sona[5], and R. Prema[6]

[1] Department of Computer Applications, Bharath Institute of Higher Education and Research, Chennai, Tamil Nadu, India

[2] Department of Computer Science, Bharath Institute of Higher Education and Research, Chennai, Tamil Nadu, India
silviaprisila.cbcs.cs@bharathuniv.ac.in

[3] Department of Computer Science, Vels Institute of Science, Technology and Advanced Studies, Chennai, Tamil Nadu, India

[4] Department of Computer Science (PG), Kristu Jayanti College Autonomous, Bangalore, India
vimala.jr@kristujayanti.com

[5] Department of Computer Science, KristuJayanti College, Bangalore, India
sona.kv@kristujayanti.com

[6] New Horizon College, Kasturinagar, Bangalore, India

Abstract. The Human Freedom Index (HFI) assesses the extent of human freedom in different nations by considering both personal and economic freedoms. It incorporates measures such as the rule of law, freedom of expression, and freedom to trade, giving a holistic picture of the freedoms people have in different countries. In order to improve the HFI analysis, this research examines three advanced clustering strategies. The first technique optimises clustering performance and minimises iterations by combining the Hamely algorithm with Mini Batch k-means (MBK-means). The second method uses chromosomes to construct cluster centroids, therefore leveraging an Improved Genetic Algorithm (GA) based K-means (IGBKM) to address concerns with initial cluster centre sensitivity and local optimisation. The third and most promising method combines Ant Colony Optimisation (ACO) and Improved K-means (ACO-IKM) to solve the shortcomings of K-means clustering by selecting cluster centroids optimally and maximising ability based on the disparity statistics. This comparative research study shows that in terms of clustering robustness and accuracy, the ACO-IKM method performs much better than the other approaches. The results of this study show that the ACO-IKM approach is a better way to analyse the HFI dataset, providing dependable and effective clustering performance using Silhouette score. This assists in identifying trends in global freedom and pinpoints areas that require improvement in order to advance human dignity and autonomy.

Keywords: Human freedom Index · Ant Colony Optimisation with Integrated K-means · Mini Batch k-means · K-Means cluster · Silhouette score · genetic algorithm · fitness function

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
S. Rajagopal et al. (Eds.): ASCIS 2024, CCIS 2428, pp. 26–39, 2025.
https://doi.org/10.1007/978-3-031-86302-8_2

1 Introduction

Human freedom is a cornerstone of human dignity and well-being, allowing individuals to act according to their own free will and pursue their personal and economic goals. The HFI serves as a comprehensive measure to evaluate the state of human freedom across various countries by incorporating both personal and economic dimensions. It includes a wide range of indicators such as trade freedom, the rule of law, and the freedom of expression. These indicators collectively offer a detailed perspective on the levels of freedom enjoyed by individuals in different nations, helping to identify regions where freedoms are upheld and areas where improvements are necessary. Analysing the HFI provides valuable insights into global freedom trends, highlighting the importance of fostering environments where human dignity and autonomy can thrive.

In the realm of data analysis, clustering techniques play a pivotal role in organizing complex datasets into meaningful groups. The aim of this research is to compare and evaluate three advanced clustering techniques to enhance the analysis of the HFI dataset. The first technique, MBK-means with the Hamely algorithm, aims to optimize clustering performance and reduce iterations, providing a more efficient clustering solution. Clustering methods are often employed in data science to find related data groups based on hidden patterns in the dataset [1]. The second technique utilizes an enhanced IGBKM, addressing common challenges include initial cluster centre sensitivity and local optimization using chromosomes to create cluster centroids.The primary objective of clustering is to arrange the data into groups of items that are both significantly similar and different from each other [2]. The third and most promising technique integrates ACO-IKM, which optimally selects cluster centroids and maximizes fitness functions through the use of gap statistics.In contrast, it is very responsive to the choice of the initial cluster centre and to the random selection of the first centroids, remaining at the local optimum [3]. The essential parameters for clustering techniques are often provided by evolutionary algorithms, which are then employed for optimisation.

This research paper concentrates on comparing the three different IKM methods and suggests the best for HFI selection of variables to generated best prediction in ranking based on the requirement for the desired user. Hence, the three different IKM are focused for better solution in identifying the best clustering fitness using HFI dataset and it can be evaluated through one of the parameters named Silhouette score.

2 Literature Review

Ramírez Gallego et al. have proposed a novel incremental and distributed categorisation system that uses the closest neighbour approach. To enable quicker searches, this technique is incorporated into Apache Spark includes distributed metric space ordering. Furthermore, an effective incremental data instance selection technique has been presented for the purpose of regularly updating massive data streams and getting rid of old instances from case libraries. This makes the original classifier more appropriate for the task by reducing its heavy computing demands. The efficacy of the suggested approach is demonstrated by experimental investigations carried out on a collection of actual mass data streams [4].

Madan and Dana., have suggested the necessity in dividing a given dataset's N items into K clusters, each of which indicates a division. Angle cosine function and Euclidean distance function are two examples of similarity measuring functions that are mostly used in partitioning techniques. In contrast to partitioning techniques, the Hierarchical Clustering approach eliminates the requirement to predetermine the number of clusters. Cure algorithm and Birch algorithm are the two primary components of standard hierarchical clustering methods [5].

Liu et al., have introduced an enhanced Kmeans clustering technique [6]. Unfortunately, none of the preceding studies focused on clustering efficiency. Harifi et al., suggested Kmeans with the Apache Spark Machine Learning Library and testing their experiments on three benchmark datasets: Internet advertisements, KDDCUP99, and Forest Cover Type. However, despite the fact that they take clustering efficiency into account, the experimental data they provide prevents us from successfully determining which K-value clustering is optimal [7]. Siddiqui et al., have recommended segmenting the provided data records by building 1000 clusters using the Kmeans clustering method over ODM. As far as researchers are aware that average clustering results does not imply that the results will get better with additional clusters [8].

Based on the study by Aydin N, Yurdakul, the optimal number of classes to categorise the nations for additional performance analysis is determined with cluster analysis by hierarchical and K-means clustering methods [9]. KUCUKEFE., have made a comparison between the GDP decreases and mortality rates in China and the OECD nations is presented in this research [10]. The present influence of account balances, fatalities, and GDP growth rate are determined by clustering analysis using K-means. The K-means clustering technique significantly affects patient care and medical-related tasks. The k-means method is used by many researchers in their investigations. Fuente-Tomas LDL et al., have classified individuals with bipolar illness using the k-means method.

Głabowski et al., have suggested an ACO technique for finding the shortest path within a network. The algorithm is composed of many sub-problems that are progressively displayed. In order to help researchers identify the best solutions for the challenges, each sub-problem is examined from a various perspective. An ideal solution may not always be found by algorithms that rely on the ant colony meta heuristic. Based on the results, it is crucial to experiment in selecting and determining parameters specific to each of the issues being studied. The usefulness and comprehensiveness of application of each technique used in this method should be examined individually. The several ways in which the procedure's separate components might be executed are directly reflected in the design of the Shortest Path ACO algorithm. In this way, the technique of approaching or arriving at optimal solutions for the shortest path issue can be improved. Determining whether to execute a particular scheme as the best option or as a substitute for more computationally expensive or time-consuming methods, consideration should be given to both the length of time required and the quality of results that are produced [11].

3 Methodology

The HFI offers a comprehensive measure of the state of human freedom across various countries by evaluating both personal and economic dimensions. This index includes indicators such as the rule of law, freedom of expression, and freedom to trade, providing a detailed overview of the levels of freedom enjoyed by individuals in different nations. By analyzing the HFI, researchers can gain valuable insights into global freedom trends, identifying regions where freedoms are upheld and areas where improvements are necessary. Advanced clustering algorithms are utilised to identify patterns and links within the HFI dataset in order to conduct an effective analysis. Based on the similarities, data can be separated into groups or clusters, using the unsupervised learning technique of clustering. This research focuses on comparing three advanced clustering algorithms are

1. MBK-means
2. IGBKM algorithm
3. ACO-IKM algorithm

The HFI dataset, which includes various indicators related to personal and economic freedoms, serves as the basis for this study. Preprocessing is done on the dataset to manage missing values, standardise the information, and maintain consistency. Feature selection techniques are employed to identify the most relevant indicators for clustering, reducing dimensionality and improving algorithm performance. The processed dataset is then split into training and testing subsets to facilitate robust evaluation.

The fundamental idea of the K-Means algorithm is simple. There is a predetermined constant k. The constant k denotes the number of final clusters. The sample points are assigned to the most similar class, the starting point is selected at random to serve as the centroid, and each class's centroid (also known as the class centre) is recalculated. The Euclidean distance is then calculated to find a level of similarity between each sample and the centroid. This procedure is then carried out again until the centroid remains unchanged, at which point each sample is identified. The each class centroid and its category are analyzed whereas the K-Means algorithm's delayed convergence on large-scale data sets is caused by the constant calculation of the similarity between each centroid as well as all samples [12, 13]. The following steps comprise the primary process of MBK-means:

- Randomly separate data sets into several batches, considering each tiny batch as the entire set.
- Decide how many initial cluster centroids to create.
- Designate the closest cluster centre for small batch data allocation.
- Continue to update the cluster centre repeatedly until it stops changing.

The Mini Batch K-means technique relies on multiple tiny sample sets rather than a single data point, in contrast to the K-means method. The small quantity of data is the operation object. The average value is utilised to calculate the update clustering centre for each small set of data, which is subsequently allocated to the new clustering centre. The cluster centre steadily becomes more stable as the number of iterations increases, eventually becoming unchanging. In order to shorten operating time and retain data

accuracy, the batch type concept was developed. This allows for faster data clustering. The MBK-means is an easy to understand and quick to implement algorithm. The speed of the conventional clustering procedure and the effects of clustering are minimal. When the data set is being processed in batches, it gets a subset of samples from different data categories for computation, avoiding the need to compute every data sample. When there is a significant reduction in the amount of data required for calculation the running time will also be shortened allowing performance to be displayed even with larger data sets. It is an improvement over the conventional clustering algorithm. Figure 1 shows the optimal cluster identification by MBK-Means cluster algorithm.

An online retail market recommender system employing GA + KMeans clustering shows that the K-Means clustering method's success is strongly reliant on the initial cluster centres' random creation. This frequently leads to clustering solutions that are trapped in a local minimum. The early process-based K-Means clustering process can be made more accurate by using genetic algorithms to determine the central point. In order to acquire the ideal starting centroid, the stages algorithm GA + KMeans performs data pretreatment by the use of a Genetic algorithm.

- In this study the chromosomes are represented through integer encoding. Genes encoded as a series of binary characters. Chromosome length is represented by the binary length.
- The length of chromosomes is significantly greater than the number of clusters and attributes.
- Initialize population using data from previous data preprocessing the centers of the cluster K are generated at random and encoded on each chromosome to create the initial population.
- Evaluation Fitness Function Calculation: At this point in the evaluation process the objective function value which is a component of the judge or measure for each chromosome of the criteria resolution is to be determined.
- Determine the separation between each original data point and the chromosomes centroid. The objective function value was computed using the following formula which uses the Euclidean distance.

When paired with GA, the K-Means method will produce superior results than each algorithm used separately. The disadvantages of using the K-Means and GA algorithms combined to get the best optimised outcome will be eliminated by this technique. The GA algorithm finds the global optimal point by using a probabilistic method to finding the ideal solution. As a result, it finds a new optimal solution formal lteration. Typically, it is advised to run the algorithm 50 times and determine its mean value before proceeding with additional processing. GA is a useful clustering technique and it struggles with huge or complicated datasets. Figure 2 illustrates the architecture of optimal cluster by Improved GA based k Means cluster.

The hybrid clustering algorithm known as ACO-IKM combines the advantages of the K-means algorithm with ACO to enhance clustering performance. This method addresses common problems like sensitivity to initial cluster centroids and local optimization by utilizing the optimization powers of ACO to improve the initialization of the K-means algorithm. ACO with K-means method is referred to as the ACOKM algorithm. The hybrid algorithm presented by the suggested technique combines the ACO method with

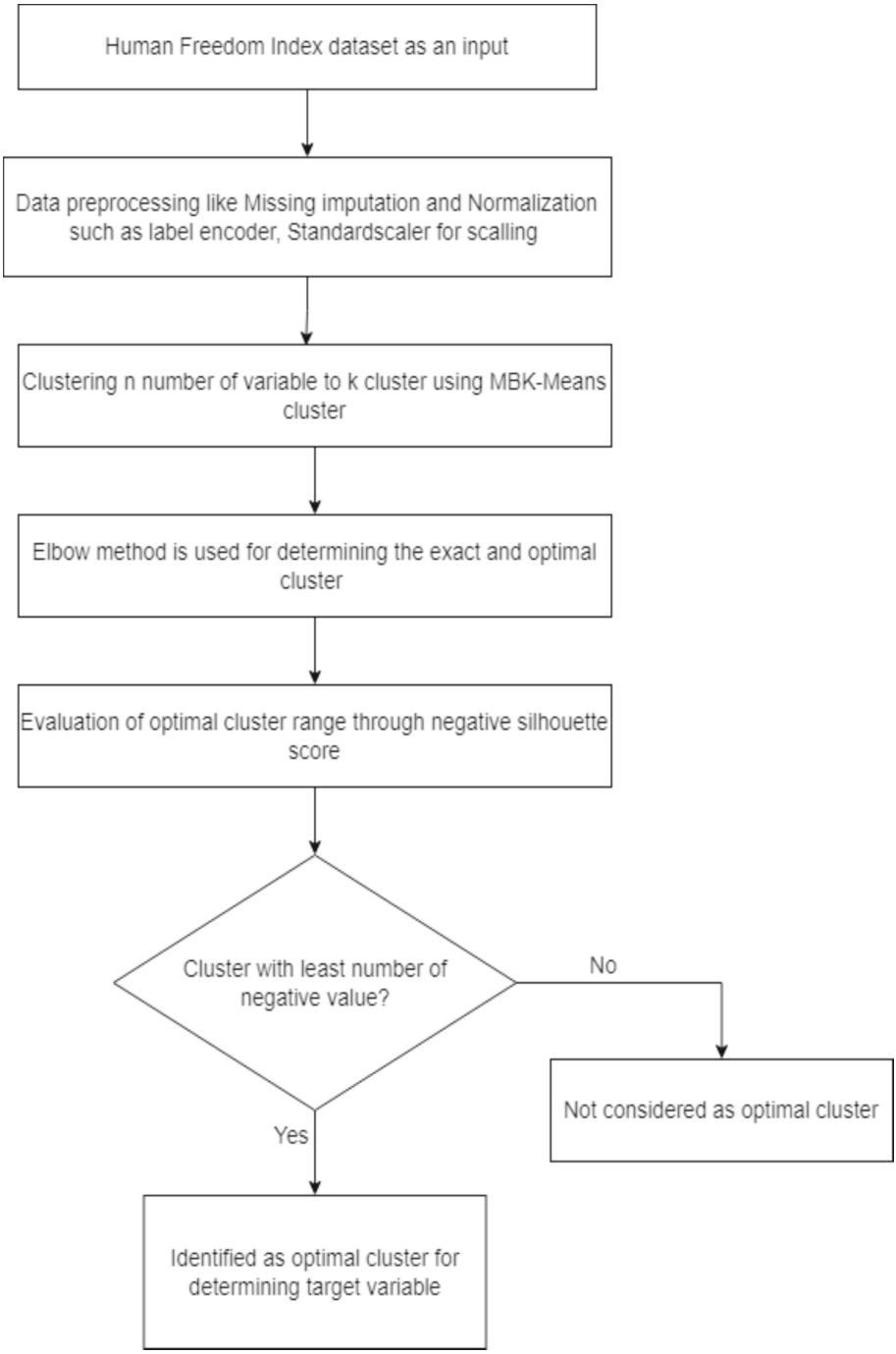


Fig. 1. Optimal cluster Identification by MBK-Means cluster

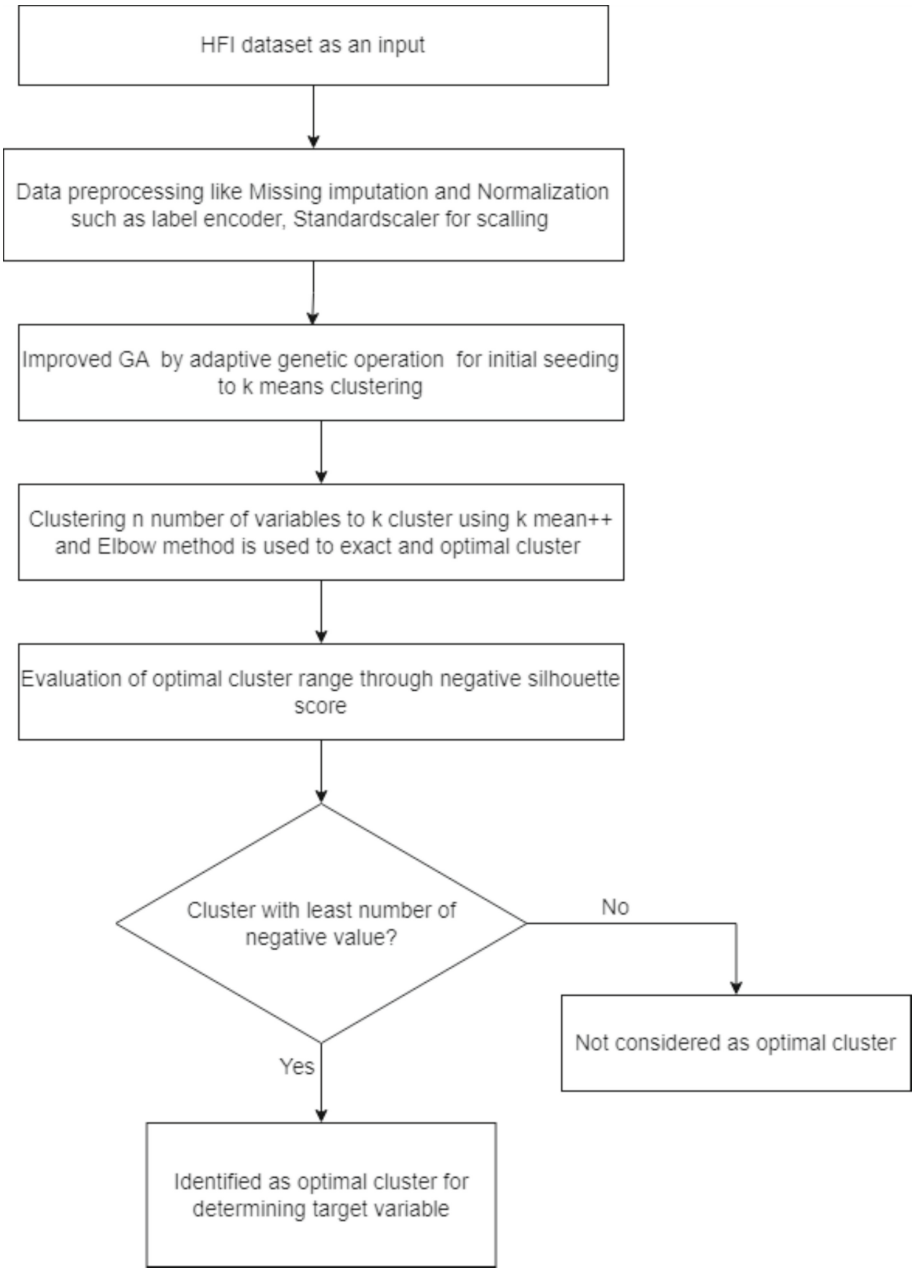


Fig. 2. Improved GA-based k-means cluster architecture.

the K-means procedure. K-means is renowned for its efficient clustering, and the ACO methodology is a well-known optimisation method. This method of preprocessing huge

data has been suggested as a way to combine the effectiveness of both techniques. The properties of the algorithms listed below are included into the projected algorithm.

3.1 Working of ACO Method

Ant's social behaviours may be utilized to tackle challenging issues and this approach known as ACO which tackles challenging combinatorial optimization issues like the workshop scheduling issue was inspired by ants. The ant colony's behavior served as the model for the ant-based clustering technique. Ants generate pheromone trails behind individuals when they go from place to place in pursuit of food. Eventually, the path with the strongest pheromone trail is selected as the primary path. The ant moving probability from node m to nis formulated in Eq. 1.

$$ACO_{mn} = \frac{\left(\tau_{mn}^{\alpha}\right)\left(\eta_{mn}^{\alpha}\right)}{\sum\left(\tau_{mn}^{\alpha}\right)\left(\eta_{mn}^{\alpha}\right)} \tag{1}$$

where,

τ_{mn}^{α} = Phenomenon amount on the edge.
η_{mn}^{α} = Edge desirability of m, n $\left(\frac{1}{d_{mn}}\right)$
d_{mn} = Edge distances.
α = Parameter for controlling the τ_{mn} influences.
β = Parameter for controlling the η_{mn} influences.

Algorithm of ACO-IKM

Input: Dataset of n features of HFI.
Output: Finding optimum cluster.
Step 1: Selecting the cluster number as k and initializing the cluster centroid as C_1, C_2, …, C_k with respect to arrangement.
Step 2: Allocating a data point to a cluster, determine its cluster centre by calculating the Euclidian distance between each data point and the cluster centroid.
Step 3: Recomposing the cluster center using the cluster data point mean and repeat until there is no new re-assignment.
Step 4: Output of the k-means cluster is utilized for calculating ACO methods. For m = 1 to I:

- Measuring the data item of ant picking with high probability in pickup
- Calculating the entropy of the high probability in pickup
- If the highest pickup probability is more than initial, the data item has shifting into another cluster

Step 5: Repeat the process until Kth cluster.

Based on the ACO-IKM algorithm, the output of the optimal clustering can be determined through gap statistics as well as silhouette average score. The gap statistic

is a DM type algorithm that focuses in improving the cluster process using an efficient estimation in obtaining best cluster number. Moreover, this technique has assisted to determine all cluster measure whereas ACO-IKM algorithm has been implemented to accomplish the clusters amount available in the dataset. It basically measures the sum of all objects distance from cluster's mean is said to be dispersion. Moreover, it has generated certain sample amount original datasets as well as provides sample datasets mean dispersion. The gap gets maximized while applying the 'k' minimum value shown in Fig. 3. The overall flowchart for ACO is shown in Fig. 4.

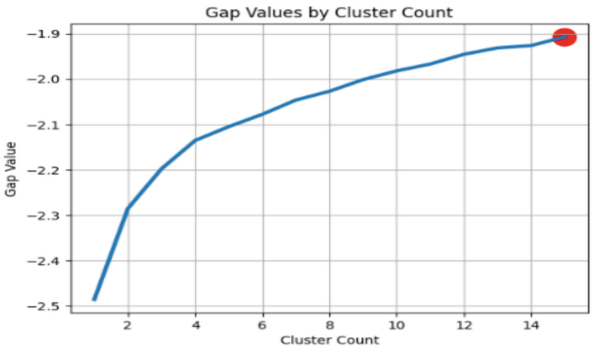


Fig. 3. Gap value by cluster count

In Silhouette method, the silhouette coefficient has been utilized to combine separation as well as cohesion. The coefficient of silhouette can be determined through division of separating measure by cohesion measure as well as subtraction of value by 1. However, the separation measure is considered to be greater than cohesion measure or subtracting the cohesion measure value by 1 that get divided by separation measure at cohesion while it is greater than separation. Hence, the highest Silhouette coefficient is considered to be better cluster for modeling.

4 Result and Discussions

The dataset used in this study is the "Human freedom Index dataset"(HFI), which is freely downloadable from Kaggle. HFI is a comprehensive metric that encompasses economic, civil, and personal freedoms. It is employed in evaluating the global status of human freedom as well as researching the connections between freedom and other social and economic elements.

The suggested ACO-IKM cluster has been impacted by this experimental study in its efforts to address the difficulty of large-scale unsupervised data. This suggested clustering technique reduces the amount of time and memory used while helping to solve complicated and huge datasets. This study evaluated the benefits of the ACO-IKM cluster by comparing it to the current k-means cluster approach and analyzing factors including peak memory, silhouette average score and time used. A cluster with a lower number is seen to have a better score than a cluster with a higher number. The silhouette

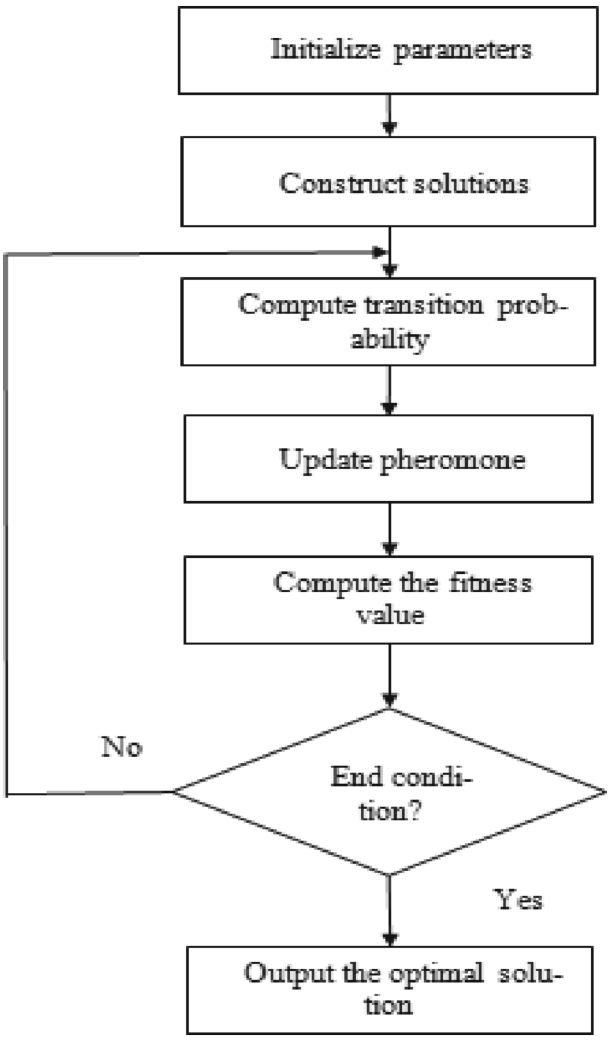


Fig. 4. Flowchart of ACO algorithms

average score is effectively calculated using both positive and negative values. Cluster 29 has the fewest negative values (132) among the clusters, with the Gap value starting at cluster 15 and continuing until cluster 29, as shown in Fig. 5.

The unsupervised data may be used to determine information by clustering similar values, as shown in Fig. 6, which helps to identify the optimum cluster. These clusters address negative value counts, and clusters 15 to 29 contain 132 negative value counts.

Table1 illustrates the performance comparison of various improved K-Means models by silhouette average score, time consumption and peak memory used in which ACO-IKM clustering has better optimum cluster with less time consumption and low peak memory usage.

```
This Many Clusters: 15 | Number of Negative Values: 132
This Many Clusters: 16 | Number of Negative Values: 132
This Many Clusters: 17 | Number of Negative Values: 132
This Many Clusters: 18 | Number of Negative Values: 132
This Many Clusters: 19 | Number of Negative Values: 132
This Many Clusters: 20 | Number of Negative Values: 132
This Many Clusters: 21 | Number of Negative Values: 132
This Many Clusters: 22 | Number of Negative Values: 132
This Many Clusters: 23 | Number of Negative Values: 132
This Many Clusters: 24 | Number of Negative Values: 132
This Many Clusters: 25 | Number of Negative Values: 132
This Many Clusters: 26 | Number of Negative Values: 132
This Many Clusters: 27 | Number of Negative Values: 132
This Many Clusters: 28 | Number of Negative Values: 132
This Many Clusters: 29 | Number of Negative Values: 132
```

Fig. 5. Number of negative values for various clusters

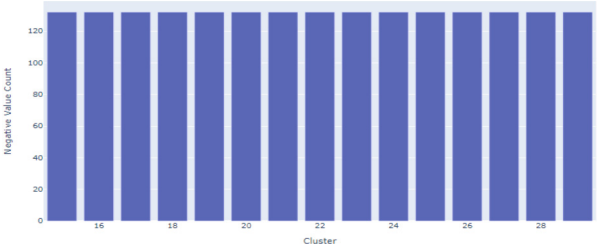


Fig. 6. Number of negative value counts for various clusters

Table 1. Performance comparison of various improved K-Means model

Performance metrics	K-means	IGBKM cluster	ACO-IKM cluster	MBK-mean cluster
Silhouette average score	0.0992	0.0953	0.113	0.0943
Time consumption	0.6812	0.32	0.19	0.7868
Peak Memory used	253.57	241.83	203.81	225.7

Figure 7 shows how the silhouette average score compares to different K-Means in which ACO-IKM cluster has 0.113 that is comparatively higher than IGBKM cluster, MBK cluster and k-mean cluster are 0.0953, 0.0943 and 0.0992 respectively.

Figure 8 illustrates the comparison of time consumption to various K-Means in which ACOIKM cluster has consumed lesser time of 0.19 s than MBK cluster, IGBKM cluster and k-mean cluster are 0.7868 s, 0.32 s and 0.681 s respectively.

Figure 9 illustrates the peak memory comparison used to various K-Means in which ACOIKM cluster has consumed 203.80 Mbs that comparatively higher than MBK means

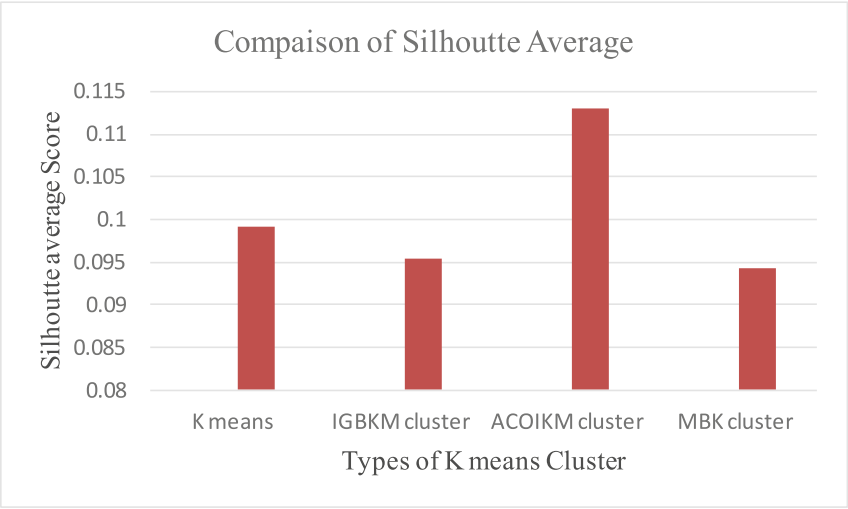


Fig. 7. Silhouette average score comparison for various K-means

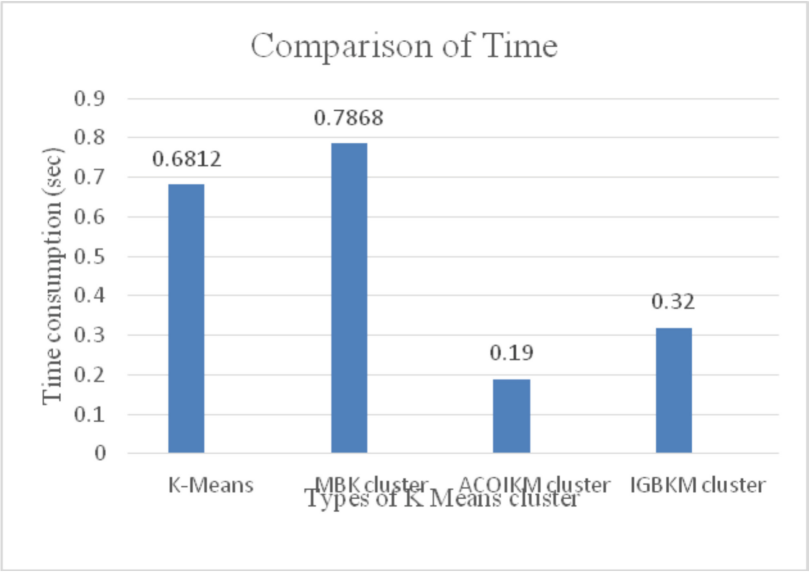


Fig. 8. Time-consuming comparison of different K-means

cluster, IGBKM cluster and k-mean cluster are 225.7 Mbs, 241.83 Mbs and 253.57 Mbs respectively.

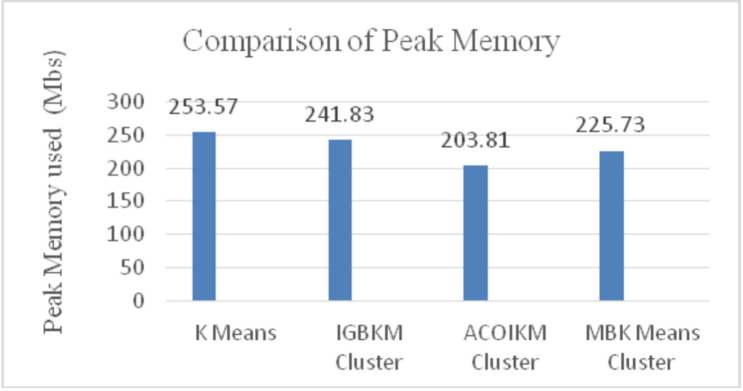


Fig. 9. Peak Memory VS Types of Clustering

5 Conclusion

The study is to develop methods to increase accuracy by selecting statistics with a particular shape using the human freedom index (HFI) after examining a large number of data. The comparison of several K-Means clustering techniques reveals the advantages and disadvantages of each technique with regard to resource consumption, computational efficiency and clustering performance. Though well-known and often utilised, traditional K-Means clustering has a random selection of initial centroids that causes it to require a lot of memory and take a long time to execute. In order to overcome these problems, the suggested MBK-Means clustering which has been improved by the Hamely algorithm focuses to increase cluster performance and decrease iteration counts by selecting centroid locations more consistently. Although MBK-Means offers improvements over methods like IGBKM and ACO-IKM, the ACOIKM cluster stands out as the most efficient and effective. ACOIKM is an excellent choice for reducing runtime and memory utilisation in addition to providing greater silhouette scores, which indicate better cluster quality. ACOIKM is the best option out of all the techniques assessed for the HFI dataset, based on the performance evaluation, which demonstrates that it produces the most effective clustering solution, even in cases with high prediction needs.

References

1. Sarker, I.H.: Data science and analytics: an overview from data-driven smart computing, decision-making and applications perspective. SN Comput. Sci. **2**(5), 1–22 (2021). https://doi.org/10.1007/s42979-021-00765-8
2. Keyvanpour, M., Serpush, F.: ESLMT: a new clustering method for biomedical document retrieval. Biomed. Tech. **64**(6), 729–741 (2019)
3. Jothi, R., Mohanty, S.K., Ojha, A.: DK-means: a deterministic Kmeans clustering algorithm for gene expression analysis. Pattern Anal. Appl. **23**(3), 649–667 (2017)
4. Ramirez-Gallego, S., Krawczyk, B., Garcia, S., Wozniak, M., Benitez, J.M., Herrera, F.: Nearest neighbor classification for high-speed big data streams using spark. IEEE Trans. Syst. Man Cybern. Syst. **47**(10), 2727–2739 (2017). https://doi.org/10.1109/TSMC.2017.2700889

5. Madan, S., Dana, K.J.: Modified balanced iterative reducing and clustering using hierarchies (m-BIRCH) for visual clustering. Pattern Anal. Appl. **19**(4), 1023–1040 (2016)
6. Liu, H.C., Hou, X.N., Yang, Z.: Design of Intrusion Detection System Based on Improved Kmeans Algorithm. Comput. Technol. Dev. **26**(1), 1–18 (2016)
7. Harifi, S., Byagowi, E., Khalilian, M.: Comparative Study of Apache Spark MLlib Clustering Algorithms. In: Tan, Y., Takagi, H., Shi, Y. (eds.) Data Mining and Big Data: Second International Conference, DMBD 2017, Fukuoka, Japan, July 27 – August 1, 2017, Proceedings, pp. 61–73. Springer International Publishing, Cham (2017). https://doi.org/10.1007/978-3-319-61845-6_7
8. Siddiqui, M.K., Naahid, S.: Analysis of KDD CUP 99 dataset using clustering based data mining. Int. J. Database Theory App. **6**(5), 23–34 (2013). https://doi.org/10.14257/ijdta.2013.6.5.03
9. Aydin, N., Yurdakul, G.: Assessing countries' performances against covid-19 via wsidea and machine learning algorithms. Appl. Soft Comput. **97**, 106792 (2020)
10. Kucukefe, B.: Clustering macroeconomic impact of covid-19 in OECD countries and china. Ekonomi PolitikaveFinansAra¸stırmalarıDergisi, **5**(Özel Sayı), 280–291 (2020)
11. Ratnakar, S., Rajeshwari, K., Jacob, R.: Prediction of heart disease using genetic algorithm for selection of optimal reduced set of attributes. Int. J. Adv. Comput. Eng. Netw. **1**(2), 2320–2106 (2013)
12. Yang, J.: Clustering analyzing of undergraduate schools based on k-means algorithm. In: International Conference on Advanced Mechatronic Systems (ICAMechS), Xiamen, China, pp. 309–311 (2017)
13. Joshi, A.; Sabitha, A. S.; Choudhury, T.: Crime analysis using K-means clustering. In: International Conference on Computational Intelligence and Networks (CINE), Odisha, India, pp. 33–39 (2017)

Advancements and Innovations in Mixed Reality: A Comprehensive Review

Jitesh Solanki(✉) and Kalpesh Popat

Marwadi University, Rajkot-Morbi Highway, Rajkot, Gujarat 360003, India
mr.jiteshsolanki@gmail.com

Abstract. This paper will cover ground-breaking developments and innovations in Mixed Reality (MR). From its evolution, to its applications, user studies, sensor systems and future implications—these will be the main topics analyzed in this research. First, the article kicks off with a general overview of MR, and consequently, goes on with a literature review. It presents the historical reveals, and discourses MR effects in different fields of application. Contrasting AR, VR, and MR, one can establish that they are basically different due to their distinguishing characteristics. User study case views which points us usability and user experience doctrines; a deep look at sensor describes are used in MR systems is another component of the knowledge structure. Based on this global overview, the post on the current MR technology advances discusses the future implications and challenges facing the field of research. It provides a comprehensive approach including the benefits and limitations of the topic.

The review also focuses on the social and industry areas that are anticipated to be influenced by MR and appreciates its ability to bring revolutionary changes. Using problematic areas and possibilities to view, such paper becomes the source of knowledge for researchers, practitioners, and anyone looking for how to handle mixed realty in a complex way. The mixed reality development and utilization is presented in this conference from an in-depth study with this knowledge being the prerequisite for understanding where the technology is coming from and where it is headed.

Keywords: Mixed Reality · Augmented Reality · Virtual Reality · Educational Game · Technology Acceptance

1 Introduction

Mixed Reality (MR) technology is a type of a technology comprising the reality with the physical realities and virtuality, which in its turn gives the users opportunity to perceive also quite an immersive experience after the colliding those elements of reality and virtuality. Within the last few years, the actual focus was in the establishment of the new MRI technicalities that can be used in different aspects. The particular review appraises such viewpoints and theories as may be based on new findings and researches described in the conferences.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
S. Rajagopal et al. (Eds.): ASCIS 2024, CCIS 2428, pp. 40–56, 2025.
https://doi.org/10.1007/978-3-031-86302-8_3

A marked phenomena is learning through MR which is one of the newest concerns in education of today. As shown in Fig. 1 the research of Liu and others ("how neuroscience research can influence the K-12 education?") [1], the students can be included in their learning environment by using concept maps embedded in VR space. That is why this kind of connection is useful as a tool for memorization and learning process. Research also produces information in line with the cited study in the footnote [6], which targets exactly why people are receptive (or not) to learning games. Success of our MR educational tools will rely on users' preferences and behaviors, which will be our major focus as we design them.

MR is arguably the new face of entertainment. For instance, Lu and his colleagues came up with a "Gaze Pinch Menu," that lets user to handle multiple tasks at a time in an MR environment. Utilization of this invention improves user experiences. The participation of MR in heritage conservation and tourism up gradation is a new field which was studied by another study by Lu and colleagues [11]. It offers a new dimension to history and culture perception too.

In health care, MRI possesses multitude of capabilities such as training medical staff and assisting in patient rehabilitation. Schieber together with others [12] have made an MR to help visually impaired people in navigation. In parallel, Banaga and O'Sullivan [17] examined the possibility of using MR for stroke rehabilitation, aiming to deliver a more engaging and effective treatment.

Fig. 1. AI Generated Visualization of Mixed Reality. Image created by the author using an AI model for research purposes

The field of engineering and design are delighted to use MR as a powerful medium for visualization, simulation and prototyping. Suleiman, and his team [20] assume that there is a process from virtual reality to mixed reality, which increase mixed reality and design precision. In addition, Li and the team developed MR user interfaces especially for engineering science with guiding the process and user-friendliness that are time-saving.

Innovative technical inventions seem to make a quite strong contribution into the MR system's development. Instances like gesture recognition [14], physics-induced user

interfaces [7], and predictive models [30] are only a few among the many algorithms for the enhancing of MR's realistic, interactive, and colorful features. The hardware improvements as discussed in the citation [9], including the writable vibro motors, are another factor that is broadening the application of MR instruments.

The critical assessment is an in-depth coverage of the current state of art of mixed reality (MR) technology used in educational, entertainment, health and engineering fields. It brings data from different projects and researches together and proceeds in an easy understandable way of explaining what phases MR has passed through and where this technology is heading to.

2 Literature Review

Mixed reality (MR) is now much more commonly used than virtual reality (VR) is and covers a wide range of research and applications. During this section, I will bring up some of the major fields, their results and how they are generating new knowledge.

In education, MR is breaking down the barriers and playing useful role now. Li and researchers [1] presented the example of using concept maps in MR course formation to illustrate their more interactive of knowledge activity. Cao and Sun fabricated a practical educational toolkit suitable for building educational environments on mobile mixed reality devices. Students are no longer being restricted to lecture-based classes. They are given a chance to actively participate in their lessons through technology. The Zhang and his team [6] found the key factors to which users are willing to try out MR for education. With the users' behaviors and learning processes now clearly displayed, we are getting the better understanding of target audience and learning environments.

When it comes to the theater, MR is considered a unique way of storytelling and an interactive experience. Lu and his team [4] were the first to put forward the "Gaze-Pinch Menu" as an interface to combine several tasks at a time within MR, where users can enjoy the operation and which is also efficient. Additionally, another person's [11] study on multi-sensory (MR) in tourism of cultural heritage showed that the novelty brought by this tech could value history by its interaction and familiarity from different perspectives.

The other one of the major areas where the MR stands out is medical imaging, where it returns astounding results. It is indeed Schieber and her staff who have proposed a road code that forbids to blind or visually impaired people to drive more freely not only giving a tribute to MR but also concerning us about improving accessibility and independence. MR methods that were studied by Banaga and Sullivan [17], can now help patients with stroke who are going through the process of motor skills rehabilitation.

MR is assuming a bigger and bigger role in engineering world as well as a great deal in highly sophisticated design. Sulaiman's group travelling the path of gas turbine designing that with assistance of MR giving them a handy transition between virtual reality and real world. Li and other researchers [10] were concerned about how to design the most comfortable user interface for MR gadgets imbedded in helmets, and they wanted to allow both scientists and designers to use these tools with greater convenience.

Technical innovations form the foundation of major progress in MRI technology. For example, techniques like a motion manipulators [14], physical affordance in user

interface [30], and model predicting [7] are some of the edge things in modalities of user interaction with MR. Combining all these together is the software which can be known as the control center for the whole space. After that the progress in the hardware also for example using wearable vibro motors [9], is increasing the usability and flexibility of MR.

In general, MR research is incredibly active and provides multiple fields with new possibilities. That is from education, entertainment, healthcare, and even engineering. The very recent developments and progress are outlining while preparing the ground for what could be the future of Mixed Reality exciting fields.

3 Evolution of Mixed Reality

Through time and several industries and the new technology, MR has changed about a lot. This section, therefore will analyze the path of MR to date, including the milestones and the categories.

Initially, basic principles were being designed and prototypes were sometimes being made. For close use, Zhang et al. come up with the physics to enable the MR to be more realistic and alike to reality while many others opted for public display which should enabled to explore some challenges and opportunities for MR to be employed in real life situations [3].

MR was bettered and people began to see it as a tool to be used in education. Liu and his associates were among the first who integrated concept maps with the mixed reality in the quest for more learning immersion through the blending of both virtual + real-life experiences [1]. This concept eventually led to the development of mixed reality games which were designed to improve the interactivity level and make the process more fun [6].

At the film's significant success, it did not take long time for the entertainment industry to realize the strengths of MR, especially in creating narratives and entertainment. In order to solve several issues all at once, Lu et al. [4] created a "Gaze-Pinch Menu" in MR which made the interface interesting and added some more paths for games and entertainment in MR.

Besides, MR's position in healthcare is very important, for rehabilitation, therapy, and medical training, in particular. The virtual reality (VR) system for visually impaired users or those with low vision proposed by Schieber et al. [12] is not only to promote their access but also to generally positively contribute in their lives; more so, Banaga and O'Sullivan [17] tried to demonstrate the efficacy of the VR therapy in stroke recovery which allowed the patients to improve their movement abilities and physical strength.

In the realistic field of engineering design and other where applicable, so far none can match MR as it mentioned by Sulaiman et al. in their discussion about gas turbine project transfer from virtual reality into mixed reality that could enhance the collaboration among engineers as well as facilitating design process [20]. The research by Li et al. [10] is more centered on technical work processes looking into user interfaces of displaying information on head mounted displays (HMD) that could be self-used with ease within technical work environments.

Hardware redesigns and software modifications are some of the changes needed to support the growth of MR. For example, gesture supported controls, predictive systems,

feedback tools for feeling virtual objects among many others have helped mixed realities to be more interactive and realistic [14, 30, 31]. Lastly, implementations of MR [5] in mobile devices yield broader coverage across beneficiaries and sectors.

In other words, moving from an idea to an innovative tool that influences different aspects of society. Which presents the kind of force that can be brought about when new techs are incorporated with multi-disciplinary work and creative applications? Mixed reality has continually been changing not only how we interact with computers but also how learning, entertainment as well as healthcare sector will look like thus shaping our future.

4 Applications of Mixed Reality

Already now the Mixed Reality (MR) technology effectively works in different spheres of life and blurs reality and fantasy. This way, we not only engage digitally with the information but also with the world around us. The next part will discuss the use of VR in education, healthcare, entertainment, and engineering in detail.

Education: In education, as shown in Fig. 2 MR contains actual settings and adds virtual elements to the real world thus making the learning environment more interesting. The researchers Liu et al. [1] proved that constructing concept maps results in better knowledge comprehension and retention of information among students. Cao and Sun [5] came up with various tools for authors to create meaningful MR experiences on all mobile devices, which made this learning fun and interactive. Besides, Zhang et al. noticed [13] that MR is on its way of entering education as it delivers fresh tools for improving training and learning activities.

Fig. 2. AI Generated Mixed Reality in Education. Image created by the author using an AI model for research purposes

Healthcare: As shown in Fig. 3 Healthcare workers' training, patient rehabilitation and psychotherapy have all been altered with the use of mixed reality, thus opening new

ways for clinical practice. As an example, the convenient MR system by Schieber et al. [12] let's nearly blind people find their way through their natural surroundings more efficiently. Banaga and O'Sullivan [17] clinically studied neurological rehabilitation using MR-based strategy, in which motor skills are regained, and cognitive areas of the brain are retrained.

Fig. 3. AI generated Mixed Reality in Healthcare. Image created by the author using an AI model for research purposes

Entertainment: Media entertainment led the way in merging MR and turning narratives into immersive stories and engaging games. The research by Lu et al. [4] referred to as "Gaze-Pinch Menu" is one that enables the users under MR environments to do several activities at one position without having to change their location. In this case, the scope of video games and other applications gets broadened as they get the players more involved. Jiddi et al. [8] suggest the development of sensors to let virtual reality become closer to real experiences and to obscure the usual reality navigation between virtual and real worlds seem more smooth.

Engineering and design: The platform of MR also conducts changes in the manufacturing processes and the sketching of concepts. Ruminski et al. [9] has shown that coupled MR system and wearable devices made virtual prototyping and simulation with provision of real sensory experiences possible. Sulaiman et al. [20], talk about the use of MR technology in consultation and decision-making. They mention the use a gas turbine engine prototype design as an example.

The findings of the empirical studies support the idea of having a virtual psychotherapy assistant, who is based in a therapy room especially designed for him and her work is to conduct CBT, which is an effective stress reducing process. The simulated resting breathing- HRV-BF (Heart Rate Variability Biofeedback) technique is also a stress reduction technique due to its effect on the Parasympathetic Nervous System [32].

Although use of mixed reality in the field of tour guiding is not well developed, the natures of the application is the ability to blend the real environments and the virtual of objects and the vectors which the end user can interact with the virtual contents through

gestures. Therefore, the outcomes of the study are to design the mixed reality digital guide system for historical places and cultural artifacts have prescribed below to compare the degree of learning and immersiveness over and above that of M-Guide and Paper guide [33].

Moreover, there is more in these neglected topics than you can imagine. Thus the heritage is the first one as it is used with both the MR together and makes the experience of the visitors of the historical site and the art pieces bigger. Let's discuss health care, particular to autism and its treatment. MR practice by researchers has been successful in the area of building social skills among individuals having autism spectrum disorder (ASD) [22]. Also worth to mention is that even the financial sector is applying MR technology for virtual banking or augmented trading platforms [15].

As the final point, most of the MR applications across many sectors significantly provide multiple benefits. Teach students how technology affects the way we learn, how patients are cared for, and more other aspects of people's lives can be quite fascinating since they are basis of the wider application of the technology into many other disciplines.

5 Comparison AR/VR/MR

Aspect	Augmented Reality (AR)	Virtual Reality (VR)	Mixed Reality (MR)
Interaction with Real World	AR overlays digital content onto the real world, enhancing real-world experiences	Limited interaction; primarily virtual environments	MR seamlessly integrates digital content with the user's environment, blending virtual and real elements
Immersion	Partial immersion; digital overlays enhance real-world perception	Fully immersive, often blocking out the physical surroundings	Varying levels of immersion, allowing users to interact with both real and virtual elements simultaneously
Environment Control	Users have limited control over real-world elements	Users have full control over virtual environments	Allows users to manipulate and interact with both digital and physical elements in their environment
Integration of Digital Content	Digital content is integrated into the real-world environment	Entirely virtual, digitally constructed worlds	Seamlessly integrates digital content with the user's physical environment, enhancing real-world experiences with virtual elements

(*continued*)

(*continued*)

Aspect	Augmented Reality (AR)	Virtual Reality (VR)	Mixed Reality (MR)
Real-time Interaction	Real-time interaction with digital overlays in the user's environment	Enables real-time interaction within immersive virtual environments	Facilitates real-time interaction with both virtual and real-world elements simultaneously
Physical Feedback	Physical feedback through haptic devices and wearables	Offers physical feedback through haptic feedback devices	Can incorporate physical feedback devices to enhance user interaction with virtual and real-world elements
Education	Enhances educational experiences by overlaying educational content onto the real world	Transforms learning experiences by integrating virtual content into real-world environments	Transforms learning experiences by integrating virtual content into real-world environments
Healthcare	Assists in medical training and visualization	Used in healthcare for therapy, training, and patient experiences	Employed in healthcare for simulations, training, and remote collaboration, enhancing patient care and medical education
Entertainment	Provides interactive entertainment experiences through AR-based games and apps	Offers immersive and interactive entertainment experiences through VR gaming	Includes interactive gaming, immersive storytelling, and virtual tours, blending real-world and virtual experiences
Engineering & Design	Enhances engineering and design processes with real-time data overlays	Revolutionizes engineering and design by enabling collaborative virtual design sessions	Revolutionizes engineering and design by enabling collaborative virtual design sessions
Cultural Heritage Preservation	Preserves cultural heritage by overlaying historical information onto physical artifacts	Preserves cultural heritage by creating immersive experiences, such as virtual tours of historical sites	Preserves cultural heritage by creating immersive experiences, such as virtual tours of historical sites

(*continued*)

(*continued*)

Aspect	Augmented Reality (AR)	Virtual Reality (VR)	Mixed Reality (MR)
Accessibility for Special Populations	Can improve accessibility for individuals with disabilities	Can provide virtual experiences tailored to special needs	Offers enhanced accessibility by integrating digital assistance tools and adaptive interfaces into the real-world environment

6 User Studies of Mixed Reality: Advancements and Innovations

The progress in the fields of VR engineering and innovations give the people ideas of the extent the projects will impact their lives and this will be achieved through phenomenology, technology integration and UX testing in diverse situations. This essay has demonstrated and responded to issues of usability tests, which is highly relevant in accomplishment of the complete literature analysis.

Education and Training:
Besides that, many researchers studied the Moon in a different aspect. To elaborate on the impact of using MR educational game that includes a concept map, Liu et al. [6] explored how students found the complexity of the subject more understandable. This research evidences that the concepts taught in those rivers can be applied into education. Zhang et al.'s [13] published study addresses the future tendencies and directions that researchers in educational MR ought to be heading.

Healthcare and Rehabilitation:
In this area, we would be in close proximity to a lot of research that reveals the use of Virtual reality for patient rehabilitation and therapy. Through this, they focused on stroke recovery by using MR and looking on the role of proxy agency and embodiment in such researches. In another research work that was done by Schieber et al. [12], they developed and tested an MR ghost-navigating device for the visually impaired showing the many other uses of MR in health conditions.

Cultural Heritage Preservation:
As well, the progress of MR in context of preserving our cultural heritage and virtual visits is also taking place. To this Borodkin [21] added eg the enhancements in historical research through virtual and augmented reality and the fully immersive/interactive experiences which may be brought about by the deployment of the mixed reality. Other than that we know of Lu, et al. [11], asserting the mixed reality use in tours to cultural heritage and showing prospects of how to make MR applications for heritage usage.

Collaboration and Communication:
Aside this, various studies have been conducted on seamless integration of MR in teamwork and communication. For example, Hart et al. [23] were focused on emotional sharing and expression in MR common zones. Furthermore, Daeijavad and Maurer [27] noted that collaborative multi-view user interfaces for which MR is used to enhance the experience of users in the MR environment interfaces.

Research at user level is the main tool to enhance the technology of MR regarding various areas, such as education, medicine, cultural heritage and collaborative communication. With the main focus on MR's convenience, efficacy and users satisfaction researchers give a vital input into development of immersive and effective mixed reality applications, and furthermore this leads to the wide spread of new MR usage.

7 Sensors Used in Mixed Reality: A Comprehensive Overview

The MR applications contain the sensors based on numerous regions so that the users get a realistic interaction and virtual environments where the physical and digital features are connected without any types of gaps. The diagnosis included the sensor types as well as the case being the one described in previously published MR applications relevant in scientific papers.

Depth Sensors:
The primary sensor like LiDAR (Light Detection and Ranging) and ToF (Time-of-Flight) cameras is used for depth sense so that the environment can be more realistic and its objects are visible. Furthermore, these sensors are also accountable for offering spatial and object feature data that are accurate. Therefore, they help the real-time immersion and the identification of the scene [29].

Inertial Measurement Units (IMUs):
Without IMUs, the determination of MR devices' orientation and movement, as well as the users, would not be possible. Through accelerometers, gyroscopes, and magnetometers, as a calibration components, allow accurate moving of digital objects without effort and with good tracking [28].

Cameras:
Sensors that wrap around under the shape of RGB cameras, for instance, or stereo can visualize and physically encode digital information into the environment. One of the advancements is the capabilities of the RGB camera, which generates full color scenes while the stereo cameras use binocular vision to perceive depth, adding the element of realism to the MR surroundings [5].

Gesture Recognition Sensors:
Gesture recognition sensors, which involve depth cameras and wearable devices, are used to sense hand gestures and body movements allowing perceived natural and simple interaction between the user and the augmented world. These sensors convert the patients hand gestures as feedback input to which the virtual objects are controlled, menus are navigated and applications are operated [14, 28].

Environmental Sensors:
Environmental sensors, such as temperature sensors and humidity sensors will act as data collectors of the physical environment. This is to improve the realism and respondent of the varied MR experiences. Through the process of simulating authentic real-world environmental conditions in the virtual mode, such sensors make scenarios more detailed and pertinent [29].

Eye Tracking Sensors:
The eye tracking sensors that are integrated into MR systems help monitor the movement of users' eyes. Such sensors give MR systems a chance to identify visual attention and optimize content delivery. These trackers make possible focused rendering in that high quality graphics are shown in the user's line of sight only; such an approach significantly reduces computational load and enhances performance [4].

Merits:
The sensors involved in Mixed Reality (MR) technologies offer relative enhancements in interaction and user experience. LiDAR and ToF cameras are also used to measure the distance of an object from the sensor so it introduces perception of real space and one can interact with objects in the virtual world. They support live experience within the MR environment and enhance the depth of perspective of the MR scenes. IMU accelerometers, gyroscopes, and magnetometers significantly provide high accuracy for the device's orientation and user's movements for perfect control over virtual objects. Examples includes RGB and stereo cameras which are prevalent to offer a three-dimensional full color images with pseudo-realistic MR scenes. Gesture recognition sensors improves the level of interactivity because using hand gestures to maneuver virtual objects is much easier than other commonly accepted methods. Similarly, environmental sensors allow conditions such as temperature and humidity to make the environment more or less realistic while eye tracking sensors improve system performance and display high quality images within the user's field view to avoid wasting resources on what the user is not seeing.

Demerits:
Yet, these sensors have some limitations associated with them. These range from expensive points cloud capturing devices which perform poorly in certain conditions such as low light or reflecting surfaces poor. An IMU has the problem of "drift", whereby the tracking accuracy reduces over time and some reorientation is necessary. They are also affected by magnet influence and might display wrong movement detection. Cameras, like any image sensing devices, must be well illuminated to function the best; the depth and image data captured have to be processed; too extensive a computation can slow down the real-time work. Gesture recognition sensors can potentially suffer from the problem of inability to recognize complex gestures or movements especially if the area that a user occupies is crowded, which would reduce the interaction to be expected. Environmental sensors provide extra parameters and can be viewed as a load that increases the demands on the work of a digital processor. Finally, eye tracking sensors are expensive, and its application may not be appropriate for people with glasses or other vision disorders thus restricting its utility.

8 Future Implications of Mixed Reality: A Comprehensive Exploration

Mixed reality or MR technologies are developing very quickly and the number of possible implementations of such technologies is increasing every day in fields as diverse as medicine, entertainment or education. The latest knowledge and developments in the MR field are given attention in this all-inclusive review. Furthermore, it strives to emphasize MR's future implications on the people's lives.

MR application in education and training is an area of research with high potential [1–3, 6, 13, 19]. Addressing MR as educational tool multiples interactive opportunities for students which increase their involvement with study process and knowledge. MR adds physical depth to previously dry material from an interactive concept map [1] to pre hospital intervention simulations [19], touching many areas of learning including STEM domain and medical training. As MR technology gets more and more available and affordable its chance to redefine learning significantly rises with the time.

In medical field, MR holds a very bright future that can be helpful for both patient care and medical training effectiveness [12–22]. Interventions based on MR can support visually impaired patients in orientation [12] as well as stroke recovery [17]. Consequently, their applications will be intended to tackle different healthcare challenges. In addition, the MR-based surgical investigation systems [25] provide doctors a better visualization, which is necessary to carry out surgery with higher accuracy and efficiency. While MR system is still in its infancy, it has a lot of future prospects being influential in how healthcare service and health outcomes are handled.

The last important area is preservation and tourism of cultural heritage which stands for [11, 21]. Connecting MR to historical sites and artifacts creates an opportunity for people to experience past times in rich, multi-layered ways, which hitherto was impossible. Whether by its symbolic rebirth of well-known landmarks or its aid to archaeological investigations, MR has the power to provide a new experience of culture as well as protect our heritage.

MR is applied in a human-computer interaction field, which enables the increase of both the interface and interaction modes [3, 26]. Change-of-interface has made possible by advanced forms of MR menu [4] and interface [14] with forces feedback tools [31]. Gaze is used in place of mouse, gesture in place of the interface and force is connected to touch, alike. Such developments not only improve user interfaces, but they empower with new communication, and collaboration mechanism, to be used in domains of diverse nature.

Additionally, MR is on track to overturn the existing order in different spheres such as finance [15], factory work [20] and online gaming [28]. On the financial side, MR technologies are generally expected to help develop counteraction of immersive data visualization and decision-making tools [15]. In manufacturing, these applications are used to spur hands-on modeling, design validation and process line efficiency [20]. Additionally, in gaming, the MR immersive [28] games cater for players' distinct requirements, ranging from diving into the fantastic world that gets the line between the virtual and physical worlds blurring.

But, the future of MR in this regard is literature dependent upon technological advancement and research [16, 24, 29]. With increasing capability of scientists to analyze

historical data, information about users' preferences and predictive modeling, usefulness and versatility of VR and MR systems will continue to go up. Besides that, the seeing of applications of these novel technologies such as the use of AI and IoT can be expected to increase the functional abilities of MR.

As a wrap-up, the prospects of MR appear promising and are pointing to many great chances. Whether as part of education and medical technologies reformation or cultural experience and human-computer interaction transformation, MR will be of utter importance to the change of the way we learn, work and interact with our surround. Through full-scale optimization and interaction, employee performance is enhanced, and a hybrid workforce with strong identity formation will be created.

9 Challenges and Opportunities in the Advancements of Mixed Reality

MR is a relatively new invention which is promising so much and can completely change education, medicine, entertainment etc. In that case, the technology comes with its swarm of problems as well as chances, which is not a walk in the park but a real task. This essay reviews these subjects including recommending solutions based on the latest academic research and innovation.

MR integration, in general, entails the use of different advanced technologies which is the biggest challenge by development. A user experience that is not interrupted or makes users feel submerged in the digital world is our major goal. One is still obliged to come up with a user interface that on the one hand can handle various inputs concurrently, and on the other hand, cope up with those inputs within reasonable limits. Moreover, determination of exactness and validity of MR in tracking and inputting [7, 24, 31] ways of performing actions are also problems that complicate things. The effective implementation of MR throughout the different sectors extensively depends on the technical solutions of the mentioned problems.

The other obstacle is particularly in the content creation and environment reconstruction [5, 8, 21, 29]. The construction of highly immersive interactive and engaging content for MR requires specialized skills and internal resources. On the other hand, skillful reconstruction of real world scenarios is a technology that requires advanced perception and modeling methods. Resolving these fake challenges is imperative for the enhancement of the quality and immersion of MR content and experiences.

Accessibility and inclusion in MR experience are two biggest problems now [12, 17, 22]. Developing MR systems that consider visually impaired peoples' demands [12] and stroke rehabilitation applications [17] requires a special attitude towards a variety of users' requirements and abilities. Furthermore, digital equity and affordability must be provided to enable the popularization of MR technologies and prevent digital gaps from further expanding.

In the same way, MR creates an ample scope for innovation and flow in publication across different fields of genetics. In the education sector, MR enables the alike situation of education though unequivocal learning experiences, also including the students' engagement and their understanding of the courses material. Making mind-maps and

educational games available as learning tools in MR education, allows for multidimensional, student-friendly learning by integrating conceptual understanding and different forms of engagement.

As a result, MR technology in healthcare can quickly become a new standard facilitating a better diagnosis and training [12, 17, 19]. Via providing orientation assistance to the visually impaired in a virtual environment [12] and augmenting rehabilitation experiences through proxy agency and embodiment [17], MR is revolutionized the way healthcare solutions are delivered and consumed. Healthcare providers could be able to take advantage of MR technologies to have an efficient diagnosis, provide great treatment and reach an optimal result.

Not only does MR pioneer the field human-computer interaction and entertainment but also it pushes the boundaries. Gesture control interface [14] allows people to get involved with a new level of interactiveness and user friendly application. Primarily, sitting in front of the screen is replaced by constant movement. MR gaming also integrates the physical and virtual world, providing novel ways of social contact and recreation.

Therefore, rapid development and novelty in MR result as both the problem and the instrument in academic life, productivity, and policy influencing. Through focusing on main obstacles like graphics and comprehensiveness, content building, and availability, we can yield MR technologies' instrumental results in education, medicine, culture and entertainment. Leveraging team work, creation, and diversity can help harness MR to manufacture better and wider human experience in the future.

10 Conclusion

The extensive analysis of the progress and the occurrences in MR technology illustrated a virtual journey of an ongoing progress and discoveries through the latest researches in this quickly changing area. Through diversifying extant MR applications we came to appreciating a variety of existing studies and approved them as a benchmark to picture the future MR prospects.

This paper is focus on the issue of the widening spread of MR and its comprehensiveness – in fields such as education, healthcare, entertainment, and industry. Humanized data: Researchers and users of MR have been building up of virtual learning experience [1, 6, 13] in order to overthrow the old way of patient care and medical training [12, 17, 19], and lead to an efficient computer-human interaction [4, 14, 28]. Addition to the list, this technology is giving people ability to utilize it for cultural heritage preserving [11, 21], autism intervention [22], and financial services [15] as well.

The stress on resolving technological merits and deploying a user oriented perspective in MR applications is the main feature of the review. By simplifying the user interface features [4] and incorporating accurate tracking and input tricks, researchers continuously exploring the possible applications of MR technology. Furthermore, focus has been on content creation in the space the current is AI-powered [5, 8, 21, 29], accessibility into the space [12, 17, 22], and enhancing social experiences [23].

In addition to this, the article stresses the interdisciplinary nature of MR analysis in terms of psychology, education, computer science and health care that participate in it. MR is multidisciplinary so it pushes the actors towards the collaboration, integration and

innovation. Thus providing the room for emergence of the new solutions within these fields.

When looking at the future things get a little bit harder. One will at least face some moments that are challenging. Hardware innovation examples including the utilization of sensors on wearable devices function in conjunction to activate a range of unique VR scenarios. In addition to that AI and ML techniques can influence the evolution of the medicine research path and even develop interrelated tools that work with single patient.

Whereas these options come with unprecedented ethical and social issues that need to be addressed timely. Policies, for example, privacy evaluation policies, data security and digital inclusion become the subject and thus they should be put first while moving forward with MR technologies development.

References

1. Liu, Y., Liu, Y., Yue, K.: Integration of concept maps into the mixed reality learning space: quasi-experimental design and preliminary results. In: 2021 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW), Lisbon, Portugal, pp. 627–628 (2021). https://doi.org/10.1109/VRW52623.2021.00196
2. Zhang, Z., Weng, D., Jiang, H., Liu, Y., Wang, Y.: Inverse augmented reality: a virtual agent's perspective. In: 2018 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct), Munich, Germany, pp. 154-157 (2018). https://doi.org/10.1109/ISMAR-Adjunct.2018.00056
3. Cheng, K., Furusawa, I.: The deployment of a mixed reality experience for a small-scale exhibition in the wild. In: 2018 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct) Munich, Germany, pp. 214–215 (2018).https://doi.org/10.1109/ISMAR-Adjunct.2018.00070
4. Lu, L., Duan, P., Shen, X., Zhang, S., Feng, H., Flu, Y.: Gaze-pinch menu: performing multiple interactions concurrently in mixed reality. In: 2021 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW), Lisbon, Portugal, pp. 536–537 (2021). https://doi.org/10.1109/VRW52623.2021.00150
5. Cao, C., Sun, J.: Perceptual MR space: interactive toolkit for efficient environment reconstruction in mobile mixed reality. In: 2019 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct) Beijing, China, pp. 353–357 (2019). https://doi.org/10.1109/ISMAR-Adjunct.2019.00-13
6. Liu, Y., Liu, Y., Yue, K.: Investigating the factors that influence technology acceptance of an educational game integrating mixed reality and concept maps. In: 2021 International Conference on Advanced Learning Technologies (ICALT), Tartu, Estonia, pp. 409–413 (2021). https://doi.org/10.1109/ICALT52272.2021.00130
7. Zhang, Z., Li, Y., Weng, D., Liu, V., Wang, Y.: Physics-inspired input method for near-field mixed reality applications using latent active correction. In: 2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), Tuebingen/Reutlingen, Germany, pp. 737–738 (2018). https://doi.org/10.1109/VR.2018.8446291
8. Jiddi, S., et al.: Probeless and realistic mixed reality application in presence of dynamic light sources. In: 2018 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct), Munich, Germany, pp. 419-420 (2018).https://doi.org/10.1109/ISMAR-Adjunct.2018.00125
9. Rumiński, D., Klinker, G.: Augmenting mixed reality applications with the vibro motors wearable. In: 2018 IEEE International Symposium on Mixed and Augmented Reality Adjunct

(ISMAR-Adjunct), Munich, Germany, pp. 401-402 (2018).https://doi.org/10.1109/ISMAR-Adjunct.2018.00116
10. Li, Y., Hu, Y., Shen, X.: Exploring the display patterns of object-centered user interface in head-worn mixed reality environment. In: 2023 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW), Shanghai, China, pp. 737–738 (2023). https://doi.org/10.1109/VRW58643.2023.00211
11. Lu, Z., Zhang, J., Shapiro, K., Numan, N., Julier, S., Steed, A.: Reviving the Euston arch: a mixed reality approach to cultural heritage tours. In: 2023 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct), Sydney, Australia, pp. 821–826 (2023).https://doi.org/10.1109/ISMAR-Adjunct60411.2023.00181
12. Schieber, H., Kleinbeck, C., Pradel, C., Theelke, L., Roth, D.: A mixed reality guidance system for blind and visually impaired people. In: 2022 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW), Christchurch, New Zealand, pp. 726–727 (2022). https://doi.org/10.1109/VRW55335.2022.00214
13. . Zhang, H., et al.: Hotspots and trends of virtual reality, augmented reality and mixed reality in education field. In: 2020 6th International Conference of the Immersive Learning Research Network (iLRN), San Luis Obispo, CA, USA, pp. 215–219 (2020).https://doi.org/10.23919/iLRN47897.2020.9155170
14. Kumar, A., Mantri, A.: Gesture-based model of mixed reality human-computer interface. In: 2020 9th International Conference System Modeling and Advancement in Research Trends (SMART), Moradabad, India, pp. 226–230 (2020). https://doi.org/10.1109/SMART50582.2020.9337130
15. Goyal, N., Chauhan, S.S., Pandey, S., Singh, R., Kumar, A., Joshi, K.: Financial services 4.0 - future perspective based on mixed reality. In: 2023 IEEE International Conference on Contemporary Computing and Communications (InC4), Bangalore, India, pp. 1–4 (2023). https://doi.org/10.1109/InC457730.2023.10263246
16. Xing, Y., Liang, Z., Shell, J., Fahy, C., Guan, K., Liu, B.: Historical data trend analysis in extended reality education field. In: 2021 IEEE 7th International Conference on Virtual Reality (ICVR), Foshan, China, pp. 434–440 (2021). https://doi.org/10.1109/ICVR51878.2021.9483828
17. Banaga, S.G., O'Sullivan, C.: Stroke rehabilitation through proxy agency and embodiment in mixed reality. In: 2022 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct), Singapore, Singapore, pp. 964-967 (2022).https://doi.org/10.1109/ISMAR-Adjunct57072.2022.00216
18. Li, T., Liu, Y., Ma, S., Hu, M., Liu, T., Song, W.: NailRing: an intelligent ring for recognizing micro-gestures in mixed reality. In: 2022 IEEE International Symposium on Mixed and Augmented Reality (ISMAR), Singapore, Singapore, pp. 178-186 (2022). https://doi.org/10.1109/ISMAR55827.2022.00032
19. Morales-Mere, J., Chessa, J.J., Palomares, R., Cornejo, J.: Mixed reality system for education and innovation in prehospital interventions at Peruvian fire department. In: 2020 IEEE XXVII International Conference on Electronics, Electrical Engineering and Computing (INTERCON), Lima, Peru, pp. 1–4 (2020). https://doi.org/10.1109/INTERCON50315.2020.9220259
20. Sulaiman, H., Yusof, A.M., Ibrahim, N., Latif, R.A.: A gas turbine virtual reality application migration to mixed reality: development experience. In: 2020 IEEE Graphics and Multimedia (GAME), Kota Kinabalu, Malaysia, pp. 19–24 (2020).https://doi.org/10.1109/GAME50158.2020.9315123
21. Borodkin, L.: Evolution of virtual and augmented reality technologies in historical and archaeological research. In: 2022 International Workshop on High Technologies History and Development (HISTHIGHTECH), Moscow, Russian Federation, pp. 9–12 (2022).https://doi.org/10.1109/HISTHIGHTECH57099.2022.10038138

22. Chen, X., Lu, Y., Chen, Y., Zhao, W.: Visual analysis of research on the application of mixed reality technology in autism intervention. In: 2022 Asia-Pacific Computer Technologies Conference (APCT), Wuhan, China, pp. 7–10 (2022).https://doi.org/10.1109/APCT55107.2022.00016
23. Hart, J.D., Piumsomboon, T., Lee, G., Billinghurst, M.: Sharing and augmenting emotion in collaborative mixed reality. In: 2018 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct), Munich, Germany, pp. 212–213 (2018)https://doi.org/10.1109/ISMAR-Adjunct.2018.00069
24. Son, Y., Yeom, J., Lim, S., Kim, D. -H., Choi, K. -S.: Method and system for evaluating tracking performance of VR/AR/MR devices. In: 2022 13th International Conference on Information and Communication Technology Convergence (ICTC), Jeju Island, Republic of Korea, pp. 2074–2076 (2022). https://doi.org/10.1109/ICTC55196.2022.9952887
25. Cheng, C. -H., Nguyen, H.A., Liu, T.-C.: Design and implementation of a pre-surgical investigation system by VR+AR+MR techniques. In: 2020 IEEE International Symposium on Circuits and Systems (ISCAS), Seville, Spain, p. 1 (2020). https://doi.org/10.1109/ISCAS45731.2020.9180957
26. Li, X., Han, Y., Wu, Y., Yue, K., Liu, Y.: User study of an AR reading aid system to promote deep reading. In: 2021 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW), Lisbon, Portugal, pp. 611–612 (2021). https://doi.org/10.1109/VRW52623.2021.00188
27. Daeijavad, P., Maurer, F.: Hybrid user interfaces for multiple views: why designer intuition is not enough. In: 2023 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct), Sydney, Australia, pp. 196–198 (2023).https://doi.org/10.1109/ISMAR-Adjunct60411.2023.00045
28. Chen, K.-Y., Lee, J.-D., Zhang, T.-J.: The body sensor suit with mixed reality interactive games. In: 2021 IEEE International Conference on Consumer Electronics-Taiwan (ICCE-TW), Penghu, Taiwan, pp. 1–2 (2021). https://doi.org/10.1109/ICCE-TW52618.2021.9603144
29. Suma, E.A., Krum, D.M., Bolas, M.: Sharing space in mixed and virtual reality environments using a low-cost depth sensor. In: 2011 IEEE International Symposium on VR Innovation, Singapore, pp. 349–350 (2011).https://doi.org/10.1109/ISVRI.2011.5759673
30. Sumithra, T.V., Ragha, L., Desai, R., Vaishya, A.: User centric predictive model for mixed reality applications. In: 2022 2nd Asian Conference on Innovation in Technology (ASIANCON), Ravet, India, pp. 1–5 (2022). https://doi.org/10.1109/ASIANCON55314.2022.9908855
31. Gonsher, I., Lei, Z.: Prototype of force feedback tool for mixed reality applications. In: 2021 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct), Bari, Italy, pp. 508–509 (2021).https://doi.org/10.1109/ISMAR-Adjunct54149.2021.00123
32. Nath, N., et al.: Integrating cognitive behavioral therapy and heart rate variability biofeedback in virtual reality, augmented reality, and mixed reality as a mental health intervention. In: 2024 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW), Orlando, FL, USA, pp. 1200–1201 (2024). https://doi.org/10.1109/VRW62533.2024.00394
33. Chu, Y.-C., Lin, C.-C., Fu, K.-L., Chen, H.-R.: Design and evaluation of mixed reality application for historical sites guide. In: 2024 10th International Conference on Applied System Innovation (ICASI), Kyoto, Japan, pp. 172–174 (2024). https://doi.org/10.1109/ICASI60819.2024.10547905

Enhancing Sleep Disorder Prediction: CNN with IQR and Lasso Preprocessing

V. Nagarajan and C. Meenakshi(✉)

Department of Computer Applications, Vels Institute of Science, Technology and Advanced Studies (VISTAS), Chennai, India
meenasi.c@gmail.com

Abstract. Millions of people worldwide suffer from sleep disorders, which spotlight the significance of precise diagnosis and prognosis. Sleep disorder problem is nowadays a major problem in adults which make them unsleep thus leading to many complications. The main scope of this research is to predict sleep disorder in a well efficient manner using machine learning approaches. To improve the predictive accuracy for the classification of sleep disorder, we present a unique method in this study that combines the application of Lasso feature extraction and Convolutional Neural Network (CNN) with Interquartile Range (IQR) preprocessing. The data quality was enhanced and the impact of anomalies on model performance was reduced by preprocessing the dataset first. The preprocessing procedure was done using IQR, a reliable technique for the identification and removal of anomalies. At the same time, Lasso feature extraction is implemented to choose the more related attributes, improving interpretability, and minimizing dimensionality. The CNN algorithm is well-known for its capability to spontaneously read hierarchical attributes derived from the dataset. The model is accelerated with the preprocessed attributes extracted from the Lasso and IQR combination. The incorporation permits CNN to acquire both global and local designs in the data, increasing its ability to discriminate between different types of sleep disorders. The algorithm is trained with the help of the ISRUC-Sleep dataset. The outcome of the algorithm is estimated concerning precision, accuracy, F1-score, and recall scales. The results obtained from the experiment illustrate that the suggested combination of Lasso and IQR preprocessing along with CNN performs well than any other individual algorithms based on exactness of 90.30%, precision value of 0.89, recall rate of 0.87 and F1-score of 0.90 respectively. The software tool used is Jupyter Notebook and the Python programming language is used.

Keywords: Healthcare · Sleep disorder · Classification · Prediction · IQR preprocessing · Lasso feature extraction · CNN · Machine learning · Performance evaluation · Accuracy

1 Introduction

Humans naturally sleep to restore, repair, and re-energize the brain and body. Sleep has a direct effect on both the physical and mental well-being of a human. As a result, sleep disturbances cause weariness and decreased productivity (Dinh-Van Phan et al., 2021).

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
S. Rajagopal et al. (Eds.): ASCIS 2024, CCIS 2428, pp. 57–68, 2025.
https://doi.org/10.1007/978-3-031-86302-8_4

The sleep disorder sickness falls under the dual categories of mental and physical issues. It levies several challenges and issues, including anxiety, sleep-wake disturbances, and discomfort during the day (Raed Alazaidah et al., 2024). Further, depression, a type of varied mental illness has a chronic tendency and is typified with symptoms like low mood, poor cognitive function, lethargy, and suicidal thoughts. Notably, sleep issues are common in those with depressive illnesses (Mi-Mi Zhang et al., 2022).

Sleep disruptions were identified as a crucial component in the present diagnostic criteria for mental disorders. In major depressive disorders, subjective sleep complaints such as sleeplessness and excessive sleepiness are quite common. A critical issue of depressive illness starts with disturbed sleep and complaints of insomnia during the night for a duration of two weeks or more. It works as a valuable indicator of the later beginning of major depression. Hence, the right treatment for sleep distraction issues exclusively lowers the chance of acquiring depression among common people (Theofanis Vorvolakos et al., 2020).

Sleep-related illness comprises a wide range of criteria that disturb general sleep patterns. It also leads to a significant impact on the mental and physical health of an individual and overall creates turmoil in the quality of life. These illnesses occur in different ways, such as nap apnea, narcolepsy, fidgety legs condition, and insomnia, among others. It also results in impaired cognitive function, and daytime fatigue, and maximizes exposure to the danger of chronic health issues like diabetes and cardiovascular disease. In the current scenario, sleep disorders are commonly detected among people throughout the world. Millions of people are affected with sleep disorders irrespective of demographic conditions and age groups. Though sleep disorders are extensively spread, their impacts often remain under-treated and under-diagnosed, emphasizing the prominence of increasing awareness, constructing efficient interventions, and promoting early prediction to manipulate complicated health challenges. Sleep chaos poses critical challenges to the health condition of common people having adverse effects on the quality of life and wellness of an individual. Accurate identification and prognosis of the illness are important for constructive treatment and management of the disease.

Significance of Machine Learning in Predicting Sleep Disorder:
ML has become a major tool in predicting sleep disorder using complex data's, pattern matching by both supervised and unsupervised learning. This will help in predicting sleep disorder with high accuracy.

2 Literature Survey

A sleep-related illness causing complete or partial termination of breathing during sleep in an individual is termed Obstructive Sleep Apnea (OSA). Different methodologies are suggested to detect the occurrence of OSA automatically, yet only limited work concentrates on the prediction of OSA in advance. The discussed disease when identified at an early stage aids in developing specialized devices to govern the breathing process of an individual during sleep. Eileen Wang et al., 2023 suggested four approaches for the prediction of sleep apnea with the use of convolutional and LSTM. The proposed models utilize raw data obtained from three breathing indications namely, abdominal without

including any artificially-engineering attributes. The study predicted OSA (hypopnea or apnea) and common breathing patterns, thirty seconds in advance with the use of data from the prior ninety seconds. The study demonstrated that various models produced an accurateness of 81%. The outcomes demonstrate that OSA events are recognized in advance using respiratory signals. It opens up various chances for the evolution of equipment to anticipatory governs the flow of air into a person during sleep for preventing the OSA event. In addition, the study renders good performance in the prediction of the event even when there are sampled respiratory signals. The performance of the model towards low sampling frequencies permits low storage space in the devices, making it efficient for home environments [5].

One of the natural processes of human beings to relax the mind and physical body is sleep. When there is a disturbance in sleep activity, it results in lower productivity and fatigue. To investigate the condition of sleep illness in an asthma unit, Dinh-Van Phan et al., 2021 tested one million samples from NHIRD. The sickness records of the asthma sufferers were converted into matrices and sequences for the identification of sleep illness through the implementation of various ML approaches and DL55 algorithms. The CNN model turned out with best performance in recognizing sleep disorders among asthma units [1].

Sleep disorders are analyzed with the help of laboratory-based polysomnography (PSG) signals. Using the automatic monitoring procedure of sleep relegation phases, the identification of neurological sicknesses was efficient and perfect. In this context, G. Naveen Sundar et al., 2021 have designed an automatic flexible DL algorithm and ML model for employing raw electroencephalogram (EEG) signals. A DeepCNN was utilized from different DL algorithms to investigate the invariant time frequency actualities and time capabilities and gather estimated modifications. In addition to capturing the degree of sleep phase relegation, the model also acquires the protected and extended brief length establishing conditions among the epochs. An innovative function was utilized in the method to estimate the data loss and during the training process; the network was identified with misclassified errors in the sleep stage. The method was utilized in conjunction with ML modules for forecasting the best applicable approach for the requirement. The effectiveness of the method was established using two open sources which are available publicly from PhysioNet. The suggested technique was found more effective than existing DL frameworks. The models' comparative analysis was used in a clinical setting and adjusted by specifications [6].

A disease caused due to improper sleep patterns is a sleep disorder. The impacts caused by sleep disorders are classified into two, namely physical and mental issues. It creates various problems and difficulties like sleep-wake disorders, anxiety, distress in daily life, and several more issues. The main aim of Raed Alazaidah et al., 2023 was to make use of the strong advantages of ML in the detection of sleep disorders. Specifically, the study was organized to encounter three ultimate goals. These goals were to recognize the best learning strategy, the preeminent classification algorithm, and the greatest regression approach which highly adapted nap illness details. Taking into consideration the two correlated databases and various evaluation metrics that are associated with the work of classification and regression, the outcomes open up the superiority of the KStar, SMOreg, and MLP (Multilayer Perceptron) algorithms when compared with the other

existing twenty-three models for regression. Moreover, RF, IBK, and Randomizable Filtered Classifier came out with superior performance when juxtaposed with other models for classification which pertained to numerous learning strategies. At last, the learning policy demonstrated the best prognostic execution between the six contemplated approaches in the given datasets [2].

In the modern world, the most common disorder identified among all age people was sleep apnea. It has become a general issue for patients who endure sleep disturbances. In association with it, Hnin Thiri Chaw et al., 2023 have suggested a deep CNN algorithm concerning oxygen saturation (SpO_2) signal received using the smart sensor. Hence, ZleepNet was called the system for sleep apnea prediction. The suggested algorithm involves three convolutional layers that include two dense layers, a ReLu activation function. This system has reduced the usage of sensors to detect the signals of sleep apnea from twenty-five to one. Experiments were carried out to estimate the performance of the suggested CNN based on a dataset obtained from real patients and the same was juxtaposed with conventional ML methods like SVM, LDA, ANN, and baggy representation tree over the publicly existing sleep databases with a similar parameter condition. The output demonstrates that the recommended algorithm performs better than other models showing an accuracy of 91.30%. When compared with LDA, the CNN model turned out with 90.33% accuracy, and over SVM, it produced 91.56% accuracy. The rate of accuracy of CNN was 91.89% when set with a bagging representation tree and with ANN, the accuracy was 91.30% [7].

Research Gap:
From the literature survey analyzed the following research gap has been determined:

- Less accuracy.
- Improper precision and recall.
- Complex model and not user friendly.
- Cannot be implemented in all datasets.

3 Proposed Model

The workflow for sleep disorder classification involves several steps to leverage CNN along with preprocessing techniques like Interquartile Range (IQR) and LASSO regularization.

3.1 Preprocessing

IQR
Initially, the dataset undergoes preprocessing where features are extracted using the IQR method to handle outliers effectively. While dealing with sleep disorder databases, IQR preprocessing plays a pivotal role in improving data quality, enhancing model performance, and accelerating more accurate detection. Often, sleep disorder databases include different kinds of data, including sleep patterns, medical history, demographic information, and physiological measurements. Still, these datasets suffer from outliers,

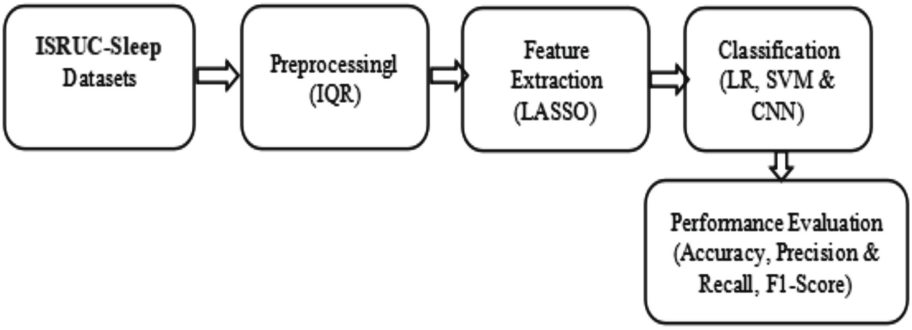


Fig. 1. Structure of the Proposed System

inconsistencies, and noise that have adverse effects on the reliability of the predictive approaches.

The process begins by sorting the data and calculating the First Quartile (QR1) and Third Quartile (QR3), representing the medians of the lower and upper halves of the dataset, respectively. Subsequently, the IQR is computed by subtracting QR1 from QR3.

$$\mathrm{IQR} = \mathrm{QR3} - \mathrm{QR1} \tag{1}$$

Here

$$\mathrm{QR1} = \mathrm{I\,Quartile}$$

$$\mathrm{QR3} = \mathrm{III\,Quartile}$$

Through effective identification and detaching of outliers, IQR preprocessing ensures the accuracy and strength of the prediction algorithms. It also assures and avoids the unduly influence of measurement errors or extreme values over the analysis. A constant stability was acquired among data retention and outlier deletion through IQR preprocessing, permitting for the conservation of significant traits and patterns during the data cleaning process. It ultimately directs to enhance model performance with more accurate predictions in managing and detecting sleep disorders.

3.2 Feature Extraction

Lasso

Lasso attribute extraction gains notable importance in the classification of sleep disorder as it can improve predictive algorithms by choosing the most related attributes and subsequently removes dimensionality. Often, sleep disorder databases include enormous attributes, inclusive of sleep patterns, medical history, demographic information, and physiological measurements. Lasso attribute extraction enhances the model's interpretability by recognizing the key predictors linked with sleep-related illness. It also eases a better perception of the prime physiological mechanisms and associated risk parameters.

The LASSO regularization process can be represented mathematically using the following equation:

$$\text{minimize}\left(\|y - X\beta\|_2^2 + \lambda[[\beta]]_1\right) \tag{2}$$

Here

y denotes the vector value of the target data.

X denotes the feature matrix

β representing the coefficient vector

$\|.\|_2^2$ describing the squared L_2 norm, defining the residual sum of square values.

$\|.\|_1$ defining the L_1 norm, encouraging sparsity in the coefficient vector.

λ illustrates the regularization argument used to control the strength of regularization.

Features extracted through IQR are further refined using LASSO regularization to select the most relevant features. This refined feature set is then fed into a CNN architecture for classification.

3.3 Classification

3.3.1 CNN (Convolution Neural Network)

Sleep illness classification includes the investigation of multi-dimensional and complicated data. From researchers, it was found that CNNs suit well for the above task well as it is capable of effectively capturing temporal and spatial patterns intrinsic in these kinds of data using hierarchical attribute learning. CNNs can derive hierarchical representations of attributes from multiple phases of abstraction using leveraging pooling layers, non-linear activation functions, and convolutional layers. Consequently, it perceives subtle designs and relationships related to sleep illness categorization.

One way to express the 1D convolution is as:

$$x_k = b_k + \sum_{i=1}^{N} w_k \times y_i \tag{3}$$

where y_i denotes the ith feature map, x_k is the k^{th} level feature map, b_k describes the bias of the k^{th} feature map, and wk represents the kth convolutional kernel value from every feature of the k^{th} feature map (Erdenebayar Urtnasan et al., 2021).

A crucial role was played by CNNs in the procedure of classification of sleep disorder by rendering adaptive and powerful tools for deriving meaningful attributes from complicated data. Hence, it enhances the efficiency and accuracy of diagnostic procedures and smoothens more personalized medication strategies for persons infected with sleep disorders.

3.3.2 LR (Logistic Regression)

LR works as a useful method for classifying sleep disorders because of its ease of use, readability, and efficiency in simulating multiclass or binary outcomes. LR is implemented in the classification of sleep disorders to detect the absence or presence of

particular sleep illnesses concerning the set of associated attributes extracted from physiological dimensions, demographic information, sleep patterns, and medical history. By calculating the likelihood that a person will fall into a specific class (e.g., absence or presence of insomnia, sleep apnea, etc.), LR renders an interpretable and clear knowledge of the correlation between the feed attributes and the probability of getting sleep illness. The threat variables for nap problems in affected people, comprisingmedical, and medical experiment characteristics, were investigated using binary LR analysis.

The possibility that a given instance goes to a specific class is predicted using the logistic regression model. It is expressed mathematically as follows and is derived from the logistic method, also called the sigmoid variable:

$$P(y = 1|x) = \frac{1}{1 + e^{-(\beta_0+\beta_1 x_1+\beta_2 x_2+.....+\beta_n x_n)}} \tag{4}$$

where:

$P(y = 1|x)$ denotes the possibility that the target identifier y equals 1 specified the features x.

e identifies the base of the natural logarithm.

$\beta 0,\beta 1,\ldots,\beta n$ are the constants connected with each feature $x_1,x_2,\ldots,x_n$.

$x_1,x_2,\ldots,x_n$ are the features.

n is the number of features.

The logistic function can be used to model probabilities since it transfers the linear combination of the characteristics and coefficients onto the range [0, 1]. Two-way stepwise regression analysis was used to create a prediction model based on the risk variables (Rongpeng Xu et al., 2022).

3.3.3 SVM (Support Vector Machine)

In sleep disorder classification, SVM was applied to differentiate among various classes of sleep disorders according to the attributes derived from sleep patterns, physiological signals, and other corresponding parameters. SVMs perform through sorting the hyperplane which segregates the data points in the best way into several classes along with maximization of margin in terms of distance between nearest data points and hyperplane. When margin maximization is done, it strengthens the classification algorithm which is less possessive to outliers and noise existing in the data.

The following equation, which defines y(a) hyperplane, is the generic SVM formula. Here, w stands for the vector, a for the input type vector, and b for the bias factor.

$$y(a) = w.a + b \tag{5}$$

The binary SVM rulescan be used to categorize four different sleep stage circumstances (Edita Rosana Widasari et al., 2020). Nevertheless, SVMs work with nonlinear relationships among attributes and class labels with the use of kernel functions that map the feed-in attributes into a higher-dimensional space in which the data are separated linearly.

4 Result and Discussion

The aim of this model is o enhance the F1-score, precision, and accuracy of sleep disorder detection by leveraging the mixture of CNN with IQR preprocessing along with Lasso attribute extraction methods. The study juxtaposed the performance of three models, namely SVM, LR, and CNN, both with or without the application of Lasso and IQR preprocessing.

4.1 Dataset Description

ISRUC-Sleep is a PSG dataset, where the data acquired from human adults are stored. The data indulges subjects with good health and subjects affected with sleep illnesses under the impact of sleep treatment. The dataset is designed to assist several research objectives, involving three clusters of data. The first group includes data with 100 subjects having one footage session per topic. In the second group, data was gathered from eight various subjects where two footage sessions for each topic that benefitted the study containing the modification in the PSG signals over a period. In the third group, data was obtained from one recording session concerning ten healthy subjects that are advantageous for the study to compare the healthy subjects with the subjects affected with sleep disorders [8].

In sleep disorder prediction models, performance assessment scales like accuracy rate, precision value, recall value, and F1 score data are usually used to assess the efficacy of the model in classifying instances correctly.

4.2 Performance Evaluation

4.2.1 Accuracy Analysis

In the prediction of sleep disorder, the term accuracy refers to the proportion of rightly classifying the cases (e.g., the person affected with or without a particular sleep disorder) to the entire amount of cases available in the given dataset. It is very important to achieve high accuracy in the estimation of sleep complaints to ensure effective treatment planning and reliable diagnosis.

$$\text{Accuracy} = \frac{\text{Number of Correctly Predictions}}{\text{Number of Predictions}} \times 100 \tag{6}$$

Table 1. Accuracy Comparison of classifiers

Algorithms	Accuracy (%)
Logistic Regression	85.60
Support Vector Machine	88.10
Convolution Neural Network	90.30

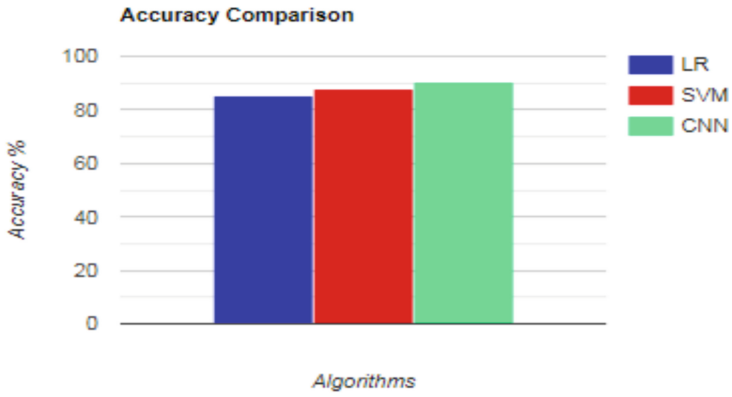


Fig. 2. Accuracy Value of Proposed CNN Graph

The Table 1 & Fig. 2 represents accuracy evaluations of recommended CNN. Outcome show that accuracy of recommended CNN is 90.30% which is more compared to SVM is 88.10% and LR is 85.60%.

4.2.2 Precision Analysis

While considering the prediction of sleep disorder, precision in this context is the percentage of precisely diagnosed cases (true positives) of a given sleep disorder out of all cases that were projected to have that illness (false positives + true positives). Precision renders an overview of the capacity of the model to flawlessly categorize individuals with specific disorders, reducing false positives and guaranteeing that positive detections are dependable.

$$\text{Precision} = \frac{\text{True Positive Cases}}{\text{True Positive Cases} + \text{False Positive Cases}} \tag{7}$$

Table 2. Precision Value Vs ML classifiers

Algorithms	Precision
Logistic Regression	0.84
Support Vector Machine	0.87
Convolution Neural Network	0.89

The above Table 2 & Fig. 3 represents precision vale assessment of proposed CNN. Outcomes show that precision values of recommended CNN are 0.89 which is more than SVM is 0.87 and LR is 0.84.

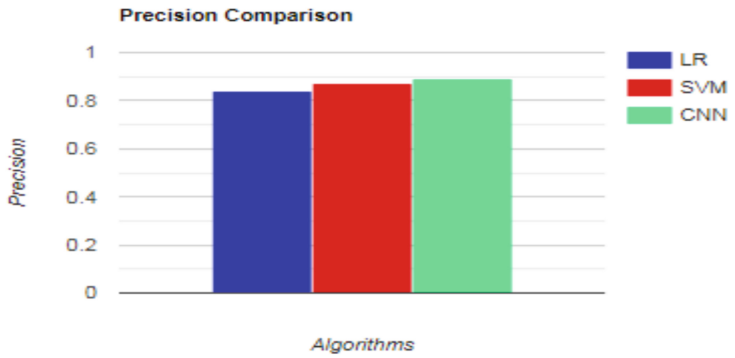


Fig. 3. Precision Comparison of Proposed CNN Graph

4.2.3 Recall Analysis

Recall term also called as TP rate or sensitivity, calculates the fraction of real positive instances (i.e. people with sleep disorder) to that of the rightly recognized through the algorithm. In the concept of sleep disorder detection, recall measures the capability of the system to acquire all associated instances with certain sleep disorders, removing false negatives and ensuring that people with the above condition are recognized properly.

$$\text{Recall} = \frac{\text{True Positive Cases}}{\text{True Positive Cases} + \text{False Negative Cases}} \tag{8}$$

Table 3. Recall Values Vs ML classifiers

Algorithms	Recall
Logistic Regression	0.82
Support Vector Machine	0.85
Convolution Neural Network	0.87

The Table 3 & Fig. 4 illustrates recall values comparisons of suggested CNN. Outcomes show that recall value of recommended CNN is 0.87 which is more compared to SVM is 0.85 and LR is 0.82.

4.2.4 F1-ScoreAnalysis

In the context of sleep disorder prediction, the F1-score works as a only metric to combine both recall and precision. It produces a balanced calculation of the model's execution. The score is specifically implemented in conditions where an imbalance among the classes exists or during false negatives and false positives with different consequences.

$$\text{F1} = \frac{2 \times \text{Precision Rate} \times \text{Recall Value}}{\text{Precision Rate} + \text{Recall Value}} \tag{9}$$

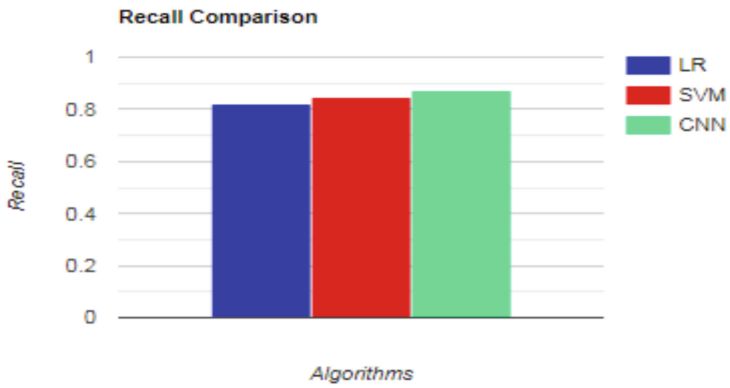


Fig. 4. Recall Value Comparison of Recommended CNN

Table 4. F1-ScoreVs ML Classifiers

Algorithms	F1-Score
Logistic Regression	0.85
Support Vector Machine	0.88
Convolution Neural Network	0.90

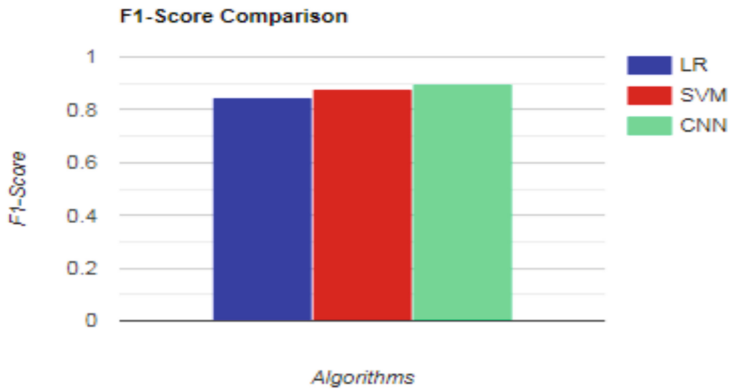


Fig. 5. F1-Score Comparison of Proposed CNN Graph

Table 4 & Fig. 5 describesF1-Scoreof proposed CNN with other two models. Outcome show that F1-Score of suggested CNN is 0.90 which is more compared to SVM is 0.88 and LR is 0.85.

The results based on the experiments showed that the combination of CNN along with Lasso and IQR preprocessing constantly outperforms individual algorithms in terms of F1 score, precision, and accuracy. Particularly, the CNN algorithm associated with Lasso and IQR preprocessing delivered the highest accuracy.

5 Conclusion and Future Work

In conclusion, findings from this proposed system have emphasized the significance of preprocessing methods like Lasso attribute extraction and IQR in enhancing the efficiency of sleep disorder predictive algorithms. Through integrating these methods with CNN, we accomplished superior outputs based on F1-score, precision, and accuracy, underlining the capability of upgraded ML algorithms for maximizing diagnostic accuracy and customized treatment methodologies in sleep medicine. In the future, research has an opening to explore additional preprocessing techniques and model frameworks to further maximize predictive performance and establish the results to more diverse and larger datasets. Finally, the performance of this combined model is compared against the individual classifiers, focusing on improvements in accuracy, precision, recall, and F1 score. This comprehensive workflow ensures an effective approach to sleep disorder classification, leveraging both traditional ML methods and DL techniques for enhanced performance. In future we can go with more deep learning algorithms such as CNN, LSTM etc. which will give better results than suggested model.

References

1. Phan, D.-V., Yang, N.-P., Kuo, C.-Y., Chan, C.-L.: Deep learning approaches for sleep disorder prediction in an asthma cohort. J Asthma **58**(7), 903–911 (2021)
2. Alazaidah, R., Samara, G., Aljaidi, M., Qasem, M.H., Alsarhan, A., Alshammari, M.: Potential of machine learning for predicting sleep disorders: a comprehensive analysis of regression and classification models. Diagnostics **14**(1) (2024)
3. Zhang, M.-M., et al.: Sleep disorders and non-sleep circadian disorders predict depression: a systematic review and meta-analysis of longitudinal studies. Science Direct. Neurosci. Biobehav. Rev. **134** (2024)
4. Vorvolakos, T., et al.: The association between sleep pathology and depression: a cross-sectional study among adults in Greece. Psychiatry Res. **294** (2020)
5. Wang, E., Koprinska, I., Jeffries, B.: Sleep apnea prediction using deep learning. IEEE J. Biomed. Health Inform. **27**(11), 5644–5654 (2023)
6. Naveen Sundar, G., Narmadha, D., Amir AntonJone, A., Martin Sagayam, K., Dang, H., Pomplun, M.: Automated sleep stage classification in sleep apnoea using convolutional neural networks. Sci. Direct Inf. Med. **26**, 100724 (2021)
7. Chaw, H.T., Kamolphiwong, T., Kamolphiwong, S., Tawaranurak, K., Wongtanawijit, R.: ZleepNet: a deep convolutional neural network model for predicting sleep apnea using SpO2 signal. Hindawi Appl. Comput. Intell. Comput. **8888004**, 1–12 (2023)
8. https://paperswithcode.com/dataset/isruc-sleep
9. Urtnasan, E., Joo, E.Y., Lee, K.H.: AI-enabled algorithm for automatic classification of sleep disorders based on single-lead electrocardiogram. Diagnostics (Basel) **11**(11), 2054 (2021)
10. Xu, R., et al.: Risk factors and prediction model of sleep disturbance in patients with maintenance hemodialysis: a single-center study. Front. Neurol. **13**, 955352 (2022)
11. Widasari, E.R., Tanno, K., Tamura, H.: Automatic sleep disorders classification using ensemble of bagged tree based on sleep quality features. Electronics **9**(3) (2020). https://doi.org/10.3390/electronics9030512

A Survey on Various Sentiment Analysis Applications Employing Opinion Mining

Meghna Bhatt(✉), Sridaran Rajagopal, and Kinjal Raval

Marwadi University, Rajkot, India
meghna.bhatt@marwadieducation.edu.in

Abstract. Opinion mining aka sentiment analysis refers to the use of natural language processing, text analysis and computational linguistics to identify and extract subjective information in source data. This, further helps to extract the subjective information into various categories that help in a decision-making process. In this survey, a novel attempt is made to provide an application-specific classification on opinion mining and sentiment analysis. The survey presented here widely covers the processes, levels and challenges of opinion mining that would help the application designers of opinion mining applications.

Keywords: Opinion Mining · Sentiment Analysis · Data Mining · Web Mining · Sentiment Classification

1 Introduction

With ever-increasing usage of internet and other technologies, there is a huge amount data being created or handled in most of the applications. This voluminous data at times creates difficulties in enhancing relevant information required in a processing step. Hence, in order to deal with large data handling many automatic mining techniques in various forms were invented.

Data Mining (DM) is defined as the process of analysing large databases, usually data ware houses or internet, to discover new information, hidden patterns and behaviours [1]. Data mining involves the use of various techniques such as machine learning, statistical analysis, and database management to extract insights and information from structure data. Many diverse areas use data mining methods for technical, commercial and research purposes. Some of the applications of data mining are Bioinformatics, Financial Banking, Education, Criminal Investigation, Market Basket Analysis, Future Health Care, Manufacturing Engineering, Health Care, banking [2, 3].

Opinion mining also known as Sentiment analysis refers to the use of natural language processing, text analysis and computational linguistics to identify and extract subjective information in source materials. [70]. Opinion mining focuses on identifying and extracting subjective information, such as opinions, emotions, and attitudes about entities, products or services aims to determine the polarity of this information, categorizing it as positive, negative, or neutral.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
S. Rajagopal et al. (Eds.): ASCIS 2024, CCIS 2428, pp. 69–94, 2025.
https://doi.org/10.1007/978-3-031-86302-8_5

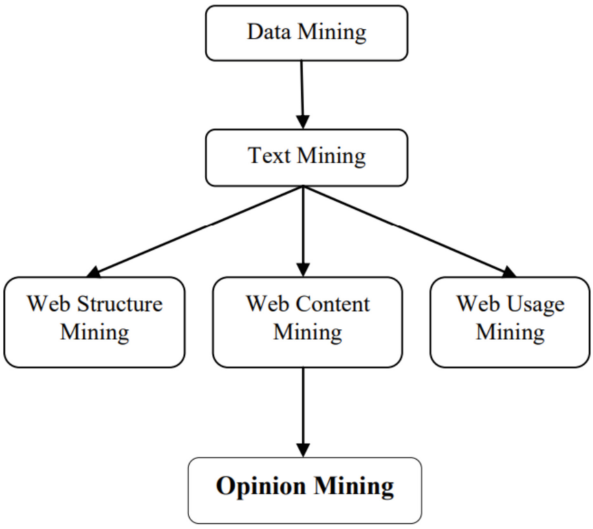


Fig. 1. Opinion Mining Taxonomy

Above figure Fig. 1 explain the Opinion Mining Taxonomy that explains proper Structure of the Data Mining, text mining, classifications in text mining are Web Structure mining, Web content mining and Web Usage mining, Opinion Mining [4, 5, 6].

Sentiment Analysis (SA) involves mining information related to sentiments or opinions from a group regarding a specific topic. Sentimental Analysis is a technique used to study people's opinions, sentiments, evaluations, appraisals, attitudes, and emotions expressed in an article, issues, events, and topics. Using this technique to determine the subjectivity of a given text and, if it is subjective, to further classify it as either negative or positive. Opinion mining techniques utilize *natural language processing* (NLP) and machine-learning algorithms to understand public opinion, monitor brand reputation, and make informed decisions based on user-generated data. These techniques help businesses, governments, and individuals monitor brand reputation and make informed decisions.

A. Opinion Mining And Sentimental Analysis Process:

Figure 2 shows the functional diagram of opinion mining [7]. A number of important steps make up the opinion mining process. The first step is **data collection**, which means getting information from places like social media, online reviews, and customer service chats to figure out how people feel. The next step is **opinion identification**, which means finding these views and putting them into different groups. Then there is **aspect extraction**, which pulls out specific features about a product or service from user reviews to learn more about how customers feel. After that, **opinion classification** uses a variety of methods and algorithms to arrange the views that have been found. In this step, different **classification methods** used to make the process better. Finally, the **evaluation** step checks the sentiment analysis model's accuracy, precision, and memory to see how well it works and to make sure it is reliable and useful.

A. Opinion Mining And Sentimental Analysis Process:

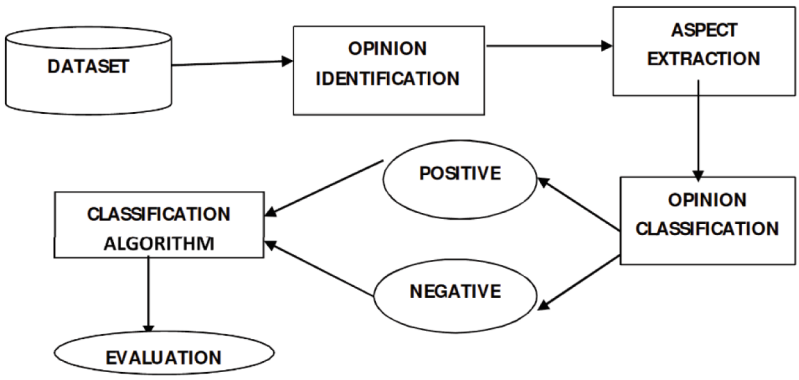


Fig. 2. Functional Diagram of Opinion mining

B. Opinion Mining and Sentimental Analysis Levels:

Opinion mining and sentimental analysis investigated on document level, sentence level and aspect level. As compare to document level & sentence level the aspect level more difficult because it performs a fine-grained investigation. A brief introduction of levels describes in this figure Fig. 3 [8].

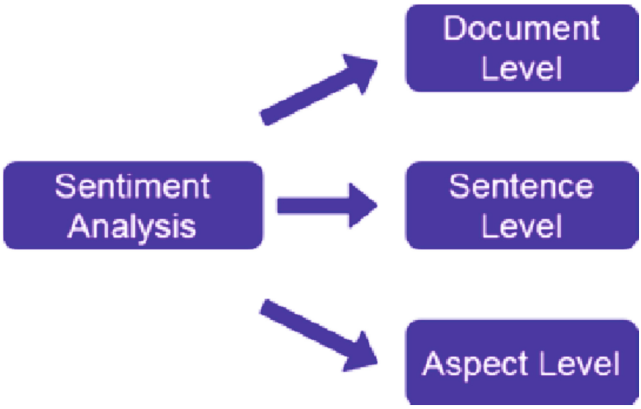


Fig. 3. Sentimental analysis level

i. Document level

The task at this level is to determine the overall opinion of the document. Each document classified based on the overall sentiment of the opinion holder about a single entity [9]. Thus, it is not suitable for documents that analyze or compare more than one entity. The goal is to classify a review as positive, negative, or neutral.

ii. Sentence level

Sentence-level sentiment analysis checks each sentence's positive, negative, or neutral tone. Neutral statements convey facts, whereas subjective ones provide views. This approach is most useful for identifying subjective or objective statements.

iii. Aspect level

Aspect-based sentiment analysis examines whether the specific component is positively or negatively mentioned [9]. Aspect-based sentiment analysis is a focused approach to sentiment analysis that aims to identify the emotions related to particular aspects or features of a product, service, or entity in written text. This is different from simply assessing the overall sentiment of a sentence or document.

C. Applications of Opinion Mining

Opinion mining is a powerful tool used in various industries, including product reviews, marketing, healthcare, and crime analysis. It helps users determine product ratings, assess customer preferences, and evaluate campaign success. It also helps identify competitors and assess marketing strategies. Its applications extend to tracking trends, managing reputations, and responding to feedback across various industries [10, 11, 71].

D. Challenges of opinion mining and sentimental analysis

There are several challenges of sentimental analysis that can affect its accuracy. Negation is one of the challenges that can alter the sentiment of a sentence, yet some methods mistakenly ignore negation words, leading to misinterpretations. Unclear language compounds the issue, especially when systems lack context or struggle with irony, sarcasm, and pronoun resolution. Additionally, Sentiment without words if very difficult to making positive or negative feedback and Sarcastic sentences, which can distort meaning and affect the overall value of a document, also require careful attention. Fake opinions or reviews, intended to mislead, undermine the reliability of sentiment analysis and pose a significant social challenge. Domain dependency is one of the challenges that adds complexity, as the relevance of sentiment varies between domains. Apart from Non-expertise, opinions from forums and blogs complicate data interpretation, while idioms, contextual nuances, and evolving language on the internet introduce ambiguity and variability [9, 13–15, 18].

In This paper Sect. 2 related work, Sect. 3 contains the classification of the applications of onion mining, and Sect. 4 contains discussions and Sect. 5 concludes the work.

2 Related Work

In order to construct a fully automated system for interpreting human emotions and feelings, a survey that addresses the challenges of opinion mining and Sentimental analysis conducted Raisa Varghese [16]. Author has reviewed and summarized relevant works in this field, they have also investigated issues such as name entity recognition, conference resolution, and domain dependency, which make it difficult to produce a comprehensive solution. However, application-specific search is still outstanding and will be subject to future study.

Mrs. Smitha et al. [17] gives a comprehensive review of sentiment analysis, which focusses specifically on how media and user-generated material influence people's opinions and behaviours. The author provides a summary of the process that is involved in sentiment analysis, as well as a brief overview of techniques and a comparative analysis of classifier algorithms based on their level of accuracy. The survey concluded by stating that the challenges that encountered in big data applications left for future work.

Apoorva Arya et al. [69] presents a review of techniques and methods in sentiment analysis with their challenges. Authors highlighting the Importance of Sentiment Analysis, review of sentiment analysis methods with algorithm performance comparison. In this paper author focus on both producer and consumers in reviewing feedback by using machine learning techniques. Proposed survey presented a comparative study of two techniques *Naive Bayes* and *SVM* only.

A review of emotional analysis and opinion mining methodologies, classification, tools that are available for sentiment analysis, and a new feature known as Product Aspect Ranking are all included in the survey study that presented by Saud Alhojelu [18]. Only the product-based aspects are taken into consideration in this poll.

Vidhisha pradhan et al. [19] made a literature review on mood analysis algorithms for opinion mining. They looked at different algorithms used in this field and talked about the problems they face and how they can be used. The writers looked at supervised techniques and dictionary-based approaches and concluded that supervised learning methods work better. Further study on alternative methods for summarizing opinions to gain a clearer understanding of sentiment classification.

3 Survey on Opinion Mining Applications

Opinion mining and sentimental analysis is burgeoning field within Natural Language Processing that primary objective to comprehend & categorized the sentiments expressed in textual data to extract valuable insights. These technical approaches enable organizations to make data-driven decisions, improve services, and enhance user experiences based on the sentiments expressed in textual content. Based on an extensive literature survey done by me, the classification of applications of opinion mining show in figure Fig. 4.

The applications of opinion mining depicted in the diagram are organized into distinct categories, each representing a specific sector or domain where sentiment analysis is widely utilized. These categories include education, hospitality, customer care & service, finance & banking, fraud detection, social media & other.

3.1 Education

Opinion mining/sentimental analysis has emerged as a valuable tool for educators and educational institutions. By leveraging natural language processing techniques, sentiment analysis enables educators to gain critical insights into student feedback, engagement, and learning experiences. The application of this technology not only improves the quality of education delivery but also informs strategic decision-making. Based on my literature survey I further classified applications such as feedback assessment, performance analysis, auto question generation & curriculum evaluation.

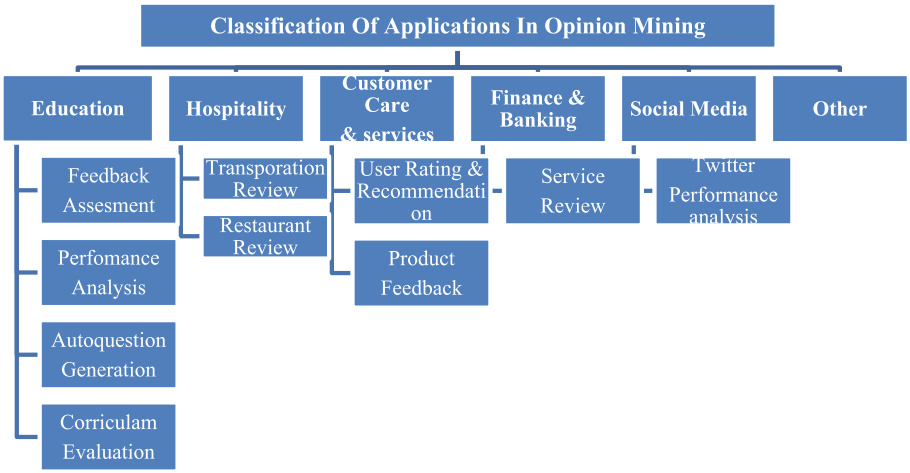


Fig. 4. Classification of opinion mining applications

3.1.1 Feedback Assessment

The study conducted by Cong Qi and Shudong Liu [20] presents a noteworthy contribution in evaluating MOOC courses through review mining. The research focuses on extracting student's concerns from their reviews, providing valuable insights for MOOC managers and universities. Author proposed model First, gather all MOOC comment text data then use short-sentence *linear discriminant analysis* (*LDA*) model to extract subject information from all reviews; then, utilising auto-encoder and *Bi-LSTM* model to classify review sentiment. After that, calculate emotional values for various subjects in course remarks for evaluation. Finally, the emotional values of all specific student comments on different topics are calculated and matched with course scores, with higher-valued courses chosen for students' reference. By utilizing *LDA and Bi-LSTM*, they effectively enhance the platform infrastructure and service experience. However, it is important to highlight that the comment sentiment analysis falls short in considering the impact of instructor performance. Additionally, the analysis only takes into account the positive and negative dimensions, disregarding the neutral dimensions.

D. Buenano-Fernández et al. [21] explores text-mining techniques for analyzing student interactions in online MOOCs, highlighting the challenges of maintaining student commitment and motivation, and the impact of globalization on education. The research primarily concentrates on calculating and analyzing term frequencies, examining concordances, and grouping n-grams. One of the disadvantages is that it presupposes emotions conveyed through words have a direct effect on our learning. It ignores other factors like individual differences and teaching methods, which also have an impact on learning. Therefore, while examining how emotions affect learning is crucial, we must take into account all these factors for a comprehensive understanding.

Kastrati, Z. et al. [22] address the challenge of manually analyzing student's feedback in large-scale educational settings, particularly in online courses like MOOCs where author proposed framework for aspect-level sentiment analysis in student review that

reduce need for manually annotated data in sentiment analysis techniques. Proposed framework first, manually tagged reviews or keywords by category. Manually annotated user reviews are pre-processed using TF-inverse document frequency weighting to extract relevant terms. Labelled reviews are used to create a CNN model for deep learning in the aspect category learning module. Next, CNN and LSTM deep networks will check aspect polarity. However, this framework not tested with no-educational data.

Student feedback is essential for faculties as it provides valuable insights into teaching effectiveness and student satisfaction. It helps faculties identify areas for improvement, adjust teaching methods, and enhance the learning experience. One of the researches done by I. Sindhu et al. [23] discusses the student's opinion for faculty teaching performance evaluation where author address the challenges in processing qualitative feedback to improve the accuracy and fairness of academic performance assessments. To tackle this, the author proposes two *Long Short-Term Memory – (LSTM)* systems for automating aspect extraction and polarity detection. Proposed model creates academic area and prepared datasets first. Author used skip-gram model for domain word embeddings to understand academic word relationships. Finally, we discuss the two-layer LSTM neural network that classifies mood-related aspects. Layer 1 includes pedagogy, knowledge, behaviour, measurement, experience, and general. The second part of our model predicts positive, negative, or neutral for each extracted aspect. The proposed model correctly identified aspects 91% and emotion direction 93%. However, a limitation of this model is that comments may be misclassified when multiple aspects are mentioned in a review sentence without connectives like 'AND' and 'BUT'.

One of the studies related with social computing feedback assessment made by Nisha Tanwan et al. [24] that presents a solution for extracting and analyzing comments made by master's students in the Facebook academic group. The proposed method utilizes the Facebook graph API to extract the comments and categorized by using a *Bayesian Network* probabilistic model. This system has the potential to assist the administration in enhancing the learning environment by offering various forms and analyzing diverse student opinions. During study come to know that research does not include significant particulars concerning the comments, such as the specific topics under discussion or the backgrounds of the students who made the comments. Without this contextual information, accurately understanding the sentiments expressed could be problematic. Classify and analyse text/comments at various levels, including document and comparative levels are left for future research.

Sentiment analysis in education using NLP and POS tagging techniques by Nandakumar R. et al. [25] focuses on student feedback analysis for teaching system improvement using NLP where Machine learning-based sentiment classification with real-time dataset is done and Tokenization and *POS* tagging method used for text pre-processing. The study on student feelings using computer tools faces challenges due to its small dataset, manual labelling, and lack of comparison to other methods. It also only looks at one school, making it difficult to determine its effectiveness. To enhance its effectiveness, larger datasets, standardized labelling, and comparative analysis recommended.

To improve teaching efficiency, Mr. D. Selvapandian et al. [26] has introduced an innovative fusion-based neural network (EF-NN) classifier for sentiment analysis on student feedback data. The primary aim of this study is to enhance teaching quality

by analysing both positive and negative feedback, ultimately refining training methods. Author proposed EF-NN model, which merges support vector machine and convolutional neural network, specifically crafted to predict common context patterns in student feedback. By conducting dataset processing, attribute selection, and data transformation, the proposed model ensures accurate classification of sentimental data. The performance of proposed model compares with k-means clustering that achieved 93% accuracy, 89% precision, and 91% recall. There are some issues with the study, such as not having enough diverse datasets, not being able to correctly identify sentiment, and not having any means to obtain immediate feedback.

An evaluation method for teaching and learning sessions, to provide feedback to tutors and recommend areas for improvement. This could enhance educational outcomes and optimize teaching strategies. To address this Karunya K. et al. [27] define automated recommendation systems have the potential to streamline this process and offer actionable insights in real-time. Implementing such a system could greatly benefit both tutors and learners by facilitating continuous improvement and personalized learning experiences. The document is first converted to lower case for easier pre-processing, followed by tokenization, which breaks the text into tokens for matrix generation. A bag-of-words model is created by splitting the document into individual words, while stop words—frequent but irrelevant terms—are removed as they do not contribute meaningful information. Lemmatization is then applied to reduce words to their base or dictionary forms. Author used Latent Dirichlet Allocation (LDA) for topic modelling that identify underlying topics within a set of documents from which opinion is extracted. The recommendation system proposed by author Collect tokens for polarity and opinion extraction, apply sentiment analysis to student feedback using a lexicon-based approach, scoring polarity from 3 to +3 and classifying opinions and Generate recommendations for tutors.

3.1.2 Performance Analysis

Performance Analysis is a specialized field that offers athletes and coaches objective information to enhance their understanding of performance. This process based on systematic observation, which provides valid, reliable, and detailed information about performance. It involves the collection, interpretation, and analysis of data to assess the functionality of something and to identify areas that can be improved.

Menna Fateen et al. [28] introduced a transparent student-based vector for forecasting student performance. The primary objective of this research is to examine the subjects covered and psychological traits pinpointed in teachers' daily written reports. Subsequently, these discoveries are integrated into a model for predicting student performance. In this method, first teacher highlights different aspects to summarize the students' development from teachers' observation report and using Topic Rank, the most used key-phrases by the teachers were identified. Comment sentences were scored by a pretrained BERT sentiment model. Author use LIWC software was to extract linguistic and psychological features by calculating the percentage of words used in over 70 categories, providing an output measure for each category.

One more author Muslihah Wook et al. [29] proposed novel method for analysing student feedback on lecturers' teaching performance in Institutions of Higher Learning (IHLs). Authors tackle the issue of extracting valuable insights from the qualitative

section of assessment questionnaires, which is often disregarded due to the unstructured nature of students' opinions. To address this challenge, the authors propose the OMFeedback system, which utilizes an opinion mining technique, specifically a lexicon-based approach, to automatically analyze and interpret students' feedback.

Kingsley Okoye et al. [30] conducted research on the effects of educational coaching programs and mentoring services on users' preferences. Authors created a Data-Driven Support and Decision Model (DSDM) and employed a combination of qualitative and quantitative approaches to collect data from different support facilities. Author's findings identified crucial factors that influence users' decisions and emphasized the significance of services such as tutoring. The objective of this study is to improve educational support services to effectively addressed users' requirements in a competitive setting. The process includes collecting data from users and analyzing it using a mix of quantitative and qualitative methods. The data is organized based on what users are doing and models are used to find patterns and insights. Educational process innovation improves performance and support for users by using insights gained from models or data analysis. This helps enhance the overall experience based on readily available information. The study's limitation lies in its inability to apply the proposed model to various organizational contexts, making it uncertain about its ability to evaluate the impact of different services or training programs.

3.1.3 Auto Text Generation

Automated text generation is the creation of natural language texts via computer systems. It leverages knowledge in computational linguistics and AI, especially in its natural language processing (NLP) subfield, to generate natural language texts with minimal human intervention. It is powered by a method of AI called a neural network, which is a computer system that is designed to mimic the way the human brain operates.

A. Perez-Castro et al. [31] have brought up a significant issue regarding the advancement of Artificial Intelligence, particularly the rise of text generators that closely mimic human writing. The author emphasizes the potential dangers these generators present to e-commerce and consumer confidence, given the prevalence of misleading reviews. By utilizing classifiers trained to differentiate between authentic and deceptive reviews, the research assesses the effectiveness of GPT2, a widely used text generator. The findings suggest that automatically generated deceptive reviews negatively impact the reliability of current systems. The research offers strategies for combating fake review generators through the examination of a single generative model and the development of classification models. However, it is important to note that the study is constrained by limited data sets and an unclear correlation between sentiment and review authenticity.

3.1.4 Curriculum Evaluation

Curriculum evaluation is the methodical assessment of how effective, relevant, and impactful an educational curriculum is. This evaluation process includes defining the curriculum's goals, setting up criteria for evaluation, gathering and examining data, engaging stakeholders, giving feedback, making necessary adjustments, and promoting ongoing enhancement. Through these procedures, educational institutions can guarantee

that their curricula align with their objectives, cater to students' requirements, and adapt to changing educational environments.

V J Chakravarthy et al. [32] designed a framework to analyse student reviews on courses and improve online course selection. The process begins with inputting a dataset of text reviews from Coursera students. After tokenising these reviews, NLP is used for text processing. Deep parsing and deep learning are used for thorough analysis. Customer feedback is analysed for sentiment and polarity using opinion mining. To determine opinion outcomes, polarity counts are used to execute recommendations. The framework compared algorithms and get 94% accuracy.

3.2 Hospitality

Nowadays, the internet plays a crucial role in allowing people to share their experiences and opinions about different products and businesses including hospitality. Hospitality reviews typically cover a wide range of aspects, such as cleanliness, staff friendliness, food quality, room comfort, amenities, and value for money. Analyzing these reviews enables businesses to identify their strengths and weaknesses, gain insights into customer preferences, and make informed decisions to enhance their offerings and improve guest satisfaction. Additionally, positive hospitality reviews can serve as powerful marketing tools, attracting new customers and fostering trust in the brand.

3.2.1 Transportation Review

The use of private transportation in DKI Jakarta is increasing annually, with 21.76 million motor vehicles in the region in 2018 leading to issues of air pollution and congestion. The government launched a public transportation integration program to address these challenges but saw limited success. A study by Fajar Fathur Rachman et al. [33] analyzed public sentiment towards public transportation in Jakarta using sentiment analysis of tweets. Negative sentiment was highest for KRL, while positive sentiment was more common for MRT and LRT. Integrating public sentiment into prediction models could improve accuracy, though it may not directly affect public transportation usage.

João Guerreiro et al. [34] use text mining to find customer recommendations. The rise in peer travel opinions and personal suggestions can help consumers make better choices, but the large number of reviews can be hard to go through. Therefore, author used NLP and SPSS modular for text mining transformation and extraction. Then part of speech used to identify uni-terms or multi-terms, extraction engine labelled them as positive or negative according to their semantic information & classified sentiment categories are used to firm result using binary logistic regression and CHAID algorithm.

Reviews of airlines are essential for guiding consumer choices by providing insights into passenger experiences, which help travellers make informed decisions based on service quality, punctuality, comfort, and overall satisfaction. By analyzing trends in reviews, airlines can identify areas for enhancement, fostering transparency and community engagement within the travel industry. Hyun-Jeong Ban et al. [35] carried out a study that aimed to assess the customer experience and satisfaction levels by analyzing online reviews from airline passengers. Authors collect large number of reviews from

Skytrax. In the analysis of airline passengers' reviews, the initial stage involves extracting keywords through text mining. Subsequently, these keywords are grouped together using the CONCOR analysis approach. Moreover, the investigation performed three consecutive linear regression analyses to explore the connection between evaluation factors, customer satisfaction, and recommendations featured on the customer review site.

Aksh patel et al. [36] also analyse the customers feedback for airline services using *BERT* model. Author proposed a framework gathered from data from Kaggle. Google's *BERT* text classification method and other machine learning approaches ML algorithms *Logistic Regression, SVM, DT, Random Forest*, and Adaboost used by author to evaluate the performance. When results compared BERT delivered top results in accuracy, precision, recall, and F1-score.

3.2.2 Restaurant Review

Xiaobin Zhang et al. [37] analyzed online reviews of Hong Kong Disneyland hotel to determine the main themes reflecting customers' experiences and their connection to customer satisfaction. For this study, Google travel was used to collect relative online customer reviews, including platform selection and content selection. The data collection process conducted by SCTM3 - web crawling program capable of simultaneously collecting and processing data on the web. The second is data analysis of collected data using three procedures. The initial step involves data pre-processing using text mining techniques, where collected sentence data is categorized into single words based on their relative frequency. The study used a semantic network analysis of top frequency words relevant to the research subject, categorizing it as qualitative analysis. Centrality measures and CONCOR analysis were used to assess the significance of these words and explore customer perceptions. Finally, factor analysis and linear regression were used to examine the factors affecting customer experience and their relationship with customer satisfaction.

The challenge of accurately analyzing sentiment in customer reviews, particularly in the Bengali language facing by bangla restaurant. To address this issue Naimul Hossain er al. [38] development of a novel technique by combined *CNN-LSTM* architecture. The data collection done using web application FoodPanda and Shohoz Food. Pre-processing this raw data includes stop word removal, punctuation removal, and contractions. Additionally, the unicode is replaced to ensure data consistency. After pre-processing, a labelled dataset is created, and word embeddings are added to represent words numerically. A combined Convolutional Neural Network (CNN) and Long Short Term Memory (LSTM) mechanism is used to build the sentiment analysis model. This model is then trained on the labelled dataset and evaluated to assess its performance. TensorFlow 2.0.1 used to train model. Finally, the model is used to make predictions on new, unseen data, enabling the system to understand the sentiment of customer feedback. The model's accuracy showing its performance limits and method might not work in other Bengali text analysis situations.

In order to acquire a true impression of a restaurant, the quality of the food, the level of service, the conduct of the staff, etc., people these days rely more on reviews than on advertisements. Elham Asani et al. [39] proposed a context-aware recommender system that collects food preferences from people's comments and then recommends nearby

open restaurants based on those preferences. This proposed model first extract users' preferences from their comments then use sentiment analysis to analyse comments. The author uses semantic method for clustering using three criteria of Wu–Palmer, Gloss, and the combination of these two criteria. It's observed that combination of Wu–Palmer and cosine methods has higher precision, so to cluster the nouns and the similarity criterion of cosine is opted to calculate the similarity between menus and user preferences & suggest nearby restaurants.

3.3 Customer Care and Services

Customer care and services play a vital role in businesses by influencing customer satisfaction, loyalty, and retention. Exceptional customer care is about more than just meeting basic needs. It means understanding customers' preferences, addressing their concerns promptly and politely, and delivering personalized experiences that exceed their expectations. According to my collected literatures, I classified customer care & service-related papers into user rating, product feedback, product recommendations & performance improvements.

3.3.1 User Rating and Recommendation

Sentiment analysis for movie reviews is a relatively new concept, and the examination of movie reviews in the Gujarati language is still an area that requires further research. Parita Shah et al. [40] conducted a study on sentiment analysis for movie reviews in the Gujarati language (MSAGL) using machine learning techniques. The authors proposed a method that utilized TF-IDF for feature extraction from the dataset. The method employed two distinct classifiers, namely multinomial naive Bayes classifier and k-nearest neighbour, to separate the features and perform sentiment analysis. After comparing the results, the authors concluded that the MNB model accurately predicted opinions using TF-IDF features compared to Count Vectorise features.

Another investigation carried out on movie reviews and recommendations, presenting a hybrid recommendation system model by Raghav Mehta et al. [43] where two databases employed to construct the recommender system. The VADER method utilized for sentiment analysis of user reviews. The Hybrid Recommendation System combines Content Based recommendation and Collaborative Filtering. Weighted cosine similarity score incorporated to enhance model accuracy, resulting in better performance compared to other models.

P. Karthika et al. [42] published a paper that provide a comprehensive analysis of flipkart products, aiding online buyers in making informed decisions. Author use random forest for the classification of product. This module has introduced various performance evaluations, such as Real Positive Reviews, Fake Positive Reviews, Real Negative Reviews, and Fake Negative Reviews, along with Performance Metrics. The *Random Forest and Support Vector Machine* algorithms compared for accuracy, with the *Random Forest Algorithm* demonstrating superior performance.

Ahlam Alrehili et al. [43] et al. introduced a model for sentiment classification. This model gathers customer reviews from Amazon and categorizes them as either positive or negative reviews. To enhance accuracy, they plan to employ ensemble learning,

a supervised machine-learning algorithm that combines five classifiers: *Naive Bayes*, *SVM, Random Forest, Bagging, and Boosting*. The ensemble method enhances classification performance by leveraging the strengths of each classifier. To determine the final classification, author classify by majority voting, choosing the label with the most classifier votes. This approach aims to enhance the accuracy of the sentiment classification model. The study found that the random forest technique achieved the highest accuracy 89.87% when using unigram with word removal, while the voting algorithm performed best in other scenarios.

Shigang Hu et al. [44] introduced the CISER model for product recommendations, which focuses on credibility, interest, and sentiment. The model combines sentiment analysis, reviewer credibility analysis, and user interest patterns to improve product recommendations. It consists of six modules: candidate feature extraction, user product review mapping, reviewer credibility analysis, user interest mining, candidate feature sentiment scoring, and ranking and recommendation. The study also suggests using sentiment analysis for product quality assessment and enhancing recommendations by considering reviewer trust, expertise, and influence. The model tested on the Amazon camera dataset and showed better performance than existing methods.

3.3.2 Product Feedback

Product reviews play an essential role in developing customer trust in ecommerce. As they are unbiased and based on the user's experience, these reviews help clear the customers' doubts and make them more confident about purchasing the services provided by e-commerce. To automatically extract sentiment or opinions from these reviews, sentiment analysis employed. Lakshay Bharadwaj et al. [47] established a framework that accurately classify emotions expressed in reviews. In this framework author combines two advance technologies feature engineering and deep learning technique that achieve higher accuracy in sentiment categorization. Feature engineering extract various lexical, system and semantic features from review. Performance metrics used to clarify accuracy and effectiveness.

A study to determine the most effective model for sentiment analysis by Manal Loukili et al. [46]. The purpose of this study was to determine whether customers would recommend a product based on the reviews they read. The researchers utilized a dataset from Amazon, specifically focusing on a women's clothing e-commerce site. They analyzed customer reviews using supervised machine learning techniques, including *Random Forest, Logistic Regression, k-Nearest Neighbor, and Catboost Classifier*. When comparing the classifiers, it found that *Random Forest and k-Nearest Neighbor* had the same accuracy and AUC score.

Pankaj et al. [47] written a paper to analyse customer feedback on amazon product reviews. The reviews are organized into four main groups: mobiles, computers, flash drives, and electronics. Pre-processing, pre-filtering, biassing, and data accuracy were all done on the collected dataset. Part-of-speech tagging is used to group words together. Lastly the analysis of sentiment identifies the opinions of customers regarding products. The paper categorizes sentiments as Positive, Negative, and Neutral.

The majority of consumers influenced by product reviews while making purchases. Syiti Liviani Mahfiz et al. [48] conducted study in which they reviewed products in

the Indonesian beauty product market. Aspects like color, cost, pigment, packaging, styling power, moisture content, and scent extracted from reviews using the *Navie Bayes* classifier. By utilizing various pre-processing methods and examining the impacts on every facet. Performance compared with sampling and without sampling.

Sobia Wassana et al. [49] developed an innovative approach for analyzing product reviews. Rather than focusing on surface-level features, they suggest examining the underlying characteristics of each product. The method involves collecting review and rating data from Amazon, then pre-processing the text using tokenizing and removing stop words - to extract meaningful information using natural language processing. Next, they perform subjectivity and sentiment analysis. Subjectivity analysis identifies subjective statements reflecting personal values, opinions, or sentiments. Sentiment analysis determines the polarity (positive or negative orientation) of subjective statements. The relationships between subjectivity and polarity represented graphically.

Cheng Yang et al. [50] build a product evaluation indicator system that accurately predicts the product's negative reviews & then proposed strategies to improve product. This methodology is particularly useful for products that undergo iterative design processes and have a large number of online user reviews. Proposed method implemented in three phases: first useful review acquisition where web crawler system collects reviews & using *LDA* product attributes and user emotions obtained. Secondly, the establishment of a product evaluation indicator system using multidimensional indicators such as user satisfaction, user attention, and polarity. In the last phase strategy analysis performed where negative reviews for product targeted to find dissatisfaction of user using *NLP* so can improve strategies.

Vandana Yadav et al. [51] created a system for analyzing customer sentiment in the Hindi language. This system utilizes machine-learning algorithms to collect data from different online platforms through web crawlers. The collected data is then pre-processed and tokenized before extracting relevant features. These extracted features then employed in classifying the data using machine-learning techniques. In particular, they utilize *SV*M, a well-known supervised machine-learning algorithm that often used for *regression* and classification tasks.

An analysis conducted on consumer reviews of amazon products by Aamir Rashid et al. [52] using multiple visualization tools. By utilizing these tools, we can observe various statistical trends that offer valuable insights into how the Amazon Review system operates. The main goal of this study is to explore the connection between product cost and positive reviews quantity. The analysis revealed a negative correlation between the two, indicating that as the price goes up, the number of helpful reviews tends to decrease.

3.4 Finance and Banking

Customer feedback on banking services is crucial for influencing the industry. It offers direct insights into customer satisfaction, pain points, and improvement areas for banks. Positive reviews establish trust and lure potential customers, while negative reviews pinpoint areas for enhancement. By resolving issues raised in reviews, banks can improve their services, innovate new products, and uphold a positive reputation.

Li Chen Cheng et al. [53] Utilize text mining techniques to assess reviews of digital banking applications and investigate the connection between text feedback and review

ratings. This paper aims to determine the primary user concerns and evaluate both positive and negative reviews. For this study, a digital banking application from the Philippines named CIMB Bank PH was chosen as the dataset & Data Cleaning and Data Pre-processing through Text-Mining. To extract the topics behind the reviews, Topic Modelling was adopted with the use of the *LDA* model tool from the Gensim package. To find the optimal number of topics, a CV Coherence analysis was conducted. *Association rule mining* was used to determine which features are generally positive or negative and their relationship to review scores. The study only examines text reviews and review scores for one bank, limiting its generalizability.

In research of Gutha Jaya Krishna et al. [54], a model to do Sentiment Analysis of the customer complaints of Indian banks proposed. The authors used machine-learning classifiers to determine the sentiment of bank customers based on online complaints and feedback. They utilized LIWC to extract psycho-linguistic features for sentiment analysis, supported by a t-test. The most effective results obtained by employing LIWC alongside *Random Forest* and *Naïve Bayes* techniques during the analysis of three banking datasets, as revealed by the study. These findings further substantiated through statistical analysis using a t-test.

One more research on banking Analysis of Feedback in Banking System using machine learning by Maheswaran M et al. [55]. This method involves utilizing opinion mining and sentiment analysis techniques to determine customer emotions regarding staff performance in areas such as speed, punctuality, and customer support. This approach integrates artificial intelligence and natural language processing systems to analyze feedback data obtained from banking system surveys. The study concludes that the *Support Vector Machine* algorithm exhibits superior accuracy compared to the other algorithms we employed, as evidenced by our results.

3.5 Social Media

Social media platforms such as Facebook, Twitter, and YouTube allow users to share their experiences, opinions, and suggestions on different subjects. These honest reviews offer genuine perspectives that help platforms promote businesses, propose enhancements, or advise against services depending on user feedback. Various papers collected for social media analysis including live tweeter data.

Sharnitha Katragadda et al. [56] made a study Using Opinion Mining helps in understanding student emotions towards teaching and learning experiences. This study categorizes student feedback collected through surveys conducted each semester. The aim is to comprehend students' opinions on various education aspects like modules, teaching methods, and assessments. The author pre-processed the data before utilizing opinion mining techniques. Results compared using *Naïve Bayes, Support Vector Machine*, *and Artificial Neural Network* algorithms. The findings showed that the developed *Neural Networks algorithm* consistently performed better in accuracy compared to the others.

Dr. Anamika Rana et al. [57] and her team developed a student responses mining system to evaluate topics and sentiments from student reviews. Author gathered feedback from Twitter using the Twitter API and generated a flat-file feedback document during the preprocessing phase. Next, cluster comments using *POS* to classify and label words and read the comments in each cluster and categorize comment as positive or negative. Precision, Recall, and F Measure were employed for comparative assessment.

A framework introduced by Andleeb Aslam et al. [58] for market analysis and customer insights. This framework focuses on gathering real-time Twitter data on various topics and converting it from unstructured to structured format. By utilizing this framework, they collect and present the most recent and popular opinions on any subject, product, personality, organization, government project, or movement. They also apply sentiment analysis to generate results in terms of polarity, subjectivity, and objectivity, along with historical data to provide recommendations.

Wenhuan Wang et al. [59] propose the OL-DAWE model, which uses Data Augmentation technology to generate opposed tweets according to the original tweet. The process begins with the utilization of tools to analyse how words or punctuation alters the message of a tweet. Subsequently, the OL-DAWE model introduced, generating contrasting versions of tweets and converting them into machine-readable code. Another program is then employed to determine the emotional tone of tweets, categorizing them as positive, negative, or neutral. Finally, the model tested on numerous tweets to assess its ability to accurately discern the emotions conveyed, demonstrating its efficacy in sentiment analysis of tweets.

Another study made on tweeter data conducted by Dr. Suwarna Gothane et al. [60] to examined the polarity of twitter posts using a deep learning model. The authors employed *LSTM* and *RNN* models for sentiment analysis on Twitter data. They used specific datasets for their research and conducted various experiments for their publication. Pre-processing was carried out to standardize the tweet format and remove unwanted symbols. The authors compared the performance of *RNN* and *LSTM* models on different datasets, with the results indicating that *LSTM o*utperformed *RNN* and *CNN* models in terms of accuracy. Additionally, they successfully enhanced the *LSTM* model's ability to interpret tweets effectively.

M. A. Al-Shabi [61] made research to analyze two specific datasets of Twitter data: one from Stanford and one from Sandars without preprocessing data. The research assesses the effectiveness of five lexicons in classifying Twitter polarity: VADER, SentiWordNet, SentiStrength, Liu and Hu opinion lexicon, and AFINN-111. Findings indicate that the VADER lexicon demonstrates superior accuracy in identifying positive and negative sentiments, rendering it a viable option for analyzing short texts for positive, negative, and neutral sentiments. This sentiment analysis approach relies on a lexicon-based methodology.

3.6 Other

The study introduced a framework for identifying spam in text messages and assessed various machine learning models using a large dataset of SMS texts by M.Rubin Julis et al. [62]. In the analysis phase, the author pre-processed the text data, converting unstructured content into a structured format, and then examined patterns using MIS. A binary classifier was developed by using feature vectors from both spam and non-spam messages. The findings indicated that *Support Vector Machines* achieved the highest accuracy score of 98% compared to other algorithms tested.

Noor Afiza Mat Razali et al. [63] publish a research paper that focuses on enhancing opinion mining techniques and the Kansei approach in cyberspace, identifying relevant sectors benefiting from opinion mining, examining techniques used in various domains, and discussing challenges and future research scope. Here a structured literature review carried out by analyzing 122 articles to investigate research on opinion mining application and the Kansei approach for addressing challenges in sentiment analysis of text in cyberspace. The study highlights a possible approach that includes machine learning, lexicon-based approach, and Kansei approach to analyze individuals' emotions in text, which can impact their information reception online.

In the research of H. H. Patel et al. [64], the application of a Hybrid approach, combining Convolutional Neural Network with a Rule-based method for conducting Opinion Mining on Gujarati Language text. The proposed framework identifying the polarity the polarity score of the sentence so as to understand the strength of the opinion. The authors utilize specific rules to address challenges in Opinion Mining posed by the presence of negation and conjunction words in the review. An experiment result shows that by utilizing the *CNN* method, a 90% accuracy rate was achieved in extracting a specific number of features from the text. The accuracy rate dropped to 70% when a negation rule incorporated.

Dattatray G. Takale et al. [65] present study recovers posts and finds emotions and polarity of posts. Author use various techniques to determine emotions and polarity of posts. The emotion package allows for the analysis of six distinct perspectives: anger, disapproval, tears, delight, low feelings, and overwhelming feelings. Pre-processing done on labelled data. The proposed method with Naive Bayes model trained using a training set, a validation set, and a test set. The dataset pre-processed into numerical characteristics using TF-IDF representation or count-based features, with the test set assessing the model's performance. The *Naive Bayes* model is evaluated using metrics like accuracy, reliability, and F1-score, and then optimized through pre-processing steps and hyper parameters. The model then tested on unseen data and if it meets performance criteria, it deployed for sentiment analysis in real-world applications.

The research paper introduces by Text based Sentiment Analysis using *LSTM* Dr. G. S. N. Murthy et al. [66] for sentiment classification method using *LSTM* for text data. The study utilizes the IMDB and Amazon Product datasets for experiments. IMDB is a vast movie review dataset. The aim is to determine whether a review conveys positive or negative sentiment. DL techniques like *LSTM* demonstrate improved sentiment classification accuracy of 85% with larger training datasets.

The use of *LSTM* neural networks and attention mechanisms is common in sentiment representation learning and text detection. However, they may not take into account

how emotions can affect feature extraction, which could impact sentiment classification performance. To address this Faliang Huang et al. [67] proposed a novel model, *AEC-LSTM*, for text sentiment detection, enhancing the *LSTM* network by integrating emotional intelligence and attention mechanism. *ELSTM*, a *novel LSTM variant*, utilizes emotional intelligence to enhance model inference by utilizing the positive interaction between emotion and cognition in the human brain. A new attention mechanism is suggested to efficiently adjust the weights of hidden text representations. The study presents a deep network model, *AEC-LSTM*, for sentiment analysis, which combines topic attention and convolutional *ELSTM* for text sentiment detection. As a result, *AEC-LSTM* has been found to be more effective than other classification tasks.

Financial market participants constantly monitor news, influencing stock prices based on past and new information. Extracting sentiments from news is crucial for investment decision-making. Sentiment analysis models offer efficient methods, but financial sentiment analysis is challenging due to domain-specific language and dataset limitations. The paper by Kostadin Mishev et al. [68] introduces a platform for assessing the efficiency of *Natural Language Processing (NLP)* techniques in extracting text features in the finance sector. The research indicates that sentiment extraction models can efficiently applied in the finance sector for predicting stock market trends, trading securities, and managing portfolios, overseeing brand reputation, and identifying fraud.

4 Discussion

Opinion mining becoming more and more important in many fields because it uses text data to look at and figure out how people feel about things. This survey has synthesized the current state of opinion mining, the process and different levels of revealing several key insights and trends in the field and challenges. This survey highlights the classification of various applications of opinion mining finds a wide and varied range of applications across a variety of fields. During literature survey different algorithms of pinion mining are tabulated as Table 1 and different sentimental analysis methods, characteristics, performance along with application are discussed as Table 2 in appendix.

5 Conclusion

Opinion mining is a rapidly growing research field processing text-based unstructured data for information systems applications like classified and real-time reviews, despite facing challenges. This research offers a thorough examination of sentiment analysis, encompassing fundamental principles, methodologies, assessment approaches, and practical applications. It also presents a detailed classification of applications across various domains, serving as a valuable resource for researchers and developers. Future work should focus on developing opinion summarization algorithms to generate summaries of user reviews.

Appendix

Table 1. Applications of opinion mining using sentimental analysis methods

Applications	Education				Hospitality		Customer Care & Services		Finance &Banking	Social Media	Other		
Algorithms	Feedback Assesment	Performance Analysis	Autoquestion Generation	Corriculam Evaluation	Transportation Review	Restaurant Review	User Rating & Recommendation	Product Feedback	Service Review	Tweeter Performance	spam detection & National Security	Language Specific Opinions	Emotions
BERT		√			√								
BI-LSTM	√		√										
LDA	√				√			√	√				
LSTM	√					√							√
Bayesian Network	√												
CNN	√					√		√		√		√	
Decision Tree				√	√						√		
K - Means									√				
K-Nearest Neighbor				√			√	√			√		
Linear Regression					√								
Logistic Regression				√						√	√		
SVM	√			√			√	√	√	√			
Naïve Bayes							√	√	√	√	√		√
Neural Network					√								
Random Forest					√		√	√			√		

Table 2. Sentimental analysis Methods, characteristics, algorithms, performance & Applications

Method	Characteristics	Algorithm Used	Performance	Applications
Rule-Based Methods	Utilizes predefined sentiment lexicons for simple, low-cost implementation, relying on heuristics and linguistic rules to identify sentiment in text	Sentiment lexicon-based analysis (e.g., AFINN, VADER)	- Effective for specific domains with clear sentiment expressions - Limited ability to handle sarcasm, irony, and complex emotions	Product Reviews (E.G., Amazon, Flipkart), Social Media Sentiment (E.G. Tweets), Basic Sentiment Detection In Customer Feedback

(*continued*)

Table 2. (*continued*)

Method	Characteristics	Algorithm Used	Performance	Applications
Machine Learning	Requires labelled training data and crucial feature extraction (e.g., n-grams, TF-IDF) for models to learn from historical data and predict sentiment	SVM (Support Vector Machines), Navie Bayes, Logistic Regression, Random Forest	- Generally high accuracy with sufficient labelled data -Performance can vary based on feature selection and the balance of classes in the training data	E-Commerce Feedback Analysis (Customer Reviews), Customer Service Analysis, Email Sentiment Classification, User Rating, Product Feedback
Deep Learning	Utilizes complex neural networks to capture intricate patterns and contextual dependencies, requiring large datasets for training to understand sentiment nuances	LSTM (Long Short-Term Memory), CNN (Convolution Neural Networks), BERT (Bidirectional Encoder Representations from Transformers), GRU (Gated Recurrent Unit)	- Often outperforms traditional machine learning methods in sentiment classification tasks - Requires significant computational resources and time for training	Sentiment Classification In Social Media (E.G., Twitter), Performance Analysis, Transportation Review, Restaurant Review, Opinion Mining In Blogs And Articles, Emotion Detection In Customer Reviews
Ensemble Methods	Combines multiple models to improve accuracy while reducing overfitting	Random Forest, Gradient Boosting Machines (GBM), XGBoost, Voting Classifier	- Typically achieves better performance than individual models - Performance is contingent upon the diversity and quality of the base models	Financial Market Prediction (Stock Sentiment Analysis), News Sentiment Analysis (Impact on Stock Prices), Comprehensive Sentiment Analysis In Surveys

(*continued*)

Table 2. (*continued*)

Method	Characteristics	Algorithm Used	Performance	Applications
Aspect-Based Sentiment Analysis (ABSA)	This method focuses on identifying and analyzing sentiments related to specific product or service aspects, such as battery life and camera quality, by combining aspect extraction with sentiment classification for more detailed insights	Aspect extraction Models (e.g., LSTM with attention mechanisms), Sentiment classification Models (e.g., CNN for sentiment)	- Provides more granular insights compared to overall sentiment analysis - More complex due to the need for accurate aspect identification	Detailed Product Reviews Analysis (E.G., Electronics, Cosmetics), Hotel And Restaurant Feedback (Specific Features), Service Review Analysis (Customer Experience), Auto generation Text, Language Specific Reviews
Sentiment Lexicon-Based	Utilizing sentiment word dictionaries to assess polarity while considering word intensity and context offers a simple and interpretable method for sentiment analysis	Sentiment lexicons (e.g., AFINN, SentiWordNet, LIWC), BERT, ReBERTa	- Can be effective for basic sentiment detection - May overlook context and nuanced expressions, leading to misclassification	Basic sentiment analysis tasks (general sentiment trends), Topic sentiment extraction in qualitative research, Keyword sentiment analysis
Unsupervised Learning	Does not require labelled data, making it suitable for exploratory analysis through clustering techniques or topic modelling to discover patterns without explicit training	K-means clustering, Topic Modelling (LDA - Latent Dirichlet Allocation), Word Embeddings (Word2Vec, GloVe)	- Can struggle with accuracy in sentiment classification; results may be exploratory and less precise - Often requires post-processing for interpretation	Discovering sentiments in large text corpora (e.g., Social media, reviews), Feedback, Service Review, Topic modelling in research papers

References

1. Rashid Al-Azmi, A.-A.: Data, text and web mining for business intelligence: a survey. In: International Journal of Data Mining & Knowledge Management Process, vol. 3, no. 2. Academy and Industry Research Collaboration Center (AIRCC), pp. 1–21, Mar. 31 (2013). https://doi.org/10.5121/ijdkp.2013.3201

2. Neha, K., Reddy, M.Y.: A study on applications of data mining. Int. J. Sci. Technol. Res. **9**(2), 3385 (2020)
3. Verma, S., Rattan, P.: Introduction to data mining tools and techniques & applications: a review. Business **57** (2021)
4. Srinath, K.R.: An overview of web content mining techniques. Int. Res. J. Eng. Technol. (IRJET) **4**(11), 1858–1864 (2017)
5. Prakash, M.K., Dinesh, G.: A survey on opinion mining: applications, challenges, tools and techniques. In: International Journal of Computer Sciences and Engineering, vol. 6, no. 4, pp. 1–10 (2018). E-ISSN: 2347-2693
6. Sharma, A.K., Kumar, A.: An investigative approach of web data mining tools for web analytics. In: International Journal of Computer Science and Technology, vol. 7, no. 1, pp. XX-XX, Jan.-Mar (2016). ISSN: 0976-8491 (Online), 2229–4333 (Print)
7. Goel, S., Gupta, R.K.: A survey on opinion mining. Int. J. Res. Appl. Sci. Eng. Technol. (IJRASET) **8**(IX), 981–987 (2020). ISSN: 2321-9653. www.ijraset.com
8. Mehta, P., Pandya, S.: A review on sentiment analysis methodologies, practices, and applications. Int. J. Sci. Technol. Res. **9**(2), 601–605 (2020)
9. Birjali, M., Kasri, M., Beni-Hssane, A.: A comprehensive survey on sentiment analysis: approaches, challenges and trends. Knowl.-Based Syst. **226**, 107134 (2021). https://doi.org/10.1016/j.knosys.2021.107134
10. Joshi, M., Prajapati, P., Shaikh, A., Vala, V.: A survey on sentiment analysis. In: International Journal of Computer Applications (Vol. 163, Issue 6, pp. 34–38). Foundation of Computer Science (2017). https://doi.org/10.5120/ijca2017913552
11. Pathak, A., Sharma, S., Pandey, R.: A methodological survey on sentiment analysis techniques and their applications in opinion mining. In: International Journal of Emerging Trends in Engineering and Development, vol. 1, no. 11. RS Publication (2021). https://doi.org/10.26808/rs.ed.i11v1.05
12. Okeke, O.C., Amaechi, C.: A review on opinion mining: approaches, practices and applications. Int. Res. J. Modernization Eng., Technol. Sci. **3**(3), 1011 (2021)
13. Chandni, C., Chandra, N., Gupta, S., Pahade, R.: Sentiment analysis and its challenges. In: International Journal of Engineering Research & Technology (IJERT), vol. 4, no. 3, pp. x-y (2015). IJERTV4IS030925
14. Jenifer, L.C., Kesavaraj, G.: A survey on sentiment analysis: levels, methods and application. In: Journal of Emerging Technologies and Innovative Research (JETIR), vol. 6, no. 3, pp. x-y (2019). www.jetir.org. (ISSN-2349–5162)
15. Mohan Kumar, A.V., Nandkumar, A.N.: A survey on challenges and research opportunities in opinion mining. SN Comput. Sci. **1**(3), 1–6 (2020). https://doi.org/10.1007/s42979-020-00149-4
16. Varghese, R., Jayasree, M.: A survey on sentiment analysis and opinion mining. Int. J. Res. Eng. Technol. **2**(11), 312–317 (2013)
17. Nayak, N., Smitha, S., Janaki, J.S.: An approach on big data for sentimental analysis applications. Eur. Chem. Bull. **12**, 1347–1350 (2023). https://doi.org/10.31838/ecb/2023.12.si6.125 2023
18. Alhojely, S.: Sentiment analysis and opinion mining: a survey. Int. J. Comput. Appl. **150**, 22–25 (2016)
19. Pradhan, V.M.: A survey on sentiment analysis algorithms for opinion mining. Int. J. Comput. Appl. **133**(9), 7–12 (2016)
20. Qi, C., Liu, S.: Evaluating on-line courses via reviews mining. IEEE Access **9**, 35439–35451 (2021). https://doi.org/10.1109/ACCESS.2021.3062052

21. Buenaño-Fernández, D., Villegas-Ch, W., Luján-Mora, S.: Using text mining to evaluate student interaction in virtual learning environments. In: 2018 IEEE World Engineering Education Conference (EDUNINE), Buenos Aires, Argentina, pp. 1–6 (2018). https://doi.org/10.1109/EDUNINE.2018.8450969
22. Kastrati, Z., Imran, A.S., Kurti, A.: Weakly supervised framework for aspect-based sentiment analysis on students' reviews of MOOCs. IEEE Access **8**, 106799–106810 (2020). https://doi.org/10.1109/ACCESS.2020.3000739
23. Sindhu, I., Muhammad Daudpota, S., Badar, K., Bakhtyar, M., Baber, J., Nurunnabi, M.: Aspect-based opinion mining on student's feedback for faculty teaching performance evaluation. In: IEEE Access, vol. 7, pp. 108729–108741 (2019). https://doi.org/10.1109/ACCESS.2019.2928872
24. Tanwani, N., Kumar, S., Jalbani, A.H., Soomro, S., Channa, M.I., Nizamani, Z.: Student opinion mining regarding educational system using facebook group. In: 2017 First International Conference on Latest trends in Electrical Engineering and Computing Technologies (INTELLECT). IEEE (2017). https://doi.org/10.1109/intellect.2017.8277622
25. Nandakumar, R., Pallavi, M.S., Harithas, P.P., Hegde, V.: Sentimental analysis on student feedback using NLP & POS tagging. In: 2022 International Conference on Edge Computing and Applications (ICECAA). IEEE (2022). https://doi.org/10.1109/icecaa55415.2022.9936569
26. Selvapandian, D., et al.: An efficient sentiment analysis on feedback assessment from student to provide better education. In: 2020 Fourth International Conference on I-SMAC (IOT in Social, Mobile, Analytics and Cloud) (I-SMAC). IEEE (2020). https://doi.org/10.1109/i-smac49090.2020.9243594
27. Karunya, K., Aarthy, S., Karthika, R., Jegatha Deborah, L.: Analysis of student feedback and recommendation to tutors. In: 2020 International Conference on Communication and Signal Processing (ICCSP), Chennai, India, pp. 1579–1583 (2020). https://doi.org/10.1109/ICCSP48568.2020.9182270
28. Fateen, M., Mine, T.: Extraction of useful observational features from teacher reports for student performance prediction. In: Lecture Notes in Computer Science. Springer International Publishing, pp. 620–625 (2022). https://doi.org/10.1007/978-3-031-11644-5_58
29. Wook, M., et al.: Opinion mining technique for developing student feedback analysis system using lexicon-based approach (OMFeedback). In: Education and Information Technologies, vol. 25, no. 4. Springer Science and Business Media LLC, pp. 2549–2560 (2019). https://doi.org/10.1007/s10639-019-10073-7
30. Okoye, K., Hosseini, S., Arrona-Palacios, A., Escamilla, J.: Impact of educational coaching programs and mentoring services on users' perception and preferences: a qualitative and quantitative approach. In: IEEE Access, vol. 9. Institute of Electrical and Electronics Engineers (IEEE), pp. 48105–48120 (2021). https://doi.org/10.1109/access.2021.3069130
31. Perez-Castro, A., Martínez-Torres, M.R., Toral, S.L.: Efficiency of automatic text generators for online review content generation. Technol. Forecast. Soc. Change **189**. 122380 (2023). Elsevier BV, https://doi.org/10.1016/j.techfore.2023.122380
32. Chakravarthy, V.J., Kameswari, M., Mydeen, H.D., Seenivasan, M.: Opinion mining from student text review for choosing better online courses. In: IOP Conference Series: Materials Science and Engineering, vol. 1070, no. 1. IOP Publishing, pp. 012067 (2021). https://doi.org/10.1088/1757-899x/1070/1/012067
33. Rachman, F.F., Nooraeni, R., Yuliana, L.: Public Opinion of Transportation integrated (Jak Lingko). In: DKI Jakarta, Indonesia, Procedia Computer Science, Volume 179, pp. 696–703 (2021). ISSN 1877-509, https://doi.org/10.1016/j.procs.2021.01.057
34. Guerreiro, J., Rita, P.: How to predict explicit recommendations in online reviews using text mining and sentiment analysis. J. Hosp. Tour. Manag. **43**, 269–272 (2020). https://doi.org/10.1016/j.jhtm.2019.07.001

35. Ban, H.-J., Kim, H.-S.: Understanding Customer Experience and Satisfaction through Airline Passengers' Online Review, pp. 309–319. School of Hospitality and Tourism Management, Kyungsung University, Busan, Korea (2019)
36. Patel, A., Oza, P., Agrawal, S.: Sentiment analysis of customer feedback and reviews for airline services using language representation model. Procedia Comput. Sci. **218**. 2459–2467 (2023). Elsevier BV, https://doi.org/10.1016/j.procs.2023.01.221
37. Zhang, X., Kim, H.-S.: Customer experience and satisfaction of Disneyland hotel through big data analysis of online customer reviews. Sustainability **13**(22), pp. 12699 (2021). https://doi.org/10.3390/su132212699
38. Hossain, N., Bhuiyan, M., Tumpa, Z., Hossain, S.: Sentiment analysis of restaurant reviews using combined CNN-LSTM. In: 2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT), pp. 1–7 (2020). https://doi.org/10.1109/ICCCNT49239.2020.9225328
39. Asani, E., Nejad, H.V., Sadri, J.: Restaurant recommender system based on sentiment analysis. Mach. Learn. Appl. **6**, 100114 (2021). https://doi.org/10.1016/j.mlwa.2021.100114
40. Shah, P., Swaminarayan, P., Patel, M.: Sentiment analysis on film review in Gujarati language using machine learning. In: International Journal of Electrical and Computer Engineering (IJECE), vol. 12, no. 1. Institute of Advanced Engineering and Science, pp. 1030 (2022). https://doi.org/10.11591/ijece.v12i1.pp1030-1039
41. Mehta, R., Gupta, S.: Movie recommendation systems using sentiment analysis and cosine similarity. Int. J. Mod. Trends Sci. Technol. **7**(1), 16–22 (2021)
42. Karthika, P., Murugeswari, R., Manoranjithem, R.: Sentiment analysis of social media network using random forest algorithm. In: 2019 IEEE International Conference on Intelligent Techniques in Control, Optimization and Signal Processing (INCOS), Tamilnadu, India, pp. 1–5 2019. https://doi.org/10.1109/INCOS45849.2019.8951367
43. Alrehili, A., Albalawi, K.: Sentiment analysis of customer reviews using ensemble method. In: 2019 International Conference on Computer and Information Sciences (ICCIS), Sakaka, Saudi Arabia, pp. 1–6 (2019). https://doi.org/10.1109/ICCISci.2019.8716454
44. Hu, S., Kumar, A., Al-Turjman, F., Gupta, S., Seth, S., Shubham.: Reviewer credibility and sentiment analysis based user profile modelling for online product recommendation. IEEE Access **8**, 26172–26189 (2020). https://doi.org/10.1109/ACCESS.2020.2971087.4o
45. Lakshay, B.: Sentiment analysis in online product reviews: mining customer opinions for sentiment classification. Int. J. Multi. Res. **5**(5) (2023). https://doi.org/10.36948/ijfmr.2023.v05i05.6090
46. Loukili, M., Messaoudi, F., El Ghazi, M.: Sentiment analysis of product reviews for e-commerce recommendation based on machine learning. Int. J. Adv. Soft Comput. Appl. **15**, 1–13 (2023). https://doi.org/10.15849/IJASCA.230320.01
47. Pankaj, P., Muskan, P., Soni, N.: Sentiment analysis on customer feedback data: amazon product reviews. In: 2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon), Faridabad, India, pp. 320–322 (2019). https://doi.org/10.1109/COMITCon.2019.8862258
48. Mahfiz, S.L., Romadhony, A.: Aspect-based opinion mining on beauty product reviews. In: 2020 3rd International Seminar on Research of Information Technology and Intelligent Systems (ISRITI). IEEE (2020). https://doi.org/10.1109/isriti51436.2020.9315350
49. Wassan, S., Chen, X., Shen, T., Waqar, M., Jhanjhi, N. Z.: Amazon product sentiment analysis using machine learning techniques. Revista Argentina de Clínica Psicológica **XXX**(1), 695–703 (2021), https://doi.org/10.24205/03276716.2020.2065
50. Yang, C., et al.: Online user review analysis for product evaluation and improvement. J. Theor. Appl. Electron. Commer. Res. **16**, 1598–1611 (2021). https://doi.org/10.3390/jtaer16050090

51. Yadav, V., Verma, P., Katiyar, V.: E-commerce product reviews using aspect based Hindi sentiment analysis. In: 2021 International Conference on Computer Communication and Informatics (ICCCI). IEEE (2021). https://doi.org/10.1109/iccci50826.2021.9402365
52. Rashid, A., Huang, C.: Sentiment analysis on consumer reviews of amazon products. In: International Journal of Computer Theory and Engineering, vol. 13, no. 2. IACSIT Press, pp. 35–41 (2021). https://doi.org/10.7763/ijcte.2021.v13.1287
53. Cheng, L.C., Sharmayne, L.R.: Analysing digital banking reviews using text mining. In: 2020 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM). IEEE (2020). https://doi.org/10.1109/asonam49781.2020.9381429
54. Krishna, G.J., et al.: Sentiment classification of Indian banks' customer complaints. In: TENCON 2019 - 2019 IEEE Region 10 Conference (TENCON), Kochi, India, pp. 429–434 (2019). https://doi.org/10.1109/TENCON.2019.8929703
55. Maheswaran, M., Benaka Santhosh, S.: A novel approach for analysis of feedback in banking system using machine learning algorithm. In: International Journal of Engineering Research & Technology (IJERT), NCCDS - 2021 Conference Proceedings, ISSN: 2278–0181
56. Katragadda, S., Ravi, V., Kumar, P., Lakshmi, G.J.: Performance analysis on student feedback using machine learning algorithms. In: 2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS), Coimbatore, India, pp. 1161–1163 (2020). https://doi.org/10.1109/ICACCS48705.2020.9074334
57. Rana, A., Malik, S., Sajwan, V., Awasthi, M.: Predicting student's mental health by analyzing their tweets. In: 2022 5th International Conference on Contemporary Computing and Informatics (IC3I). IEEE (2022). https://doi.org/10.1109/ic3i56241.2022.10072816
58. Aslam, A., Qamar, U., Khan, R.A., Saqib, P., Ahmad, A., Qadeer, A.: Opinion mining using live twitter data. In: 2019 IEEE International Conference on Computational Science and Engineering (CSE) and IEEE International Conference on Embedded and Ubiquitous Computing (EUC). IEEE (2019). https://doi.org/10.1109/cse/euc.2019.00016
59. Wang, W., Li, B., Feng, D., Zhang, A., Wan, S.: The OL-DAWE model: tweet polarity sentiment analysis with data augmentation. In: IEEE Access, vol. 8. Institute of Electrical and Electronics Engineers (IEEE), pp. 40118–40128 (2020). https://doi.org/10.1109/access.2020.2976196
60. Gothane, S., et al.: Sentiment analysis in social media using deep learning techniques. Int. J. Intell. Syst. Appl. Eng. **12**(3), 1588–1597 (2024)
61. Al-Shabi, M.A.: Evaluating the performance of the most important lexicons used in sentiment analysis and opinion mining. IJCSNS Int. J. Comput. Sci. Netw. Secur. **20**(1), 1–7 (2020)
62. Ara, J., Hasan, M.T., Al Omar, A., Bhuiyan, H.: Understanding customer sentiment: lexical analysis of restaurant reviews. In: 2020 IEEE Region 10 Symposium (TENSYMP), Dhaka, Bangladesh, pp. 295–299 (2020). https://doi.org/10.1109/TENSYMP50017.2020.9230712
63. Razali, N.A.M., et al.: Opinion mining for national security: techniques, domain applications, challenges and research opportunities. In: Journal of Big Data, vol. 8, no. 1. Springer Science and Business Media LLC (2021). https://doi.org/10.1186/s40537-021-00536-5
64. Patel, H.H., Patel, B.C., Lad, K.B.: Opinion mining of gujarati language text using hybrid approach. In: United International Journal for Research & Technology, vol. 3, no. 4, pp. [Page Numbers] (2022). ISSN: 2582–6832
65. Takale, D.G., et al.: Sentiment analysis through the application of machine learning algorithms. J. Comput. Netw. Vis. **2**(1), 219–14 (2024). https://doi.org/10.48001/JoCNV.2024.219-14
66. Murthy, G.S.N., Allu, S.R., Andhavarapu, B.: Text based Sentiment Analysis using LSTM. Int. J. Eng. Res. Technol. (IJERT) **9**(5), 290 (2020). http://www.ijert.org. IJERTV9IS050290
67. Huang, F., Li, X., Yuan, C., Zhang, S., Zhang, J., Qiao, S.: Attention-emotion-enhanced convolutional LSTM for sentiment analysis. IEEE Trans. Neural Netw. Learn. Syst. **33**(9), 4661–4672 (2022). https://doi.org/10.1109/TNNLS.2022.3167714

68. Mishev, K., Gjorgjevikj, A., Vodenska, I., Chitkushev, L.T., Trajanov, D.: Evaluation of sentiment analysis in finance: from lexicons to transformers. In: IEEE Access (Vol. 8, pp. 131662–131682). Institute of Electrical and Electronics Engineers (IEEE) (2020). https://doi.org/10.1109/access.2020.3009626
69. Arya, A., Shukla, V., Negi, A., Gupta, K.: A review: sentiment analysis and opinion mining. In: Proceedings of International Conference on Innovative Computing & Communications (ICICC) (2020). https://ssrn.com/abstract=3602548 or https://doi.org/10.2139/ssrn.3602548
70. Management Association, I.: Opinion Mining. IGI Global (2024). https://www.igi-global.com/dictionary/opinion-mining/21327#google_vignette. Accessed 17 Aug 2024
71. TechTarget: What is opinion mining (sentiment mining)? SearchBusinessAnalytics (2024). https://www.techtarget.com/searchbusinessanalytics/definition/opinion-mining-sentiment-mining. Accessed 11 Aug 2024

Mitigating the Technological Challenges of Regenerative Farming with an Integrated Framework Merging IoT, Ontologies, and GIS

Tshepiso L. Mokgetse[1(✉)], Rajagopal Sridaran[2], and Hlomani Hlomani[1]

[1] Botswana International University of Science and Technology, Boseja, Palapye, Botswana
mt21100029@studentmail.biust.ac.bw, hlomanihb@biust.ac.bw
[2] Marwadi University, Rajkot, Gujarat, India
sridaran.rajagopal@marwadieducation.edu.in

Abstract. Regenerative farming offers a sustainable approach to agriculture by promoting ecosystem restoration, soil health, and biodiversity conservation. However, the adoption of regenerative practices faces significant technological challenges, including data collection, management, interoperability, integration, mapping, and spatial analysis. These challenges hinder the scalability and effectiveness of regenerative farming, particularly in terms of monitoring, decision-making, and resource optimization. This article addresses these challenges by proposing an integrated framework that combines the Internet of Things, Ontologies, and Geographic Information Systems. IoT is leveraged for real-time data collection and management, Ontologies ensure seamless data interoperability across systems, and GIS enables advanced spatial analysis and mapping. Using a meta-synthesis of existing literature, this study identifies the key technological obstacles to regenerative farming and presents a solution that integrates these emerging technologies. The proposed framework aims to enhance the effectiveness of regenerative practices, support ecosystem health, and improve the sustainability of agricultural systems. Farmers will also leverage from sharing practices across different regions. The data will allow farmers to get practical results of the effectiveness of regenerative farming.

Keywords: Internet of Things · Geospatial Information Systems · Regenerative farming · framework

1 Introduction

Author [1] states that regenerative farming is currently the transformative and holistic best farming practice. This is because it has the objective of sustaining the world's food production while seeking to revitalize and restore the health of the ecosystem. The main principles of regenerative farming include carbon sequestration, soil health, water management, and biodiversity conservation.

Scaled-up regenerative farming whose key principles are no-tillage, cover cropping, reduced pesticide use, and diverse rotations has shown to improve soil health and

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
S. Rajagopal et al. (Eds.): ASCIS 2024, CCIS 2428, pp. 95–106, 2025.
https://doi.org/10.1007/978-3-031-86302-8_6

sequester carbon in many parts of the world [2]. However, there are threats to Regenerative Farming adoption, including socio-political barriers, scalability questions, and, at the farm level, machinery constraints and herbicide-resistant weeds. However, this research article is focused on the technological challenges. Integrating technology in regenerative practices is important for collecting and managing the data, sharing information across different regions, and monitoring the best practices for effective land management. [3] agrees with the benefits of technology in the food system, especially stakeholders like commercial farmers and political participants.

The meta-synthesis methodology is used to carry out a review of literature that is in alignment with the research question. The review searches, selects, analyses, and syntheses literature to determine the main technological challenges in adopting regenerative agriculture. Some of the challenges include data collection, data management, interoperability, integration, and finally analysis and mapping of data. Based on the findings this research proposed a regenerative farming framework integrating Internet of Things (IoT), Ontologies, and Geographical information system (GIS). Each of the emerging technologies functionalities forms a triangulation of regenerative farming practices for the good of ecosystem health. IoT is used for data collection and management, ontologies for interoperability and integration, and GIS for spatial data analysis and mapping.

This article is presented as follows; next section is related work which gives an overview of regenerative farming. The next section is the methodology explaining how meta-synthesis was conducted to get relevant information. The findings section follows presenting the technological challenges extracted from the methodology section. From the challenges a solution is provided in the proposed regenerative farming framework, with a triangulation of the three emerging technologies. The article concludes in the last section.

2 Related Work

A comparison of the types of farming systems was carried out [4] to validate the benefits of regenerative farming. The types of farming systems identified are traditional agriculture, conventional agriculture, organic agriculture, and regenerative farming. Mitigation measures have been reported by the Intergovernmental Panel on Climate Change [5] and recommended [4] to ensure that there is resilience to future challenges in farming. The types of farming systems are compared according to the mitigation challenges predicted for the future; Fig. 1 depicts the comparison.

The traditional farming system has existed for decades and is the commonly known type of farming. It still exists in many countries but has substantially reduced in Europe [4]. Traditional farming only provides half of the mitigation measures as seen in Fig. 1 the quality of the seeds is compromised, no legume promotion, and no biodiversity increment. The next type of system in comparison with regenerative farming is the conventional farming system also known as industrial farming. Businesses are in favour of the industrial system because it yields crop output that is quite significant. The crop output makes it favorable for it effectively provides food in a short time. Although there is a short time turnaround it takes a heavy toll on the environment and our health. The disadvantages include soil degradation, deforestation at a large-scale, water quality and quantity problems, and greenhouse gas emissions is increased [6]. Even though

climate change threatens the future of industries in agriculture there are still big contributors, Fig. 1 shows how conventional farming is in no way practicing measures for the preservation of agriculture as compared to regenerative farming.

The goal of organic farming is to sustain rather than regenerate the land. Regenerative and organic farming have similar priorities [4]. However, organic farming measures do not fully take into account the future farming mitigation measures that are crucial, as seen in Fig. 1. In organic farming, there is a decrease in water uptake, biodiversity, and soil fertility [7]. While organic prioritize biodiversity and soil conservation they forget the community's social, material, and economic needs. An effective type of farming system is needed that covers all the mitigation measures ensuring that the people in the world will still have plenty of food in the future. That system is a regenerative farming system, in comparison with other farming systems, Fig. 1, it can improve water quality, provide a long-term increase in biodiversity, improve the soil, consider local context also sustain the ecosystem and community without erosion of the standard of living [4]. Farmers earn a decent living while regenerating the land for future use, making them feel secure for their future income.

IPCC Measure	Traditional agriculture	Conventional agriculture (IPCC, 2019)	Organic agriculture (European Commission)	Regenerative agriculture (ROC, 2019)
Reduced fertilizer use	✓	✗	✓	✓
Increased crop diversification	✓	✗	✓	✓
Increasing the use of quality seeds	✗	✗	✓	✓
Promotion of low energy production systems	✓	✗	✗	✓
Avoiding burning of crop residues	✗	✗	✗	✓
Promotion of legumes in crop rotations	✗	✗	✗	✓
Increasing biodiversity	✗	✗	✗	✓
Integrated crop/ livestock	✓	✗	✗	✓

Fig. 1. Comparison of the types farming systems [4]

According to [8] regenerative farming aims to promote environmental health by the restoration of the soil through practices that minimize tillage, fertility contribution, and improvement of soil quality. Research by Khangura [9], also lists similar practices that focus on soil health. He states that the soil should be covered year-round, soil

disturbance should be minimized, and plants and roots that are live should be kept in the soil to the maximum extent. Authors [10] are in accord with [8] [9], He provides the same lists and adds that using synthetic pesticides and fertilizers should be greatly reduced. Moreover, the decline of soil health can be reversed by reduction in compaction, carbon sequestration increment, and soil organic matter enhancement.

Biodiversity is defined [11] as the planet's resilience and health cornerstone. Since the year 1970 places like the United Kingdom have experienced a loss of 50% biodiversity [10]. As the cornerstone of the planet, it is one of the elements promoted by regenerative farming. It is imperative to enhance biodiversity both below and above ground. [10] recommend using diversified crop rotation, integrating animals, adding legumes to pastures, and planting buffer strips. Researchers [8, 9] agree that integrating those agricultural practices will promote and enhance biodiversity. In some studies, it shows that regenerative farming may cause long-term yield loss. However, [10] found that after the cost reductions for the long-term are accounted for there will be a profit increase after three to six years.

3 Methodology – Meta-synthesis Methodology

The methodology adopted from [12] for this research study is the meta-synthesis which is used. Figure 2 is a summary of the steps that were taken to search and review data from the literature to answer the research question. The first step is to define the research question which will determine the search criteria when searching in the multiple search engines. The research question set by the research is:

Research Question: What are the technological challenges to widespread adoption of regenerative farming practices and how can emerging technologies be integrated to support regenerative farming practices?

The second step (Fig. 2) is the study selection used to retrieve results from the database searchers. Microsoft Excel was used to systematically organise the data from different literature. During the process, duplicates were scanned for and removed. Two authors collaborated to individually screen all the abstracts from the remain articles and select studies that aligned with the research question. If the title and abstract of the article did not provide enough clarity the full text was read. Upon the completion of the review process by both authors the data was systematically written in the excel document with relevant topics that will make presenting the data efficient. The last step of the meta-synthesis as shown in Fig. 2 was analysing the document and synthesis of the data. Themes were developed in the analysis in first and second constructs to help build a synthesis of the research. The themes were ordered accordingly and interpreted to present the results in the findings section of this research.

4 Findings - Challenges of Implementing Emerging Technologies in Regenerative Agriculture

Technologies implemented in regenerative agriculture will solve a lot of the challenges they are currently facing. However, there are challenges that have been identified that hinder the success of regenerative agriculture. The identified challenges will impact the overall adoption and effectiveness of emerging technologies (Table 1).

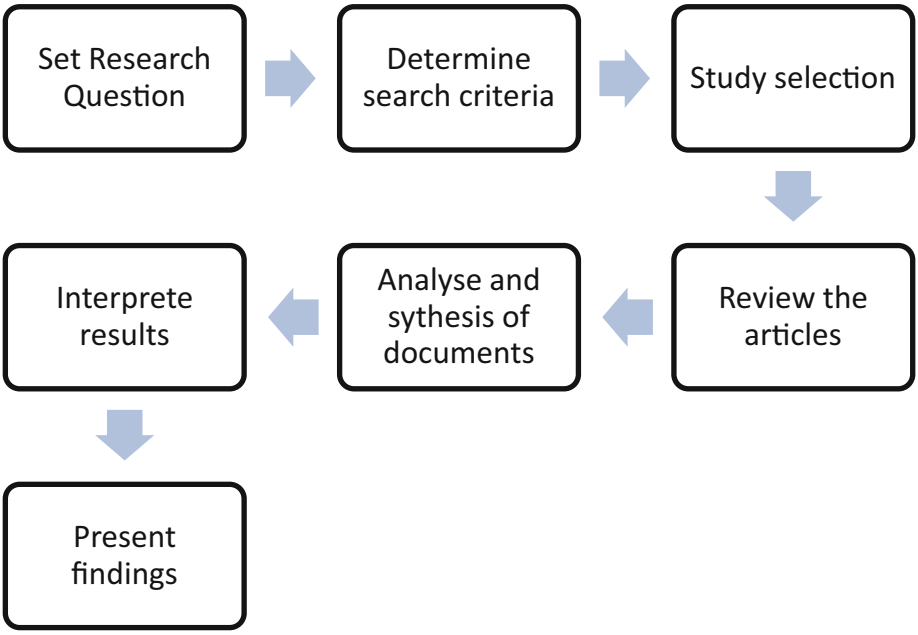


Fig. 2. Meta-synthesis methodology for collection of data.

Table 1. Identified technological challenges of regenerative agriculture.

Technological Challenges	Details
Data collection	Real-time collection
	Automatic collection
	Reliability of data
Data management	Data sharing
	Data quality
	Data accuracy
Interoperability and Integration	Systems interoperability
	Expert Knowledge
	Standardization – communication protocols
	Decision support
Analysis and mapping	Data analysis
	Data interpretation
	Data reports

According to [13] farmers who do not have technical assistance are unable to plan, implement, then monitor climate-smart practices. Moreover, farmers fail to apply for and

track grant incentives from the government because they still use paperwork. Researcher [14] agrees with [13], that the lack of farmers to implement emerging technologies does not provide the full benefits of regenerative agriculture. [14] presents the image below Fig. 3 which depicts the regenerative farm that is digitally enabled. Figure 3 is a proposed mixed farming system that uses regenerative farming practices. It uses basic technology and does not use effective emerging technology tools. The large data quantity of data needed to implement the digital farm has presented another challenge. Furthermore, regenerative environments that have less technological practices present more barriers for farmers and consumers. The following is a list of categorized technological challenges found in regenerative farming.

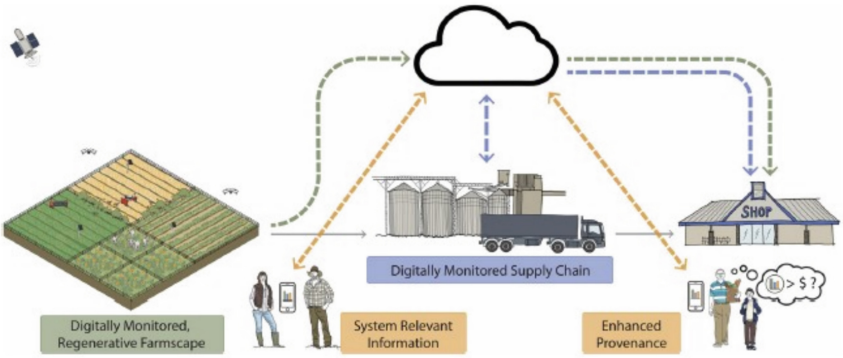


Fig. 3. Proposed regenerative farming system [14]

4.1 Data Management

The complexity and volume of data that is received by farmers on a daily basis has presented a challenge. With the adoption of emerging technologies in regenerative farming there will be large volumes of diverse data from tools like satellite imagery and sensor data, which may be challenging to effectively manage and analyze. Author [15] explains the strong advancements in scientific and farming knowledge, has caused lagging behind of large-scale implementation. It is important to have all the various stakeholders of regenerative farming involved in the data and knowledge advancements. A participatory process is recommended [15] as to tackle the data collection of important regenerative farming data. This process will ensure the data is accurate and good quality. O'Connor [13] stressed the importance of information sharing especially financial management for young farmers. The lack of sharing data collected has caused some failures for young farmers venturing into regenerative farming. It is a costly process that gives returns after a couple of years. However, it is the best farming system to ensure there is still food production in 100 years to come.

Researchers [3] identified geographies and specific farming systems as limitations to effective data collection. Benson [3] Encourages the development of an outcome-based framework that will provide data collection that is robust. The framework should

focus on tackling creating a methodology with standardized data collection and reports with farmer practice outcomes. Author [14] concurs, that standardized data collection is essential and also adds that the lack of scientific validation should be tackled also to ensure data quality.

4.2 Interoperability and Integration

The interoperability of a system deals with standardization and device compatibility. So far we have not found any device used in regenerative farming that uses the same standards and communication protocols. This makes it challenging to provide data exchange and interaction that is seamless between devices. The main challenge of interoperability and integration is the lack of metrics and standardized protocols used in data collection and analysis. Research carried out by Glaros [16], listed challenges of interoperability in regenerative farming. The first is the binding user agreements that farmers find themselves locked in with companies because of the lack of interoperability among software and hardware platforms. The challenge has posed inefficient and time-consuming business management practices. Another challenge comes from market vendors and farmers not able to perform cross-platform sales [16]. Furthermore, a pole taken on the lack of inter-platform interoperability from participants (farmers and marker vendors) indicates that 60% accept the difficulty in navigating the multiple platforms across food systems.

There is limited choice for farmers in adopting emerging technologies because of the semantic and structural interoperability [17]. This creates barriers in the efficiency of utilizing software that comes from various providers and devices. Researcher [17] found that individual technology solutions are unable to share information easily with their stakeholders, which leads to more repetitive and manual data management. Soto [15] agrees that interoperability and integration can be beneficial to regenerative farmers by learning scientific knowledge from expertise.

5 Implementation - Proposed Theoretical Framework for Regenerative Agriculture Integrating IOT, Ontologies, and GIS

The objective of regenerative agriculture is to bring enhancement and restoration to the wellbeing of the ecosystem by practicing sustainable farming. The farming practices include climate change resilience, sustainable farming promotion, and soil biodiversity improvement. However, this objective can be achieved through implementation of innovative technologies that will assist the farmers to manage, optimize, and monitor agricultural systems. The triangulation of ontologies, geographic information systems, and the Internet of Things offers a breakthrough opportunity for regenerative farming advancements.

This research proposes a framework that integrates emerging technologies IoT, GIS, and Ontologies that will serve as a guiding principle to developing robust interoperable systems to manage regenerative agricultural practices. Figure 4 illustrates the integrated IoT, GIS, and Ontologies framework for effective regenerative farming. IoT provides

sensors with real-time data collection and monitoring of essential variables like temperature, moisture of the soil, and health of the crops. The sensors will generate data in large volumes in need of organizing and analyzing. It is important for the data to be shared among the diverse stakeholders and systems, for interoperability of the data, and data integration. These functionalities can be facilitated by the provision of semantics relations and standardized vocabulary provided by Ontologies. After the sensors have collected the data and ontologies have effectively managed the data, GIS is used for the assessment of the geographical context of the data through its visualization and spatial analysis tools. The results will allow researchers and farmers to easily identify trends and patterns in the areas.

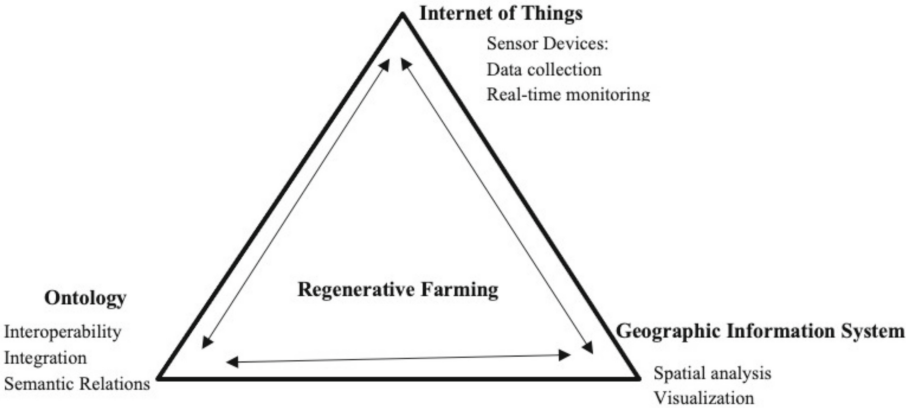


Fig. 4. Integrated Framework for regenerative farming.

5.1 Internet of Things

The main contributions of IoT to regenerative farming alongside ontologies and GIS are seen in Fig. 4. IoT will be used to collect data and real-time monitoring, the sensors will continuously monitor the moisture content, humidity, soil health, growth of the crops, and temperature. The data collected in real-time will help farmers in deciding where and when to add resources like watering and applying nutrients. IoT provides alerts to farmers that can inform them early of problems such as water stress, disease outbreaks, and pest invasions before it is critical and difficult to solve.

Farming practices such as cover cropping, reduced tillage, and crop rotation are essential to improving soil health [18]. These practices can be implemented at the right time by using IoT devices to track the soil organic matter, nutrients, and pH levels. Good management of soil health will improve the soil so much that the land will become resilient to weather conditions that are extreme.

Climate adaptation is critical to maintaining good farming practices for the future production of food. IoT can serve regenerative farming in farm monitoring of microclimatic conditions, helping farmers identify custom-made management practices for

adaptation to environmental situations that are specific. The reduction of carbon footprint by IoT systems will contribute greatly to lowering greenhouse gas emissions, thereby promoting mitigations to climate change.

5.2 Ontology

Data interoperability and information that is standardized is provided by using ontologies. The data that is collected by IoT sensor devices is first organized then shared across different systems. Ontologies as depicted (Fig. 4) are useful in establishing shared vocabulary specific to regenerative farming, that will allow effective communication among stakeholders. The standardized vocabularies will ensure data from IoT sensors and GIS systems is accurately interpreted and understood. Ontologies will facilitate in the integration of heterogeneous data, making it easy to analyze data from different domains. The multiple regenerative practices will be uniform and comprehensive. Young farmers new to regenerative farming will have platforms where they can adopt practices that will effectively work for them. The knowledge-sharing provision by ontologies will encourage good practice recommendations across regions, thus allowing for global advancements.

5.3 Geographic Information System

To maintain a resilient and sustainable farming system insights into adaptation of the climate, management of land, biodiversity, and conservation of water resources are essential. GIS provides the necessary insights that are needed for regenerative farming with the incorporation of IoT and Ontologies. GIS offers farmers magnified spatial data analysis for informed decision-making in resource allocation and land use. This is done through mapping of water availability, quality of the soil, and topography.

One of the main challenges of regenerative farming was monitoring the effectiveness of regenerative farming practices on the ecosystem. Implementing GIS will track changes that occur over time in the growth of vegetation, habitat conditions, and erosion of the soil.

Mapping and assessing the value the ecosystem will promote regenerative farming, GIS maps water filtration, pollination, and carbon sequestration. It allows farmers to measure how effective is regenerative farming in enhancing soil carbon stocks and capturing atmospheric carbon. The mapping is also beneficial in assessing the risk of natural hazards like wildfires and drought. Even after the disaster has happened GIS will be able to map areas that are affected and prioritise areas for ecological restoration. Furthermore, providing guidance to efforts of recovery while maintaining the regenerative way.

6 Future Work

This research proposes an integrated framework with IoT, Ontologies, and GIS to improve regenerative agricultural practices. However, the research did not carry out the practical work for validating the integrated framework. Future research includes a

demonstration of the framework's effectiveness through empirical validation. The integrated framework will be accurately assessed to demonstrate its potential to enhance regenerative farming activities while promoting sustainability throughout various agricultural fields. The empirical validation will be carried out by (a) identifying the key performance indicators. (b) Implementing practical pilot studies with relevant agricultural stakeholders, that include data collection, real-time data acquisition using IoT devices, creation of domain-specific ontologies for standardization and seamless interoperability between different systems, and analyzing spatial patterns also developing GIS-based decision support systems. (c) Evaluation of the practical pilot results. (d) Documenting and reporting all activities and results of the integrated framework validation process. Addressing the practical implementation limitation of the integrated framework is important to realize the full potential of the framework in the regenerative agricultural context.

7 Conclusion

Throughout the years farming land has been depleted of nutrients essential to producing food. More and more companies are using chemicals to quickly grow crops and meet the demand for food by people. However, the world has been witnessing a growing need for urgent transition to sustainable and regenerative farming practices. The global population is increasing each day putting pressure on natural resources to meet the demand. It is critical to explore the challenges slowing regenerative farming and ways the challenges can be addressed for environmental sustainability and food security.

This research posed a research question "What are the technological challenges to the widespread adoption of regenerative farming practices and how can emerging technologies be integrated to support regenerative farming practices?". The question was first addressed carrying out meta-synthesis of literature to identify the main technological challenges in regenerative farming. The identified challenges presented include the current data collection, data management, interoperability, integration, analysis, and mapping techniques in use. To counter the challenges effective techniques have to be adopted and implemented. According to [3] an outcome-based framework in regenerative farming is long overdue to overcome the everyday challenges. Researcher [10], agrees that socio-economic relies on sustainable and regenerative agriculture but the field lacks an implementation framework. The research recommended a comprehensive framework implementing a triangulation of three emerging technologies IoT, Ontologies, and GIS.

A combination of the three technologies in the regenerative framework will help in reduced consumption of resources, improved resilience and sustainability, and better processes in decision-making. The framework presents a novel contribution in regenerative agriculture by addressing the need for adaptive and robust countermeasures to a complex landscape in agriculture. IoT supports regenerative data collection and management with sensor devices. While ontologies provide standardized communication, interoperability and integration. A call for interoperability has been made [16] in the agriculture field to enhance agricultural practices. GIS assists in the essential decision-making processes by analysing and mapping the data.

References

1. Rai, A.K., Sengupta, S., Kundu, M., Rani, R.: Regenerative agriculture -a new beginning for tomorrow. Food Sci. Rep. **4**(7), 89–94 (2023). https://www.researchgate.net/publication/372786925_Regenerative_Agriculture_-A_new_beginning_for_tomorrow
2. McLennon, E., Dari, B., Jha, G., Sihi, D., Kankarla, V.: Regenerative agriculture and integrative permaculture for sustainable and technology driven global food production and security. Agron. J. **113**(6), 4541–4559 (2021). https://doi.org/10.1002/agj2.20814
3. Ewer, T.: Aligning regenerative agricultural practices with outcomes to deliver for people, nature and climate. The food and land use coaliation, Norway, Educational (2023). https://www.foodandlandusecoalition.org/knowledge-hub/regenag-people-nature-climate/
4. The Regeneration Academy and TheRockGroup. Regenerative Agriculture: A Manual. Europen Union, pp. 1–23 (2023). https://www.eitfood.eu/media/documents/FR_edited_Intro_15_11.pdf
5. Stendahl, K., Lynn, J.: Summary for Policymakers — Special Report on Climate Change and Land. University of Canterbury, New Zealand, Issue report (2024). https://www.ipcc.ch/srccl/chapter/summary-for-policymakers/
6. Lemke, S., Smith, N., Thiim, C., Stump, K.: Drivers and barriers to adoption of regenerative agriculture: cases studies on lessons learned from organic. Int. J. Agric. Sustain. **22**(1), 2324216 (2024). https://doi.org/10.1080/14735903.2024.2324216
7. Moyer, J., Smith, A., Rui, Y., Hayden, J.: REGENERATIVE AGRICULTURE and the SOIL CARBON SOLUTION. Rodale Institute, 611 Siegfriedale Road, Kutztown, PA 19530 (2020). https://rodaleinstitute.org/wp-content/uploads/Rodale-Soil-Carbon-White-Paper_v11-compressed.pdf, https://doi.org/10.1080/14735903.2024.2324216
8. Jayasinghe, S.L., Thomas, D.T., Anderson, J.P., Chen, C., Macdonald, B.C.T.: Global application of regenerative agriculture: a review of definitions and assessment approaches. Sustainability **15**(22), 15941 (2023). https://doi.org/10.3390/su152215941. https://www.mdpi.com/2071-1050/15/22/15941
9. Khangura, R., Ferris, D., Wagg, C., Bowyer, J.: Regenerative agriculture—a literature review on the practices and mechanisms used to improve soil health. Sustainability **15**(3), 2338 (2023). https://doi.org/10.3390/su15032338. https://www.mdpi.com/2071-1050/15/3/2338
10. Goss, D., Bush, L.: Sowing Resilience: Unlocking the potential for regenerative farming. McCain Foods, 15 Whitehall, Westminster London SW1A 2DD, Issue report (2023). https://demos.co.uk/research/sowing-resilience-unlocking-the-potential-for-regenerative-farming/
11. Rathoure, A.K.: Biodiversity and its importance. Biodivers. Int. J. **7**(1), 2 (2024). https://doi.org/10.15406/bij.2024.07.00204. https://medcraveonline.com/BIJ/biodiversity-and-its-importance.html
12. Lachal, J., Revah-Levy, A., Orri, M., Moro, M.R.: Metasynthesis: an original method to synthesize qualitative literature in psychiatry. Front. Psychiatry **8**, 269 (2017). https://doi.org/10.3389/fpsyt.2017.00269, https://doi.org/10.3389/fpsyt.2017.00269/full
13. O'Connor, J.: Barriers For Farmers & Ranchers to Adopt Regenerative Ag Practices In The US. Patagonia, Agricultural (2020). https://forainitiative.org/barriers-for-farmers-ranchers/
14. O'Donoghue, T., Minasny, B., McBratney, A.: Digital regenerative agriculture. npj Sustain. Agric. **2**(1), 5 (2024). https://doi.org/10.1038/s44264-024-00012-6, https://www.nature.com/articles/s44264-024-00012-6
15. Luján Soto, R., Cuéllar Padilla, M., Rivera Méndez, M., Pinto-Correia, T., Boix-Fayos, C., De Vente, J.: Participatory monitoring and evaluation to enable social learning, adoption, and out-scaling of regenerative agriculture. E&S **26**(4), art29 (2021). https://doi.org/10.5751/ES-12796-260429, https://www.ecologyandsociety.org/vol26/iss4/art29/

16. Glaros, A., Thomas, D., Nost, E., Nelson, E., Schumilas, T.: Digital technologies in local agri-food systems: opportunities for a more interoperable digital farmgate sector. Front. Sustain. **4**, 1073873 (2023). https://doi.org/10.3389/frsus.2023.1073873, https://doi.org/10.3389/frsus.2023.1073873/full
17. Shahan, M.: Data Interoperability in Agriculture. Issue Report (2021). https://www.farmfoundation.org/wp-content/uploads/2021/10/IssueReport_DataInteroperability.pdf
18. Schreefel, L., Schulte, R.P.O., De Boer, I.J.M., Schrijver, A.P., Van Zanten, H.H.E.: Regenerative agriculture – the soil is the base. Global Food Secur. **26**, 100404 (2020). https://doi.org/10.1016/j.gfs.2020.100404, https://linkinghub.elsevier.com/retrieve/pii/S2211912420300584

Leveraging Blockchain Technology: Decentralized Finance (DeFi) in Health Insurance

D. Susana(✉) and V. Srividya

PSG Institute of Management, PSG College of Technology, Coimbatore, India
{susana,srividya}@psgim.ac.in

Abstract. Decentralized Finance (DeFi) is an emerging technology that leverages blockchain technology to facilitate peer-to-peer financial transactions without the need for intermediaries. While DeFi has primarily been applied widely in the traditional finance arena, there is growing interest in its potential application in the healthcare industry, specifically in the field of health insurance. Health insurance is a complex and highly regulated industry on the verge of disruption. DeFi has the potential to streamline the claims process, reduce costs, and increase transparency in the insurance sector. However, significant challenges must be addressed before DeFi can be widely adopted in healthcare insurance. This study aimed to explore the scope of implementation of DeFi in the health insurance field by conducting interviews among industry experts and administering surveys among traditional healthcare insurance users on their perception of acceptance of DeFi. The expert interviews highlighted transparency, accuracy, cost efficacy, accessibility, efficiency, and customization as major advantages. The drawbacks of DeFi in the health insurance sector included regulatory challenges, integration issues, data privacy and security, and technical expertise. The survey findings using SEM analysis revealed that accessibility and transparency were the significant determinants influencing individuals' readiness to adapt to DeFi. Key insights on the potential benefits and limitations of DeFi in the healthcare sector and the implications of implementation of DeFi in the healthcare insurance sector are discussed.

Keywords: Decentralized Finance (DeFi) · Blockchain · Smart contracts · healthcare · insurance

1 Introduction

In the centralized financial system, financial institutions play a vital role as intermediaries to mediate and control financial transactions between facilitators and users [1]. The financial intermediaries include market makers, dealers, financial aggregators, and insurance companies, and they play a significant role in mediating and structuring economic transactions that are difficult to execute [2]. The recent disruption in the financial sector with the proliferation of technology has revolutionized the role of financial intermediaries. Decentralized Finance, also known as DeFi, is a modern financial architecture

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
S. Rajagopal et al. (Eds.): ASCIS 2024, CCIS 2428, pp. 107–120, 2025.
https://doi.org/10.1007/978-3-031-86302-8_7

that eliminates the need for central banks or government agencies to approve financial transactions [3]. The financial products and services including payments, lending and borrowing, trading and investments, crowd-funding, and insurance can be executed through DeFi without a trusted central intermediary [3].

The architecture of DeFi (Fig. 1) includes publicly available open protocols, smart contracts, and decentralized applications or dApp [4]. However, the foundation of DeFi lies in smart contracts, which are self-executing computer programs that run on blockchain technology. Blockchain comprises interconnected blocks, each housing a comprehensive transaction history, which can be shared publicly or privately among network users [5]. Smart contracts activate automatically upon meeting specific conditions, eliminating the necessity for intermediaries [6], and are employed to auto-mate agreement execution, ensuring transparency and minimizing the risk of manipulation [3]. Consequently, transactions occur without the need for a central authority, establishing a trust-free and flexible protocol stack [5]. Smart contracts are used for developing decentralized applications (dApps) offering a range of financial products and services [3]. A dApp is a software application functioning on a peer-to-peer blockchain network, often built on the Ethereum platform. These applications are decentralized, enhancing transparency and democracy in their operations. dApps can be developed for diverse purposes such as decentralized finance, web browsing, gaming, and social media [7].

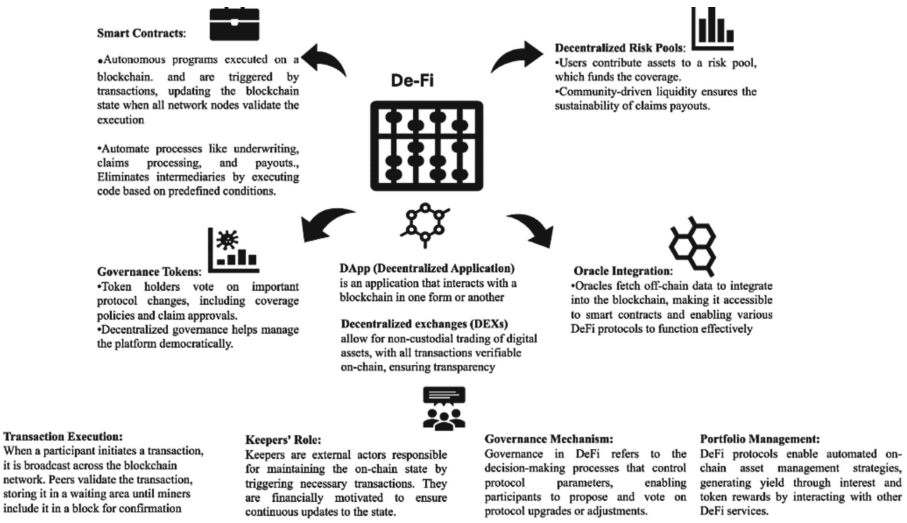


Fig. 1. Architecture of DeFi (adapted from [8, 9])

Decentralized finance (DeFi) offers a host of benefits. First, the efficiency of financial transactions can be improved by decentralized finance [6]. Using smart contracts in place of intermediaries, DeFi can improve efficiency in transactions [3]. In a decentralized financial environment, bilateral token transactions can be used to transfer digital assets [5]. Second, transparency in finance can be greatly improved by decentralization [4, 10]. The codes of smart contracts can be examined on the blockchain that are openly

observable [11] and hence all DeFi financial transactions are open and transparent [5]. The transparency of DeFi will facilitate remote monitoring by lawmakers and financial regulators, enabling them to understand the impact of decentralized finance protocols on the stability of the financial system more easily [12]. Third, decentralized finance provides accessibility to the decentralized finance protocols. DeFi protocols would enable a free and open financial system for all and facilitate transaction execution by all individuals. It can eliminate income discrimination between the poor and wealthy which is predominant in the conventional financial system [5]. Furthermore, DeFi facilitates financial services to become more advantageous than traditional financial service providers, which are associated with high costs, lengthy processes, and a lack of transparency [4]. DeFi enables users to participate as service providers and consumers and can automate and execute applications namely bank loans, derivatives, and trading in a transparent and secure manner [13].

The insurance sector requires the coordination of a large number of users for the purchase of contracts, and payment of insurance premiums, defaults, and claims. DeFi facilitates congruence of a large number of users in the network and enables users to purchase insurance coverage from other users and make payments through smart contracts in the event of a clearly defined claim [14] and facilitates members to pool and share risk [15].

In India, the health insurance industry is experiencing growth due to the expanding middle class, escalating hospitalization expenses, costly healthcare services, digitalization, and heightened user awareness [16]. However, the health insurance sector faces barriers both on the supply and demand side. Insurance companies encounter structural challenges, including a high claim ratio and evolving customer needs, necessitating innovation in products to ensure customer satisfaction [17]. The health insurance companies are also confronted by the misperception of users on the timely claims settlement, transparency in policy guidelines, and usability of technology [18]. On the other side, health insurance users are burdened by surging premiums, insufficient information, lack of proper documentation and plan validation of claims, and unavailability of the network of hospitals. Furthermore, medical costs are rising rapidly in India, with a 14% inflation rate (as of 2021), the highest among Asian countries, and a further 10% rise in inflation rate is anticipated during 2023 [16].

The key advantages offered by DeFi over traditional financial services systems in particular in the insurance sector encompass the reduction of counterparty risks by utilizing smart contracts, allowing transactions to occur without intermediaries like clearing houses. It promotes greater financial inclusion by providing open, permission-less access to services, bypassing the limitations of traditional finance. DeFi also enhances transparency through fully accessible, on-chain transactions, while improving security by minimizing single points of failure. Furthermore, DeFi empowers users to participate in governance via DAO or DEXs, unlike centralized decision-making in traditional financial institutions [19]. A conceptual process flow of DeFi health insurance is illustrated in Fig. 2.

Given the breadth of applying DeFi to the entire health insurance industry, this study focuses specifically on the potential of DeFi for the health insurance sector, with specific reference to the Indian context. Furthermore, due to a lack of knowledge and experience

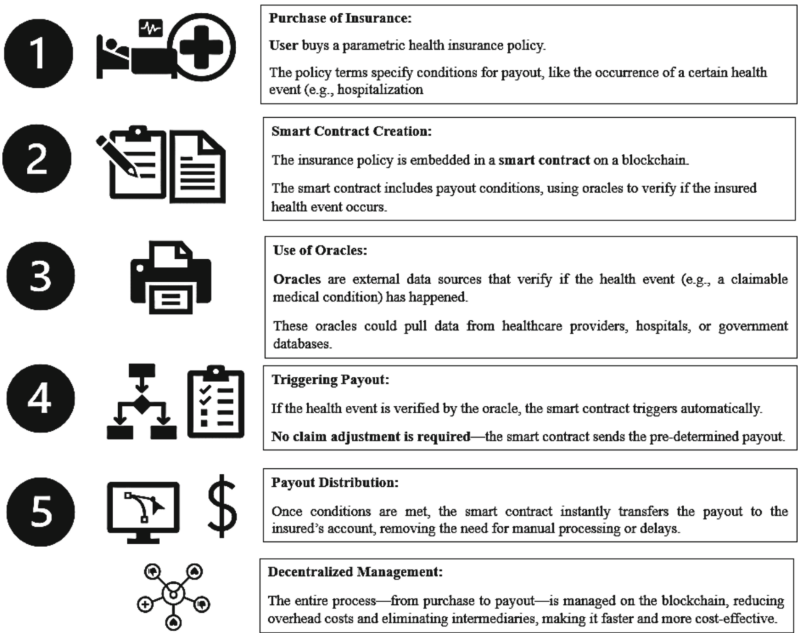


Fig. 2. DeFi Health Insurance Process flow (author compilation adapted from [20])

in understanding DeFi in the context of health insurance, this study seeks to investigate the potential advantages obstacles, and user expectations of DeFi for the health insurance sector in India.

2 Literature Review

Decentralized Finance (DeFi) is a new financial service concept that operates on decentralized blockchain networks without intermediaries such as banks or other financial institutions. Numerous researchers have conducted extensive research on Blockchain Technology and Decentralized Finance in recent years. Examination of the function of blockchain and smart contracts in DeFi intermediation was explored [6]. Studies have provided a thorough examination of the applications and potential benefits of Blockchain Technology in various sectors [21], and discussed the changing aspects of Blockchain Technology and its business applications [12]. Studies also explored the opportunities and difficulties presented by DeFi [1, 11] and applications of DeFi [11].

Since DeFi has grown in popularity in recent years as a means of creating a more inclusive and accessible financial system, researchers have investigated various aspects of DeFi, including decentralized oracles [22], Systematization of Knowledge (SoK), [11] and suggested taxonomies and future research for DeFi [23]. Decentralized oracles are an important element of DeFi [24], supplying secure information feeds for decentralized applications, and analysis of the opportunities and limitations of decentralized oracles in DeFi [24] was carried out. The benefits and drawbacks of Decentralized Oracles

in DeFi were also examined [24]. Various aspects of DeFi in the Systematization of Knowledge (SoK), including its definition, characteristics, architecture, security [9], technical boundaries, and identity integration in DeFi [25] were also investigated.

DeFi has a critical yet potential effect on financial regulation [21]; financial infrastructure [5] and also scope to disrupt traditional financial regulation [21]. In this vein, few studies investigated the implications of Decentralized Finance (DeFi) for the financial sector and its potential for boosting financial inclusion [9], the role of middlemen in DeFi, and the wasted resources associated with their search process [26]. In addition to all these focus areas, researchers have suggested categorization for DeFi [24], the possibility of collaboration between centralized and decentralized financial services [27], and the future impact on financial services [14]. These studies provided useful perspectives on the present state of DeFi research and suggested future re-search and development directions. Financial Infrastructure & Financial Applications applies to the fundamental systems, processes, and technologies that enable financial transactions and facilitate the functioning of financial markets. The review on the development of decentralized finance (DeFi) showed that it has enabled the creation of new financial infrastructure based on blockchain and smart contracts, with the potential to boost financial inclusion and decrease dependence on conventional financial intermediaries [11] Collaboration between centralized and decentralized financial services, as suggested by [28], could offer the "best of both worlds" by incorporating centralized services' efficiency and reliability with DeFi's transparency and accessibility.

Decentralized Finance (DeFi) is reshaping the insurance landscape by offering an alternative to traditional, centralized systems. Unlike conventional insurance, which relies on companies and intermediaries to handle policies, claims, and payout often resulting in slow and expensive processes [29]. DeFi uses blockchain technology to streamline these operations. Through smart contracts, claims and payments are automatically executed when preset conditions are met, eliminating the need for intermediaries and reducing processing times [9, 14]. This decentralized approach also enhances transparency, as blockchain allows users to track how their funds are managed, making premium calculations and claim settlements more accessible [30]. Furthermore, DeFi enables decentralized risk management through collective risk-sharing pools, leveraging real-time data via oracles to improve risk assessments [30]. However, this new model is not without challenges, as regulatory uncertainty and the risk of security breaches, such as hacking, remain significant concerns [11]. Despite these risks, DeFi is expanding insurance accessibility, especially in underserved regions, offering a more efficient and transparent solution compared to traditional models [31].

In India, health insurance is a critical component of healthcare. The impact of the Indian Government's publicly financed health insurance scheme on public sector district hospitals has been extensively studied. Health index scores and insurance policies are important indicators of a population's health status. The current status and challenges of commercial health insurance in India were studied and several issues, including a lack of awareness and affordability [32] were identified. State-level spatiotemporal analysis discovered that health insurance coverage varied significantly across Indian states [33]. From a health financing perspective, the health insurance scheme increased revenue for district hospitals, enabling them to offer higher-quality healthcare services that were

accessible to those who could afford them [34]. The impact of community-based health insurance schemes provided to households especially in the state of Gujarat for catastrophic healthcare expenditures was prominent which reduced such expenditures among poor households [35]. Studies have emphasized the importance of health insurance in India and its potential to improve population health outcomes [33–36]. The examination of factors influencing insurance participation was studied and factors identified included the reimbursement fees, minimal paperwork, swift processing of reimbursements, and a reduced likelihood of claim reduction and denial [37]; policy characteristics (comprehensiveness and variety), agent characteristics (knowledge and accessibility), individual preferences and the availability of alternative options for health financing [36]; investor relationships, the company's reputation, and an affordable premium [38] and financial security, seamless cashless transactions, tax advantages, continuous coverage, and assistance during medical emergencies [39].

Based on the review of the literature on catchphrases such as "Decentralized Finance", it can be inferred that there is limited research evidence in the current body of academic knowledge of DeFi in the health insurance sector. Despite the review providing a general understanding of decentralized finance mechanisms, research in the intersection of decentralized finance and health insurance was found to be limited. In this context, this study seeks to investigate the key aspects surrounding the scope of implementation of decentralized finance technology in the medical insurance process through the lens of the consumers.

3 Methodology

The research utilizes a descriptive methodology to investigate the key determinants influencing the readiness of people to adapt to DeFi in the health insurance sector and identify the advantages and challenges of implementing DeFi in the health insurance sector. Interviews from industry experts and surveys among insurance users were employed in the study.

Experts for interview selection must possess specialized knowledge on the subject [40] and outcomes of interviews can include the aggregation of multi-faceted information [40]. Structured interviews were carried out among the industry experts in the Decentralized Finance system, and the Indian health insurance sector using a predetermined set of fixed questions to capture the advantages and challenges of implementation of DeFi. Representatives from academia, health care, and industry professionals with considerable expertise or knowledge in Decentralized Finance system, Blockchain, and smart contracts were chosen as experts as suggested by [41] using a convenience sampling technique. The responses of the experts identified the key aspects of the advantages and challenges of DeFi in the health insurance sector.

The factors influencing the readiness to use the DeFi by traditional health insurance users were examined by conducting a survey. The instrument constituting the variables accessibility, duration, transparency, surging insurance premiums, and responsiveness of the agents identified from the literature review [37] and from the structured interviews was developed and used to examine the readiness of individuals to adapt to DeFi. Data was collected through an online survey among traditional health insurance policy users

to capture the key factors perceived as important for DeFi as an alternative for healthcare insurance using the snowball sampling technique and 211 responses obtained were used for analysis. The adequacy of the samples for the study was verified using the Gpower3 sample size determination method. Five-point Likert scale (strongly agree to strongly disagree) was used for items of each construct for the participants to rate their level of agreement with each statement. The instrument of the constructs was tested for reliability and validity by employing Cronbach alpha, Composite reliability, Average Variance Explained, and Factor loading. Structural Equation Modelling (SEM), a multivariate analysis was used to examine causal relationships between the constructs and identify the significant variable influencing the readiness of users to adapt to DeFi. The model fit was verified using R square and Goodness of Fit statistics [41].

4 Analysis and Findings

4.1 Summary of Interview

The excerpts gleaned from the interviews on the benefits and challenges of the integration of DeFi into the health insurance sector of India are summarized below.

Advantages of DeFi in Health Insurance:
Transparency and Accuracy: DeFi's blockchain technology ensures tamper-proof records, reducing fraud and improving the accuracy of claims processing.

Cost-efficiency: DeFi eliminates intermediaries, automating processes and potentially reducing premiums for policyholders.

Accessibility: DeFi enhances healthcare accessibility, reaching underserved populations through inclusive financial services.

Efficiency and Customization: DeFi platforms process claims faster, offer transparent policies, and allow customization based on real-time data, aligning insurance coverage with individual needs.

Challenges and Concerns:
Regulatory Challenges: The lack of comprehensive regulations poses risks to policyholders, healthcare providers, and insurers, hindering widespread adoption.

Integration Issues: Integrating DeFi with existing healthcare systems faces technical challenges and requires significant investment in infrastructure and training.

Data Privacy and Security: DeFi platforms demand sensitive personal and medical data, raising concerns about data breaches and cyberattacks.

Technical Expertise: Implementing and maintaining DeFi solutions necessitates advanced technical expertise, posing a barrier to widespread adoption.

4.2 Analysis of Survey

The demographic profile of the traditional health insurance users who participated in the survey showed that out of 211 participants, 118 are males (56%) and 93 are females (44%). In terms of age, 47 individuals were aged 22–30 (22%), 49 were in the 31–40

age range (23%), 14 were between 41–60 years old (7%), and 1 individual was above 60 years (0.5%). On the professional front, 54 individuals were public and private sector employees (26%), 24 were in business (11%), and 33 were retired employees (16%).

The validity and reliability of the instrument comprising of the variables Accessibility, Transparency, Duration, Surging insurance premium, and Responsiveness were examined using the measures Cronbach alpha, Composite Reliability (CR), and Average Variance Extracted (AVE), and the results are presented in Table 1. The ideal level of standardized loadings for reflective indicators is 0.7, but 0.6 is considered to be an acceptable level [43] and the factor loadings of items for each construct were found to be greater than 0.7 in the study. The Cronbach alpha of the constructs was found to be above the threshold value of 0.7 [44] and hence the items fully described the constructs. The Composite reliability (CR) of the constructs was found to be greater than 0.8 [45]. And the reliability of the set of items loaded on each latent construct satisfied the threshold criteria. The Average Variance Extracted (AVE) measures the variance in the items that can be explained by the latent factor and the AVE of the constructs was found to be greater than the threshold value 0.5 [46] and the validity and reliability of the constructs were established.

Warp Partial Least Squares (PLS)-Structural Equation Modelling (SEM) approach was used to estimate the relationship of the variables. The PLS approach was used to examine the measurement model by conducting confirmatory factor analysis which can handle both small and large sample sizes equally. The hypothesis results of SEM (Table 2) showed that hypotheses H1 (Accessibility influenced Readiness to adapt DeFi), and H3 (Transparency influenced Readiness to adapt DeFi), were supported, and the t values for the paths were found to be greater than the cut-off value of 1.96 at a 5% significance level. Hence the constructs Accessibility and Transparency were found to have a significant influence on the readiness of users to adapt to DeFi.

The model fit was examined using R square and Goodness of Fit values. R-square should be higher than the permissible range of 0.5 [42]. The model's prediction accuracy was confirmed with R square value of 0.611 (Fig. 3). The degree to which the given model matches the observed or sample data is referred to as "goodness of fit (Gof)", 'Small = 0.1, medium = 0.25, big = 0.36 [47]. The Tenenhaus Goodness of Fit (Gof) of this study was 0.603, implying that the values are large and desirable.

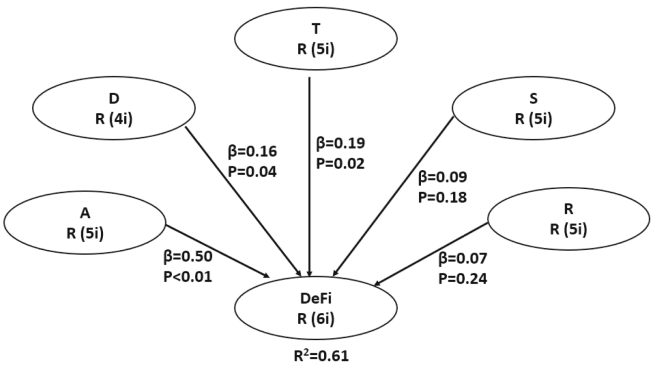


Table 1. Results of Factor Loadings, Cronbach's Alpha, Composite Reliability and Average Variance Extracted of constructs

Constructs	Items	Factor Loadings	Composite Reliability (CR)	Cronbach's alpha	Average Variance Extracted (AVE)
Accessibility	A1	0.719	0.875	0.821	0.584
	A2	0.791			
	A3	0.708			
	A4	0.775			
	A5	0.823			
Duration	D1	0.681	0.872	0.796	0.642
	D2	0.775			
	D3	0.932			
	D4	0.931			
Transparency	T1	0.725	0.932	0.907	0.735
	T2	0.934			
	T3	0.734			
	T4	0.936			
	T5	0.929			
Surging Insurance Premium	S1	0.716	0.842	0.762	0.524
	S2	0.694			
	S3	0.861			
	S4	0.738			
	S5	0.869			
Responsiveness	R1	0.672	0.888	0.837	0.625
	R2	0.712			
	R3	0.900			
	R4	0.880			
	R5	0.899			
Readiness to adapt DeFi	DeFi_1	0.869	0.823	0.742	0.457
	DeFi_2	0.712			
	DeFi_3	0.851			
	DeFi_4	0.712			
	DeFi_5	0.698			
	DeFi_6	0.707			

Table 2. Hypothesis Results of Constructs

Hypothesis	Relationship	Path coefficient	T- ratio	Results
H1	Accessibility→ Readiness to adapt DeFi	0.504	6.052	Supported
H2	Duration→ Readiness to adapt DeFi	0.159	1.748	Not Supported
H3	Transparency→ Readiness to adapt DeFi	0.193	2.137	Supported
H4	Surging insurance premium→ Readiness to adapt DeFi	0.086	0.925	Not Supported
H5	Responsiveness→ Readiness to adapt DeFi	0.067	0.717	Not Supported

It can be inferred from the findings of the study that Transparency [14]; [30] and Accessibility [31] are the key factors perceived as important for using DeFi in the health insurance sector. The blockchain technology underlying DeFi platforms creates a tamper-proof record of all transactions. This could potentially reduce fraud and improve the accuracy of claims processing in the health insurance sector. Medical professionals could have access to the patient's medical history in a secure manner through the DeFi platform. Insurance companies could use blockchain technology to create a tamper-proof record of customer data and transactions, which can help prevent fraud and enhance customer trust. With blockchain technology, policyholders can verify transactions and check the terms and conditions of their policies without intermediaries, reducing the possibility of fraud and corruption DeFi platforms have the potential to enhance the accessibility and inclusivity of health insurance, particularly for those excluded sections of the society by conventional health insurance systems. Through digital wallets and tokens, DeFi facilitates micro-insurance and peer-to-peer (P2P) insurance, ensuring people can access health insurance plans customized to their requirements with improved terms and conditions.

The survey was conducted in Coimbatore city among the urban population using traditional health insurance schemes and the findings revealed that transparency and accessibility were the primary factors influencing the readiness to adopt DeFi. However, the study's limitation lies in the lack of insight into the perspectives of rural and semi-urban health insurance users, which could differ significantly and this presents an area for future research.

DeFi offers several advantages, including greater inclusivity, reduced costs, improved transparency, and the ability to function without geographic or regulatory restrictions [14, 30]. Despite these benefits, challenges such as scalability issues, liquidity constraints, regulatory uncertainties, and the risks tied to poorly programmed smart contracts remain [48]. While DeFi brings innovation, establishing clear regulatory frameworks and robust cybersecurity measures is essential to promote sustainable growth and trust in these emerging models [29].

The survey was carried out in Coimbatore city among the urban population availing traditional health insurance schemes and has identified Transparency and Accessibility as the key factors influencing readiness to use DeFi, however, the perspective of rural and semi-urban health insurance users on using DeFi would differ and this forms the limitations of the study and scope for future research.

It can be surmised that greater inclusivity, lower costs, enhanced transparency, and the ability to operate without geographic or regulatory boundaries are the major benefits. [14, 30] offered by DeFi, the challenges would include scalability, liquidity constraints, regulatory ambiguity, and potential risks associated with incorrectly programmed smart contracts [48]. While DeFi introduces innovation, the need for clear regulatory frameworks and strong cybersecurity measures to ensure sustainable growth and trust in these new models [29] is also vital.

5 Conclusion

This study aimed to explore decentralized finance (DeFi) and its possible applications in the health insurance sector. Over the past decade, blockchain technology, the backbone of DeFi, has found its way into various industries, including finance, real estate, and logistics revolutionizing the way businesses operate.

To explore the potential of DeFi in health insurance, a structured interview was carried out among a diverse group of healthcare experts, including academics and professionals from the industry. The insights from the interview were examined with existing literature, to understand the key factors influencing people's readiness to embrace DeFi in the realm of health insurance. Survey was conducted and the findings reveal that accessibility and transparency are the key determinants influencing their readiness to adopt DeFi in the health insurance sector.

DeFi can provide more transparency by using blockchain to record and verify transactions, policies, and claims, which can improve trust and accountability. DeFi can make health insurance more accessible to excluded populations in India who have limited access to traditional financial services or live in remote or rural areas, who may have difficulty accessing healthcare services. By leveraging blockchain technology, DeFi can create a more accessible and inclusive healthcare financing system.

Based on the findings of this study, DeFi can use blockchain technology to create a decentralized, transparent, interoperable, accessible, and borderless financial system for the health insurance industry. This can lead to a reduction in transaction costs, improved transaction efficiency, and enable peer-to-peer transactions, addressing key challenges in the health insurance industry. Decentralized business models that rely on blockchain technology can enable disbursements for insurance claims and payments of insurance premiums from the insured which would lower the cost of the premium and increase accessibility and transparency through paperless and hassle-free systems, and also a benefit for the insurance company which could reduce their agent and administrative cost ending up in a win-win situation for both the insured and the insurer. With the elimination of intermediaries and the lower administrative requirements, DeFi can enable the public to access health insurance from anywhere in the world. However, certain challenges such as governance risks, smart contract risks, operational security, and, scalability must

be addressed. Research on DeFi indicates that it is still an evolving and distinct technology, and due to its early stage, it is challenging to predict its success in the health insurance sector. Literature findings on technology implementation indicate that it often takes decades or even longer for innovative technologies to realize their maximum potential. Currently, there is a debate about whether DeFi can completely replace centralized finance transactions in the health insurance sector. However, the fundamentals of DeFi are on the right track. Integrating DeFi into the health insurance sector of India could have significant benefits. However, it is important to have proper regulations, risk management protocols, and security measures in place to ensure that DeFi can be integrated successfully and safely. The following implications and policy recommendations may be suggested. Policy-makers should focus on creating a regulatory framework that balances innovation and consumer protection, fostering DeFi adoption. Collaboration between the health insurance sector and the fintech industry can bridge technical gaps and promote innovation in DeFi solutions. Educating consumers, businesses, and professionals about the benefits and risks of DeFi can drive acceptance and facilitate a smooth transition.

References

1. Chen, Y., Bellavitis, C.: Blockchain disruption and decentralized finance. The rise of decentralized business models. J. Bus. Ventur. Insights **13** (2020)
2. Williamson, S.D.: Costly monitoring, financial intermediation, and equilibrium credit rationing. J. Monetary Econ. **18**, 159–179 (1986)
3. Carapella, F., Dumas, E., Gerszten, J., Swem, N., Wall, L.D.: Decentralized Finance (DeFi): Transformative Potential and Associated Risks (2022)
4. Amler, H., Eckey, L., Faust, S., Kaiser, M., Sandner, P., Schlosser, B.: Defining defi: challenges & pathway. In: 2021 3rd Conference on Blockchain Research & Applications for Innovative Networks and Services (BRAINS), pp. 181–184 (2021)
5. Schär, F.: Decentralized finance: On blockchain-and smart contract-based financial markets. FRB of St. Louis Review (2021)
6. Aramonte, S., Huang, W., Schrimpf, A.: DeFi risks and the decentralisation illusion (2021)
7. Antonopoulos, A.M., Wood, G.: Mastering ethereum: building smart contracts and dapps. O'reilly Media (2018)
8. Auer, R., Haslhofer, B., Kitzler, S., Saggese, P., Victor, F.: The technology of decentralized finance (DeFi). Digital Finance **6**, 55–95 (2024)
9. Werner, S., Perez, D., Gudgeon, L., Klages-Mundt, A., Harz, D., Knottenbelt, W. Sok: Decentralized finance (defi). In: Proceedings of the 4th ACM Conference on Advances in Financial Technologies, pp. 30–46 (2022)
10. Michalikova, K.F., Poliakova, A.: Decentralized finance. In SHS web of conferences Vol. 129, p. 03008. EDP Sciences (2021)
11. Ozili, P.K.: Decentralized finance research and developments around the world. J. Banking Financ. Technol. 117–133 (2022)
12. Attaran, M., Gunasekaran, A.: Blockchain-enabled technology: the emerging technology set to reshape and decentralise many industries. Int. J. Appl. Decis. Sci. **12**, 424–444 (2019)
13. Holotiuk, F., Pisani, F., Moormann, J.: The impact of blockchain technology on business models in the payments industry (2017)
14. Schueffel, P.: DeFi: Decentralized finance-an introduction and overview. J. Innovation Manag. **9**, I–XI (2021)

15. Grassi, L., Lanfranchi, D., Faes, A., Renga, F.M.: Do we still need financial intermediation? The case of decentralized finance–DeFi. Qual. Res. Account. Manag. **19**, 323–347 (2022)
16. Dutta, M.M.: Health insurance sector in India: an analysis of its performance. Vilakshan-XIMB J. Manag. **17**, 97–109 (2020)
17. Binny, D.M.G.: Health insurance in India–Opportunities and challenges. Int. J. Latest Technol. Eng., Manag. Appl. Sci. (IJLTEMAS) **6**, 36–43 (2017)
18. Mathur, T., Paul, U.K., Prasad, H.N., Das, S.C.: Understanding perception and factors influencing private voluntary health insurance policy subscription in the Lucknow region. Int. J. Health Policy Manag. **4**, 75 (2015)
19. Jeng, L., Lansang, C. R., Lam, K., Lee, S., Peltekci, T.: Key Elements of an Effective DeFi Framework (2023)
20. HEDERA Homepage. https://hedera.com/learning/decentralized-finance/defi-insurance
21. Zetzsche, D.A., Arner, D.W., Buckley, R.P.: Decentralized finance (DeFI). J. Financ. Regul. **6**, 172–203 (2020)
22. Sriman, B., Kumar, S.G.: Decentralized finance (DeFI): the future of finance and DeFI application for Ethereum blockchain based finance market. In: 2022 International Conference on Advances in Computing, Communication and Applied Informatics (ACCAI), pp. 1–9 (2022)
23. Jensen, J.R., von Wachter, V., Ross, O.: An introduction to decentralized finance (DeFI). Complex Syst. Inform. Model. Q. **26**, 46–54 (2021)
24. Perera, S., Nanayakkara, S., Rodrigo, M.N.N., Senaratne, S., Weinand, R.: Blockchain technology: Is it hype or real in the construction industry?. J. Ind. Inf. Integr. **17** (2020)
25. Harwick, C., Caton, J.: What's holding back blockchain finance? On the possibility of decentralized autonomous finance. Q. Rev. Econ. Finance **84**, 420–429 (2022)
26. Momtaz, P.P.: How Efficient is Decentralized Finance (DeFi)?. SSRN 4063670 (2022)
27. Meegan, X.: Identifying key non-financial risks in decentralised finance on ethereum blockchain. MIP Politecnico di Milano (2020)
28. Meegan, X., Koens, T.: Lessons learned from decentralised finance (DeFi) (2021)
29. Feng, R. .: DeFi Insurance. In: Decentralized Insurance: Technical Foundation of Business Models, pp. 217–256. Cham: Springer International Publishing (2023)
30. Bekemeier, F.: A primer on the insurability of decentralized finance (DeFi). Digital Finance **5**, 643–687 (2023)
31. Piesse, D..: Decentralized Finance (DeFi)–Risks and opportunities for the insurance industry. International Insurance Society Report (2021)
32. Prasad, S., Ghosh, I.: Commercial Health Insurance in India–Status and Challenges (2020)
33. Kamath, R., Lakshmi, V., Brand, H. Health index scores and health insurance coverage across India: a state-level spatiotemporal analysis. Clinical Epidemiology and Global Health, vol. 18 (2022)
34. Prinja, S., et al.: Impact of India's publicly financed health insurance scheme on public sector district hospitals: a health financing perspective. Lancet Reg. Health-Southeast Asia **9** (2023)
35. Ranson, M.K.: Reduction of catastrophic health care expenditures by a community-based health insurance scheme in Gujarat, India: current experiences and challenges. Bull. World Health Organ. **80**, 613–621 (2002)
36. Tati, R.K., Baltazar, E.B.B.: Factors influencing the choice of investment in life insurance policy. Theor. Econ. Lett. **8**, 3664 (2018)
37. Shobiye, H.O., Dada, I., Ndili, N., Zamba, E., Feeley, F., de Wit, T.R.: Determinants and perception of health insurance participation among healthcare providers in Nigeria: a mixed-methods study. Plos One **16** (2021)
38. Gajula, S., Dhanavanthan, P.: Exploration on consumer's perception and buying behavior of health insurance policies in Hyderabad city. Int. J. Eng. Adv. Technol. **9** (2019)
39. Mathur, T., Das, G., Paul, U.K.: Convenience, satisfaction, and post-purchase behavior in India's health insurance market. J. Glob. Mark. **29**, 218–232 (2016)

40. Von Soest, C.: Why do we speak to experts? Reviving the strength of the expert interview method. Perspect. Polit. **21**, 277–287 (2023)
41. Maestas, C.: Expert Surveys as a Measurement Tool: Challenges and New Frontiers. In: The Oxford Handbook of Polling and Survey Methods, ed. Lonna Rae Atkeson and R. Michael Alvarez, pp. 583–608. Oxford: Oxford University Press (2018)
42. Hair, J.F, Black, W.C., Babin, B.J., Anderson, R.E.: Multivariate Data Analysis, illustrated, revised', Pearson Education Limited, USA (2013)
43. Fornell, C., Larcker, D.F.: Evaluating structural equation models with unobservable variables and measurement error. J. Mark. Res. **18**, 39–50 (1981)
44. Thompson, R., Barclay, D.W., Higgins, C.A.: The partial least squares approach to causal modeling: personal computer adoption and use as an illustration. Technol. Stud. Spec. Issue Res. Methodol. **2**, 284–324 (1995)
45. Hair, J.F., Black, W.C., Babin, B.J., Anderson, R.E., Tatham, R.L.: Multivariate data analysis, 6th Edition, Pearson-Prentice Hall, Upper Maddle River (2006)
46. Hair, J.F., Anderson, R.E., Tatham, R., Black, W.: Multivariate Data Analysis With reading, Englewood Cliffs. Prentice Hall, NJ (1998)
47. Diamantopoulos, A., Siguaw, J.A.: Introducing LISREL: A guide for structural equation modeling. Sage Publications, California (2000)
48. Weingärtner, T., Fasser, F., Reis Sá da Costa, P., Farkas, W. . Deciphering DeFi: A comprehensive analysis and visualization of risks in decentralized finance. J. Risk Financ. Manag. **16**, 454 (2023)

Analysis of Security with Countermeasures in IoT Based Applications – A Survey

V. Sini[1(✉)] and Sridaran Rajgopal[2]

[1] Smt. Devkiba Mohansinhji Chauhan College of Commerce and Science, U.T of DNH & DD, Silvassa, India
sini.v25@gmail.com
[2] Faculty of Computer Application, Marwadi University, Rajkot, India
Sridaran.rajagopal@marwadieducation.edu.in

Abstract. The Internet of Things (IoT) has gained significant popularity in this decade due to its capability to facilitate data exchange between devices and systems via the Internet. Since data plays a crucial role in IoT Applications, it is necessary to secure them from the threats. There exist many surveys on security threats of IoT applications but most of them either fail to produce the necessary integration of technologies or lightweight solutions. In the proposed survey described in this paper, it is attempted to classify security threats with their countermeasures from the available literature so that future solutions can be thought of layer-wise, which will be essentially lightweight. It is believed that the proposed survey would certainly help IoT application designers to a greater extent.

Keywords: IoT · Security threats · IoT attacks · Countermeasures · Sensors · Actuators · IoT applications · IoT challenges

1 Introduction

The Internet of Things (IoT) [1] refers to a network of interconnected physical devices and objects embedded with sensors, software, and various technologies. These components allow the devices to collect and exchange data over the internet with other systems and devices. This technology enables the extensive collection of data from connected devices, which can be analyzed to gain valuable insights and improve efficiency in sectors like manufacturing, healthcare, agriculture, and transportation. IoT devices also support task automation, remote system control, and real-time condition monitoring. As IoT continues to evolve, the number of connected devices is expected to grow rapidly, with ongoing advancements in their capabilities. However, despite its significant advantages, IoT also raises important concerns about data privacy and security that must be carefully managed in any IoT implementation.

The IoT is a sophisticated network composed of multiple elements that collaborate to facilitate data exchange between devices and systems. Figure 1 below illustrates the components of typical IoT Applications. Sensors are essential components of IoT devices, responsible for collecting data on various physical parameters like temperature,

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
S. Rajagopal et al. (Eds.): ASCIS 2024, CCIS 2428, pp. 121–131, 2025.
https://doi.org/10.1007/978-3-031-86302-8_8

humidity, pressure, motion, and other environmental conditions [23]. IoT devices use a range of connectivity options, including Wi-Fi, Bluetooth, cellular networks, and other wireless communication technologies, to transmit data to other devices or the cloud. Processing and analysis of data extracted by the sensors are carried out in the cloud or edge computing devices. Storage solutions are necessary to hold the collected data for further analysis and processing. Actuators, such as motors, valves, and switches, are components that execute actions considering the insights gained based on the sensor data. A scalable, secure, and dependable platform for processing and storing the massive amounts of data collected by IoT devices is provided by cloud computing. Edge computing, a distributed computing model, which permits IoT devices to analyse and process data nearby, reducing the dependency on transmitting large amounts of data to the cloud. The massive amount of data gathered by the IoT devices is examined using Machine learning algorithms. It can also be used to identify patterns and perceptions that can enhance efficiency and optimize processes. These components together create a scalable and robust system that enables the sharing of data between devices and systems, which leads to a growth in automation, productivity, and effectiveness across numerous industries.

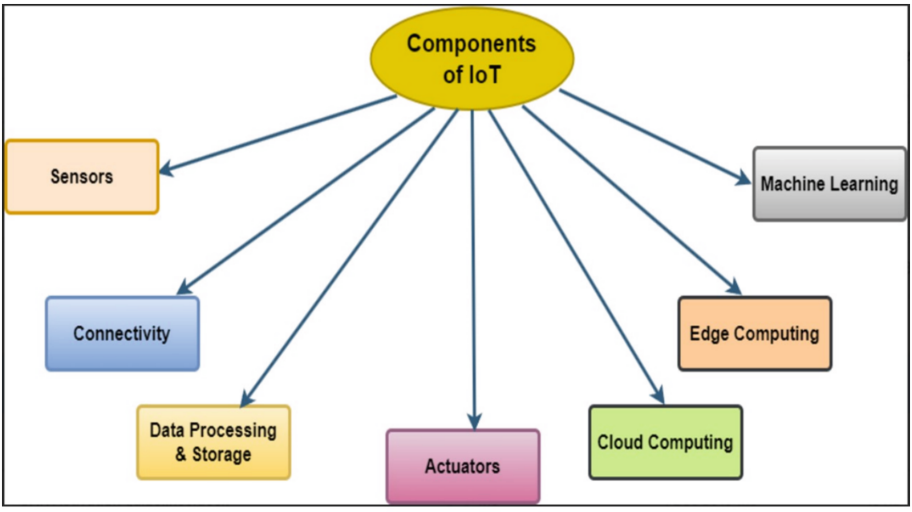


Fig. 1. Components of a Typical IoT Applications

The IoT hitches the "real world" with the "digital world", permitting for flexible communication between objects, peoples, and processes. Physical items can be remotely controlled and activated with the proper use of an internet architecture, enabling secure integration between the physical and digital realms.

IoT applications are revolutionizing various societal functions. A key concept, "smart cities" or "Cities of the Future," leverages IoT technology to improve urban infrastructure, aiming to make cities more efficient, livable, and cost-effective. This also supports economic development and environmental sustainability.

Smart homes are a popular IoT application that allows control over household items like lighting, alarms, and water flow. Advanced security systems enhance home safety. Homeowners can manage and oversee various aspects of their homes through smartphones, tablets, and computers. For instance, if you forget to turn off your oven, you can easily do so using your phone. A smart home enables centralized control of all household devices.

IoT also has significant applications in healthcare, including chronic illness management, ingestible sensor tracking, mobile health, smart hospitals, and remote monitoring of health. The main advantage of remote monitoring of health systems is that they allow individuals to live normal lives while being continuously monitored for health issues. These systems are especially useful for monitoring patients with diabetes, heart conditions, neurological issues, and elderly individuals at risk of falling [8]. In the realm of military technology, IoT has numerous critical applications. Wireless network technologies such as ZigBee and GSM networks are used for tracking and monitoring the health of armed personnel. Military uniforms and helmets embedded with sensor devices can provide data on the positions and vital signs of soldiers. This allows the command center to respond swiftly to save lives. Intelligent military gear incorporates sensors such as temperature sensors, pulse rate monitors, oxygen level sensors, accelerometers, ECG sensors, and movement sensors [9].

IoT and 5G are distinct technologies, yet they are interrelated and can enhance each other's functionalities. 5G, the latest generation of cellular network technology following 4G LTE, offers enhanced capacity, reduced latency (reduced delay), and quicker data transfer speeds compared to previous generations. These improvements are necessary for the effective implementation of certain IoT applications that demand real-time data processing and speedy reaction periods, like in remote surgery, smart traffic management, and industrial automation. The increased capacity of 5G networks enables them to ensure a dependable and robust network for IoT implementations, to accommodate the enormous number of IoT devices that are projected for the future. IoT devices, in turn, can benefit 5G network users by transmitting data more efficiently and effectively due to the enhanced connectivity. In summary, IoT and 5G are transformative technologies shaping the future of connectivity, automation, and data-driven applications. While IoT focuses on connecting devices and collecting data, 5G provides the high-speed, low-latency, and high-capacity network infrastructure necessary to maintain the extensive ecosystem of IoT devices and enable innovative new services [21].

1.1 Security Issues in IoT

Because of the increasing number of IoT devices and the rise in cyberattacks, ensuring robust IoT security has become critical. IoT developers must invest considerable time and effort to guarantee these device's security. The Internet's evolution into the "Internet of Things" has introduced complexities such as information fraud, data loss, and integrity issues. These difficulties nowadays involve vehicles, service customers, animals, supply chain links, and electronic devices. Agriculture, production, transport, and energy distribution are some domain applications of IoT [6].

1.2 Security in Handling of Threats

As per current studies, by 2025 the number of IoT devices could reach 75.44 billion [25]. With the rapid development of IoT devices, security attacks against the IoT domain will increase further and impact serious problems in IoT devices, networks, and applications. IoT devices gather large numbers of personal data, consequently, it is vital to prevent it from being misused and to protect it from attacks. The IoT domain contains a massive diversity of user devices, hardware, firmware, software applications, networks, etc. The IoT threats are likely to cause major shocks and problems if they are not handled properly. These possibilities will be equally affected both personally, and in monopolistic and industrial situations. Increasing Cyber-attacks, Data breaches, and lack of privacy are the major consequences that may occur. Unfortunately, most, of the IoT applications and devices are not designed to grip security and privacy attacks and, it increases a lot of problems like authentication, confidentiality, and integrity of data. For that reason, an active mechanism is tremendously required to ensure the security of the IoT domain.

The various IoT applications and how IoT and 5G are connected are explained in the first part of the paper. The second part includes the surveys related to the proposed work. The layer-wise analysis of security threats and countermeasures is included in the third part of the paper. Later included the discussions, conclusions, and future directions.

2 Related Work

In 2020, Anand S. et al. [18] led a widespread investigation of privacy issues and security risks in IoT applications. They recognized the main weakness together with device manipulation, data breaches, and unauthorized access and pointed out common issues like weak authentication, inadequate encryption, and deficient update protocols. Some mitigation techniques like robust encryption, regular software updates, advanced authentication protocols, blockchain for security, decentralized networks, and machine learning for threat prediction are proposed in their research. This research delivers a vital reference for understanding privacy and security in IoT, highlighting the necessity for integrating these advanced techniques for making the IoT applications more secure.

N.M. Kaire et al. [19] observe progressions in IoT threat detection and cybersecurity, prominence of dangerous vulnerabilities like data integrity, device authentication, and network security. Challenges in securing diverse, maintaining privacy, resource-constrained devices, and implementing robust security protocols are reported in their research. Their study deliberate edge-cutting techniques like AI, machine learning for anomaly detection, and blockchain for tamper-proof records.

P. Williams et al. [20] stretch a detailed introduction to IoT security, architecture, and the issues across network layers. Their research shows to identify vulnerabilities and propose countermeasures for secure communication, data encryption, and the authentication of the devices. Some real-world attacks like DoS, MitM, and malware infiltration are pointed in this study. The authors promote advanced technologies such as AI and machine learning for better threat identification and response.

Haque. A. B. et al. [22] critically inspect the current literature on IoT security and privacy providing perspectives on the state of IoT cybersecurity. By investigating suggested solutions, the paper gives a detailed summary of the art of addressing these issues. A

comprehensive range of approaches from encryption techniques and secure communication protocols to progressive threat detection methods using AI are also mentioned in this manuscript. The paper recognized key areas for future research, highlighting the need for continued innovation to retain step developing threat landscape. This progressive viewpoint is vital for the advancement of IoT technology, as it guides researchers to new or underdeveloped paths that might yield significant improvements. They emphasize the technological difficulties faced by resource-constrained IoT systems, highlighting the need for lightweight security measures. Their work guides future research and development, aiming to enhance IoT security and resilience without compromising device performance.

2.1 Gaps in Existing Related Work

The surveys mentioned above describe the architecture of IoT and general security challenges that occur in IoT applications. Countermeasures for various threats in IoT applications are also mentioned in many surveys. The numerous threats that affect different IoT layers are very significant. The above surveys lack the proper analysis of layer-wise security challenges and countermeasures. This forms the motivation behind the development of the proposed survey.

3 Classification of Security Challenges vs. Countermeasures

The word “things” in IoT refers to a network of physical objects or individuals equipped with software, electronics, networks, and sensors that enable them to gather and exchange data. The core objective of IoT is to provide internet connectivity for everyday devices like mobile phones, computers, and tablets. The architecture of the IoT provides a framework outlining the physical components, network setup, operational protocols, and data formats used. The perception layer, network layer, and application layer comprise the three-layer model, which is a well-acknowledged framework in IoT architecture [24].

Figure 2 illustrates challenges and solutions for IoT security across the layers of perception, network, and application. The IoT architecture’s physical layer is called the perception layer. In these systems, embedded systems and sensors are commonly utilized to collect extensive data based on specific requirements. Additionally, this includes edge devices, sensors, and actuators that interact with the environment. The sensors in the perception layer are designed to detect and collect environmental data. The perception layer identifies specific spatial parameters or nearby intelligent objects/things. In the IoT framework, the network layer functions as the primary hub for data processing and transmission. This layer includes components including the network management centers, communication and Internet convergence networks, and information centers, for intelligent processing. Its primary function is to facilitate data processing and transmission gathered from the layer of perception [5]. Application-specific services must be provided to the user by the application layer. A few examples of the various uses for IoT are smart cities, smart homes, and smart health care [4]. The application layer represents a crucial convergence point where IoT technology integrates deeply with industrial needs, paving the way for the development of intelligent industries that can profoundly impact society,

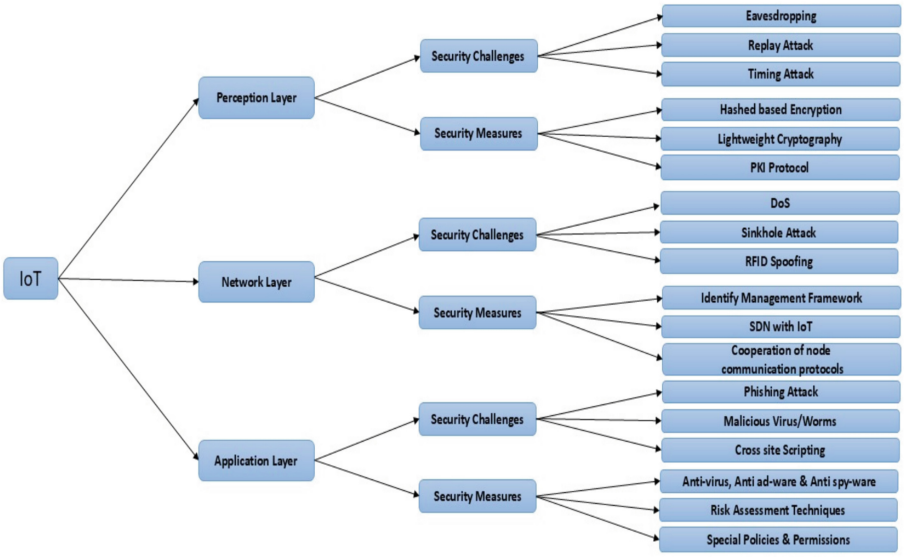


Fig. 2. Classifications of Security Challenges and Countermeasures.

akin to the social division of labor among individuals. Here, the specific requirements of industries merge with IoT capabilities to achieve extensive intellectualization and efficiency. Meanwhile, connectivity between computers, network infrastructure, and smart devices is managed by the network layer. It is essential for processing and transmitting sensor data, supporting the operational backbone of IoT applications [4].

3.1 Layer-Wise Analysis

The different threats are classified into Perception Layer, Network Layer, and Application Layer which are discussed in the subsequent sections below:

3.1.1 Perception Layer

Hardware assaults are frequently the biggest dangers at the perception layer [10] especially when they target sensors like RFID [11], Zigbee, and wireless sensor networks (WSNs). These devices are susceptible to physical attacks, especially if they remain stationary in one location for extended periods. IoT gadgets including wearable health monitoring, smart TVs, smart appliances, and security systems, assemble vast amounts of information about us. However, this wealth of information also makes them potential targets for hackers who may exploit or access this data for unethical purposes.

An eavesdropping attack involves an intrusion where someone attempts to intercept the data transmitted by devices. These attacks exploit insecure transmission methods to gain access to both sent and received data, commonly referred to as snooping or sniffing attacks [12]. Replay attacks occur when unauthorized parties eavesdrop on communications from the sender to the recipient, stealing legitimate data from the sender. These attacks often target authentication procedures to compromise validity certificates. Timing attacks are another type of attack that deliberately introduces unnecessary delays between interactions of IoT devices. Their primary objective is to take advantage of these delays to obtain entry to sensitive information [14].

3.1.2 Network Layer

Sinkhole attacks, DoS attacks, and RFID spoofing are notable examples of network layer attacks in IoT environments. Sinkhole attacks occur within networks when a malicious actor compromises a network node to initiate disruptive actions like DoS attempts, packet manipulation, or dropping [14]. DoS attacks involve overwhelming an IoT device accompanied by an overload of service requests, leading to delays in service or outright interruptions for legitimate users. These attacks aim to deny service to genuine users by inundating devices or network resources with excessive traffic [14]. RFID spoofing attacks target IoT systems by infiltrating and injecting malicious data into the network. It also leads to IP spoofing and exploits vulnerabilities in IoT networks and disrupts operations.

3.1.3 Application Layer

Software attacks are the main cause of vulnerabilities in IoT-enabled applications because these applications are defined at the application layer. The attackers copy genuine users to obtain email addresses, usernames, and passwords is known as a phishing attack. Users will be attracted by the fraudulent website, and attackers can steal their authorizations once they log in. The worms which are self-propagating malware which can compromise the user data and are a significant threat to IoT devices are the Malicious worms. Trojan horses, worms, and other malware are capable of stealing and alteration of the sensitive data. To fight against these attacks effective measures like virus detection systems and firewalls are required. Cross-site scripting (XSS) is a type of injection attack in which the attackers add scripts to trusted websites viewed by the other users. It permits attackers to hypothetically expose users to malicious activities and manipulate the website content [15].

3.2 Layer Wise Countermeasures

The countermeasures for the threats discussed above are summarized below:

3.2.1 Perception Layer

The implementation of security controls in IoT architecture is very critical, particularly in the perception layer where data transmission and data collection occur. Robust encryption mechanisms play a vital role in safeguarding against malicious attacks. There is a

security method to transform messages into ciphertext called Hash-based encryption, it ensures that only approved users can use a decryption key for the original information. This encryption method generates keys based on the length of the message, it is very challenging to make the decryption key. Incorporating encryption with authentication protocols, permission management, and intrusion detection improves security within the perception layer of the IoT architecture and looks like a protocol mechanism similar to Public Key Infrastructure (PKI). Implementing multiple security mechanisms separately further fortifies these defenses. Within IoT networks, where multiple interconnected nodes form a network responsible for secure message delivery, trust issues are addressed through cryptographic techniques, including lightweight cryptography tailored for resource-constrained devices. Lightweight cryptographic algorithms such as symmetric-key encryption are specifically designed to ensure secure communication across IoT networks, demonstrating effective data transfer security [16].

3.2.2 Network Layer

The safeguard the Network Layer from hacker attacks, several preventive measures are crucial. Each device must undergo authentication before communicating with others, ensuring the legitimacy of transmissions. An identity management framework is recommended to manage device authentication effectively. Within the IoT environment, components include sensors embedded in objects (the environment), which collect data transmitted to computers for processing. The receiver then interprets the processed data, and the network facilitates data transmission between sensors, computers, and receivers [12].

Software-defined networking (SDN) is commonly used for managing network security across various sectors, including businesses, smart homes, and healthcare systems. Unlike traditional networking where switches and routers integrate data and control planes, SDN separates them. The centralized control plane, operated via software, directs traffic flow, while the hardware-based data plane handles actual data forwarding. SDN enhances network security by enabling comprehensive traffic monitoring and identification of malicious activities. It can isolate compromised nodes, preventing internal network attacks.

Cooperative node communication protocols, like security-conscious ad-hoc networks, are crucial for identifying and mitigating improper node behavior that could jeopardize network integrity. These protocols incorporate components such as monitors, reputation systems, path managers, and trust managers. The trust manager plays a pivotal role by gathering information on network nodes and detecting suspicious activities. When anomalous behavior is detected, the trust manager issues ALERT messages to notify all nodes within its scope, information about packet loss, and the attacker's node address, among other things.

3.2.3 Application Layer

To protect the application layer from malicious attacks in IoT environments, several security measures are essential. Establishing specific rules and permissions is crucial for regulating access to the IoT framework. Antivirus, anti-adware, and anti-spyware technologies play vital roles in guaranteeing the security, integrity, privacy, and dependability of IoT systems. Conducting regular risk assessments is also important for identifying potential vulnerabilities within the application layer. Another critical measure is that the system device's firmware should be updated. This practice strengthens security by addressing known vulnerabilities and improving overall system resilience [10]. These combined efforts help mitigate risks and enhance the protection of IoT applications against various types of cybersecurity threats. Consolidation of various security threats and countermeasures discussed here are available in Table 1 of the appendix.

4 Discussion

Amongst many different threats, Eavesdropping, Replay Attacks, and Timing Attacks are the main ones that affect the perception layer. The security challenges occurring in the network layer are Dos, Sinkhole Attack, and RFID Spoofing. Phishing Attacks, Malicious viruses/worms, and Cross-site scripting are the security challenges that affect the application layer. The various threats as well as countermeasures have been consolidated in Table 1 of the Appendix.

5 Conclusion and Future Work

In this paper, an attempt has been made to analyze and present the threats of various IoT applications classified layer-wise. Since IoT technology enables data exchanges between the devices and computer systems, they are vulnerable to many threats. Hence it becomes essential to learn about various threats with their countermeasures. It is believed that the survey presented here would greatly contribute for building threat-free IoT applications. Our future work would involve developing a framework of lightweight solutions for the most common threats on IoT applications.

Appendix

Table 1. Consolidation of Security Threats and Solutions.

IoT Layers	Security Threads	Nature	Solutions
Perception Layer	Eavesdropping	Passive	Hashed based Encryption
	Replay Attack	Passive	Lightweight Cryptography, PKI Protocols
	Timing Attack	Passive	Consent Time Algorithm

(*continued*)

Table 1. (*continued*)

IoT Layers	Security Threads	Nature	Solutions
Network Layer	DoS	Vary	Traffic Filtering, Network Filtering
	Sinkhole Attack	Vary	Encryption Protocols, Network Segmentation
	RFID Spoofing	Active and Passive	Signal Monitoring, Encryption
Application Layer	Phishing Attack	Passive	Firmware, Software updates
	Malicious virus/worms	Active	Authentication, Firmware, Software updates
	Cross-site Scripting	Active	Security Testing, Security Policies

References

1. Patel, K., Patel, S., Scholar, P., Salazar, C.: Internet of Things-IOT: definition, characteristics, architecture, enabling technologies. Appl. Future Challenges (2016)
2. Ezechina, M.A., Okwara, K.K.: Ugboaja3 1,2,3 Department of Computer Science, Akanu Ibiam Federal Polytechnic, Uwana P'. Akanu Ibiam Federal Polytechnic, Uwana P.M.B, vol. 1007
3. Garg, H., Dave, M.: Securing IoT devices and securely connecting the dots using REST API and middleware. In: 2019 4th International Conference on Internet of Things: Smart Innovation and Usages (IoT-SIU), pp. 1–6 (2019). https://doi.org/10.1109/IoT-SIU.2019.8777334
4. Litoussi, M., Kannouf, N., El Makkaoui, K., Ezzati, A., Fartitchou, M.: IoT security: challenges and countermeasures. Procedia Comput. Sci. **177**, 503–508 (2020). https://doi.org/10.1016/j.procs.2020.10.069
5. Wu, M., Lu, T.-J., Ling, F.-Y., Sun, J., Du, H.-Y.: Research on the architecture of Internet of Things. In: 2010 3rd International Conference on Advanced Computer Theory and Engineering (ICACTE), Chengdu, pp. V5-484–V5-487 (2010). https://doi.org/10.1109/ICACTE.2010.5579493
6. Davis, M., Abey, R.: Security in IoT (2022)
7. Ramson, S.R.J., Vishnu, S., Shanmugam, M.: Applications of Internet of Things (IoT) – an overview. In: 2020 5th International Conference on Devices, Circuits and Systems (ICDCS), Coimbatore, India, pp. 92–95 (2020). https://doi.org/10.1109/ICDCS48716.2020.243556
8. Ramson, S., Moni, D.J.: A case study on different wireless networking technologies for remote health care. Intell. Decis. Technol. **10**(4), 353–364 (2016)
9. Iyer, B., Patil, N.: IoT enabled tracking and monitoring sensor for military applications. Int. J. Syst. Assur. Eng. Manag. **9**(6), 1294–1301 (2018)
10. Rao, T.A., Haq, E.U.: Security challenges facing IoT layers and its protective measures. Int. J. Comput. Appl. **975**, 8887 (2018)
11. Kannouf, N., Douzi, Y., Benabdellah, M., Azizi, A.: Securityon RFID technology. In: International Conference on Cloud Technologies and Applications (CloudTech), IEEE, pp. 1–5 (2015)
12. Burhan, M., Rehman, R.A., Khan, B., Kim, B.S.: IoT elements, layered architectures and security issues: a comprehensive survey. Sensors **18**(9), 2796 (2018)

13. Deogirikar, J., Vidhate, A.: Security attacks in IoT: a survey. In: 2017 International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), Palladam, India, pp. 32–37 (2017). https://doi.org/10.1109/I-SMAC.2017.8058363
14. Gautam, S., Malik, A., Singh, N., Kumar, S.: Recent advances and countermeasures against various attacks in IoT environment. In: 2019 2nd International Conference on Signal Processing and Communication (ICSPC), Coimbatore, India, pp. 315–319 (2019). https://doi.org/10.1109/ICSPC46172.2019.8976527
15. Andrea, I., Chrysostomou, C., Hadjichristofi, G.: Internet of Things: security vulnerabilities and challenges. In: 2015 IEEE Symposium on Computers and Communication (ISCC), IEEE, pp. 180–187 (2015)
16. McKay, K.A., Bassham, L., Turan, M.S., Mouha, N.: Report on Lightweight Cryptography. National Institute of Standards and Technology, USA (2017)
17. RajeshKandaswamy, D.F.: Blockchain-based transformation. https://www.gartner.com/en/doc/3869696-blockchain-based-transformation-a-gartner-trend-insight-report/. Accessed 5 Jun 2018
18. Anand, S., Sharma, A.: Assessment of security threats on IoT based applications. Materials Today: Proceedings (2020). https://doi.org/10.1016/j.matpr.2020.09.350
19. Karie, N.M., Sahri, N.M., Haskell-Dowland, P.: IoT threat detection advances, challenges and future directions. In: 2020 Workshop on Emerging Technologies for Security in IoT (ETSecIoT), Sydney, NSW, Australia, pp. 22–29 (2020). https://doi.org/10.1109/ETSecIoT50046.2020.00009
20. Williams, P., Dutta, I., Daoud, H., Bayoumi, M.: Security aspects of Internet of Things – a survey. In: 2020 IEEE 6th World Forum on Internet of Things (WF-IoT), New Orleans, LA, USA, pp. 1–6 (2020). https://doi.org/10.1109/WF-IoT48130.2020.9221429
21. Chettri, L., Bera, R.: A comprehensive survey on Internet of Things (IoT) toward 5G wireless systems. IEEE Internet Things J. **7**(1), 16–32 (2020). https://doi.org/10.1109/JIOT.2019.2948888
22. Haque, A.B., Tasmin, S.: Security threats and research challenges of IoT - a review. J. Eng. Advancements **1**(04), 170–82 (2020). https://doi.org/10.38032/jea.2020.04.008
23. Alam, S., Siddiqui, S.T., Ahmad, A., Ahmad, R., Shuaib, M.: Internet of Things (IoT) Enabling Technologies, Requirements, and Security Challenges, 8–981–15–0694–9 (eBook) https://doi.org/10.1007/978-981-15-0694-9
24. Kakkar, L., Gupta, D., Saxena, S., Tanwar, S.: IoT architectures and its security: a review. In: Goyal, D., Gupta, A.K., Piuri, V., Ganzha, M., Paprzycki, M. (eds.) Proceedings of the Second International Conference on Information Management and Machine Intelligence. Lecture Notes in Networks and Systems, vol. 166. Springer, Singapore (2021). https://doi.org/10.1007/978-981-15-9689-6_10
25. Van Tanh, N., Tri, N.Q., Giang, N.L., Duy, T.-L.: A Solution to improve the security of the Internet of Things network with lightweight encryption methods. J. Phys. Conf. Ser. **1933**(1), 012042 (2021). https://doi.org/10.1088/1742-6596/1933/1/012042

Efficient Train Traffic Prediction Using Support Vector Machines with Advanced Preprocessing Techniques

C. Radhika[1(✉)] and D. Kerana Hanirex[2]

[1] Department of Computer Applications, Bharath Institute of Higher Education and Research (BIHER), Chennai, India
raaddhika66@gmail.com

[2] Department of Computer Science, Bharath Institute of Higher Education and Research (BIHER), Chennai, India
keranahanirex.cse@bharathuniv.ac.in

Abstract. Train delays have a big impact on customer happiness and railway management. Predicting train delays accurately may enhance both passenger satisfaction and operational effectiveness. Although many other algorithms have been used to solve this issue, Support Vector Machines (SVM) in conjunction with strong preprocessing methods may provide better results. This study investigates how several preprocessing methods, namely the Modified Z-score method (MZ-Score), Interquartile Range (IQR) method, and Z-score method, affect the effectiveness of SVM in predicting train delays. This study aims to improve train delay prediction models' recall, accuracy, and precision by using powerful preprocessing methods in conjunction with SVM. Testing the suggested system using the Train_delay_Prediction.csv dataset shows that combining SVM with the MZ-score technique yields better results in terms of accuracy, precision, and recall.

Keywords: Train Delay Prediction · Support Vector Machine · Z-score Method · Interquartile Range · Modified Z-score Method · Data Preprocessing · Machine Learning

1 Introduction

Train delays affect passenger happiness and operational efficiency, which is a major problem for railway operations globally. The main obstacle facing the railway industry is operating trains on time. Therefore, forecasting train delays can aid in better train movement and more accurate estimations of a train's running condition, which in turn helps railway sections make decisions regarding other trains and communicate the same communication to travelers. Train delays can occur for various causes, the most common being commuter traffic, maintenance issues, malfunctions, inclement weather, and train movement tactics. The departure time is the primary factor that determines a train's delay; standard techniques of measuring delay take this into account alone. However,

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
S. Rajagopal et al. (Eds.): ASCIS 2024, CCIS 2428, pp. 132–144, 2025.
https://doi.org/10.1007/978-3-031-86302-8_9

two more elements affect train delays in addition to departure time: the distance/space and the amount of time between two trains [1].

Predicting train delays is essential to improving the effectiveness and dependability of railway operations. Pupils may make more educated travel selections with accurate projections, which lessens inconvenience and raises consumer happiness. Predicting delays can help railway operators to get the most out of allocation and management in order to have a performing timetable with reduced disruption. By projecting delays, the operators may also employ measures to address issues earlier - for example by adding staff or making changes in maintenance schedules- resulting in cost savings of a more reliable service. Additionally, real-time delay forecasts would also help a better inter-modality with other modes of transport making the public transportation system more integrated and efficient.

Reliable predictions of train journey times not only improve the quality of passenger services but also are essential for real-time railway traffic control that may potentially reduce knock-on delays from spreading [2]. The higher the prediction accuracy when it comes to predicting train delays, the better resource allocation can be achieved and combined with more effective scheduling will increase service quality. Due to the presence of noise, outliers, and different distributions in real-world data, traditional techniques of predicting train delays sometimes encounter difficulties. Because ML (Machine Learning) approaches can handle high-dimensional data and establish optimal decision limits, they have shown promise in solving these difficulties.

Because of the ability to handle big data and complex variables, ML approaches have a significant impact on predicting train delays. Existing methods are frequently based on deterministic models or simplified statistical approaches, which may not take into account the complex operation of railway systems. ML algorithms, on the other hand can access different types of data e.g., historical train schedules, weather conditions and track maintenance records—thus uncovering more complex patterns that lead to delays. They can process and learn from vastly complex data sets, making their end prediction much more precise to that of other outdated methods. The novelty in this research article is the combination of MZ-Score + SVM. In train delay prediction this combination has not been implemented.

This reserch article has been divided into following sections. Section 1 represents a detailed introduction part, Sect. 2 represents a detailed survey on latest papers, Sect. 3 represents proposed model, Sect. 4 represents results obained by proposed work compaed with existing algorithms and finally Sect. 5 represents conclusion and future work.

2 Literature Survey

The enhanced architecture of the train transportation management system was designed in [3]. The railway computerization and remote-control apparatus monitoring approaches for infrastructure facilities serve as the foundation for this design. Because control complexes get data on the state of infrastructure structures immediately, without the need for a “human–operator” interface in the monitoring system, such systems have a better level of safety. The rapid development of control actions for the timely development and counteraction of emerging flaws is made possible by data from organization facilities

that are not directly in contact with railway automation equipment. The railway infrastructure services states are taken into consideration in the dependability models that the authors created for the train transportation control systems. The accounting of the comprehensive safe state infrastructure is presented by the authors for rail traffic management model. The created dependability models are recommended for use in the construction of safe train traffic management model and digital twins for railway structure assets, which incorporate real-time process modeling.

The security handling systems synthesis challenge for crucial technical procedures on the railway signaling and computerization systems is examined by D. V. Efanov et al. in 2023 [3]. It is demonstrated that there is no 100% safety in the implementation of current management systems for complicated dispersed systems, such as train transport. The safety of such systems is restricted to taking into account only their failures, external failures of control systems and their elements, and failures of infrastructure items that communicate directly with control devices. During autonomous control and data transmission to on-board computerization, other infrastructural facilities are not taken into account in any manner. The theoretical underpinnings of the safety rail traffic management system synthesis are presented, taking into account the possibilities of outfitting infrastructure facilities with extremely dependable and secure technical monitoring and diagnostic tools. Taking into account the diagnosis and result monitoring of every component involved in the transportation process, a streamlined architecture for the central train transportation control complex is provided. The prerequisites are developed for the synthesis of entirely safe rail traffic control systems. The jobs associated with the problem of safe control system synthesis are mentioned. A significantly improved degree of train traffic safety will be possible thanks to the extensive monitoring of rolling stock attributes and railway facilities.

An essential problem in train scheduling and dispatching is delay prediction. Accurate delay predictions, based on actual operation data, are crucial for dispatcher choices and train operations. In 2019, Chao Wen et al. developed a system that uses train describer recordings on a Dutch railway line to show how train delays and their contributing variables interact. They identified the independent and dependent variables based on the primary determinants of train interruption and the time series trend. Python was used to create a LSTM (Long Short-Term Memory) forecast model in which the dependent variable was the actual delay time. Lastly, a comparison was made between the ANN (Artificial Neural Network) model and the RF (Random Forest) model's estimation accuracy. The LSTM model performed better than the other two models, based on the data [5].

Forecasting the state of train delays is a complicated and ever-changing issue that affects both railway companies and commuters. To anticipate train delays in railway systems, Dawei Wang et al. 2024 present a unique hybrid DL model called the CNN + TCN model, which is made up of CNN (Convolutional Neural Networks) and TCN (Temporal Convolutional Networks). Initially, they create three-dimensional data that has the spatiotemporal properties of actual train data. The suggested DL model then uses a 3D CNN element to mine the spatiotemporal features in the built 3D data and a TCN component to capture the temporal features in the railway operating data. Additionally, the characteristic factors that match the two parts are chosen. In the end, the model is

assessed using information from two UK train lines. The CNN + TCN model performs better in terms of accuracy and convergence when it comes to predicting train delays, based on numerical data [6].

Train delays frequently happen, even though timeliness is a benefit of train travel over other long-distance modes of transportation. Pu Wang et al., 2019 gathered and examined a three-month dataset of climate, train delay, and plan records for the present research to identify trends in train delays and forecast train delay times. They discovered that whereas train delays in normal weather are mostly influenced by the historical delay time and train delay frequency, in severe climate conditions the delays are mostly affected by the sort of poor climate. They created a ML model to forecast the delay duration of each train at each stop after identifying the parameters that were highly connected with train delays. The prediction model helps railway operators create more economical pricing schemes and more effective train timetables in addition to helping customers plan their travels with greater reliability [7].

Train spatiotemporal graph Convolutional Network (TSTGCN), a DL architecture, was proposed by D. Zhang et al. in 2022 to forecast the whole collective effect of train delays in a particular station for train posting and emergency preparations. Most of the proposed model refers to recent, daily, and weekly components. Each of these elements is combined spatio-temporal convolution, which can potentially exploit the spatial and temporal characteristics effectively as well as a spatio-temporal attention mechanism. The final expected output is the result of a weighted fusing of these three. Experiments on the China Railway Passenger Ticket System train-operating dataset demonstrate that TSTGCN significantly outperforms state-of-the-art baselines in predicting delays of high-speed trains [8].

3 Proposed Model

The most important benefit of ML in the prediction of train delays is that it can grow to become more accurate as time progresses. ML algorithms to recognize patterns and trends in the extensive amounts of historical data fed into them, which may not be obvious at first. This results in better predictions, which is important to reduce delays and improve overall railway operations efficiency. These models can be constantly refreshed with fresh data in real-time, they are able to adapt quickly as conditions change and provide predictions that better reflect the latest available information.

Data source, data pre-processing and model development and testing are the main challenges in building SVM based train delay prediction system. And its stages are all designed to ensure the trustworthiness and precision of prediction model, meanwhile utilizing some high-performance preprocessing techniques. The first diagram shows the workflow of a train delay prediction system with SVM as learning algorithm and different preprocessing. In the following work present a workflow how this could be achieved by giving you guidance on using structured steps between building a train delay prediction pipeline and performing data preprocessing in order to have better accuracy for SVM classification model. SVM is choosed as proposed algorithm because of the concept of hyperplane. The data will be divided into positive and negative data from that classification is done which gives better result.

3.1 Workflow

Process begins with a "Train_delay_Predication_training dataset as an input which contains the historical data of train schedules and delays. This dataset eventually becomes the training data for further analysis and modeling. In a preprocessing step, the IQR approach is used with Z-score and MZ-score for further data refinement. Outliers are detected and treated by using the IQR method, which believe that lowest 25th percentile to be outliers or highest 75th percent and should do smoothing at outlier values of servant. In z-score method, it scales the data with attributes so that they have zero mean and unit SD to get a uniform contribution from all features. The MZ-score method, a variant of Z-scores that is able to handle outliers robustly, this can help for example with skewed data preprocessing.

After the data is preprocessed, refined dataset is fed to SVM classifier. The SVM algorithm divides the data into two separate classes of "Delayed" or "Non-delayed". It is based on the relationship and patterns discovered from the historical data in training phase. SVM is great for this, given that it can create a hyperplane based on the data points and maximize the distance between two different classes (binary classification.)

Performance evaluation of SVM model using accuracy, precision and recall. These measures provide a complete assessment of how effectively the algorithm predicts delays in trains. While recall tells us what fraction of the true positives were successfully captured, accuracy measures performance at a global level and precision helps in finding out how many positive class predictions actually belonged to the positive-class. These metrics will allow to verify the reliability and precision of prediction models (Fig. 1).

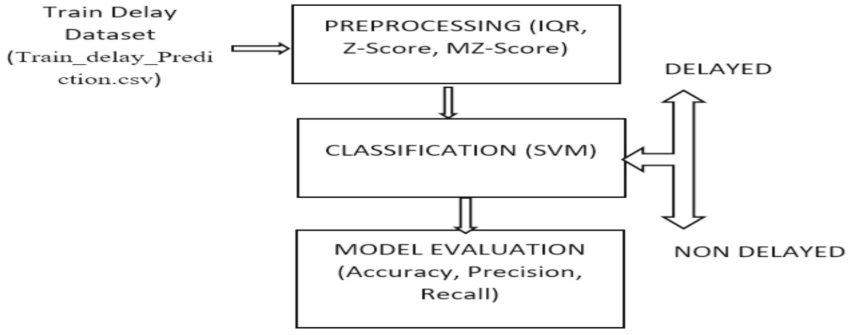


Fig. 1. Proposed Train Delay Prediction Model

This workflow will output a binary classification stating either "Delayed" or "Non-Delayed," helping in coordinating more effectively railways operation. This well-structured methodology utilizing strong preprocessing and a highly potent classification algorithm demonstrates the system's effectiveness in improving train delay prediction efficiency and accuracy.

Collecting Data: Getting the Train_delay_Prediction.csv. This dataset includes historical information on train delays, as well as relevant data such as: Train schedule Weather and other operational aspects. The prediction model is then built using this data [9]. In

real time the data has to be collected from big servers so that it will be challenging task in dealing with such huge datasets.

3.2 Data Preprocessing

Preprocessing plays an important role in train delay prediction: it increases data quality and usability. Noise and inconsistencies from raw data can be reduced by preprocessing, thus giving way to better-informed predictive models. It also constitutes feature selection and engineering that processes the input variables to extract generally applicable patterns or relationships. Moreover, data is often normalized and scaled so it can produce same level of contribution to the model without creating unwanted bias. In continuation with the prior point, preprocessing tidies data so that becomes more dependable and analysis ready which paves way for predictive of train delays to be as practical as possible. In order to get the data ready for model building needed preprocessing at the first step.

3.2.1 Z-score Method

The Z-score method is a preprocessing technique that focuses on normalizing the characteristics of a particular data set. It is also known as sometimes with the short name of z-score This technique is used to re-scaling the data in which mean and SD (Standard Deviation) are one for each of these features. Standardization is very useful for ML algorithms, as they will assume that all features have the same scale and typically not work correctly if this is not the case. In this method, we ensure that each feature equally plays a role in the model training process using the Z-score method.

The features are normalized using Z-score normalization i.e., mean = 0 and SD = 1. This is very important as SVM can also be greatly affected by the scale of input features. This way, they will receive same weight by making properties consistent in the training process.

The Z-score method enables the integration of training data with different units and scales, such as time, distance or numerical encoded variables. By doing so, it also help to prevent the model from giving preference over one feature simply because of how large your values are. The formula of Z-score for a value x in any feature goes as follows.

$$Z = \frac{x - \mu}{\sigma} \tag{1}$$

where:

- The value that has to be standardized is x.
- The feature's mean is denoted by μ.
- The feature's SD is represented by σ.

Z-score normalization in train delay prediction standardizes features with different scales to ensure each feature contribute equally to the model. Normalizing the inputs is important because this will result in each input having a mean of zero and variance of 1, meaning no one feature can dominate learning. In addition to making the features less heavy and thus balancing their impact, normalization improves efficiency and stops

problems that can slow or accelerate convergence of gradient based models for better predictions. Standardizing the data would allow to train the better, and predicting delays exactly.

3.2.2 IQR (Interquartile Range) Method

IQR is the method which we use to find outlier. The high and low of the given range are used to identify possible outliers using this method which could skew a model if present. The Preprocessing method will help to find and treat with Outliers in a dataset. Outliers can influence the outcome and also ruin the results of other ML models. The IQR based approach helps to reduce the noise in data by identifying and if required, eliminating these outliers.

The IQR is the range of a dataset between its first and third quartiles. Represents middle 50 percent of the data Values that seem way out of range can be identified as candidate outliers by inspection around the middle of these values. To detect and remove outliers in metrics as delay times, the IQR approach could be utilized for train delay prediction. If done this way, the model will be trained on a small dataset that is more common to real situations rather than one heavily biased by outliers.

The IQR value is assessed using the formula (2):

$$IQR = Q_3 - Q_1 \tag{2}$$

where:

- The first quartile, or 25th percentile of the data, is denoted by Q1.
- The third quartile, or 75th percentile of the data, is denoted by Q3.

To recognize outliers, need to measure the lowerbound and upper bound values:

$$LowerBound = Q_1 - 1.5 \times IQR \tag{3}$$

$$UpperBound = Q_3 + 1.5 \times IQR \tag{4}$$

Outlier values are those that fall outside of either the lower or upper limit.

The IQR is the efficient method for preprocessing outliers in train delay data. It also takes care of outliers in the data set and resolves them using this method, which enhances accuracy in prediction models.

IQR help in predicting train delays by assisting to handle outliers within a dataset. The IQR tells the spread of delays between the 25th and 75th percentiles, which may catch some really long (or short) times that would skew our model outright. Dealing with these outliers, either by exclusion or through special handling, helps in making the predictive model learn from typical data patterns and provide more accurate predictions. This will make the model more durable, allowing a better performance to real-world scenarios where extreme values could distort its learning.

3.2.3 Modified Z-score Method (MZ-Score)

If the data is not normally distributed and there are a large number of outliers in the specified variables, one good way to identify it. Because it uses the median and MAD

instead of the mean and SD, MZ-score is more robust to outliers than normal Z-score. The Modified Z-score approach works as follows:

Steps to Compute Modified Z-scores

a) **Calculate the Median:** Measure the median value of the dataset.

$$Median = med(x) \tag{5}$$

b) **Calculate the MAD:** Calculate the absolute deviations from this median and find their own median.

$$MAD = med(|x_i - Median|) \tag{6}$$

Here $|x_i - Median|$ denotes the absolute deviation of every data point from the median value.

c) **Compute the MZ-score:** For every data point, determine the MZ-score using the following equation:

$$MZ - Score = 0.6745 \times \frac{x_i - Median}{MAD} \tag{7}$$

The constant 0.6745 is used here to ensure that the MZ-score becomes equivalent to a Standard Z-score in cases where the data truly have normally distributed scores.

The MZ-score is a robust statistical method capable of detecting outliers within datasets, which can help in train delay predictions. The MAD and there is a MZ-score that uses this in place of SD to account for skewness. This makes it robust against outliers and better for detecting anomalies in data with non-normal distributions. With using the contextual MZ-score on historical train delay data, one can already detect abnormal behavior in missing schedules that are able to be used later for improving accuracy with predicting schedule of trains.

3.2.4 Support Vector Machine (SVM)

The resource optimization using ML for train delay prediction, is another major advantage. This permits railway operators to make resources available in a much more targeted manner—adjusting timetables, deploying maintenance crews where they are needed most and informing passengers accordingly. Not only would it increase operational efficiency, but also brings customer satisfaction – uncertainty during the train delays due to collisions or unfortunate accident.

SVM is a ML algorithm, and has been used in many fields from economics to pattern recognition. MalekSarhani et al., 2024 describe it as a convenient tool for regression and classification related issues. Testing the above method is tested with a TheTrain_delay_Prediction.csv dataset. Predictions are that combining SVM with sophisticated preprocessing steps (i.e. MZ-score) will result in higher accuracy, specificity and sensitivity. Results with baseline models are shown in this paper to evaluate the proposed method. In the problem of train delay prediction, SVM can serve as a great tool for classifying whether or not trains will be delayed based on various factors [10].

Where x is a feature vector A Decision function for SVMs - The decision question of an input has been determined to be belonging to one class or another Delayed vs Not Delayed.

$$f(x) = w^T x + b \tag{8}$$

where:

- x = Feature vector for a train (e.g., delay-related features)
- w = Weight vector (coefficients of the features, learned during training)
- b = Bias term (offset of the hyperplane)

The equation of the hyperplane that separates the two classes (delayed vs. not delayed) is:

$$w^T x + b = 0 \tag{9}$$

This hyperplane is selected to exploit the margin between the various classes.

The margin illustrates the distance between the hyperplane and the nearest training samples (support vectors). For a linear SVM, the margin is:

$$Margin = \frac{2}{|w|} \tag{10}$$

where ‖w‖ indicates the Euclidean norm of the mass vector w.

The trade-off between minimizing model complication and gaining a short error on the training set is accomplished by the regularization variable C. In SVM, the objective function that has to be minimized is:

$$Minimize \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{N} \xi_i \tag{11}$$

where:

- ξi = Slack variable for each training example iii (measures the degree of misclassification for that example)
- N = Number of training examples

$\|w\|^2$ = Squared norm of the weight vector, representing model complexity

Moreover, the practice of ML contributes to models that scale effectively. With the expanding of rail networks and growing volume of data ML models can be further enhanced to scale accurately without needing a proportionate increase in computer power. This scalability preserves the applicability of machine learning models in train delay prediction as railway systems increase their complexity and size over time.

4 Result and Discussion

For effectual railway traffic development and management as well as for delivering acceptable commuter quality of the service, accurate train delay prediction is essential. Predicting train delays is vital to ensure effective traffic management processes and

enhance the traveler experience. Predicting delays helps transportation agencies make necessary schedule adjustments, better allocate resources and reduce potential disruptions. This proactive process of managing expectations also lowers the risk that passengers will miss their connection, allowing usually improved customer service. Again, pin-point delay prediction allows to plan strategically given the incoming operating conditions improving overall service efficiency.

4.1 Dataset Description

Data on train trips, including start and end timings, status updates, lengths of delays, and actual reach times, may be found in the Train_delay_Prediction.csv file. Every train journey is shown as a row, and the "Started On" column gives the date-time of start of that particular train trip, shows whether it's delayed or on time in Status (with 0 indicating punctuality), duration of any delay faced from its actual reach timing at home station to destination arrival etc. This can then be utilized to study train performance, understand delay trends and predict future delays.

4.2 Evaluation Analysis

In train delay prediction using ML, as performance metrics are precision, recall and accuracy is major to find it out how well predictive model have worked. These are important because of the way each metric leads to knowledge about different aspects of the performance, so that not only accurate but reliable and actionable predictions.

4.2.1 Accuracy Analysis

The precision of any train delay prediction model is a significant one, even while it may in some instances be looked at as just an end metric to examine. It shows the ratio of correct predictions to all of the models forecasts. In terms of train delay prediction accuracy, this means how well the model can distinguish between delayed and non-delayed trains. The percentage of accurate forecasts—both delayed and non-delayed—out of all the predictions is known as accuracy.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{12}$$

4.2.2 Precision Analysis

Out of all trains anticipated as delayed, precision represents the percentage of accurately predicted delayed trains.

$$Precision = \frac{TP}{TP + FP} \tag{13}$$

4.2.3 Recall Analysis

The percentage of accurately anticipated delayed trains among all actual delayed trains is called recall.

$$Recall = \frac{TP}{TP + FN} \tag{14}$$

True Positives (TP): Total Number of properly forecast delayed trains.
True Negatives (TN): Total number of exactly predicted non-delayed trains.
False Positives (FP): Number of non-delayed trains incorrectly predicted as delayed.
False Negatives (FN): Number of delayed trains wrongly predicted as non-delayed.
The following Table 1 and Fig. 2 represents the results obtained after classification.

Table 1. Results Obtained after Classification

ML Methods	Accuracy %	Precision	Recall
IQR + SVM	88.80	0.87	0.85
Z-Score + SVM	90.25	0.89	0.87
MZ-Score + SVM	94.50	0.93	0.91

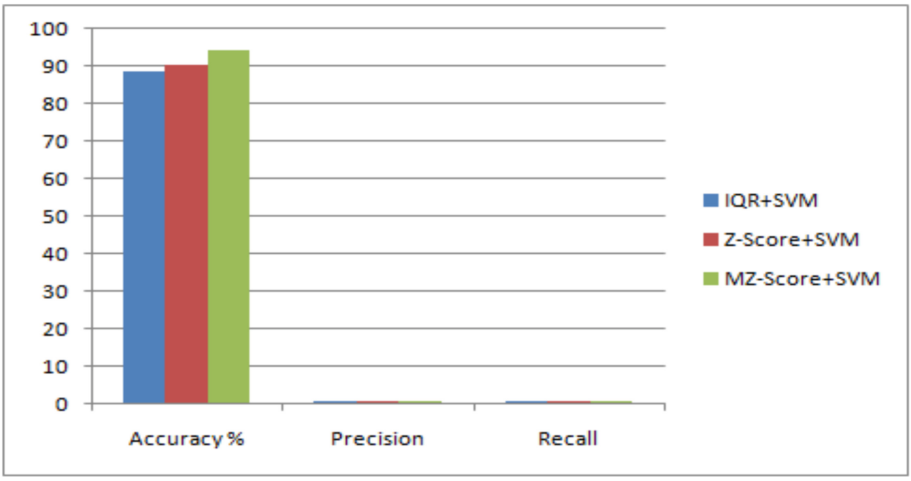


Fig. 2. Results Obtained after Classification Graph

From the above Table 1 and Fig. 2 its clearly proved that existing IQR + SVM gave accuracy of 88.80%, precision of 0.87 and recall of 0.85. Another existing Z-Score + SVM gave accuracy of 90.25%, precision of 0.89 and recall of 0.87 which is less performance compared to the proposed MZ-Score + SVM which gave accuracy of 94.50%, precision of 0.93 and recall of 0.91 respectively.

In the process of using ML for train delay prediction; precision, recall and accuracy are very important metrics. Precision reflects the credibility of model's prediction, recall guarantees it catches most delay and accuracy is the whole picture on how well did a job. Combining these two metrics side by side gives a robust model that can easily integrate with day to operations for the railway companies, resulting in better resource planning and passengers experience.

5 Conclusion and Future Work

Predicting train delays is a crucial component of contemporary railway management, to raise both passenger happiness and operational effectiveness. Schedule delays on trains may have a lot of effects; they can interfere with travel arrangements, raise operating expenses, and impair the coordination of linked transportation networks. Proactive measures are made possible by precise and early projections of train delays, which enable operators to minimize negative consequences, allocate resources optimally, and enhance service dependability. In this work, the Z-score method, IQR, and MZ-score method were among the preprocessing approaches used in conjunction with SVM to forecast train delays. The combination of SVM and the MZ-score technique outperformed the others in terms of recall, accuracy, and precision. Through efficient management of anomalies and data normalization, the suggested approach produced more dependable and resilient forecasts. The increased predictive capabilities of this technique were proven by testing using the Train_delay_Prediction.csv dataset, indicating its potential to improve customer satisfaction and operational efficiency in railway systems. For future enhancement more machine learning algorithms such as Convolution Neural Network (CNN), Artificial Neural Network (ANN) and Gradient Boosting can be used which will give more results than proposed SVM.

References

1. Dawale, N.N., Nandgave, S.: Train delay prediction using machine learning. In: Yadav, A., Nanda, S.J., Lim, M.H. (eds.) Proceedings of International Conference on Paradigms of Communication, Computing and Data Analytics. PCCDA 2023. Algorithms for Intelligent Systems. Springer, Singapore (2023). https://doi.org/10.1007/978-981-99-4626-6_54
2. Spanninger, T., Trivella, A., Büchel, B., Corman, F.: A review of train delay prediction approaches. J. Rail Transp. Plann. Manag. **22**, 100312 (2022). https://doi.org/10.1016/j.jrtpm.2022.100312
3. Efanov, D.V., Mikhailiuta, E.M., Khóroshev, V.V.: Reliability models for a safe train traffic control systems accounting the railway infrastructure states. In: 2023 International Russian Automation Conference (RusAutoCon), Sochi, Russian Federation, pp. 266–270 (2023). https://doi.org/10.1109/RusAutoCon58002.2023.10272854
4. Efanov, D.V., Khóroshev, V.V., Osadchy, G.V.: Principles of safety signalling and traffic control systems synthesis on railways. In: 2023 International Conference on Industrial Engineering, Applications and Manufacturing (ICIEAM), Sochi, Russian Federation, pp. 634–638 (2023). https://doi.org/10.1109/ICIEAM57311.2023.10139292
5. Wen, C., Mou, W., Huang, P., Li, Z.: A predictive model of train delays on a railway line. Wiley Online Library (2019). https://doi.org/10.1002/for.2639

6. Wang, D., Guo, J., Zhang, C.: A Novel Hybrid Deep Learning Model for Complex Systems: A Case of Train Delay Prediction. Advances in Civil Engineering (2024). https://doi.org/10.1155/2024/8163062
7. Wang, P., Zhang, Q.: Train delay analysis and prediction based on big data fusion. Transp. Saf. Environ. **1**(1), 79–88 (2019). https://doi.org/10.1093/tse/tdy001
8. Zhang, D., Peng, Y., Zhang, Y., Wu, D., Wang, H., Zhang, H.: Train time delay prediction for high-speed train dispatching based on spatio-temporal graph convolutional network. IEEE Trans. Intell. Transp. Syst. **23**(3), 2434–2444 (2022). https://doi.org/10.1109/TITS.2021.3097064
9. https://www.kaggle.com/datasets/vipulpisal/train-delay-prediction/data
10. Sarhani, M., Voß, S.: Prediction of rail transit delays with machine learning: how to exploit open data sources. Multimodal Transp. **3**(2), 100120 (2024). Science Direct, https://doi.org/10.1016/j.multra.2024.100120

A Comparative Analysis of Nerve Vessels Enhancement and Segmentation Techniques Through Edge Detection on Retinal Fundus Images

N. Durga[1(✉)], D. KeranaHanirex[1], and A. Muthukumaravel[2]

[1] Department of Computer Applications, Bharath Institute of Higher Education and Research, Chennai, Tamil Nadu, India
durga90.n@gmail.com, keranahanirex.cse@bharatuniv.ac.in

[2] Department of MCA, Bharath Institute of Higher Education and Research, Chennai, Tamil Nadu, India

Abstract. Diagnosing wider disease in retina is more significant in Ophthalmology. Using Image processing techniques to identify the retinal diseases has proven more precise than other technology implementations, rapidly. Certain disorder identifications lowers the quality with respect to the limiting capacities and diagnosing them. Irrespective of several algorithms and researches the issue remains unsolved on the retinal disease identifications using image processing. Where huge number of observations turned negative with retinal color images, which in turn formulated an enhanced technique that is solely focuses on the green component of the retinal image. Corona virus spread during the year 2019 is predicted to have allied with the variations in the nerve vessel, that is seen into the partitioned eye fundus with the patients in the process of detecting the variations in the retinal via image processing which is also used to assess the potential correlation over medical constraints.Initially, the retinal image of the eye is pre-processed which is further processed by identifying the presence of the pictorial particulars using contrast improvisation and histogram equalization, CLAHE. Secondly by removing the noise from the image, using WMF (www mmmm ffff). Upon comparison with filters like median, wiener etc., WMF is proved to be the best from the results obtained in this proposed research work. Finally, retinal images were segmented with the nerve cells in the retinal images of Covid-19 infected individuals using SOBCAN (SSSooobbb ccc aaa nnn) methodology. The fundus images and performance was evaluated based on parameters including MSE, PSNR and SSIM. The results were compared with other technologies like sobel, canny etc., and concludes that the proposed methods' formulated outcome outperforms the equivalent enhancement method.

Keywords: Retinal fundus image enhancement · Noise filtering · Eye nerve vessel · Edge detection · Improved eye disease visualization

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
S. Rajagopal et al. (Eds.): ASCIS 2024, CCIS 2428, pp. 145–160, 2025.
https://doi.org/10.1007/978-3-031-86302-8_10

1 Introduction

Experimentation on automatic computer systems for medical images evaluation, particularly over retinal fundus image concerning ophthalmology, have increased dramatically. Image of retina was useful in identifying a variety of illnesses including hypertensions, glaucoma, retinopathies of prematurity, diabetic retinopathies, macular degenerations as well as cardiovascular diseases [1–4]. In pathology, retinal image is useful for measuring minor degenerative changes that occur gradually as the diseases progress. Many of these minor alterations are difficult for specialists to identify manually. As a result, employing automated procedures increases the capability in noticing as well as evaluating the variations together with diagnosis [5–7]. Lower-quality retinal images impair machine and vision of human, emanating in faulty diagnosis, improper blood vessel extraction, also the inefficacy of computerized retinal imaging systems [9–11]. As a result, there is a significant demand for enhancing the existing approaches. Digital image augmentation procedures improve pictorial qualities of images by eliminating undesired effect, thereby resulting larger usability when compared with cover images in order to achieve preferred objectives [8–12]. Good-quality retinal image, in particular, increase diagnosing accuracies [13] in enhancing perceptibility in fundus feature as well as inevitably decreases noise deprived of addition or subtraction of any important data. Image enhancement procedures were typically applied during pre-processing stage for aiding illness detection, computerized clinical workflows as well as biometric verification system [2].

A majority of research available in literatures [14, 15] were focusing the grayscale retinal image other than color retinal imaging. Most retinal image improvement approaches use the notion concerning min–max color value towards histogram equalization (HE) [16, 17].This method is a useful approach for a variety of image types. Conversely, this produces negative observation for variety of retinal picture instances, particularly color retinal images [7, 9].

Adaptive histogram equalization (AHE) is a local histogram equalization that has been generalized from HE. This method was formulated on the basis of HE, and this flexible methodology will be formulating all the sub-image histograms for dispersing the intensity value of images. The hybrid cumulative histogram equalization (HCHE) technique will be improving the enhancement effects at hot object other than backgrounds [18]. This gap adjusting histogram equalization could answer over-enhancement issue while together with alleviation of feature loss problems at dark portions of images [19]. For addressing aforementioned issue, contrast limited adaptive histogram equalization (CLAHE) has been innovated.

CLAHE was considered as one of the processing methods based on blocks, and these on combination with elementary histogram equalizations, shall eliminate the problems concerning over-amplification of interferences at homogenous zones of images. CLAHE varies from regular HE as this operates in smaller sections of images commonly referred as tile, every equivalent corresponding dissimilar sections of images, thereby employs them for redistributing brightness value of images [20–22].

CLAHE enhancement method can be used in a variety of colour spaces, including RGB, YIQ and HSI. The colour space was represented as red (R), green (G) as well as blue (B) constituents in RGB colour model. These three constituents were considered as image with monochromatic intensities. As a result, the images being taken using

colour video cameras or presented on colour monitor screens, thereby RGB model was found to be an ideal tool for color production [23]. CLAHE is applicable over each of these three constituents independently in RGB colour model. By merging individual RGB constituents, the fully-colored RGB images can be produced. In this paper, green component of the retinal image was employed in this proposed technique.

China has experienced an outbreak of new strain of corona virus called SARSCoV-2 (Severe Acute Respiratory Syndrome Coronavirus -2), in the month of december of 2019. World health organization (WHO) termed this disease to soon grow till the extent of pandemics in early 2020 [24]. Retinal changes are also common in persons suffering from viral infections. These pathologic measures can develop as a result of the microorganism's direct cytotoxic effect over retinal neuron, likewise concerning cytomegalovirus [25], or resulting from impairment of microvasculature on virus targeting the vessel endothelium, as in HIV retinopathy [26]. SARS-CoV-2, interestingly, can infect endothelial cells [27] and has been found in the retina [28].

The proposed research aims towards screening individuals' fundus corresponding to SARS-CoV-2 infections concerning the abnormalities in retinas or retinal vasculature associated with COVID-19 and for assessing potential relations amid medical parameter with the help of dissimilar segmentation approaches discussed in detail in further sections.

2 Image Analysis-Enhancement Techniques

Improvising the resolution of digital image using certain algorithms without affecting the important information is called Image enhancement. For example, brighten, sharpen, or eliminate noise from an image to make it easier to identify crucial details. Enhancing retinal structure images may aid in the diagnosis of several eye diseases. Diagnosing and tracking retinal illnesses frequently need highly sensitive analysis, which is only possible with the appropriate equipmentsuch as an OCT device or high-resolution fundus camera medical expertise, and, frequently, picture enhancement. Segmentation,measurement,and quantification are among the delicate analyses, the accuracy of which is dependent on the quality of the input image. Medial opacity, defocus, and the presence of artifact have an impact on image quality [2]. To boost contrast or brighten the images, retinal image enhancement techniques were employed. Techniques for retinal image enhancement are applied to enhance contrast and brightness in images. Enhancement of fundus pictures. Using image histogram stretching techniques on fundus pictures is strongly advised.

3 Proposed Model

Fundus imaging was used to assess changes in retinal properties in patients with Covid-19 and diabetes. The Mean arterial diameter (MAD) and mean venous diameter (MVD) of patients were assessed relative to age, sex, cultural background, body mass index, alcohol consumption, hypertension, hyperlipidemia and diabetes status [12]. Regarding the Covid-19 patient, the results of laboratory data related to retinal studies were confirmed.

The proposed method was divided into three steps according to the flowchart in Fig. 1. Retinal details of the eyes, together with the contrast level of the fundus images, were initially enhanced by improving the appearance of structural details and reducing unwanted noise using a WMF filtering method. Final appropriate retinal color images were produced in a second step after uneven illumination was corrected using color enhancement techniques [19]. The neuron in the retinal image of the Covid-19 patients was segmented in the last step using the GSOBCAN method.

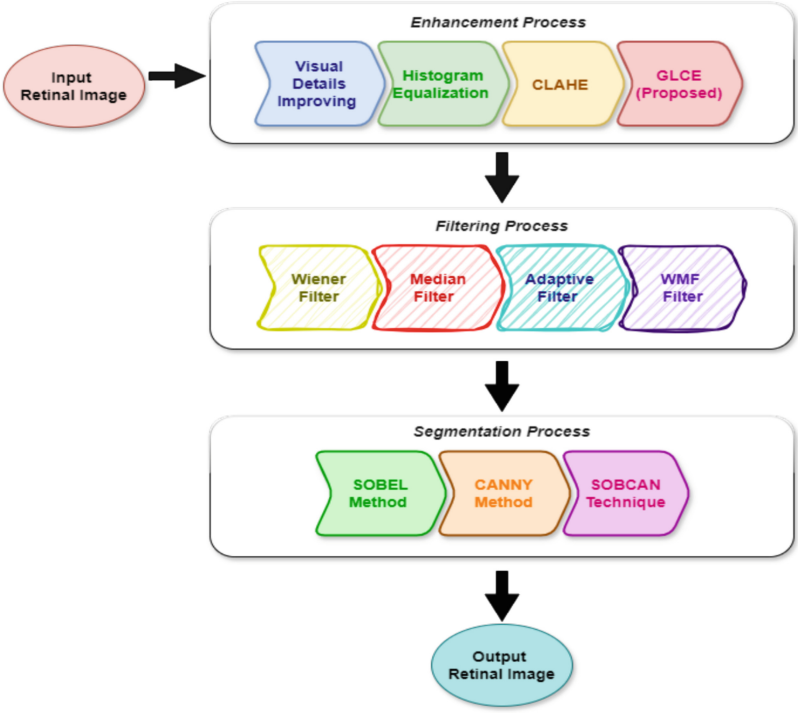


Fig. 1. General Architecture of the Proposed Methodology

The in the proposed methodlogy were detailed in subsequent sections below.

Improved Visual Details: The structural particulars from retinal image displays edge, vessel and spot with variable contrast below fundus backgrounds. To improve these types of visual particulars, a variety of techniques were employed [21]. Smoothing filters work well on backgrounds, but they do not keep adequate features like light edge or little vessel. Edge filtering preserves edge as well as vessel, however it was ineffective when it comes to background details.

Histogram Equalization (HE): HE stands for one of the image processing methods that uses an image's histogram to adjust contrast. In order to improve contrasts, this method extends the intensities of images or spreads the greatest frequent pixel intensity value. HE is used to modify the contrast of an image in its pixel. To increase the contrast

of the image, it either spreads out the most common pixel intensity values or expands out the image's intensity range [12]. By improving the contrast of the image at lower contrast areas, HE achieves these goals. It is helpful when processing is required for photographs that appear to have lost contrast. Both lighter and darker areas come together in these pictures to produce. Using histograms, this technique distributes the intensity levels of pixels in an image to increase contrast. Thus, adaptive histogram equalization outperforms conventional histogram equalization when it is necessary to improve local contrast and strengthen the edges of certain parts of images. One of the limitations of HE is that it tends to over-enhance contrast in near-contrast areas of the image. For intensity levels ranging from 0 to (L-1), the histogram can be expressed as

$$h(rk) = nk \tag{1}$$

In Eq. (1) nk denotes the number of pixels in the digital image at that intensity level and rk is the kth intensity level.

The image can be normalized by dividing the histogram with the sum of pixels in the image. The definition of the normalized histogram function for an N × N image is:

$$p(rk) = \frac{nk}{N2} \tag{2}$$

The likelihood that a pixel with an intensity level of rk will occur is represented by the p(rk) function. Evidently,

$$P(rk) = 1 \tag{3}$$

The histogram of an image provides important information about the contrast and gray scale of the image. An image is assumed to be dark if its entire histogram is oriented toward the left end of the x-axis. A histogram tilted to the right end shows a white or bright image. A low-contrast image with a few grayscale levels can be seen in a narrow histogram plot centered on the intensity axis. On the other hand, an image with a histogram evenly distributed along the entire x-axis looks very contrasty. In image processing, there is often a desire to improve image contrast. In these situations, we use an intensive transformation method called histogram equalization [21]. By choosing an appropriate intensity transformation function, the histogram of the image is evenly distributed over the entire intensity axis, a process known as histogram smoothing. Histogram smoothing is therefore an intensity transformation method.

CLAHE: For the last few years several algorithms were evaluated on image enhancement in contrast improvisation inorder to obtain the objective. The 2 basic goals of contrast enhancement are (i) improving an image for visual interpretation and (ii) making it easier to perform following tasks (including object detection, image segmentation, and image analysis), basically modifying the local or global histogram [22]. CLAHE technique circumvents the shortcomings of global solutions by enhancing local contrast. It is one method of image equalization. Moreover, CLAHE can be used on color images, usually on the luminance channel. When an HSV image is simply its brightness channel equalized, the results are superior to when a BGR image is equalized in all of its channels. CLAHE was initially used to improve low-contrast medical photographs, as it distinguishes itself from standard AHE. To address the issue of noise

amplification, CLAHE implemented a clipping limit. Before computing the cumulative distribution function (CDF), the CLAHE restricts augmentation by trimming the histograms in specified CDF values. This CLAHE method splits the original images to non-overlapping contextual segments termed as tile. This CLAHE possess Block size (BS) as well as Clip limit (CL).

Contrast Enhancement by Gamma Correction (CEGC): Gamma correction is a vital technique for adjusting photographs that are too dark or bleached out since it may be used to manage the brightness of an image overall. The primary colour spaces were signified as RGB model. The colour spaces were described with respect to RGB constituents, which are represented as primary colors in the RGB model. These three components are images of monochromatic intensity. Digital images according to this paradigm, is made up of three planes over individual image, all of this store these RGB values. Individual pixel Gamma values were locally optimized in this method by decreasing the homogeneity of the source image's co-occurrence matrix. The RGB model was considered to be based on hardware colour paradigm, where RGB components were comparable as well as possesses stronger relationship. As result, a change in one component will have an effect on the others. The Gamma correction method enhances retinal image's contrast by increasing dynamic ranges of the images.

GLSCE-Green-Component Based Luminosity Spanning Sharpening Contrast Enhancement: The main objective of this proposed processing of colour images is prompted by two fundamental factors: the likeness towards visions which is entirely chromatic, as well as availability of colour image processing software. Moreover, increasing the amount of data will chromaticity adds in analyzing the image. These two objectives have been achieved using this method by having in mind the advantages of the existing algorithms.

The steps involved in this proposed method are as follows:

- Pick green component from color image.
- Transform into L*a*b*color space.
- Span luminosity value for specific range and process the data.
- Substitute luminosity layer using processed information, subsequently converting images back to RGB color space.
- Image edges were sharpened using the unsharp masking method.
- Control the sharpness of image by using the controlling parameter radius and amount.

Figure 2 shows the results of each stage in which this proposed technique GHC method provided greater visual effects on retinal images.

Filtering Process: The preceding phase increases pictorial/structural details of the retinal images considerably. Nevertheless, it exhibits a notable increase in the amount of irregular lighting as well as darkness. Such circumstances provide significant issues as they obscure critical features as well as render objects unrecognizable at the retinal image, producing analysis to be extremely problematic. As a result, this suggested filtering technique is utilized for improving the appearances of absolute retinas image, particularly in dark areas. Restoration filters were considered as the different type of filter

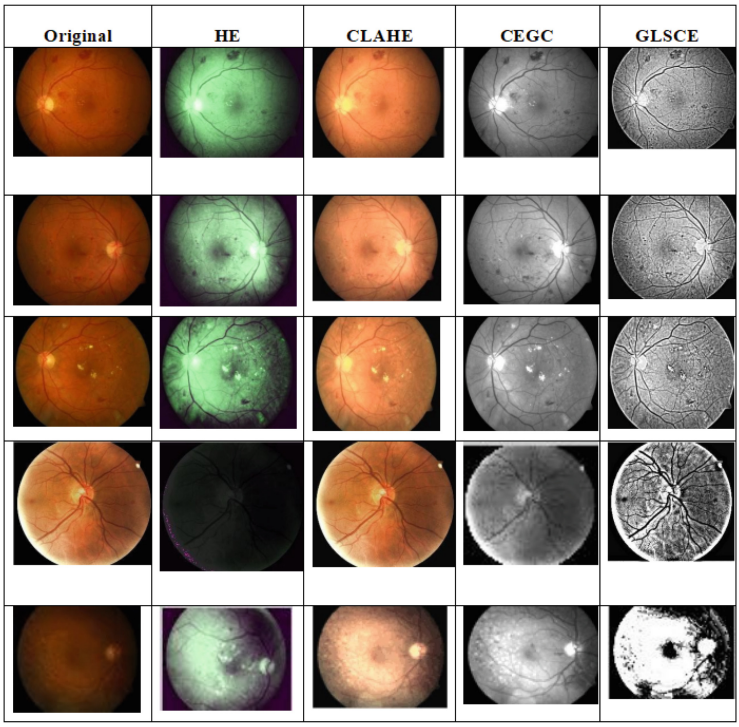


Fig. 2. Retinal Images using Various Enhancement Technique

which was employed for operation on noisy images by adding salt and pepper noise to the retinal images and estimating the original image operated on the Green Component of the images. It could include blurring methods or reverse blurring processes. The filter that was employed for restoration procedure varies from the filter that was employed in enhancing stage.

Wiener Filter: One of the earliest techniques to minimize additive random noise in photographs was Wiener filtering, where the MSE (Mean Square Error) between the original and the reconstructed image is minimized, under the presumption that independent of pixel's location the additive noise is fused in random process. Similar to a low-pass filter, Wiener filtering uses a space-varying cutoff that is intended to utilize a minimum cutoff values in the areas with little details and a higher cutoff values to maintain detail in regions characterized by edges or other features with significant variation. The overall frequency cutoff is determined by the size of the window; larger windows result in lower cutoff frequencies, which in turn provide more blurring and noise reduction.It is also called as Minimum means as filter. This achieves optimum trade-off amid filtering as well as noise flattening. This eliminates additive noises as well as input in blurring concurrently. This diminishes total mean square error by,

$$H(u, v) = H'(u, v)^2 + S_n(u, v)/S_f(u, v) \tag{4}$$

In Eq. (4), H(u, v) represents degradation function's transform, Sn(u, v) corresponds to power spectrum concerning noise, and Sf(u, v) represents power spectrum of the original images'. Wiener filtering can be applied in a variety of ways. Lee (1980) suggested a 2D adaptive filtering that uses a space-varying filter assuming white and zero-mean additive noise, which is used in PTC Mathcad function.

Median Filter: The median filtering is widely used technique for noise reduction in signals and images, distinguished for its ability to preserve edges, making it highly valuable in the field of image processing. Another name for it is nonlinear filtering. It is employed to get rid of pepper and salt noise. In this case, the nearby pixel's median value is used in place of the original pixel value. Median filtering represents the non-linear process which could be employed for reducing impulsive noises, commonly referred as salt and pepper noise. Moreover, it is helpful to retain border of images as well as decreases random noise. Random bit errors in communication lines shall result in impulsive noise. Median filters move the windows across the images,also the median intensity values of pixel inside these windows become the output intensities of pixels being processed. A foremost problem in these median filters is that they are comparatively costly as well as possess increased computational complexity, and shall be represented as,

$$a_i = \begin{cases} 1, i = \frac{(N+1)}{2} \\ 0, otherwise \end{cases} \tag{5}$$

For image processing applications, N is almost always odd. Figure 3 represents an example for the median filter calculation of a retinal image. In order to function as a median filter, it must first replace each input in the signal input-by-input area with the median of the same input and its surrounding inputs. The concept is quite similar to a motion measurement filter, where instead of each individual record, the arithmetic mean of each object and its neighbors is used. "Window" refers to a neighborhood pattern that moves during the signal, one input at a time. The most obvious window for one-dimensional signals is just the first few notations before and after; however, for two-dimensional (or higher) data, the window must contain all records within a given radius, ellipsoid, or rectangular region.

123	125	126	130	140	Neighborhood values
122	124	126	127	135	115,119,120,123,124
118	120	150	125	134	125,126,127,150
119	115	119	123	133	Median value: 124
111	116	110	120	130	

Fig. 3. Calculation of median filter for a Retinal Image

Wiener and Median Filters: A filter is principally employed over processing images for suppressing any higher frequencies of images, i.e., smoothing images. Image noise represents random variations of image's brightness or colour information, thereby it is a typical component in electrical noises. It can be generated by a scanner's or digital

camera's image sensor as well as circuitry. The advantages of both Wiener and Median filters are kept in mind and the hybrid of these two algorithms has been proposed. The results are very effective using this proposed method. This process first removed the noise which were applied, subsequently passes the results into wiener filters. Moreover, it eliminates the blurredness as well as additive white noise from images. Figure 4 represents the output of various filters in which this proposed method has outperformed the existing approaches in the Green Component of the images.

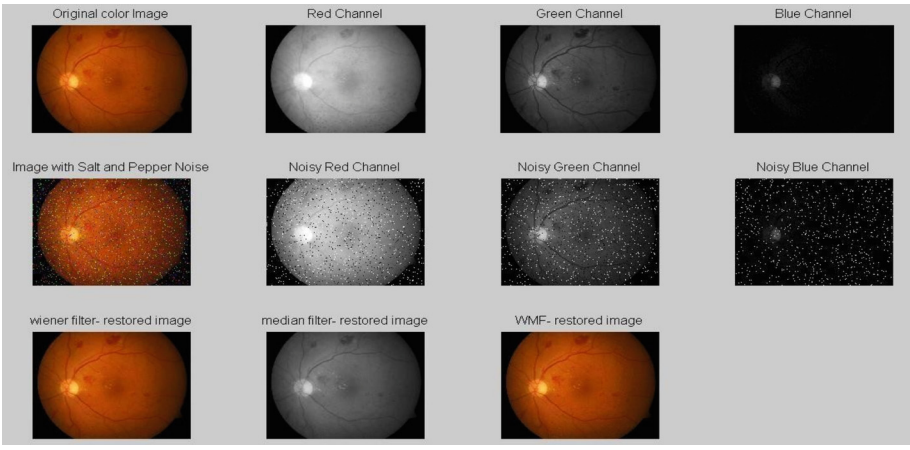


Fig. 4. Image filtering Output

Edge Detection Process: An image processing method called edge detection is used to identify the borders of objects in pictures based on the brightness discontinuities. Also it is utilized for data extraction and image segmentation. This phenomenon was considered as significant as the edges shall capture local characteristics in an image and offer relevant information. Discontinuities or large fluctuations in intensity or greylevel indicate the borders of an image, indicating the position of the object contour. Edge detection is dominant as well as noteworthy tasks in low-level image processing, and it is critical for the development of a comprehensive vision-based monitoring system for autonomous scene analysis as well as monitoring. A collection of connected pixels that separates two distinct regions is called an edge. Generally edges are horizontal, vertical and diagonal and are segmented into discontinuous areas, involving procedures like pattern recognition, image morphology, and feature extraction.

With edge detection, gray level features the changes in the image and the texture replicates the beginning and closing end of image regions. It preserves an image's structural properties while simultaneously reducing the amount of data contained within it. This data aids in the achievement of high-level goals including segmentation, object recognition, scene analysis, etc.

There are two types of edge detection operators:

- Gradient-based administrator which registers first-request deductions in a computerized picture like, Sobel administrator, Prewitt administrator, Robert administrator.
- A Gaussian-based operator that calculates digital image second-order derivations like the Canny edge detector and the Laplacian of Gaussian.

Edge detection as well as extraction methods comprise of Sobel, Roberts operators and the Canny algorithm, among others. These algorithms were employed in this paper which was further applied on the green component of the retinal images, and its results were shown in Fig. 5.

Sobel Edge-Detector: Sobel operator is predominantly common as it is one of the nonlinear filter possessing image smoothing producing reduced fragment edges in the image. Sobel operator functions by calculating the approxima-tion of retinal fundus image's gradient of intensity function. For computing gradient approximation, Sobel filter uses dual 3×3 spatial masks which are convolved in the input eye fundus images. It uses a spatial gradient measurement to highlight region of higher spatial frequencies which corresponds to edge of the eye images. This was usually employed for calculating the assessed absolute gradient degree in all the pixels in gray-scale input images. Moreover, this shall extract vertical component as well as horizontal component of the gradient corresponding to all the pixels in images with the help of convolution by two 3×3 spatial masks represented by S1 and S2 matrices

$$S1 = \begin{matrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{matrix} \qquad S2 = \begin{matrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{matrix}$$

where, S1 signifies vertical spatial mask concerning 90° direction, whereas S_2 represents horizontal spatial mask concerning 0° direction.

Canny Edge Detector: Canny is a typical edge detection technique. An image is made up of various details about scenes like sizes, colors as well as orientations of several items of that scene. The item must first be isolated from the backdrop, and then all of the edges must be identified in order to obtain the object's outline. Canny edge detector is derived from previous algorithm of Laplacian of Gaussian operator [19]. It is widely deployed as an optimal technique for edge detection, operating based the following three distinct criteria:

Low error rate
Edge points must be accurately localized
There should be just one single edge response

As a result, edge detection has become increasingly crucial in computers. The algorithm primarily possesses five steps as given below.

Step-1: Compute horizontal (Gx)as well as vertical (Gy) gradient soft distinct image pixels. The proposed edge detection strategy corresponds to a series of edge detection algorithms performed on a noise-free background image. In the future, additional image

filters and other methods of segmentation algorithms, which provide good visual effects of points and nerve vessels in the visual image with additional accuracy, should be evaluated.
Step-2: Further calculate the magnitude (G) as well as direction.
Step-3: Every non-maximum was made to be zero.
Step-4: The higher threshold as well as lower threshold was assessed with the help of histogram of gradient magnitudes of images.
Step-5: For getting appropriate edge map hysteresis, threshold was used that links weaker as well as stronger edges.

Proposed GSOBCAN Edge Detector: This proposed edge detector algorithm was constructed with the help of hybrid Sobel and Canny algorithm. The principal features of this proposed edge detector is given below.

Reduced error rate: It is significant that edges that occur in images may not miss and also have no response in non-edges.

Better Localization: The distance amid edge pixel and original edge has to belesser.

Single Response: Possessing a single response for single edge.

These criteria were achieved by this proposed method in an effective way offering a satisfying visual of the retinal nerve cells of the Covid-19 affected persons and it has outperformed the other two existing algorithms.

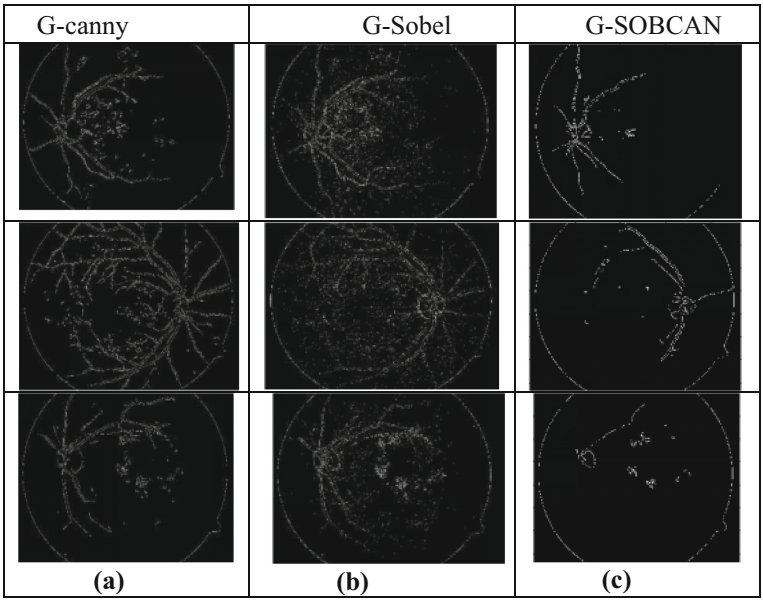


Fig. 5. Edge Detection Algorithms

4 Results and Discussions

The original image is various patients' eye retinal fundus images. The quality of images was not good. A complete analysis was conducted. Various types of image enhancement techniques were performed on retinal image to obtain better visual of an image. Then, various noises were added on retinal image and image was passed into hybrid noise filters to get the desired results. Then, various types of edge detection algorithms were performed to obtain good segmentation of nerve vessels of retina image, since it will be very useful for determination of statistical analysis of nerve vessels of diseased eye [28]. The performance was assessed in terms of PSNR, MSE and SSIM of original retinal images corresponding to the proposed output.

Peak Signal to Noise Ratio (PSNR): PSNR of the maximum possible signal power to the corrupting noise power is calculated using the below given expression, as,

$$PSNR = 20\,log\,10\left(\frac{MAX}{\sqrt{MSE}}\right) \tag{6}$$

Mean Square Error (MSE): MSE corresponds to the dissimilarity amid original image as well as recovered image, which shall be expressed mathematically as,

$$MSE = \frac{1}{mn}\sum_{0}^{m-1}\sum_{0}^{n-1}||f(i,j) - g(i,j)||^2 \tag{7}$$

Structural Similarity Index (SSIM): SSIM was employed to measure the similarity amid two images. SSIM was intended for improving the conventional approaches like PSNR as well as MSE that proved to be unpredictable for human visual perception. SSIM can be expressed mathematically as,

$$SSIM\,(x,y) = \frac{(2\mu_x\mu_y)2(\sigma_{xy} + C2)}{\left(\mu_x^2 + \mu_y^2 + C1\right)(\sigma_x^2 + \sigma_y^2 + C2)} \tag{8}$$

The qualitative metrics concerning various image enhancement techniques was performed and comparative analysis was obtained. The metrics such as PSNR, contrast, entropy was measured. The performance of image enhancement techniques was assessed over various eye fundusimagesofCovid-19 patients and other diseased patients from DIARETDB1 and NCBI databases.

Table 1. Performance of Enhancement Algorithm

Enhancement Algorithm	PSNR	Contrast	Entropy
HE	39.72	214	6.72
CLAHE	43.22	177	7.49
CEGC	38.19	220	6.89
GLSCE	44.76	236	7.53

Table 1 displays the performance of various image enhancement algorithms such as HE, CLACHE, CEGC and the proposed enhancement algorithm GLSCE on fundus images. The proposed GHE outperforms the existing methods based on the parameters such as PSNR, contrast and entropy.

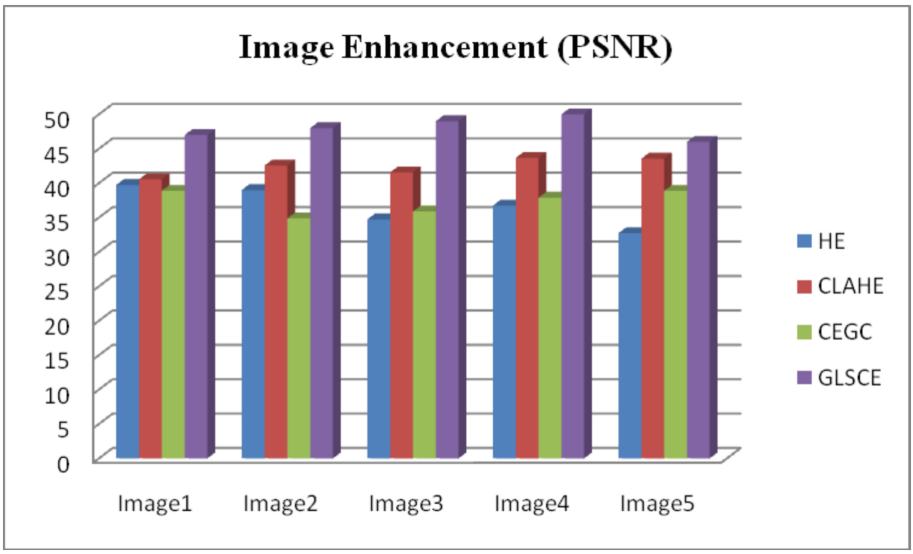


Fig. 6. PSNR Analysis

The performance of image enhancement method for dissimilar images with comparative assessment has been shown in Fig. 6.

Table 2. Performance evaluation of PSNR, MSE and SSIM on various images

Filters	PSNR	MSE	SSIM
MedianFilter	44.43	4.6E−05	0.9801
AdaptiveFilter	42.31	6.5E−05	0.9743
MeanFilter	35.89	3.1E−06	0.9720
Wiener Filter	40.12	4.7E−05	0.9817
GWMF	46.24	2.5E−05	0.9822

Table 2 shows performance evaluation of PSNR, MSE and SSIM on various images with salt and pepper noise filtered by median, wiener and proposed GWMF filtering. It shows mean value for PSNR, MSE and SSIM of 5 images with various noise density (0.02–0.06) using these filters. The proposed WMF algorithm outperforms the existing filtering techniques. Figure 7 shows the comparative analysis of various image noise

filters performed on eye fundus image database with different noise densities for two categories of noises namely salt and pepper noise and Gaussian noise. The filters such as median, adaptive, mean, wiener and proposed WMF was applied on five fundus images. Their comparison is based on qualitative metrics namely PSNR, SSIM and MSE. The proposed WMF outperforms the peers.

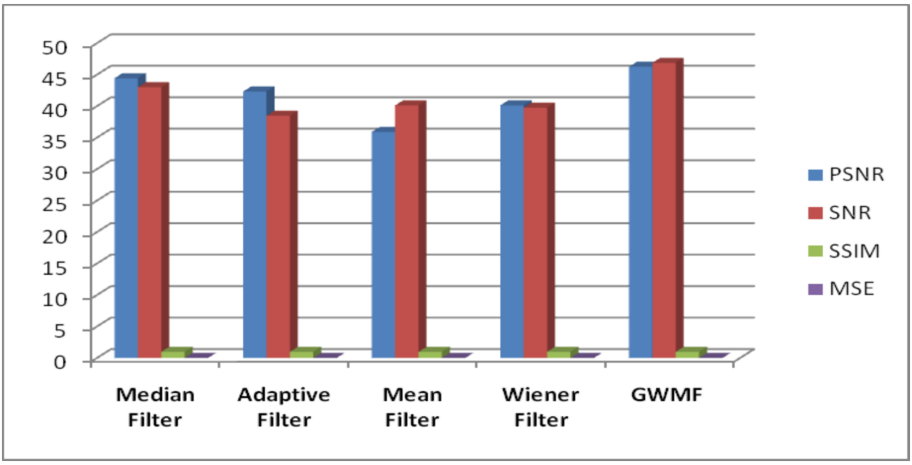


Fig. 7. Performance of Various noise filters on fundus images

5 Conclusion

The different image enhancement methods and outcomes applied to the 50 base images were compared with each other on quality metrics such as PSNR, contrast, and entropy has been summarized. All things considered, however, are discussed. In these cases, the proposed improvement technique outperforms other existing techniques. Five spatial filters were then applied to the fundus images and performance was evaluated based on parameters including MSE, PSNR and SSIM. The experimental results showed excellent performance at different noise densities where the proposed hybrid filters outperformed the other filters. The uniqueness of this paper is that, two categories of noises over fundus image were assessed, similarly dissimilar categories of filters were explained that operated under green component enhanced image with enhancement techniques and for removing the noise for increasing the resultant images' quality. After removing the noise, the nerve vessels were segmented with the help of edge detection methods. The proposed edge detection strategy corresponds to a series of edge detection algorithms performed on a noise-free background image. In the future, additional image filters and other methods of segmentation algorithms, which provide good visual effects of points and nerve vessels in the visual image with additional accuracy, should be evaluated.

References

1. Mvoulana, A., Kachouri, R., Akil, M.: Fully automated method for glaucoma screening using robust optic nerve head detection and unsupervised segmentation based cup-to-disc ratio computation in retinal fundus images. Comput. Med. Imaging Gr. **77**, 101643 (2019)
2. Mitra, A., Roy, S., Roy, S., Setua, S.K.: Enhancement and restoration of non-uniform illuminated fundus image of retina obtained through thin layer of cataract. Comput. Methods Progr. Biomed. **156**, 169–178 (2018)
3. Almotiri, J., Elleithy, K., Elleithy, A.: Retinal vessels segmentation techniques and algorithms: a survey. Appl. Sci. **8**, 155 (2018)
4. Zhu, C., et al.: Retinal vessel segmentation in colour fundus images using extreme learning machine. Comput. Med. Imaging Gr. **55**, 68–77 (2017)
5. Wong, T.Y., Loon, S.C., Saw, S.M.: The epidemiology of age related eye diseases in Asia. Br. J. Ophthalmol. **90**, 506–511 (2006)
6. Kälviäinen, R., Uusitalo, H.: DIARETDB1 diabetic retinopathy database and evaluation protocol. In: Medical Image Understanding and Analysis, p. 61. CiteSeerX (2007)
7. Reddy, P.S., Singh, H., Kumar, A., Balyan, L.K., Lee, H.-N.: Retinal fundus image enhancement using piecewise gamma corrected dominant orientation based histogram equalization. In: 2018 International Conference on Communication and Signal Processing (ICCSP), pp. 124–128. IEEE (2018)
8. Burger, W., Burge, M.J.: Digital Image Processing: An Algorithmic Introduction Using Java. Springer, Berlin (2009)
9. Joshi, G.D., Sivaswamy, J.: Colour retinal image enhancement based on domain knowledge. In: 2008 Sixth Indian Conference on Computer Vision, Graphics and Image Processing, pp. 591–598 (2008)
10. Jain, K., Arya, I.B.: A survey of contrast enhancement technique for remote sensing images. Int. J. Electr. Electron. Comput. Eng. **3**, 1 (2014)
11. Raj, A., Tiwari, A.K., Martini, M.G.: Fundus image quality assessment: survey, challenges, and future scope. IET Image Process. **13**, 1211–1224 (2019)
12. Rashmi Choudhary, S.G.: Survey on image contrast enhancement techniques. Int. J. Innov. Stud. Sci. Eng. Technol. **2**, 21–25 (2016)
13. Westerweel, J.: Fundamentals of digital particle image velocimetry. Meas. Sci. Technol. **8**, 1379 (1997)
14. Yuan, X., Gu, L., Chen, T., Elhoseny, M., Wang, W.: A fast and accurate retina image verification method based on structure similarity. In: 2018 IEEE Fourth International Conference on Big Data Computing Service and Applications (Big Data Service), pp. 181–185 (2018)
15. Liao, M., Zhao, Y.Q., Wang, X.H., Dai, P.S.: Retinal vessel enhancement based on multi-scale top-hat transformation and histogram fitting stretching. Opt. Laser Technol. **58**, 56–62 (2014)
16. Haldar, R., Aruchamy, S., Chatterjee, A., Bhattacharjee, P.: Diabetic retinopathy image enhancement using vessel extraction in retinal fundus images by programming in raspberry pi controller board. In: 2016 International Conference on Inter Disciplinary Research in Engineering and Technology, p. 37 (2016)
17. Xiong, L., Li, H., Xu, L.: An enhancement method for color retinal images based on image formation model. Comput. Methods Prog. Biomed. **143**, 137–150 (2017)
18. Lin, C.L.: An approach to improve the quality of infrared images of vein-patterns. Sensors **11**, 11447–11463 (2011)
19. Chiu, C.C., Ting, C.C.: Contrast enhancement algorithm based on gap adjustment for histogram equalization. Sensors **16**, 936(01–18) (2016)
20. Zuiderveld, K.: Contrast Limited Adaptive Histogram Equalization. Academic Press Inc., Pittsburgh (1994)

21. Pisano, E., et al.: Contrast limited adaptive histogram equalization image processing to improve the detection of simulated spiculations in dense mammograms. J. Digit. Imaging **11**, 193–200 (1998)
22. Jintasuttisak, T., Intajag, S.: Color retinex image enhancement by Rayleigh contrast limited histogram equalization. In: International Conference on Control, Automation and Systems, vol. 10, pp. 692–697 (2014)
23. Abuturab, M.R.: Multiple color-image authentication system using HSI color space and QR decomposition in gyrator domains. J. Mod. Opt. **12**, 1–16 (2015)
24. Jee, Y.: WHO international health regulations emergency committee for the COVID-19 outbreak. Epidemiol. Health **42**, e2020013 (2020)
25. Port, A.D., Orlin, A., Kiss, S., Patel, S., D'Amico, D.J., Gupta, M.P.: Cytomegalovirus retinitis: a review. J. Ocul. Pharmacol. Ther. **33**(4), 224–234 (2017)
26. Agarwal, A., Invernizzi, A., Acquistapace, A., et al.: Analysis of retino choroidal vasculature in human immunodeficiency virus infection using spectral-domain OCT angiography. Ophthalmol Retina **1**(6), 545–554 (2017)
27. Varga, Z., Flammer, A.J., Steiger, P., et al.: Endothelial cell infection and endotheliitis in COVID-19. Lancet **395**(10234), 1417–1418 (2020)
28. Casagrande, M., Fitzek, A., Püschel, K., et al.: Detection of SARS-CoV-2 in human retinal biopsies of deceased COVID-19 patients. Ocul. Immunol. Inflamm. **28**, 1–5 (2020)

Solidarity Linking Women and Technology: An Analysis

Anuradha Acharya(✉) and Kenil Thakkar

LJ School of Computer Applications, LJ University, Ahmedabad, Gujarat, India
anuradha.acharya@ljku.edu.in

Abstract. In today's dynamic landscape, the intersection of education, technology, and women's empowerment stands as a beacon of progress and opportunity. This paper explores the transformative potential of technology in empowering women, especially in societies like India where gender norms often pose significant barriers. We aim to dissect the nuanced relationship between technology and women's empowerment, focusing on its impact on various aspects of their lives, from mind set to communication skills. Our analysis spans urban and rural settings, shedding light on both the challenges and opportunities presented by technological advancements. Through a comparative analysis, we unveil the distinctive role technology plays in reshaping women's trajectories, offering insights into how platforms like WhatsApp, Instagram, and LinkedIn serve as catalysts for social and professional growth. Earlier studies focused on how women can be uplifted through technology and what are the challenges she faces and is based on women empowerment. This paper not only tried to explore the role of technology in women's life and in women empowerment, but also tested two important hypotheses based on education and traditional barriers. The results shown that the high cost of education and traditional blocks are the hurdles in the progress of women. Ultimately, the findings underscore the indispensable role of technology in driving comprehensive progress and empowerment for women in today's world and existence of very close relationship between both.

Keywords: Technology · Education · Women empowerment · Literacy rate · Traditional barriers

1 Introduction

In contemporary society, technology is a pivotal force influencing various sectors, playing a crucial role in individual and collective development on a global scale [29]. It has brought unprecedented convenience to our lives, offering facilities that were once unimaginable. Technology encompasses scientific knowledge tailored for specific purposes, significantly enhancing daily tasks, industrial processes, and commercial activities, making them more conducive, faster, and efficient [30]. Its pervasive influence is evident across diverse sectors such as healthcare, education, finance, transportation, and communication, revolutionizing the quality of life, work processes, and interaction styles

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
S. Rajagopal et al. (Eds.): ASCIS 2024, CCIS 2428, pp. 161–182, 2025.
https://doi.org/10.1007/978-3-031-86302-8_11

[30]. Women in engineering education experience a lot more stress than men because of their jobs, the outcomes of their work, and the roles they have to take on. This is especially true with technology changing things all the time [1].

Women possess the capacity to contribute significantly to societal progress across social, economic, educational, and industrial domains, necessitating empowerment. Education, technology, and the implementation of fundamental rights are key factors in empowering women. Technology, in particular, plays a crucial role in the lives of women, enhancing their skills and providing opportunities for diverse participation in society, ultimately contributing to increased economic benefits for the nation. Examining women's representation in computing through a life course perspective reveals disparities across various stages, from pre-high school access to post-baccalaureate employment, highlighting gender-related challenges in interest, attitudes, and major choices [2].

The potential of technology extends to fostering positive changes in freedom of expression, education, and employment opportunities. Despite strides in accessibility and numerous initiatives addressing gender imbalance in technology, the abstract emphasizes the need for longitudinal studies to comprehensively understand women's pathways into computing fields and the intersecting impact of interventions across different life stages [2]. Women, like men, require technology for skill development, informed decision-making, networking, self-promotion, and recreational purposes, highlighting its universal importance in empowering individuals.

1.1 Technologies in Today's Era

Technology has made it possible for people from all walks of life to have access to different resources. Technology is also being used to equip people who need assistance for improving their quality of life and help them avail opportunities that would otherwise be inaccessible [31]. Technology is the branch of knowledge and it has brought many changes in our lives in terms of saving time, helping us through instant communication and interactions, easy access to information and ensuring safety also. From education to health, career opportunities to safety, and transportation to entertainment, technology has brought immense benefits and changes [13, 31]. However, it is must to know how it has impacted the women's life too.

1.2 Woman's Role in the Society

In our society, women play an important role in terms of love, care and nurturing the family. They work for the well-being of their family and community. Their contribution can't be denied in the enhancement of every individual. Women have important roles to perform in all walks of life. Beyond this her presence can be seen and felt in sectors like social, political, economic, and religious spheres. They are capable enough with proper skills and abilities and thus can render their participation everywhere in an effective manner. But several factors in the background are affecting their participation and roles. Somewhere they need to overcome these challenges they are facing and this problem must be given importance and should be solved. Reaching to the roots of these issues can enable them in overcoming the barriers that are arising in their performance and roles. Since technology is helpful in bridging various gaps as can be seen in many fields

of life, it can be known whether technology can be helpful for them to play a better role for them.

1.3 Background Study

As per various research, conducted yet, it has been found that women are less exposure to the working environment and especially related to the IT industry. In developing countries like India, women status is considered minor despite of her total contribution everywhere. Normally they are abused or misused in every sector. Thus, women empowerment is an important issue. Different NGO's and even Government is trying to uplift her through various programs. However, ICT can provide women security, knowledge, confidence, and employment etc. [4]. Education is one of the greatest weapons for bringing drastic changes in the globe. Information and communication technology (ICT) can be one of the medium to help in empowering the women especially, in developing countries like India. ICT can empower women in various areas like education, personal, political, economic and many more [5]. Women empowerment and development are correlated to each other. The empowerment of women can lead to the development of any nation and in bringing down the inequality between man and women. Despite facing barriers, women possess untapped innovation potential, particularly in the realm of renewable energy, where their contributions can significantly enhance technology advancements and improve overall quality of life [3]. in this view there is an increasing trend in the employment of women in the light of technology also. The role of technology can be also having some impact in the increasing literacy rate for women. Thus, the role of technology can be felt in uplifting the literacy rate in women, thus leading to the overall improvement of the society.

1.4 Technological Aspects of Woman's Life

Technology is related to the scientific knowledge that can be applied practically for any type of application. Normally women don't have strong link with technology if understood in overall. But if provision can be made to acquire the advantages for women for her better education and thus to empower the family, community, and the nation. The integration of a mobile application, combining legal guidance on violence against women, health tips, and an emergency call system, exemplifies how technology plays a crucial role in empowering women by providing accessible tools for their well-being and safety [14].Technology has been a game-changer in many aspects of our lives, and one area where it has a significant impact is in advancing the rights and opportunities of women [33]. In this paper, it is tried to explore the ways in which technology can be helpful in women's life.

Following are the varied ways in which technology can said to be supportive in reforming to the lifestyles of women.

1.4.1 E-Learning

Learning, a transformative process, enriches one's knowledge through diverse experiences. Modern technology has simplified this process, making education more accessible through online platforms and mobile learning systems. Education is a fundamental

human right, and gender should never be a barrier to accessing it. However, in India, female literacy rates still lag behind those of males, as seen in recent data (Table 1, Fig. 2). Girls, especially in rural areas, are less likely to attend school or complete their education compared to boys. Jawaharlal Nehru's famous words, "If you educate a man, you educate an individual; if you educate a woman, you educate a whole family," highlight the transformative potential of women's education for society. Women's empowerment is crucial to national progress, as economic development depends on the contributions of both men and women.

In 2021, India's female literacy rate was 91.95%, marking a 14.4% increase from 2010 to 2022, with an average annual increase of 0.6%. However, a deeper examination of progress in women's safety, societal consciousness, and literacy rates across different periods provides a more nuanced understanding. Along with literacy improvements, greater focus on women's safety and increased societal awareness around gender equality have emerged as critical factors in ensuring lasting empowerment. Comparing literacy trends across decades shows that while female education is advancing, challenges like dropouts in rural areas persist, indicating that education policies must also address social and infrastructural barriers.

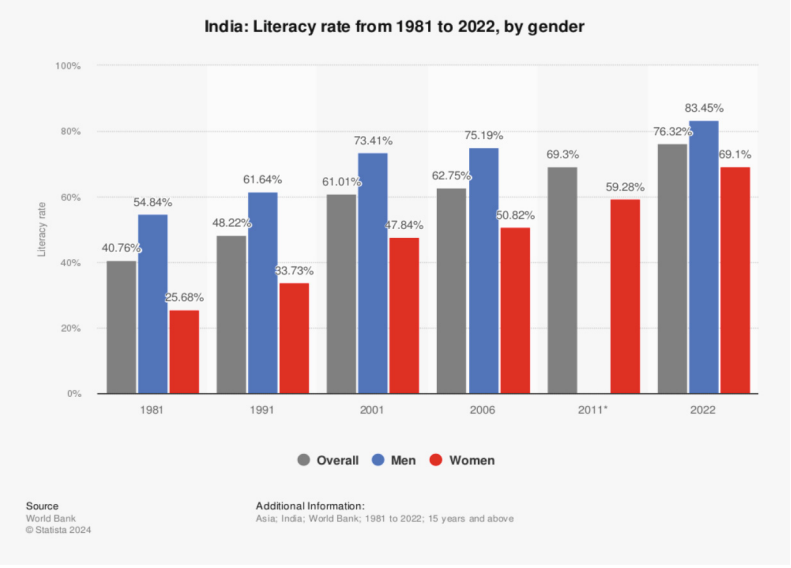


Fig. 1. Female literacy rate in India (2010–2021) [32]

Only education can empower women to create their own identity, to build a good career and to make her independent. In the past, women were not allowed to go outside for perusing their education and gain knowledge but technology can make the things possible in this regard. With the help of technology, women can have online education and thus making it easier for them to avail this opportunity to pursue knowledge without leaving their safe premises [34]. It can provide access to online educational resources like E-Learning platforms, virtual classrooms and so on which can help them to build their

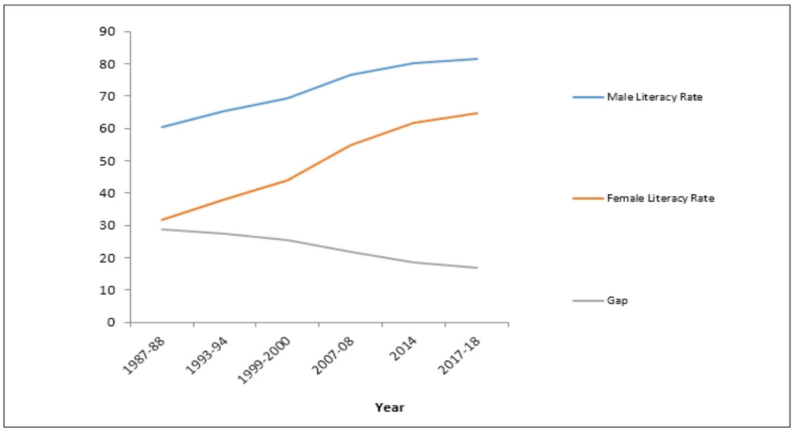


Fig. 2. Trends in male and female literacy [43]

Table 1. Yearly Gender Wise Literacy Rate in India [43]

	1987-88	1993-94	1999-00	2007-08	2014	2017-18
Male Literacy Rate	60.5	65.5	69.2	76.6	80.3	81.5
Female Literacy Rate	31.7	37.9	43.8	54.9	61.8	64.6
Gap	28.8	27.6	25.4	21.7	18.5	16.9

skills and to enhance their knowledge. E-Learning provides flexibility to them in learning as per their customization [35], by helping them to deal with their busy schedules, child, and family responsibilities. E-Learning can provide training opportunities to them, which were previously difficult to access. Through this learning mechanism, women can bring positive impacts in the national economy and a lot of potential in changing the world.

The bar chart in Fig. 1 presents India's literacy rates from 1981 to 2022, categorized by gender. Overall literacy has significantly increased from 40.76% in 1981 to 83.45% in 2022. Men's literacy rate saw a consistent rise, going from 54.84% to 90.11%, while women's literacy experienced a remarkable improvement, jumping from 25.68% in 1981 to 76.25% in 2022. A comparative analysis shows that the gender gap in literacy has been steadily narrowing, with the 29% disparity in 1981 shrinking to 13.86% by 2022. These trends underscore the positive impact of educational reforms and efforts to promote gender equality in education, particularly for women.

1.4.2 Transportation Service

Transportation service is very important in these days for a smooth and firm connectivity in every way and it has made the travel easier and convenient than ever before. The advancement of technology and infrastructure has made the transportation fast, efficient and more accessible. It is possible to travel long easily with the help of modern transportation [9]. Transportation service can be a great help for women, especially for ensuring their safety and providing them many opportunities.

The various ways in which transportation service is beneficial to women are:

Safety: Transportation service must provide safety to women, especially during late-night hours when public transportation may not be available or not safe. Online cabs booking provide special features of safety for women to ensure the safety of passenger such as GPS tracking, emergency buttons, background-checked-drivers, rating and reviews.

Flexibility: Women are facing many challenges and having multiple responsibilities, including work, family and as a caretaker. So time management has become another challenge for women to handle everything. The transportation service offers flexible scheduling and many more options that can be a great help to women. This can especially beneficial to working women, where they can book cabs as per their flexibility. According to one of the surveys, only 12% of women in Ahmedabad, Gujarat, India is using the public transportation services [10].

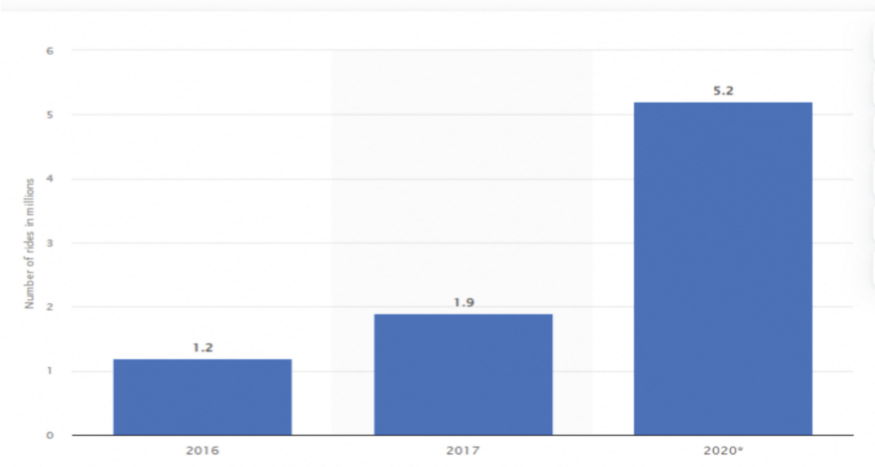


Fig. 3. Growth rate of transportation services [36]

As per data in Fig. 3, the global online taxi services market size will grow from $34.82 billion in 2022 to $38.54 billion in 2023 at a compound annual growth rate (CAGR) of 10.7%, and it is expected that the global online taxi services market size is expected to grow to $56.37 billion in 2027 at a CAGR of 10.0% [36].

1.4.3 Improved Communication

The evolution of new technologies has revolutionized the communication process. The technology makes it possible for humans to communicate instantly with each other with the help of multiple platforms and channels such as messaging, text, emails and video conferencing [37]. It gives opportunities to collaborate projects with each other irrespective of their locations. Technology provides many tools to collaborate online, like Google docs and Trello that allows people to collaborate their work on their projects [38]. As technology continues in evolves, then the communication will continue to become more efficient and effective in the future. This opportunity can also be accessed by women for a better communication purpose through the adoption of technology by women.

1.4.4 Social Media

Social media has become a primary means of communication and source of information, entertainment for everyone. Social media is a platform where the people share information, communicate with each other, spread awareness and many more. Especially, using social media women can share their voice and stories [39]. It allows women to speak up their opinions and to be aware of some important issues. With the help of this online media, women can get information on online jobs and opportunities to earn. Social media has opened up new opportunities for women to start and grow up their business. Platform like Instagram and Facebook provide a means for them to do small businesses and to earn from home [11]. According to reports, the female users are always less in social media also. However, there is a hike in the usage rate from 2018 to 2022 which is a positive indication in terms of the role of technology in the women's empowerment (Fig. 4).

1.5 Technological Challenges Faced by Women

The involvement of women in decision-making, leadership and management can bring lots of differences especially in the national economy. However their participation with reference to the technological challenges should be understood and explored for a global success. Few challenges they normally faced in technological world have been discussed below.

1.5.1 Lack of Access

The access of technology has become an essential requirement for any participant in every area. But there is a gender-gap in access and usage of technology in our society. One of the reasons for lack of access of technology for women is social-economic status. Even today women are facing problems because of limited access of education and poor infrastructure and limited access of technologies [12].

Additionally, women have to face many more challenges such as gender-based discrimination, cultural and social barriers. These barriers restrict them from education, professional career and have limited access for training and education in technology related fields [12]. These further limit their economic opportunities, and thus affecting

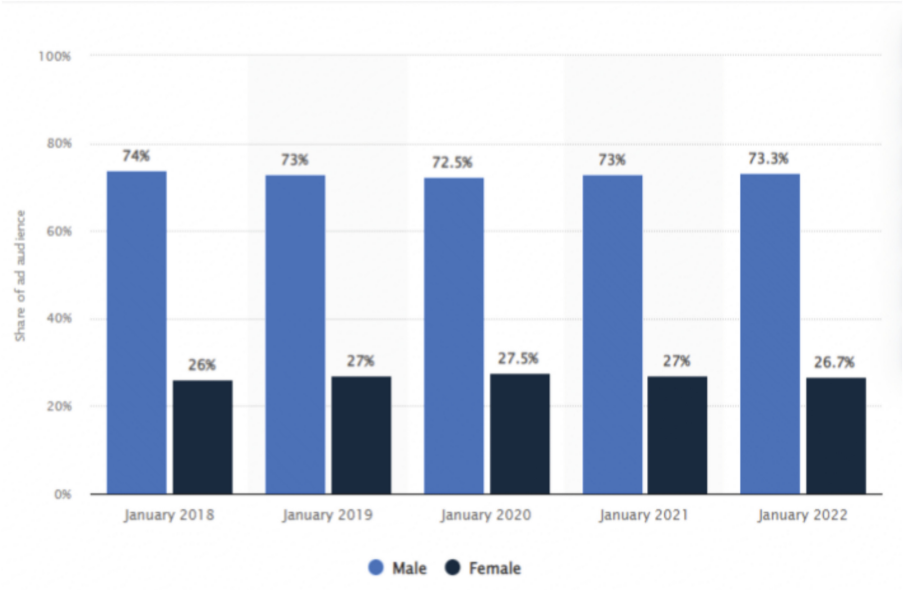


Fig. 4. Year wise report on the usage of social media by men and women [40]

the national economy. Tech events and conferences are often male dominating, making it challenging for women to connect with peers and industry leaders [42].

The challenges faced by women that restrict them to avail the benefits of technology can be depicted as in the Fig. 5.

Fig. 5. Barrier experience by women while using technology [41]

1) ***Privacy Concerns***

Women are concerned for their privacy and safety more than man in social media [45]. However technology can have an access to the personal data thus putting a question mark for the safety aspects of women. Some step should be taken for protecting

such data through strong passwords, keeping away third-party sites login through social credentials, and regularly reviewing their privacy settings [46]. Women should be aware of the privacy settings on their social media accounts and take steps to protect their privacy. This includes limiting the amount of personal information they share online, being selective about who should be added as friends, and using privacy settings to restrict who can see their posts etc. By being vigilant and proactive about their privacy and security, women can help to ensure that they are safe and secure online.

i. ***Anonymous Sharing***

Anonymous sharing is one of the communication practices that youngsters use through apps like Snapchat. Such apps allow the users to temporarily remove the posted images and messages, leading to privacy concerns, as in the world of internet nothing is temporary. These data can be under the trap of cyber thieves and bullies who can steal the private data like screenshots, photos or images etc. some of the apps like Whisper keeps the user's identity and thus can provide controlled revealing of data to be used for interaction purpose [15].

ii. ***Affordability***

Affordability can be considered as one of the major reasons because of which the women are departed from the access of education and technology. Being the backbone of house, women have to take lots of responsibilities. So there is less chance that they can afford for learning or using latest technology as they don't have the ability or chance to work on full-time or part-time for earning purpose. Thus, they are away from the benefits that are available through the technology.

Being a weak position in the society, social discrimination, gender inequality, violence, security and economic exploitation are some of the major causes because of which women are facing lots of problem to avail the opportunity of education and accessing the technology. They are thus behind the men in all senses [47].

iii. ***Challenges for woman: Urban vs. Rural***

Literacy is the biggest challenge in India which can be understood from the Fig. 6, where females are found to be less literate than males.

India has reached to such a point of success, where the success can be seen in terms of information technology, economics, peace, and science. However, it is facing one of the biggest challenges as rural India [16, 43]. This sector should not be neglected in any way where rural India women are facing lots of difficulties, struggling through problems like financial independence, stability, gender inequality, domestic work, social norms and many more [49].

Literacy rate is one of the major cause because of which women from rural area are facing more challenges. Somewhere they have been departed from technology and education which is putting the biggest barrier in their up gradation. Figure 7 shows the gender bias in terms of the literacy rate difference between men and women. If the further comparison can be made between rural and urban women in the literacy rate the kind of situation and unprivileged conditions of rural women can be understood. A comparison of the literacy rate between men and women in rural vs. urban areas can be seen in Fig. 8.

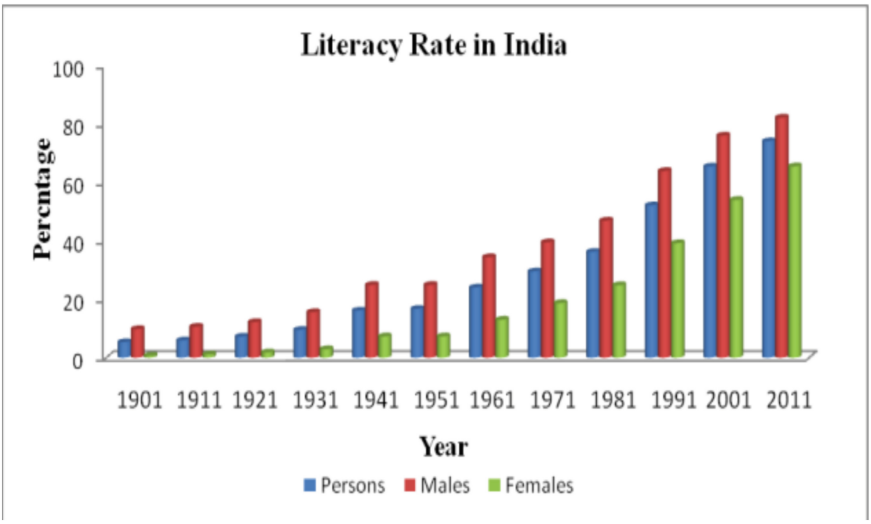


Fig. 6. Literacy rate in India [22]

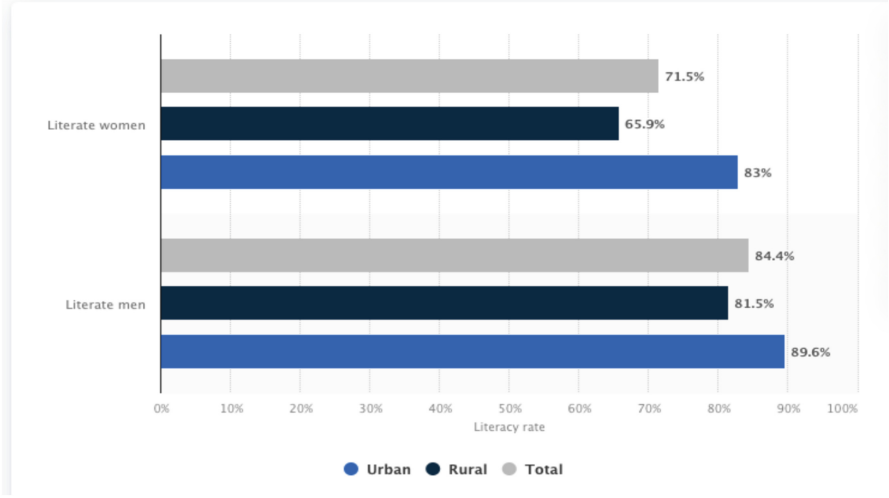


Fig. 7. Literacy rate among women and men in India between 2019 and 2021 [44]

The Fig depicts that the rural female is facing the problem of education severely than others.

Some of the biggest challenges that the rural women face can be listed as.

- Personal challenges like household work, cleaning, cooking, bringing up children. [50]
- Social attitude, constraints, gender-based violence and age inequality [17, 45]
- Poverty, cost of education, inadequacy financial supports [50]

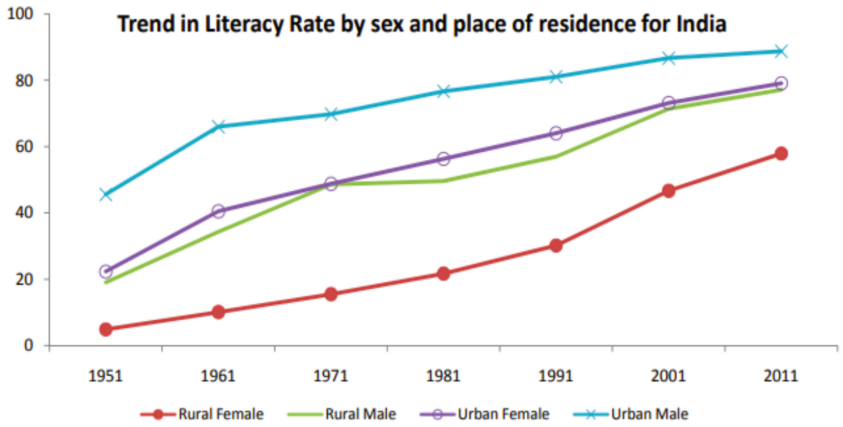


Fig. 8. Trend in literacy rate among women and men (Rural vs. Urban) [51]

- Education and mobility challenges [50]
- Lack of resources, disproportionate in the paid amounts [17]
- Risk of abuse and sexual harassment [17]
- Less decision-making power, political, economic, and social status [17, 18].

2) ***Research Objective***

We all are aware of this fact that technology has a great role in shaping the future generation. Also, if woman is educated, she can make the entire family an educated one. Technology is contributing in a varied ways towards the success and prosperity through education. In this context, if it is understood how, it can be helpful for women, it would be easier to build the nation with a high end through technology up gradation. Few things that have been tried to find are like how women are using technology in their day-to-day lives, whether employed or uneducated, what are the difficulties they are facing to use modern technology, what opportunities she is getting through technological platforms and gadgets. Accordingly, some of the objectives of this paper are as follows.

- How has technology improved the lifestyles of women?
- What are the types of services women are availing through technology?
- What are the difficulties they are facing in using technology?
- To know the main motive of using technology by women
- What are the various ways and to what extent technology is helping women to withstand?
- What are the benefits they are finding by using technology?

2 Literature Review

Recent literature underscores the profound impact of rising female literacy rates, particularly in India, on the digital landscape and women's participation in technology-related fields. In the past three years, several studies have highlighted the emergence of the

technical sector as a significant employer of women, following agriculture, showcasing substantial strides in women's engagement with technology [21]. Furthermore, case studies such as the Women Empowerment Cell demonstrate how technology can amplify initiatives aimed at advancing women's status in India [6].

The transformative potential of Information and Communication Technology (ICT) in women's lives has been extensively explored. Studies emphasize ICT's role as a powerful tool for affording women greater agency in decision-making processes, political advocacy, and economic empowerment across various dimensions [52]. However, the persistent gender divide in ICT usage remains a critical concern, limiting women's active involvement and representation across sectors.

Mobile technology, integral to bridging the digital gap, has garnered significant attention, especially in rural areas. Recent research highlights the affordability of mobile phones and improved network facilities, facilitating the establishment of mobile-based healthcare systems nationwide [20, 52]. These systems promise to transparently enhance healthcare access, transcending geographical barriers and improving overall population well-being.

Challenges faced by women in technology, particularly in the IT sector, have also been examined in recent literature. Studies identify issues such as remuneration disputes, security concerns, and gender-based disparities that hinder women's ability to leverage IT for personal development and realization of their full potential [53].

Authors underscore the multifaceted roles women can play in various domains and their potential contributions to the Indian economy. However, they assert that addressing prevailing issues such as violence and exploitation is crucial for women to harness their capabilities fully [23]. Technology-driven interventions tailored to overcome barriers are deemed essential for fostering increased representation of women in science and technology.

The main barrier for women is the gender bias, and the potential of women for using technology can never be denied [25]. Women shows interest in science and technology but their low representation in this regard are the social and economic concerns [26]. Women empowerment and technology are closely related. Based on the research related to women's empowerment and technology between 2017 and 2021, it has been found that the studies are not prompt everywhere and is concerned with areas like health, agriculture, environment, entrepreneurship, and communication [9]. None of the studies were done in the area of education or cost of education. The education is one of the major way of uplifting women and to make them empowered. Thus this paper has based on two major hypothesis related to education, cost of education and self-independence of women in this era. In recent papers there is talk on impact of information and communication technology in women's life [27]. But no vigorous analysis has been done. Studies have been done how the technology and ICT can change women's life [28].

However some recent studies highlight the role of technology in advancing women's education, particularly through platforms like Massive Open Online Courses (MOOCs). These platforms enable women to access high-quality education regardless of geographical constraints, empowering them to contribute meaningfully to societal advancements [19].

In examining the empowering potential of women through IT, recent literature emphasizes its role in the long-term development of both rural and urban areas in India [24]. The need to identify and address barriers, particularly in rural areas, is emphasized, underlining the transformative potential of empowering women in the context of comprehensive national development.

When examining social media trends, there are clear parallels in how men and women engage with technology. As of 2023, women tend to be more active on social media platforms than men, particularly on platforms like Instagram, Pinterest, and Snapchat, which have larger female user bases. Women also tend to use social media for personal connections and brand engagement, while men more often leverage it for business and information gathering. For example, 76% of women versus 72% of men engage on social media platforms, and women are more likely to interact with brands and promotions online [54, 55].

As women's literacy rates rise and their digital engagement expands, their growing presence in online platforms is likely to have a profound impact on both social and economic sectors. The intersection of educational advancements and increased access to digital tools highlights the critical role of promoting literacy and technological inclusivity, especially for women. By addressing these areas, broader gender inequalities can be reduced, fostering a more equitable landscape across various fields.

3 Methodology

In this study, a survey methodology is employed to comprehensively examine the multifaceted role of technology in women's lives, encompassing education, employment, health, and social interactions. The rationale for selecting the survey method is rooted in its efficacy in capturing diverse perspectives and obtaining quantitative data on various factors directly or indirectly influencing women's lifestyles and broader societal sectors.

The survey instrument is meticulously designed to elicit insights into women's technological inclinations, utilization patterns, and preferences. Carefully crafted questions probe participants about their interests, preferences, and usage of technology, with a focus on social media engagement and employment status. By targeting young women, both employed and housewives, the survey aims to capture a wide spectrum of experiences and perspectives.

The survey design ensures thoroughness and reliability in gathering structured data to understand the intricate dynamics of women's technological engagement. Statistical analysis techniques are applied to the collected data to uncover trends, correlations, and insights regarding women's utilization of technology across various domains.

The experimental procedures involve the following steps:

Survey Design: Development of a comprehensive survey instrument with carefully crafted questions to capture diverse aspects of women's technological engagement, including social media habits, employment status, and preferences.

Participant Recruitment: Targeted recruitment of young women, comprising both employed individuals and housewives, to ensure a representative sample reflecting diverse experiences and perspectives.

Data Collection: Distribution of the survey among participants through various modes, ensuring accessibility and convenience for respondents to provide their thoughts.

The collected data are analyzed thoroughly for understanding how both technology and women are correlated and what can be done to bring the technology into implementation especially inclined towards the empowerment of women.

4 Hypothesis

Technology plays an important role in everyone's life though it also has some negative impacts and risk associated with it. Several factors are responsible in restricting women to get the advantages of technology. Technology can be considered costly and it may not be that much ease for everyone to access it. Especially it can be more challenging for women in this respect. Today gender inequality has become one of the major factors affecting the development of women. In today's modern era, in some places and villages women are not allowed to do jobs or to get education because of traditional, family and gender based restrictions. In India people who are orthodox and following the existing rituals, keeps women more busy in household works and are thus kept away from the fundamental rights of education and freedom to prove themselves in the society.

Thus, two hypotheses can be formulated here are:

I. Women are not getting the opportunity for education somewhere because of the cost of education is high.
II. The existing culture and rituals are keeping the women away from education and independence.

5 Result Analysis

The research scrutinized information obtained from various participants to examine the contemporary role of technology in enhancing women's lifestyles and how both the technology and women are entangled with each other. Various questions were utilized to determine the extent to which technology has played a role in improving their lifestyles, identify the challenges and obstacles they encounter in accessing these technological benefits, and understand the specific ways in which technology proves advantageous for them (Fig. 9).

The graph illustrates the literacy rate trends in India from 1981 to 2022, segmented by overall literacy, men, and women. The data [32] shows a steady increase in literacy over the years, with significant improvements for both genders. Using a linear regression model, the future literacy rates for 2030 were predicted based on historical trends. The Python code used for this involved fitting a linear regression model to the existing data, with the year as the independent variable and literacy rates as the dependent variable. The predicted literacy rates for 2030 indicate that the overall literacy rate is expected to reach 84.92%, with men's literacy projected at 91.93% and women's at 78.40%. This suggests a continued closing of the gender gap, although disparities between men and women in literacy levels will likely persist, underscoring the need for ongoing educational efforts, particularly aimed at women.

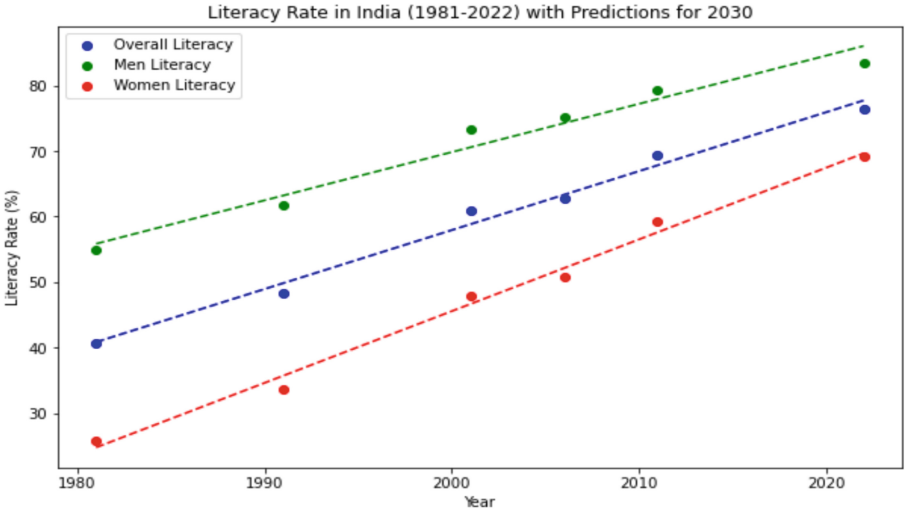


Fig. 9. Literacy Rate in India with Predictions

In the survey it is observed that technology impacts and is beneficial for 97.1% of women (Fig. 10). It is found in the study that most of the cases women use it for services like UPI transactions, online shopping, entertainment, and transportation services. Technology is being used for UPI transactions by 74.3% women, which is a great achievement in the field of online transaction services. Around 72.9% females are using technology for online shopping, 57.1% for online grocery, 72.9% are using for entertainment and 64.3% are using transportation services, The use of technology for transportation services is a good indication regarding the secured transportation for women these days (Fig. 11). The Fig further gives us the sign that women who are unemployed or housewives, are also using some of the obvious facilities provided by technology. The study also found that 56.5% of women are employed.

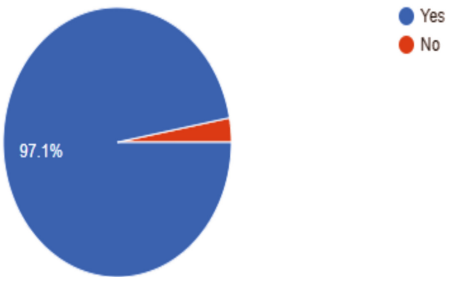


Fig. 10. Technology contributing in the development of women's life.

According to Figures, 81.4% of teenaged girls are using technology for the purpose of education, 80% for social media, 70% for entertainment, 60% are using it for better communication skills and 58.6% for transportation services. These data shows

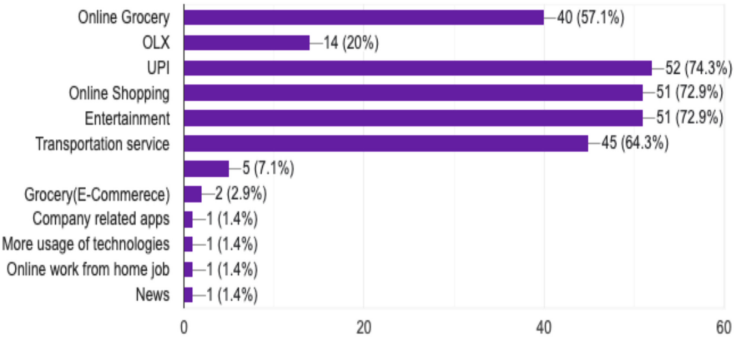


Fig. 11. Distribution of different services used by women

that females are more inclined towards education and are dependent on technology for this. One of the other responses stated that 87% women are preferring offline education which somewhere indicates the type of restrictions they might be facing because of traditional beliefs to attend college or institutes for higher education.95.7% women agreed that technology is helpful for them in getting education in many ways. 97.1% feel that technology is helpful for them in improving their communication skills. They analysis also indicates that, being a social animal, women are interested social interactions, and technology is a greater media for them to be connected, with their friends and relatives for a healthy social life (Fig. 12).

Fig. 12. Purpose of using technology by women

It is confirmed somewhere that technology helped women in improving their communication skills. Through some questionnaire it is found that 72.5% women believed that improved communication skills further helped them in dealing with daily tasks more easily, 59.4% agreed that it helped them in facing some challenges. 50.7% feeling more confident in social gatherings and 49.3% women are using this for getting better jobs and to become independent somewhere (Fig. 13). The various ways through which it is possible for them to improve their communication skills are online classes, various applications and online translation apps and dictionaries. The various responses in terms of percentage is a clear indication regarding this as shown in Fig. 14.

The study tried to find out the various difficulties faced by women. Few factors considered in terms of the barriers are cost, complexity, traditional restrictions, lack of time, lack of access etc. 88.4% women said that technology could remove few problems

Fig. 13. Benefits of communication based technologies

Fig. 14. Means of improving communication skills

in their day-to-day life in some response. 85.7% women said that they normally don't face any difficulties in using the technology. Through this study it is tried to understand which of the factors played vital role in restricting them to use the technology. The survey told that 30% women agreed that cost is the biggest factor in availing technology. Whereas 37.1% females says that lack of access to technology and lack of time were the major barriers for them. 27.1% are having traditional restrictions and complexity as the major cause for using it (Fig. 15). These Figs indicates that almost all the factors contribute equally as a barrier in the usage of technology by women.

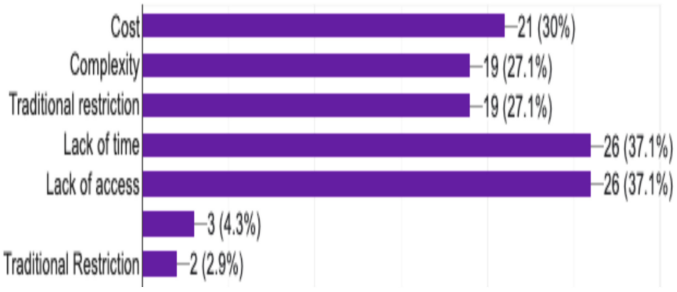


Fig. 15. Barriers faced by women while using technology

As the resultant data has already given the evidence that women want to have a healthy social life and is keen to have entertainment too. In this regard, it is found that 94.2%

of respondents are using WhatsApp, 69.6% are using Instagram. Beyond this 39.1% are using snapchat and 31.9% twitter. All these data are really a very good indicators about the social interactions of women these days (Fig. 15). Also, not only for social interactions they are using the social media apps but also for education and learning, entertainment, news and for their personal business and advertisement purposes as in Fig. 17. The Fig shows that most are using for entertainment. 78.3%. 75.4% are using for learning and education, 66.7% for social interactions and 53.6% for news. Only 33.3% are using for business purposes, which is again an alarm situation about their initiative in business and self-dependency (Fig. 16).

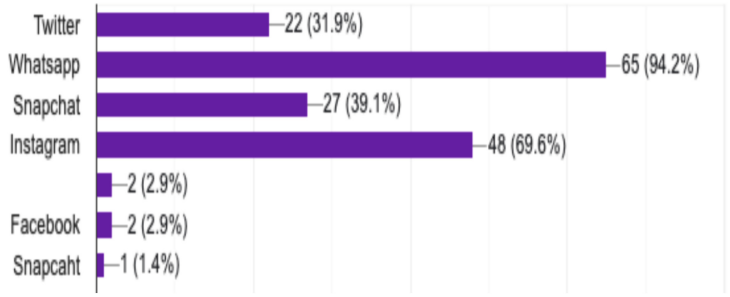


Fig. 16. Social media apps used by women

Fig. 17. Purpose of using social media app by women

Finally coming to the point, in what ways the technology can be helpful for women and especially mothers can be a greater help for upgrading the technology from women's point of view. Indian mother is normally engaged in their household work and in taking care of their child and technology has made their lives easy enough in many ways, according to the current study. It is found that technology has made a big impact in Indian mother's life. 87% women believe that technology helps them in upgrading their knowledge, 62.3% mother says child up gradation one of the ways in which technology is helping them and 60.9% says it has uplifted their life styles. Only 32% says it helps them to build their career and make their professional lives robust enough. 63.8% agreed that it made their lives easier, which is a very good Figure in this modern era of technology (Fig. 18).

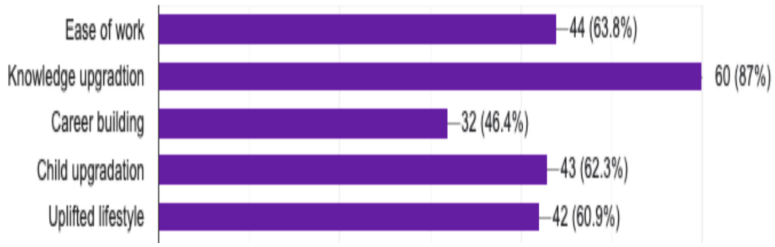


Fig. 18. Ways technology is helpful in mother's life

As per the survey 81.2% female are getting E-learning benefits with the help of technology and 76.8% females are using it for online transactions. Rest of the Figures are also very good signs regarding how technology is helping them in transportation, ticket booking and GPS (Fig. 19). All these results definitely mention that women is using technology for almost all traits of life.

Fig. 19. Benefit/s of modern technology in women's life

The above results are found to be very sound and informative in understanding to what extent technology can impact the women. Directly or indirectly, women are trying her best to avail the benefits of it. The Figures shows that she is equally capable like man to use technology and to contribute her participation in uplifting the economy of the great nation. They are using it for several purposes even after several restrictions and responsibilities they are having.

Studies have been done to understand rituals, gender gaps, role of technology for uplifting women etc. But none of them has tested any hypotheses especially related to education or cost of education nor the exact impact of traditional customs, which has been studied in this paper with a statistical approach. Further the studies have proved how technology and women are in parallelism for an overall growth of the women's life and the economy of the nation.

6 Conclusion

We have investigated the multifaceted role of technology in empowering women, particularly focusing on the challenges faced by rural women in accessing education and independence. Our hypotheses cantered around the high cost of education and entrenched cultural norms as significant barriers hindering women's progress. Through our survey and research, we aimed to assess women's current inclinations towards technology and its potential to alleviate their struggles.

Our findings indicate that while women, both married and unmarried, exhibit a keen interest in leveraging technology to enhance their lives, practical obstacles such as high education costs and entrenched cultural norms persist, particularly in rural areas. Despite these challenges, our survey suggests that women are willing to overcome these barriers to attain independence and contribute meaningfully to society.

Moving forward, future research efforts should extend their focus, especially towards rural women, and increase sample sizes to gain a more comprehensive understanding of the current situation. Addressing these challenges collectively is essential to empower women and grant them equal influence in nation-building efforts. By recognizing the potential of technology to bridge societal gaps and realizing women's aspirations, we can pave the way for a more inclusive and equitable future for everyone.

References

1. Rani, T.S., Prathiba, L., Basha, M.S.A., Khangembam, V., Sucharitha, M.M.: Are Women Employees in Engineering
2. Institutions Suffer from the Stress: An investigation approach. In: 2023 International Conference for Advancement in Technology (ICONAT), Goa, India, pp. 1–6 (2023). https://doi.org/10.1109/ICONAT57137.2023.10080369
3. Main, J.B., Schimpf, C.: The underrepresentation of women in computing fields: a synthesis of literature using life course perspective. IEEE Trans. Educ. **60**(4), 296–304 (2017). https://doi.org/10.1109/TE.2017.2704
4. Chopra, A., Aggarwal, P., Kumar, A., Lakra, S., Porwal, R.K., Narasimhan, S.R.: Women in engineering – role of women in renewable energy. In: 2023 IEEE PES Innovative Smart Grid Technologies - Asia (ISGT Asia), Auckland, New Zealand, pp. 1–4 (2023). https://doi.org/10.1109/ISGTAsia54891.2023.10372625
5. Kaur, N., Chawla, P.: Education and gender empowerment. In: 2021 9th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), Noida, India, pp. 1–4 (2021). https://doi.org/10.1109/ICRITO51393.2021.9596110
6. Beena, M., Mathur, M.: Role of ICT education for women empowerment. Int. J. Econ. Res. **3**(3), 164–172 (2012)
7. Priyadharshini, A., Thiyagarajan, R., Kumar, V., Radhu, T.: Women empowerment towards developing India. In: 2016 IEEE Region 10 Humanitarian Technology Conference (R10-HTC), Agra, India, pp. 1–6 (2016). https://doi.org/10.1109/R10-HTC.2016.7906859
8. Bhat, R.A.: Role of education in the empowerment of women in India. J. Educ. Pract. **6**(10) (2015). www.iiste.org. ISSN 2222-1735 (Paper) ISSN 2222-288X (Online)
9. Jayashree, S., Ramya, E., Malarvizhi, K., Kumaresan, N.: Women in science and technology in India: a view. In: Electrical Electronics, Communication, Computing and Automation (ICAECA), 8–9 October 2021. https://doi.org/10.1109/ICAECA52838.2021

10. Kochina, M.L., Chumakov, V.I., Budianskay, E.N., Kochin, O.V.: Advantages and disadvantages of informational technologies. In: 2006 16th International Crimean Microwave and Telecommunication Technology, vol. 1, pp. 51–53. IEEE (2006)
11. Shettar, R.M.: A study on issues and challenges of women empowerment in India. IOSR J. Bus. Manage. (IOSR-JBM) **17**(4), 13–19 (2015). Ver. I. e-ISSN 2278-487X, p-ISSN 2319-7668
12. Dixit, D.K.: The role of social media in the women empowerment. Int. J. (IJHSSI) **9**(6), 62–66 (2020). www.ijhssi.org. ISSN (Online): 2319 – 7722, ISSN (Print): 2319 – 7714. Ser. V
13. Cortesi, G., Lazzeroni, M.: Women and the access to knowledge and new technologies: the case Pisa. GeoJournal **61**, 229–237 (2004)
14. Bozieva, A.M., Zhurtov, A.V., Bozieva, M.M.: Modern information society and its impact on the education. In: 2018 International Conference "Quality Management, Transport and Information Security, Information Technologies" (IT&QM&IS), St Petersburg, Russia, pp. 587–588 (2018). https://doi.org/10.1109/ITMQIS.2018.8524904
15. Mahmud, S.R., Maowa, J., Wibowo, F.W.: Women empowerment: one stop solution for women. In: 2017 2nd on Technology, Information Systems and Electrical Engineering (ICITISEE), Yogyakarta, Indonesia, pp. 485–489 (2017). https://doi.org/10.1109/ICITISEE.2017.8285555
16. Desai, V., Kavita, S., Oza, K., Gaikwad, A.S.: Impact of internet technology on safety in international conference (2022)
17. Shambunath: Problems and challenges faced by urban working women in India. IJCRT **5**(3) (2017). ISSN 2320-2882
18. Women, U.N., Kaizer, F.A.: COVID-19 and women's rights organizations: bridging response gaps and a more future. Policy Brief (20) (2021)
19. Saravanakumar, P., Elizabeth Varakumari, J.: A comparative study on women empowerment in urban and rural setting in Nadu. Int. J. Commun. Med. Public Health **6**(3), 1108 (2019). pISSN 2394-6032
20. Ayodele, A.E., Ahmed, M.: Analysis of MOOC (Massive open Online Courses): effectiveness challenges (case study: low middle income countries). Econ. Educ. (2017). https://doi.org/10.13140/RG.2.2.20889.26722
21. Rahar, U.S.: Mobile based primary health care system for rural India. Int. J. Nurs. Educ. **3**(1), 61 (2011). https://www.w3.org/2008/02/MS4D_WS/papers/cdac-mobile-healthcare-paper.pdf
22. Kakade, O., Tattimani, D.: Impact of digital India on women: a literature review. Res. Rev. Int. J. Multidisc. **05**(03) (2020). ISSN 2455-3085
23. Suguna, M.: Education and women empowerment in India. Int. J. Res. **1** (2011). ISSN 2231-5780
24. Jayashree, S., Ramya, E., Malarvizhi, K., Kumaresan, N.: Women in science and technology in India: a view. In: International Conference on Advancements in Electrical, Electronics, Communication, Computing and (ICAECA), Coimbatore, India, pp. 1–6 (2021). https://doi.org/10.1109/ICAECA52838.2021
25. Singh, M.H., Kumar, D.: The role of technology and women empowerment in rural India
26. Siddamsetti, S., Pallavi, M., Breaking the barriers: women and technology. Rev. Sci. Int. Multidisc. J. **3**(4) (2023). ISSN 2583-1380
27. Niroo, W.T., Crompton, H.: Women's empowerment through learning using technology. Asian Distance Educ. **17**(2) (2022)
28. Nadaf, R.A.: A study on women's empowerment through the using of technologies. IJCRT **12**(1) (2024). ISSN 2320-2882
29. Bhat, M.I.: Women Empowerment and Technology: An Overview, Ethics and Society, chap. 5 (2019)

Web references

30. https://www.thecable.ng/the-role-of-technology-in-nation-building [05/05/23]
31. https://community.thriveglobal.com/how-technology-has-made-an-impact-in-making-our-lives-better/ [05/05/23]
32. https://www.simplilearn.com/importance-of-technology-article. [05/05/23]
33. https://www.statista.com/statistics/271335/literacy-rate-in-india/ [027/09/24]
34. https://press.un.org/en/2023/wom2224.doc.htm [06/05/23]
35. https://elearningindustry.com/the-use-of-technology-in-online-education [06/05/23]
36. https://spark.school/flexible-online-learning [06/05/23]
37. https://www.thebusinessresearchcompany.com/report/online-taxi-services-global-market-report [06/05/23]
38. https://www.techwalla.com/articles/how-technology-has-improved-communication [06/05/23]
39. https://lucidspark.com/blog/5-reasons-why-teams-need-collaboration-software [06/05/23]
40. https://timesofindia.indiatimes.com/blogs/voices/e2-80-8b-e2-80-8bhow-social-media-has-given-more-power-to- [06/05/23]
41. https://www.statista.com/statistics/868974/india-share-of-instagram-users-by-gender/ [15/04/23]
42. https://www.fierceelectronics.com/components/women-face-challenges-technology-careers [15/04/23]
43. https://www.indiatoday.in/education-today/featurephilia/story/5-challenges-faced-by-women-in-technology-sector-23440192023-03-08(IndiaToday,Newspaper) [15/04/23]
44. https://www.orfonline.org/research/literacy-in-india-the-gender-and-age-dimension-57150/ [15/05/23]
45. https://www.statista.com/statistics/1303303/india-gender-literacy-rate-by-area/ [15/05/23]
46. https://www.cio.com/article/284419/privacy-women-take-social-privacy-more-seriously-than-men-do.html [15/05/23]
47. https://www.neumetric.com/how-to-conFigure-privacy-settings-on-social-media-accounts/ [15/05/23]
48. https://varthana.com/school/7-barriers-and-enablers-for-accessing-quality-education-for-girls-in-rural-areas. [15/03/23]
49. https://www.tutorialspoint.com/what-are-the-problems-faced-by-rural-women-in-india-today [17/05/23]
50. https://www.unwomen.org/en/news-stories/feature-story/2022/10/three-challenges-for-rural-women-amid-a-cost-of-livingcrisis [17/05/23]
51. https://www.insightsonindia.com/social-justice/issues-related-to-women/women-in-india/problems-faced-by-women-in-rural-areas/ [17/05/23]
52. https://mospi.gov.in/sites/default/files/reports_and_publication/statistical_publication/social_statistics/.WM16Chapter3.pdf [17/05/23]
53. https://wpmu.mah.se/nmict181group1/icts-women-economic-empowerment-grameen-phones-bangladesh/ [17/03/23]
54. https://www.indianyouth.net/challenges-faced-by-working-women-in-india/ [18/03/23]
55. https://www.tintup.com/blog/social-media-habits-men-vs-women/ [27/09/24]
56. https://www.brandwatch.com/blog/men-vs-women-active-social-media/ [27/09/24]

Detection and Classification of Epileptic Seizures Using Machine Learning and Deep Learning: A Systematic Review and Challenges

Shirish Mohan Dubey[1(✉)], Kamlesh Gupta[2], and Shikha Agrawal[1]

[1] UIT, RGVP, Bhopal, India
dubeys78@gmail.com
[2] RJIT, Tekanpar, Gwalin, India

Abstract. Nowadays epileptic seizure detection and classification has been an important concern as the central nervous system is affected in epilepsy due to brain disorders that cause changes in neural activity. Seizures result from unexpected brain activity, excitement, and sometimes unconsciousness. Epilepsy is a profound condition, with approximately 1% of people worldwide affected by it, and 10% reporting that they deal with this issue daily. It has the potential to impact life prospects significantly. The primary purpose of an EEG report is to record the electrical brain activity generated within the brain. Using EEG, doctors and scientists can analyze brain activity and identify abnormalities. Several research studies using machine learning and deep learning techniques have been successful in recognizing EEG patterns associated with epileptic seizures. The proposed study explains and compares various existing algorithms and methods for detecting and classifying epileptic seizures based on certain parameters. The analytical results indicate that the KNN, SVM, and CNN classifiers are most preferred, and the Bonn dataset and CHB-MIT datasets are the most used choices.

Keywords: ESDC · EEG · ECoG · Machine Learning · Deep Learning · Challenges

1 Introduction

Epilepsy refers to a common and serious neurological disorder characterized by seizures. Although the exact cause of epilepsy is still unknown, it is often linked to genetic abnormalities that result in abnormal brain activity. Early diagnosis is beneficial for treatment, as seizures can be distressing and impact a patient's quality of life. Detecting and predicting seizures is crucial for preventing their onset through medication [1]. Epilepsy is typically diagnosed by analyzing the electrical signals produced by brain neurons. Seizures are transient disruptions in brain activity that can manifest as various physical and sensory symptoms. Electroencephalograms (EEG) are used to record brain activity and detect neuronal abnormalities. However, analyzing EEG signals is time-consuming and demands the constant availability of neurologists. Additionally, a major challenge lies in selecting the right classifiers [2–5] and features for accurately analyzing

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
S. Rajagopal et al. (Eds.): ASCIS 2024, CCIS 2428, pp. 183–194, 2025.
https://doi.org/10.1007/978-3-031-86302-8_12

EEG data. EEG recordings are often contaminated with signal artifacts, necessitating automated artifact removal methods for clinical and research purposes [6–10].

Recent developments in Deep Learning (DL) and Machine Learning (ML) have demonstrated the potential to automatically extract pertinent features from unprocessed EEG data. Several research works used ML and DL techniques [11–14] for Epileptic Seizure Detection and Classification (ESDC) from EEG signals, offering a promising approach for early detection. While the advantages of DL over traditional ML methods [15, 16] are still being debated, it has the potential to improve the analysis of EEG signals and the detection of epileptic activity. This paper illustrated a detailed and systematic study of the existing ESDC algorithms using ML and DL techniques. This systematic review described the year-wise evolution and development of ESDC methods. These contributions were discriminated based on the data source, data domain, classification or clustering technique used, and performance metrics. Furthermore, their findings, challenges, and issues are elaborated. Lastly, it provides various analytical results based on the results of the existing ESDC approaches.

The structure of this paper is as follows: Sect. 2 demonstrates the ESD concepts and framework. Section 3 discusses the literature review of various existing ESD approaches, compares them on some parameters, and provides their findings and limitations. Section 4 provides the analytical results on the usage of classifier distribution and data domain distribution. Finally, Sect. 5 concludes the paper with future recommendations.

2 Concepts and Working Procedure of ESDC System

This section describes the basic concepts of ESDC, EEG with ECoG, and the working procedure of the ESDC system.

2.1 Epileptic Seizure Detection and Classification

Epilepsy is a common neurological disorder that presents a scientific challenge when it comes to detecting epileptic seizures, as patients may not always experience warning signs. Epilepsy is a chronic disorder that affects daily life by putting individuals at risk due to repeated seizures. Since seizures occur randomly, patients may not be aware of them, increasing the risk of physical injury. Detection and prediction of seizures are crucial topics in medical signal processing because they play a key role in automatic medical diagnosis. Distinguishing between epileptic and non-epileptic signals is an important step and can assist doctors in diagnosing and treating epileptic seizures. Epilepsy is typically detected by EEG signals. Seizure detection using EEG recordings can enhance the quality of treatment and reduce medical costs. Manual examination of epilepsy seizures is time-consuming, which is why numerous automatic ESD algorithms have been proposed. Efforts have been made to detect and prevent epileptic seizures through EEG signal analysis, using ML and DL techniques.

2.2 Electroencephalogram and Electrocorticography

Epilepsy is considered a hypersensitive disease by clinical associations. Neurons communicate through complex connections to generate signals, and these brain signals are

commonly monitored using EEG and ElectroCorticoGraphy (ECoG). Neurons generate a lot of data and are noisy, complicated, non-linear, and non-stationary. Asymmetry in the EEG signal can result from an increase in electrical activity in one hemisphere of the brain during a seizure, and patients with epilepsy may exhibit asymmetric spikes or sharp waves in their interictal EEG. Because it records brain electrical activity, EEG is one of the most widely used techniques for seizure detection. EEG is widely used for capturing brain recordings and impulses, making it useful for identifying not only epilepsy related abnormalities but also other neurological disorders. The distinctive features of EEG have provided researchers with a reliable method for identifying certain health conditions in patients. However, manual inspection of EEG signals can be time-consuming for neurologists, and visual examination by experts may lead to contradictory diagnoses.

It is possible to use electrical impulses from EEG to control prosthetics or devices, according to recent studies. Nevertheless, noise can disrupt these signals and impact anticipation time as well as the True Positive Prediction (TPR) rate. EEG signals are complex, prone to noise, and require preprocessing for noise removal and feature extraction, furthermore, processing the lengthy EEG recordings can be challenging too. Even while EEG recordings are highly portable, inexpensive, and have good temporal resolution, they are frequently tainted with artifacts from sources other than neurons. As a result, before evaluating EEG data, these artifacts must be eliminated during the preprocessing phase. There is no one best way to eliminate the various kinds of artifacts that have been found in tainted EEG data, which makes more research both appealing and difficult.

2.3 Working Procedure of an Epileptic Seizure Detection and Classification System

The Epileptic Seizure Detection and Classification (ESDC) system works through a series of steps. First, brain data is collected and transformed. The data then undergoes pre-processing, followed by feature selection, classification, and performance evaluation. Figure 1 shows the basic layout of an ESDC system. This system consists of training and testing stages. In the training stage, a set of known signals is collected from public sources and used as inputs. These signals are processed and transformed into a structured form, such as tabular data, using monitoring tools like EEG and ECoG. Features are then selected from these signals and used to train an ML classifier. During the testing stage, a set of unknown signals undergoes a similar process. They are pre-processed, transformed, and have their features extracted. The dimensions of the features are then reduced, neighborhoods are selected, and the signals are classified using the previously trained classifier. Finally, the results are analyzed to evaluate the performance of the system.

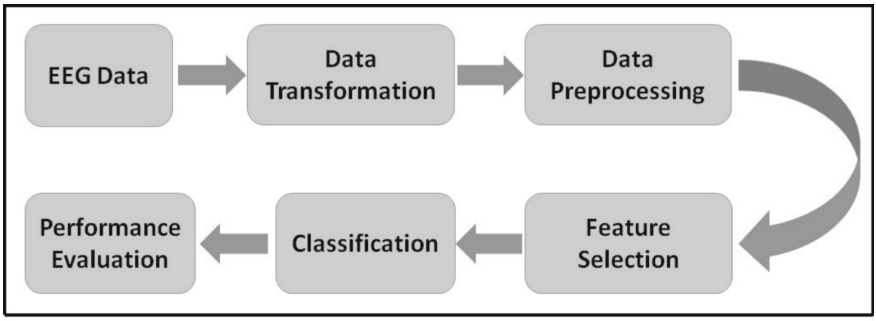


Fig. 1. ESDC Framework.

3 Discussion on Existing ESDC Approaches

As such, various researchers have developed several approaches for ESDC using ML and DL classifiers and statistical features. This review summarizes the most significant findings and important concepts involved. These research contributions are compared using a set of parameters, and their respective limitations are further addressed.

3.1 Literature Review

Many researchers have developed different methods for detecting seizures using ML and DL classifiers using statistical features. This review summarizes the most significant findings and discusses important concepts related to ML and DL for seizure detection, prediction, and classification. Intensive research has been conducted to create ML methodologies for automated epilepsy detection and enabling the analysis of brain signals with high accuracy. To aid in these studies, several academic institutions and healthcare facilities have released databases of epileptic EEG signals. This comprehensive analysis is required due to the abundance of research on the identification of epilepsy using EEG data. In one study [1], EEG occipital signals were clustered using k-means based on their frequency and temporal characteristics. To assess whether preprocessing was required to eliminate noise components from electrical grids or other potential sources, a statistical analysis of the distribution of normal EEG signals was conducted.

Another study [3] decomposed empirical mode for preprocessing the scalp EEG CHB-MIT dataset, extracted time and frequency domain features, and predicted epileptic seizures in advance of their onset. The EEG signal classification approach in another study [4] involved extracting wavelet-based features such as multi-resolution decompositions and computing relative wavelet energy. Principal Component Analysis (PCA) and Fisher's discriminant ratio were then used to optimize the normalized features. This study used 128 channels and classified using Raven's advanced progressive metric and the baseline task. Another study [5] used non-linear Independent Component Analysis (ICA) to reduce the dimensions of the EEG signals, followed by classification using Logistic Regression (LR). Another study [6] focused on classifying different human

brain states related to visual perception of ambiguous images using a Multi-Layer Perceptron (MLP) classifier to analyze multichannel EEG. It trained on one subject and then used the classifier to classify the EEG traces of other subjects. Additionally, a review [7] discussed methods for identifying and removing ocular, muscle, and cardiac artifacts from EEG data and compared their advantages and limitations. Furthermore, the review validated the performance of algorithms using both simulated and real EEG data. Another review [8] provided a discussion on the classification of seizures and epilepsy and its significance for patients, clinicians, and researchers. The following study [9] retrieved a collection of variables, including kurtosis, variance, Shannon's entropy, and range of amplitude, from EEG data and used Wavelet-Independent Component Analysis (ICA) and Support Vector Machine (SVM) to identify eye blink artifacts. Several ESD methods employing biomedical signals, including time, frequency, wavelet (time-frequency), empirical mode decomposition, and rational function, were investigated in a different review [10]. In a study [12], a time-efficient method was used to identify the brain hemisphere most affected by seizures. This approach utilized Decision Tree (DT) algorithms such as SysFor and Forest CERN.

In another analysis [13], EEG data was examined using different machine learning (ML) techniques including SVM, DT, K Nearest Neighbor (KNN), Naïve Bayes (NB), and Random Forest (RF) for bioengineering applications spanning from 1988 to 2018. A review [14] covered various existing approaches to seizure detection and classification, using statistical features and both 'black-box' and 'non-black-box' ML classifiers. Another review [17] focused on the use of DL techniques and neuroimaging modalities for diagnosing epileptic seizures, as well as analyzing rehabilitation systems for epilepsy using DL. In a specific application [18], data pre-processing was carried out using 6.274 k parameters and 12.825 M floating-point operations. A 1D Convolutional Neural Network (CNN) and a Bi-Directional Long Short-Term Memory (Bi-LSTM) were used for seizure prediction, along with cross-validation to prevent overfitting. Furthermore, a survey [19] looked at 91 existing research contributions and identified challenges related to seizure detection, prediction, and classification using EEG data and ML/DL methods. Additionally, a study [20] detailed three different methods for seizure detection and prediction, including scale-invariant feature transform, Fast Fourier Transform (FFT), and an Artificial Neural Network (ANN). This study also discussed channel selection and patient-specific approaches for seizure prediction.

The CAD system used for diagnosing epileptic seizures [21] utilized a combination of Empirical Wavelet Transform (EWT) and a Multi-Fuse Reduced Deep CNN (MF-RDCNN). It involved extracting three distinct band-limited modes from recorded EEG signals, computing the most informative unsupervised signatures using classifiers, and conducting ten-fold cross-validation with False Positive Rate per Hour (FPR/H). In a different study [22], pre-processing involved minimax scaler normalization and feature extraction, followed by the application of the BiLSTM technique for seizure detection. Another study [23] presented a systematic review of Signal Transformation techniques, classification methodologies, and databases for epilepsy seizure detection. Additionally, an early detection approach for epilepsy [24] decomposed the signals, extracted features using the Discrete Wavelet Transform (DWT), reduced feature dimensions using PCA and t-distributed Stochastic Neighbor Embedding (t-SNE) methods, employed K-means

clustering + PCA and K-means clustering + t-SNE for dataset division, and classified the data using Extreme Gradient Boosting (XGB), KNN, DT, RF, and MLP. An EEG-based seizure detector [25] used the stationary wavelet transform to analyze EEG signals. Nine statistical and Hjorth parameters were extracted to differentiate normal, interictal, and ictal signals. The features were further reduced by 13% using the binary dragonfly algorithm BDFA. Classification was performed using a Deep Neural Network (DNN). Another robust epileptic seizure detector [26] analyzed six features and five channels in the temporal domain using a 1-s time window. It reduced the number of features and channels required by simpler classifiers before classifying the datasets. Yet another epileptic seizure detector [27] extracted features using 1-D convolutional layers, BiLSTM, Gated Recurrent Unit (GRU), and average pooling layer. It then predicted the class of the EEG waveform using dense layers and employed five-fold cross-validation with five subsets of the dataset.

The next method [28] involved the development of a mobile application for monitoring EEG signal classification. This app would notify all associated users visually. A review of 56 articles in [29] investigated different methods for detecting and predicting epileptic seizures using ML, DL, and Internet of Things techniques. Another method for analyzing epilepsy EEG signals [30] involved determining various entropy measures and standard deviation using DWT decomposition. Then, important features were selected using the random forest technique, followed by interictal and ictal classification using CNN on Bonn EEG datasets and New Delhi datasets. Another method [31] detected seizure and non-seizure events by analyzing brain region connectivity using weighted directed graphs. It then used graph theory-based measures to extract features for patient-specific and cross-patient models and classify the datasets.

3.2 Comparison of Various Existing Techniques

Table 1 compares existing ESDC algorithms using ML and DL techniques from 2016 to 2023, many of which focus on feature reduction. This comparison is based on the dataset used, the classification techniques applied, and the accuracy obtained.

It is evident from Table 1 that most of the research works utilize KNN, SVM, and CNN classifiers. Additionally, the Bonn dataset and CHB-MIT datasets are the preferred choices among research contributors. The application of these ML and DL techniques has significantly resulted in high performance of seizure prediction and classification.

3.3 Findings and Challenges in Existing ESDC Approaches

The following are the limitations and gaps of the current ESDC methods:

- The cluster was not improved by using Pearson correlation between the occipital region electrodes. [1].
- A need to enhance the sensitivity of seizure prediction by improving the preprocessing of EEG signals [3].
- Small public dataset used. Difficult separation of EEG signals recorded in two different cognitive tasks. A need to expand using clinical datasets, such as the classification of ictal vs. interictal or normal EEG patterns [4].

Table 1. Comparison of several existing ESDC methods

Ref. no.	Year	Dataset used	Classification/Clustering techniques	Accuracy obtained
[1]	2016	2500 electrode samples per electrode with 1 visual stimulus & 1 volunteer.	K-means.	Success rate: 36%-80%. 90% accuracy with frequencies of 6, 7, 8, 26, 27, 28 & 29 Hz.
[3]	2017	Scalp EEG CHB-MIT dataset of 22 subjects. Data pertained to 24 patients aged between 3 & 19 years	KNN, SVM & NB.	92.23% TPR, max. Anticipation time: 33 min & avg prediction time: 23.6 min.
[4]	2017	EEG public dataset.	KNN, SVM, MLP & NB.	Accuracy (%): 99.11 (SVM) with 0–3.90 Hz; 98.57 (SVM) & 98.39 (KNN) with 3.90–7.81 Hz; 97.11–89.63 (MLP) & 91.60–81.07 (NB); 93.33 (KNN).
[5]	2017	Recordings of 20 epileptic patients for 50 min. Took from Sri Ramakrishna Hospital, Coimbatore, India.	LR.	Avg. Performance index: 91.39%, Avg. Accuracy of 95.88%, & time delay of 2.085 s.
[6]	2017	Necker cube image.	ANN-MLP.	95%.
[9]	2018	Public EEG data from EEGLAB.	SVM & Wavelet ICA.	99.1% accuracy, 97.1% sensitivity & 99.3% specificity.
[12]	2019	ECoG brain dataset of a single patient.	DT.	100%.
[18]	2021	CHB-MIT dataset.	1D convolutional layers & BiLSTM.	77.6% accuracy, 82.7% sensitivity & 72.4% specificity.
[20]	2022	Multi-channel scalp EEG recordings of 24 patients from CHB-MIT.	FFT & ANN.	100% success rate.
[21]	2022	Dataset A: Bonn-University, Germany single-channel EEG; Dataset B: Neurology and Sleep Center, New Delhi single-channel EEG; & Dataset C: Boston Children's Hospital multichannel scalp EEG.	EWT-MF-RDCNN.	Accuracy (%): 100 (2-class), 99.82 (3-class). 99.29% specificity, 99.86% sensitivity & 99.29% accuracy with 0.71% FPR/h for Dataset-C.
[22]	2022	CHB-MIT scalp EEG dataset.	BiLSTM.	99.55% accuracy, 99.64% precision & 99.52% F1-score.
[24]	2023	UCI ML repository.	XGB, KNN, DT, RF & MLP.	RF + DWT + PCA (%): 97.96 (accuracy), 99.1 (precision), 94.41 (recall) & 97.41 (F1-score). RF + DWT + tSNE (%): 98.09 (accuracy), 99.1 (precision), 93.9 (recall) & 96.21 (F1-score). MLP + PCA + K-Means (%): 98.98 (accuracy), 99.16 (precision), 95.69 (recall) & 97.4 (F1-score).
[25]	2023	Bonn dataset.	DNN.	100% for each of accuracy, sensitivity, specificity & F1-score

(*continued*)

Table 1. (*continued*)

Ref. no.	Year	Dataset used	Classification/Clustering techniques	Accuracy obtained
[26]	2023	Patient cohort dataset.	DL.	95%.
[27]	2023	Bonn dataset.	CNN & BiLSTM.	99%–100% (seizure & normal waveforms), 97.2%–99.2% (normal-interictal-seizure waveforms), 96.2%–98.4% (4 classes) & 95.81%–98% (5 classes).
[28]	2023	Bonn dataset.	SVM, KNN & ANN.	98%.
[30]	2023	Bonn EEG & New Delhi datasets.	CNN.	Bonn dataset (%): 99.9 (accuracy), 100 (sensitivity), 99.81 (precision) & 99.8 (specificity). New Delhi datasets: 100% for all 4 metrics.
[31]	2023	CHB-MIT dataset.	SVM, RF & KNN.	99.97% (patient-specific model) & 98.29% (cross-patient model).

- Need to improve classification accuracy [6, 18, 21] with modified LR [5]. Need to improve the accuracy using complex architectures and applying MinMax scaler normalization to individual patients' data. Need to apply DL techniques on the Bonn dataset, Germany EEG dataset, and other Indian patients' datasets. Need to use cross-validated DL models across all ages and geographical locations [22].
- A need to identify features to best describe artifactual components [9].
- A requirement to evaluate the effectiveness of the suggested work with alternative datasets [27] and to lower the dataset's dimensions [12].
- Need to apply with other ML techniques [20, 24].
- Need to use an edge device to detect the disease that causes seizures and notify family members and medical professionals to start taking medicine. Need to use many EEG datasets to test the effectiveness. [25].
- Need to work on this since many patients know the precise location of their seizures. A requirement to develop a comprehensive, universal model [26].
- Extend the proposed method with epileptic patients' dataset in real-time. Extend with different types of seizure detection [28].
- Requirement to expand with an effective seizure detection technique that lowers the number of false positives by integrating with the temporal correlation between EEG signal frames [30].
- Extend with automatic seizure detection of multi-channel signals [31].

4 Analytical Results from Existing Approaches

This section presents various analyses of the existing systems, comparing them graphically. The two analyses include the usage of classifier distribution, and data domain distribution, respectively.

4.1 Analysis I: Classifier Distribution and Usage

The Analysis-I compares existing algorithms for ESDC based on the classifier used, as depicted in Fig. 2. The classifiers used in the comparison are KNN, SVM, CNN, ANN, BiLSTM, DT, NB, RF, K-Means, MLP, LR, FFT, Wavelet, XGB, and DNN. The analysis revealed that the most used classifiers were KNN, SVM, and CNN, which accounted for 27.77% of the usage. Other commonly used techniques included ANN, which was used 22.22% of the time, and BiLSTM, which was used 16.60% of the time.

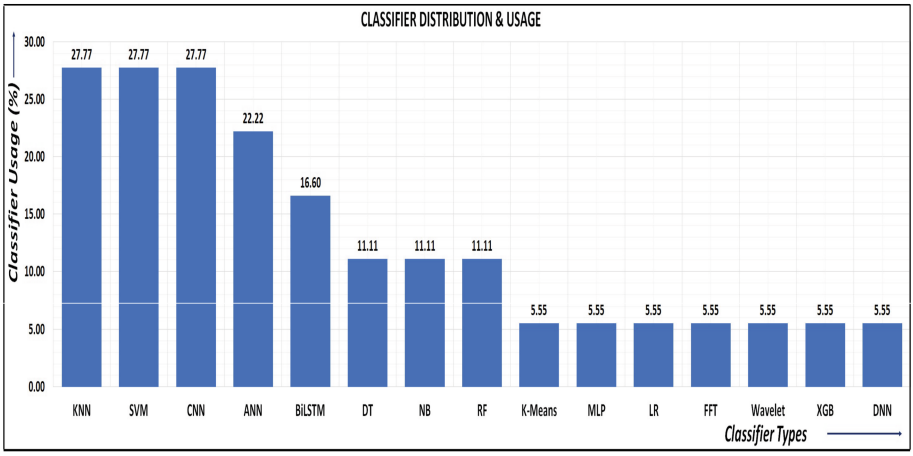


Fig. 2. Classifier-based distribution and comparison among existing approaches

4.2 Analysis II: Data Domain Distribution and Usage

Analysis II describes the use of various datasets in existing algorithms for detecting epileptic seizures, as illustrated in Fig. 3. These datasets include the Bonn, CHB-MIT, EEG Public, UCI, Boston Children's Hospital, Neurology & Sleep Center, ECoG, Necker Cube, Sri Ram Krishna Hospital, Patient Cohort, and other supplementary datasets. Most contributors focused on the Bonn dataset, which was used 27.7% of the time, followed by CHB-MIT at 22.2% and EEG Public at 16.66%.

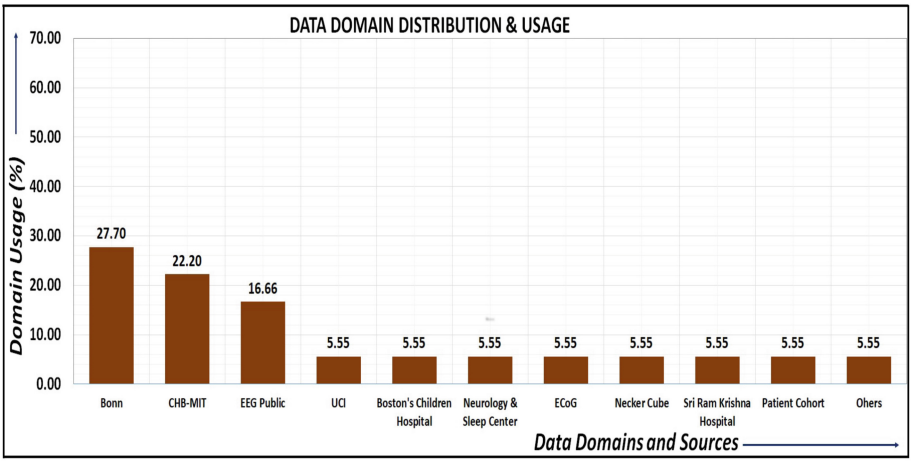


Fig. 3. Data domain-based distribution and comparison among existing approaches.

5 Conclusions and Future Recommendations

This paper provided a comprehensive review of existing algorithms for detecting and classifying epileptic seizures using EEG signal analysis with various ML and DL techniques. The study successfully demonstrated the automatic detection and classification of high-precision EEG data associated with epilepsy. In addition to concentrating on integrating several techniques with several processing phases to effectively eliminate interference, researchers also worked on creating uniform standards for verifying recorded EEG signals. The study presented concepts and frameworks for ESDC and compared various existing approaches based on the datasets, classifiers employed, and accuracy obtained. The analytical results showed that KNN, SVM, and CNN classifiers were the most preferred, and the Bonn dataset and CHB-MIT datasets were the preferred choices. However, the limitations of these methods need to be addressed, and there is a need to build a more efficient model that can differentiate neurological diseases based on proper signal analysis. The proposed research can be further extended and explored in many dimensions.

- Implementation of a system utilizing reviews from multiple data domains.
- Extend using a greater number of supervised learning ML techniques.
- Extend and implement using a combination of multiple supervised learning techniques.

Disclosure of Interests. No funding was received for this article. There is no Conflict of Interests in my article.

References

1. Asanza, V., et al.: Clustering of EEG occipital signals using k-means. In: 2016 IEEE Ecuador Technical Chapters Meeting (ETCM), Guayaquil, Ecuador, pp. 1–5. IEEE Press (2016)
2. Khan, N., Puri, S.: A study on text detection techniques of printed documents. In: International Conference on Wireless Communications, Signal Processing and Networking, pp. 2478–2482. IEEE Press, India (2016)
3. Usman, S.M., Usman, M., Fong, S.: Epileptic seizures prediction using machine learning methods. Comput. Math. Methods Med. Math. Methods Med. **2017**, 1–12 (2017)
4. Amin, H.U., Mumtaz, W., Subhani, A.R., Saad, M.N., Malik, A.S.: Classification of EEG signals based on pattern recognition approach. Front. Comput. Neurosci. Comput. Neurosci. **11**, 1–12 (2017)
5. Rajaguru, H., Prabhakar, S.K.: Non linear ICA and logistic regression for classification of epilepsy from EEG signals. In: 2017 International Conference of Electronics, Communication and Aerospace Technology, pp. 577–580. IEEE Press (2017)
6. Hramov, A.E., et al.: Classifying the perceptual interpretations of a bistable image using EEG and artificial neural networks. Front. Neurosci. Neurosci. **11**(674), 1–18 (2017)
7. Mannan, M.M.N., Kamran, M.A., Jeong, M.Y.: Identification and removal of physiological artifacts from electroencephalogram signals: a review. IEEE Access **6**, 30630–30652 (2018)
8. Falco-Walter, J.J., Scheffer, I.E., Fisher, R.S.: The new definition and classification of seizures and epilepsy. Epilepsy Res. Res. **139**, 73–79 (2018)
9. Sai, C.Y., Mokhtar, N., Arof, H., Cumming, P., Iwahashi, M.: Automated classification and removal of EEG artifacts with SVM and wavelet-ICA. IEEE J. Biomed. Health Inform. **22**(3), 664–670 (2018)
10. Paul, Y.: Various epileptic seizure detection techniques using biomedical signals: a review. Brain Inform. **5**(2), 1–19 (2018)
11. Puri, S., Singh, S.P.: Toward recognition and classification of Hindi handwritten document image. In: Hu, Y.C., Tiwari, S., Mishra, K., Trivedi, M. (eds.) Ambient Communications and Computer Systems. AISC, vol. 904, pp. 497–507. Springer, Singapore (2019). https://doi.org/10.1007/978-981-13-5934-7_44
12. Siddiqui, M.K., Islam, M.Z., Kabir, M.A.: A novel quick seizure detection and localization through brain data mining on ECoG dataset. Neural Comput. Appl. Comput. Appl. **31**(9), 5595–5608 (2018). https://doi.org/10.1007/s00521-018-3381-9
13. Hosseini, M.-P., Hosseini, A., Ahi, K.: A review on machine learning for EEG signal processing in bioengineering. IEEE Rev. Biomed. Eng. **14**, 204–218 (2021)
14. Siddiqui, M.K., Morales-Menendez, R., Huang, X., Hussain, N.: A review of epileptic seizure detection using machine learning classifiers. Brain Inform. **7**(1), 1–18 (2020)
15. Puri, S., Singh, S.P.: Advanced applications on bilingual document analysis and processing systems. Int. J. Appl. Metaheuristic Comput. **11**(4), 149–193 (2020)
16. Puri, S., Singh, S.P.: A fuzzy matching based image classification system for printed and handwritten text documents. J. Inf. Technol. Res. **13**(2), 155–194 (2020)
17. Shoeibi, A., et al.: Epileptic seizures detection using deep learning techniques: a review. Int. J. Environ. Res. Public Health **18**(11), 1–33 (2021)
18. Wang, Z., Yang, J., Wu, H., et al.: Power efficient refined seizure prediction algorithm based on an enhanced benchmarking. Sci. Rep. **11**(1), 1–9 (2021)
19. Nafea, M.S., Ismail, Z.H.: Supervised machine learning and deep learning techniques for epileptic seizure recognition using EEG signals—a systematic literature review. Bioengineering **9**(12), 1–35 (2022)
20. Emara, H.M., Elwekeil, M., Taha, T.E., et al.: Efficient frameworks for EEG epileptic seizure detection and prediction. Ann. Data Sci. **9**(2), 393–428 (2022)

21. Rout, S.K., Sahani, M., Dora, C., Biswal, P.K., Biswal, B.: An efficient epileptic seizure classification system using empirical wavelet transform and multi-fuse reduced deep convolutional neural network with digital implementation. Biomed. Signal Process. Control **72**, 103281 (2022)
22. Deepa, B., Ramesh, K.: Epileptic seizure detection using deep learning through min max scaler normalization. Int. J. Health Sci. **6**(S1), 10981–10996 (2022)
23. Miltiadous, A., et al.: Machine learning algorithms for epilepsy detection based on published EEG databases: a systematic review. IEEE Access **11**, 564–594 (2023)
24. Alalayah, K.M., Senan, E.M., Atlam, H.F., Ahmed, I.A., Shatnawi, H.S.A.: Effective early detection of epileptic seizures through EEG signals using classification algorithms based on t-distributed stochastic neighbor embedding and k-means. Diagnostics **13**(11), 1–24 (2023)
25. Yogarajan, G., et al.: EEG-based epileptic seizure detection using binary dragonfly algorithm and deep neural network. Sci. Rep. **13**(1), 1–16 (2023)
26. Vieira, J.C., Guedes, L.A., Santos, M.R., Sanchez-Gendriz, I.: Using explainable artificial intelligence to obtain efficient seizure-detection models based on electroencephalography signals. Sensors **23**(24), 1–21 (2023)
27. Mallick, S., Baths, V.: Novel deep learning framework for detection of epileptic seizures using EEG signals. Front. Comput. Neurosci. Comput. Neurosci. **18**, 1–17 (2024)
28. Lasefr, Z., Elleithy, K., Reddy, R.R., Abdelfattah, E., Faezipour, M.: An epileptic seizure detection technique using EEG signals with mobile application development. Appl. Sci. **13**(17), 1–25 (2023)
29. Jahan, S., et al.: AI-based epileptic seizure detection and prediction in Internet of healthcare things: a systematic review. IEEE Access **11**, 1–36 (2023)
30. Chen, W., et al.: An automated detection of epileptic seizures EEG using CNN classifier based on feature fusion with high accuracy. BMC Med. Inform. Decis. Mak.Decis. Mak. **23**(1), 1–17 (2023)
31. Sun, Q., Liu, Y., Li, S.: Weighted directed graph-based automatic seizure detection with effective brain connectivity for EEG signals. SIViP **18**, 1–11 (2024)

Application of Smart Computing in Steel Structural Health Monitoring: Sensor Based Damage Detection for Smart Infrastructures

Husain Rangwala(✉) and Tarak Vora

Marwadi University, Rajkot, Gujarat, India
{husain.rangwala,tarak.vora}@marwadieducation.edu.in

Abstract. Steel structures are robust and enduring, however susceptible to various damages including loss of the grips of bolts at connections. Such secondary damage is difficult to monitor manually. With the help of applying smart computing techniques for structural health monitoring methods, namely use of Piezoelectric sensors (PZT) sensor by the Electromechanical Impedance Technique the monitoring of a bolt's grip becomes effective and efficient. This paper represents results of experimental investigations on a bolted steel connection specimen prepared and tested for the damage detection conditions. The study focuses on identifying the response due to looseness with and without applied axial tensile load using different statistical damage indexes like Root Mean Square Deviation (RMSD) and Mean Absolute Percent Deviation (MAPD). The damage assessment is analyzed based on the signature of impedance concerning frequency, variation of torque in the bolts, and axial loading. Using the different number of repetitive experimental output data, the feasibility and effectiveness of this method are verified. The results show that both RMSD and MAPD damage indices increase with the decrease in the design torque in the bolt establishing an effective way of monitoring of the bolts' grip. The looseness indices were found to be highly sensitive to small damage which showed early warning for repair or maintenance helpful for monitoring process. The results showed good agreement on application of smart computing for structural health monitoring using sensor-based damage detection necessary for the smart infrastructures.

Keywords: Smart Computing · Smart Infrastructure · Piezoelectric Sensor · Damage Detection · Steel Structures · Bolted Connections

1 Introduction

1.1 Background

A phenomenal rise in the construction of structures have unconventional structures worldwide to showcase the technological capabilities in a cut-throat competitive scenario. In the modern construction of tall structures, bridges, airport terminals, transmission towers, railway structures, etc., steel is one of the most consumed materials due

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
S. Rajagopal et al. (Eds.): ASCIS 2024, CCIS 2428, pp. 195–211, 2025.
https://doi.org/10.1007/978-3-031-86302-8_13

to its strength, mouldability and ductility. Also, the ease of faster construction due to factory-made structural components is one of the greatest advantages of steel. Onsite construction processes require either bolting or welding of different structural elements onsite. Bolted connections are preferable connections not only for their ease of application and fabrication but also it has higher resilience to shocks, more damping against external vibrations, and enhanced endurance. Along with conventional and special loading environmental conditions should also be considered as an additional type of loading these days. Degradation of the structural elements and joints due to environmental conditions and fatigue are more critical than their conventional loadings. In many of the structures, the joints are also more susceptible to fatigue damage at the joints which can be identified based on the equivalent stiffness of the joint using piezo-impedance transducers [1, 2]. During the lifespan, the structures may be damaged due to various reasons. The damage is generally defined as the changes in a system which has an adverse effect on the current or future performance of the structure. From the various literature, the damage in the structures may be identified based on the comparison of the two different states of the system such as the undamaged condition which is the initial condition, and the damaged condition. [3] Damage in existing structures decreases its load carrying capacity and eventually, it leads to catastrophic failure if it is not observed at an early age. Some minor damage. Material deterioration or damage at inaccessible areas causes the failure of entire structures or structural elements. Easily identifying such damage can prevent unexpected failures and extend the service life of the structure. In many of the times, only visual inspection is not sufficient to identify the damage location and extent of the damage. In such cases structural health monitoring using reliable non-destructive damage detection methods helps maintain the safety and integrity of the structures. Structural Health monitoring offers the potential to prevent structural failures, enhance maintenance costs, and extend the service life of the structure by periodic assessment. [4] Structural responses and changes resulting from the damage are used to assess structures using various methods and different signatures in each method. Based on recent developments in sensors, signal processing, and analysis techniques have improved SHM by eliminating environmental conditions variation and increasing the damage sensitivity and performance of structures [5, 6].

The steel structural elements relate to either welding or high-strength bolts. The skills of the operators and environmental conditions are the major factors in the welded connections. But when it is connected to high-strength bolts, the joints allow safer connection. Over a long period, the strength of the bolts decreases due to continuous loading, impact, vibration, cyclic loads, thermal effect, and structural aging. [7, 8] Vibration in the structures causes external forces to overcome the friction between the bolts and joints, which leads to relative slippage between two surfaces because of which the bolts may loosen. [9] Vibration-based damage detection methods ae depend on the dynamic characteristics of the structures.

The dynamic characteristics of damaged structures are compared based on the changes in the parameters such as natural frequency, mode shape corresponding to each natural frequency, strain energy, damping values, etc. with undamaged structures. [10, 11] In many of the structures the frequency gives the variation in the behavior in damaged structures but in general, the frequencies alone are not the only sufficient parameter for

effectively identifying the behavior of joints. [12] In Some of the studies, the strength or behavior of the connection in steel structures are explored. However, there is limited research to review the quantitative evaluation or quantitative performance of steel structural elements [13, 14].

Different smart materials have undergone a revolutionary transformation, contributing to the development of efficient and smart structural health monitoring systems of structures. Recently, emerging techniques such as Wave propagation methods and Fiber Bragg Grating methods have gained popularity. Despite their ability to provide valuable health monitoring data, these technologies often face challenges in automation and the detection of time-sensitive events. However, among the diverse range of smart materials, piezoelectric materials have garnered significant attention globally. Lead Zirconate Titanate, commonly employed in the form of thin rectangular patches, serves as a notable example of a piezoelectric material. PZT exhibits strain or deformation in response to an electric field which generates a voltage when it is subjected to strain. These materials show low-strain characteristics across a range of frequencies [15, 16]. The electromechanical impedance (EMI) method works on the principle of changing mechanical impedance caused by alterations in structural properties due to damage. In this method, PZT transducers are easily integrated and engage with the host structure. When it is subjected to high-frequency excitation under an electric field, the transducers give different behavioral signatures in terms of structural impedance. The EMI Method uses the inverse function of structural impedance as a key parameter as it effectively identifies and assesses the structural damage [17–19].

Based on various reviews, the EMI method measures looseness in the bolts based on qualitative damage identification for different frequency ranges. Based on the method proposed the damage index is considered in terms of impedance shift and the degree of looseness is assessed by integrating this method with various frequency ranges. [20] Extracting the mechanical impedance is a challenging task in the structures. However, the use of smart materials like PZT has created new opportunities to establish the correlation between the impedance of the structures and PZT [21]. This correlation is made possible through the electromechanical coupling property [22, 23]. Out of different methods, the EMI method has emerged as a powerful and effective technique for local damage detection as it harnesses the piezoelectric properties of PZT patches to sense any minor damage in the structures [24–26]. In this study damage assessment of steel structures at the element level due to bolt loosening is carried out using the EMI method by considering impedance with variation in torque using Piezoelectric sensors (PZT). The results are considered in terms of different statistical damage indices like Root Mean Square Deviation (RMSD) and Mean Absolute Percent Deviation (MAPD).

1.2 Research Significance

The study focuses on effective and accurate damage detection of the secondary elements of the steel structures namely Bolts at the connection. Bolted connections are responsible for intact and firm joints between the structural members. However, with time, the bolts may show loosening of the grip upon the members at a connection and may lead to a weak joint. Being a small unit, it is difficult to monitor such damages on the bolt individually by manual checking. To overcome such challenges, the modern smart computing systems

may be applied. In this study the use of smart computing method via sensors has been demonstrated. The group of bolts are monitored for loss of their frictional grip using Piezoelectric Sensors. Piezoelectric materials such as Lead Zirconate Titanate (PZT) are widely used in structural health monitoring due to their electromechanical properties The innovative use of PZT supported by smart computing methods namely Root Mean Square Deviation (RMSD) and Mean Absolute Percent Deviation (MAPD) showed excellent agreement to the expected outcomes. As Compared to previous literature, the RMSD was used as a primary damage detection index, while in present study demonstrates higher sensitivity to minor bolt looseness through the combined use of RMSD and MAPD along with the frequency range is expanded to increases detection accuracy, particularly in early-stage damage scenarios. It is well expected that the extension of the presented work will contribute to the accurate and efficient monitoring of losses in the strength of bolted connections of steel structures. Eventually, structural health monitoring will become smart and effective to overcome such improvements of steel structural connections. This research presents a novel integration of the Electromechanical Impedance (EMI) method with combined damage indices, RMSD and MAPD, for more sensitive detection of bolt loosening in steel structures with and without considering the load conditions. These findings extend the applicability of sensor-based smart computing techniques and provide a framework for future advancements in structural health monitoring.

2 Methodology

2.1 Theoretical Background

The Electromechanical Impedance (EMI) method employs Lead Zirconate Titanate (PZT), a piezoelectric ceramic material, as both a sensor and an actuator. PZT exhibits the piezoelectric effect, meaning it generates an electrical charge when mechanically deformed (direct effect), and conversely, it deforms when subjected to an electric field (converse effect). These unique properties make PZT ideal for use in structural health monitoring. The interaction between PZT and the host structure is analyzed using the electromechanical impedance (EMI) technique. When it is excited by the impedance analyzer at a higher frequency range, the interaction of PZT and host structure can reflect by electric impedance signature which is known as electromechanical impedance. [18, 19] The application of PZT sensors in structural health monitoring has been widely explored in recent literature. Bhalla et al. demonstrated the use of piezo-impedance transducers for assessing the residual fatigue life of bolted joints. [1] PZT is known for its high sensitivity and strong electromechanical coupling, making it an excellent choice for detecting small mechanical changes, such as bolt loosening, in structural health monitoring applications. In this study, thin PZT patches are attached to the steel structure. When damage, such as bolt loosening, occurs, the mechanical impedance of the structure changes, which is detected through variations in the electrical impedance of the PZT sensor.

Based on the various literature reviews, the fundamental concept of the EMI method is represented as a one-dimensional electromechanical impedance system using a schematic diagram as shown in Fig. 1. The system is modeled with PZT attached to the host structure to measure the mechanical impedance. When the system is subjected to voltage in an alternate z direction, the behavior of PZT is as in a one-dimensional thin

bar undergoing axial vibration in the x direction. The electromechanical impendence of the coupled system is expressed as admittance which is the inverse of impedance as per Eq. 1 [21, 22]. By applying this method, the damage of the structure can be effectively monitored and provides valuable insights into the various structural element conditions. The ability to detect and quantify changes in impedance which gives precise damage identification.

$$\overline{Y} = \omega j \frac{wl}{h}\left[\left(\overline{\varepsilon_{33}}^{T} - d_{31}^{2}\overline{Y}^{E}\right) + \left(\frac{Z_a}{Z_s + Z_a}\right) d_{31}^{2}\overline{Y}^{E}\left(\frac{\tan kl}{kl}\right)\right] \tag{1}$$

where $\overline{Y}$ is electromechanical admittance in Siemens, ω is excitation frequency in rad/sec, j is an imaginary unit to show complex part, w, l, and h are the width, length and height of PZT respectively, $\overline{\varepsilon_{33}}^{T}$ is a complex dielectric constant, d_{31} is piezoelectric strain constant, $\overline{Y}^{E}$ is elastic modulus for PZT material at the constant electric field, Z_a is mechanical impedance of actuator, Z_s is the mechanical impedances of the host structure, k is the wave number in rad/m.

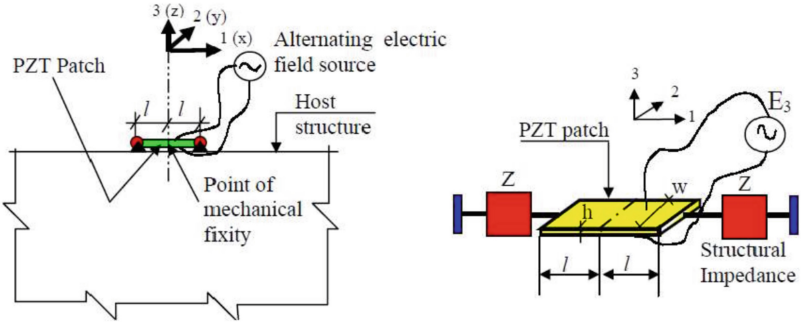


Fig. 1. A Schematic Illustration of Structural Impedance

Here in the above equation the variation in the Z_s is due to parameters like mass, stiffness, and damping of structure. Due to any damage in the structure, mechanical impedance Z_s is changed resulting in the overall impedance admittance will be changed, making this method effective for the assessment of structural damage. The above equation shows the fundamental understanding of the interaction between the electrical and mechanical impedance of PZT and host structure. Hence by monitoring the change in impedance signature for baseline measurement, the structural damage is assessed effectively using the EMI method and PZT sensor. In this study, the PZT sensor are attached on the steel plates sample which is connected to LCR meter to measure the impedance of the system. The process of the EMI methodology considered in the present study is as shown in Fig. 2.

PZT-5H sensors are suitable for low-power systems and applications having higher voltage sensitivity, higher dielectric constant along with a wide range of temperature range. PZT-5H is the most used piezoelectric material known for its higher sensitivity and electromechanical properties. The fundamental properties of the PZT sensors are shown in Table 1.

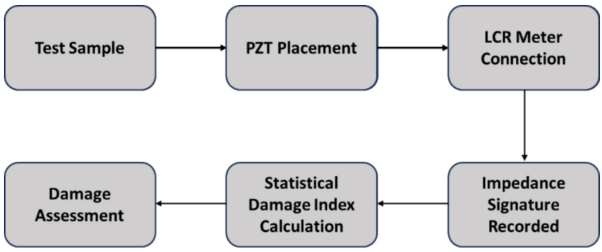


Fig. 2. Schematic Illustration of EMI method

Table 1. Standard Properties of PZT-5H

Parameters	Values
Density (ρ)	7500 kg/m^3
Poisson's Ratio (υ)	0.3
Young's Modulus (E)	56 GPa
Curie Temperature	210 °C
Relative Dielectric Constant at 1 kHz (ε)	3250
Electromechanical Coupling Co-efficient (K_{31})	0.34
Electromechanical Coupling Co-efficient (K_{33})	0.68
Planar Coupling Co-efficient (K_p)	0.56
Piezoelectric Charge Constant (d_{33})	580 pC/N

The impedance is measured over a high-frequency range of 30 to 400 kHz using an LCR meter or impedance analyzer. In this frequency range, the accuracy of change in electromechanical impedance is more, which indicates the presence of changes in any of the parameters in the structure. By comparing the electromechanical impedance of the undamaged state and potentially damaged state, it is possible to identify the damage in the structure using the EMI method. The high-frequency ranges are also more sensitive toward detecting even minor damage due to which this method is more effective for structural health monitoring.

2.2 Statistical Damage Index

In this method, structural damage is identified by the shift in the impedance. This shift may occur horizontally or vertically in the impedance signature with the variation in the frequency range. These changes in the impedance give the key indicator of the performance of the structure. The horizontal shift in the impedance signature indicates the

change in the mass distribution or structural stiffness while vertical shift indicates the change in material properties. Apart from the visual and statistical observation of the impedance signature, the quantification of damage may be identified using the statistical damage index which is an indicator of the impedance signature of the undamaged and damaged condition of the sample. Here in this study, the damage is considered in terms of bolt loosening. So, the undamaged condition is in which the bolt is in fully tight condition in which the design torque is 80 N.m. while the damage condition is in which the bolt is loosened by the variation of torque. The damage index like Root means square deviation (RMSD), Mean Absolute Percent Deviation (MAPD) and Correlation Coefficient Deviation (CCD) is used to quantify the extent of damage and assess damage severity. RMSD is identified by the comparison of the impedance magnitude over a specific frequency range for damage as well as undamaged state. The higher value of the RMSD index shows more changes or damage in structures. MAPD measures the percentage deviations of the impedance of undamaged and damaged states by providing a normalized indication of changes or damage. CCD evaluates the correlation between impedance signatures by analyzing the changes in the pattern or shape of the signature over different frequency ranges. Various literature shows that CCD may not accurately reflect small changes in impedance if the overall pattern of impedance signature remains almost similar, which leads to less precise in to detect minor damage. Hence, in the present study, RMSD and MAPD damage index are considered or the quantifying or assessment of damage due to bolt looseness. The use of RMSD and MAPD as statistical damage indices has been validated in various studies, including Thoriya et al. (2022) and Wang et al. (2022), where these indices were employed to quantify bolt looseness and corrosion levels. These metrics are particularly effective in detecting minor deviations in impedance, making them suitable for the detection of early-stage damage in bolted connections. The following equations are for damage indices RMSD and MAPD respectively [19, 20].

$$RMSD\,(\%) = \sqrt{\frac{\sum_{i=1}^{N}\left(Z^*(\omega_i) - Z(\omega_i)\right)^2}{\sum_{i=1}^{N}\left(Z(\omega_i)\right)^2}} \times 100 \tag{2}$$

$$MAPD\,(\%) = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{Z^*(\omega_i) - Z(\omega_i)}{Z(\omega_i)}\right| \times 100 \tag{3}$$

where, $Z(\omega_i)$ represents impedance signature of undamaged condition, $Z^*(\omega_i)$ represents the impedance signature obtained from actual or damaged conditions at an i^{th} frequency over N number of impedance spectrum sampling points.

3 Experimental Program

The tests are carried out on lap joint assembly made up of two steel plates of 100 mm × 8 mm having a length of 800 mm with an overlapping length of 250 mm. The plates are connected by six numbers of 16 mm diameter bolts in two rows as shown in Fig. 3. PZT sensor of size 10 mm × 10 mm with a thickness of 0.2 mm is attached the surface of the plates. The PZT patch is connected to an LCR meter (Keysight Technologies,

USA) which is used to obtain impedance data from the PZT sensor. The LCR meter applies the alternating current voltage to the PZT sensor. The impedance gets changed by the variation in mechanical properties of the structures. LCR meter collects the data of impedance over the range of applied frequencies. Hence, LCR meter is used to get the electric impedance of the sensor. Impedance signatures are recorded and analyzed for mapping the damage levels using EMI method.

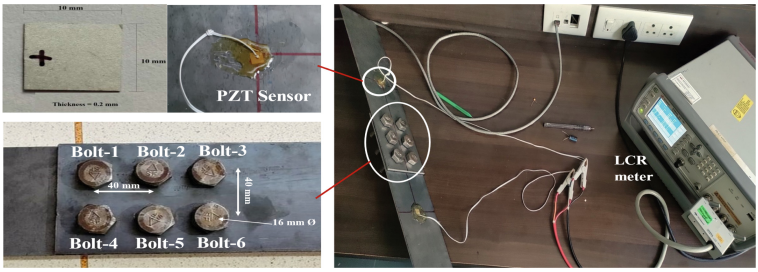


Fig. 3. Experimental Setup

In the present study, different damage conditions are considered for the bolt loosening based on single damage cases. All six bolts are tightened by applying the initial torque of 80 N.m. which is considered as an undamaged case of the test sample. The selection of the appropriate range of frequency plays an important role in the EMI method. In this study, the most significant peak frequency is observed in each case between 30 kHz to 60 kHz. Hence, the range of the frequency for the assessment of damage conditions considered in this study is from 30 kHz to 100 kHz. The impedance response for the undamaged condition of the test sample represents that the peak frequency is 50.1 kHz as per Fig. 4.

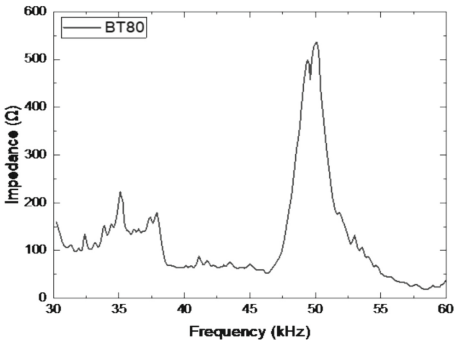


Fig. 4. Impedance Response for Undamaged Condition

4 Assessment of Bolt Loosening in Steel Connection Using EMI Technique

In this experimental study different damage cases are created by changing two conditions amount of torque released and application of loading. Mainly these damage cases are considered as single damage cases in which the amount of torque is released at the increment of 20 N.m. Details of such cases are explained in subsequent subsections. Impedance signatures of all the damage cases are compared with undamaged condition. The extent of damage derived through these comparisons are presented using statistical damage index.

4.1 Assessment of Bolt Loosening in Steel Connection Without Loading

As mentioned earlier, undamaged case is all six bolts are tightened by applying torque of 80 N.m. Subsequently, each of the bolts is given the loosening of torque at the increment of 20 N.m. up to 80 N.m. Bolt loosening of 80 N.m. is considered as bolt removal condition. In the first phase of single bolt loosening case each of the bolts are loosened up to removal conditions and 24 cases are considered for the study without applying the load. Sampled ID of each damage condition is mentioned in Table 2.

Table 2. Number of Samples for Single Bolt Loosening Cases

Sr. No.	Damage Case		Sample ID	Sr. No.	Damage Case		Sample ID
	Bolt No.	Torque in Bolt (N.m.)			Bolt No.	Torque in Bolt (N.m.)	
1	1	60	B1T60	13	1	20	B1T20
2	2	60	B2T60	14	2	20	B2T20
3	3	60	B3T60	15	3	20	B3T20
4	4	60	B4T60	16	4	20	B4T20
5	5	60	B5T60	17	5	20	B5T20
6	6	60	B6T60	18	6	20	B6T20
7	1	40	B1T40	19	1	80	B1T0
8	2	40	B2T40	20	2	80	B2T0
9	3	40	B3T40	21	3	80	B3T0
10	4	40	B4T40	22	4	80	B4T0
11	5	40	B5T40	23	5	80	B5T0
12	6	40	B6T40	24	6	80	B6T0

Impedance signatures are computed through LCR meter for each case in the range from 30 kHz to 100 kHz. However, significant variation is not observed after 60 kHz. Hence, impedance signatures presented in Fig. 4 are up to 60 kHz only. Impedance

signatures of all 24 cases are captured and compared with undamaged cases in two different ways. In the first case, comparison is done for same amount of loosening of torque for different bolts and in second case comparison is done for same bolt with different damage conditions.

4.2 Comparative Analysis of Different Bolts Loosened by Same Amount of Torque

In this case different bolts are loosened by the same amount of torque and impedance signature of each case are compared with undamaged case which is presented in Fig. 5. As presented in comparative charge there is change in peak frequency if location of bolt loosening is changed; however, that change is not significant as difference in amount of torque is also less. It is worthy to note that impedance signature of all the specimens may differ due to various factors which include the variations in the PZT properties or thickness of the epoxy layer used for bonding.

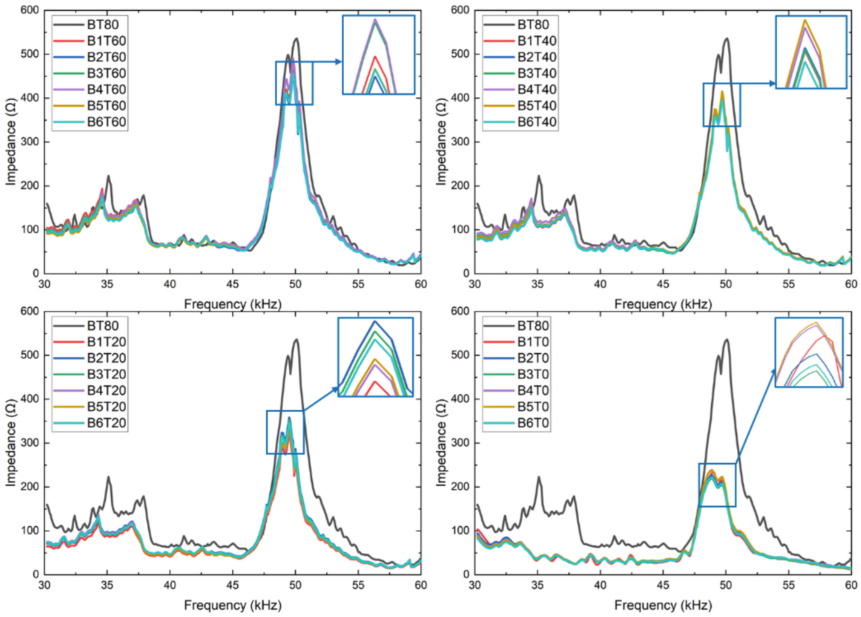


Fig. 5. Comparison of impedance signature of bolts loosened by equal torque

4.3 Comparative Analysis of Bolted Subjected to Varying Torque

In this case impedance signatures of the same bolt at different amounts of bolt loosening are done with the signature of undamaged case as presented in Fig. 6. Comparison shows the clear-cut trend of decrease in peak frequency with the increase in amount of loosening in each of the bolts. Such trend proves that this technique is useful for mapping

the damage levels. However, these trends can be seen in qualitative form. To map the damage level in quantitative form statistical tools are used as mentioned in subsection.

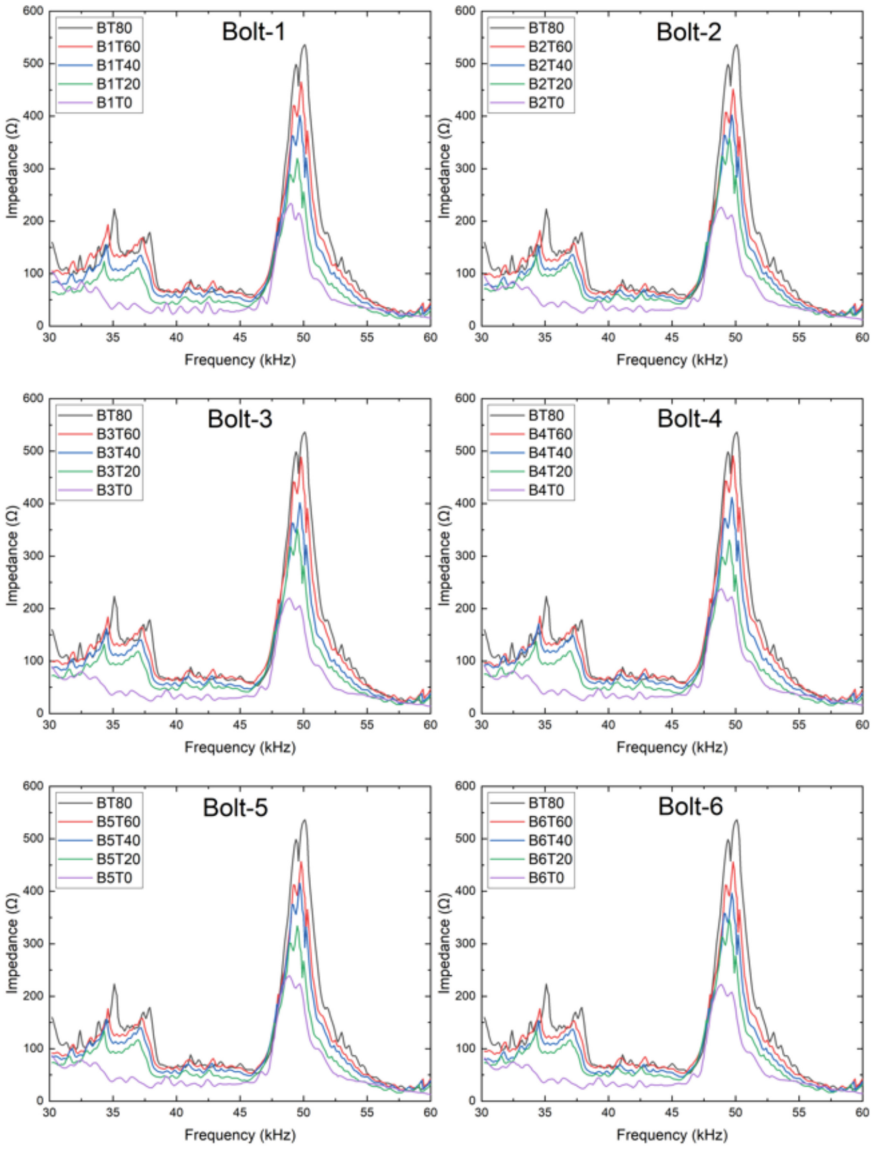


Fig. 6. Comparison of impedance signature of same bolt loosened by different amounts of torque

4.4 Relation Between Applied Torque and Frequency

The damage in the structures is due to a change in the structural parameters like mass or stiffness of the structure. In case of bolt loosening the damage is due to a change in its stiffness. As the stiffness of the connection decreases due to the looseness in the bolt, its frequency decreases resulting in the damage increasing [19]. The peak frequency of each single bolt loosening is obtained with the corresponding torque value as shown in Table 3. The average frequency response of each bolt is carried out and the relationship is drawn concerning the change in torque as shown in Fig. 7. The relation is derived using regression analysis and the frequency can be identified for the individual looseness of the bolt with change in a torque.

Table 3. Average Peak Frequency with Change in Torque in each bolt

Torque (N.m.)	Average Peak Frequency in Bolts
80	50.1
60	49.8
40	49.7
20	49.5
0	48.9

The darker pink portion in the figure indicates the range in which the true mean frequency values are expected to lie with 95% confidence curve while the lighter pink area indicates the range in which future individual frequency measurements are expected to fall with 95% confidence. It also shows that the torque is a good predictor of frequency due to a higher R-Square value. However, the R-square value being less than 50% shows that other additional variables may be beneficial to influence the frequency. Based on the statistical analysis of frequency vs torque supported by confidence intervals shows the reliable prediction of frequency based on torque.

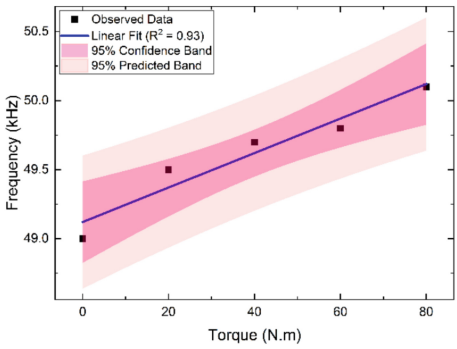


Fig. 7. Relation of Change in Frequency with Variation of Torque in Each Bolt

Damage identification indices RMSD and MAPD as mentioned in Eqs. 2 and 3 are identified based on the statistical analysis of impedance responses of different single damage cases. From the equation it should be noted that the as looseness in the bolt increases the damage in the connection increases as a result of the $Z^*(\omega_i) - Z(\omega_i)$ shows the difference in response between undamaged and damaged will decrease and hence, the damage indices increase with the increase in the looseness. [22] For each bolt, the damage indices RMSD and MAPD are found with a change in torque as shown in Fig. 8. The statistical damage indices show a similar trend in each bolt that, the RMSD and MAPD values also increase with the increase in the bolt looseness.

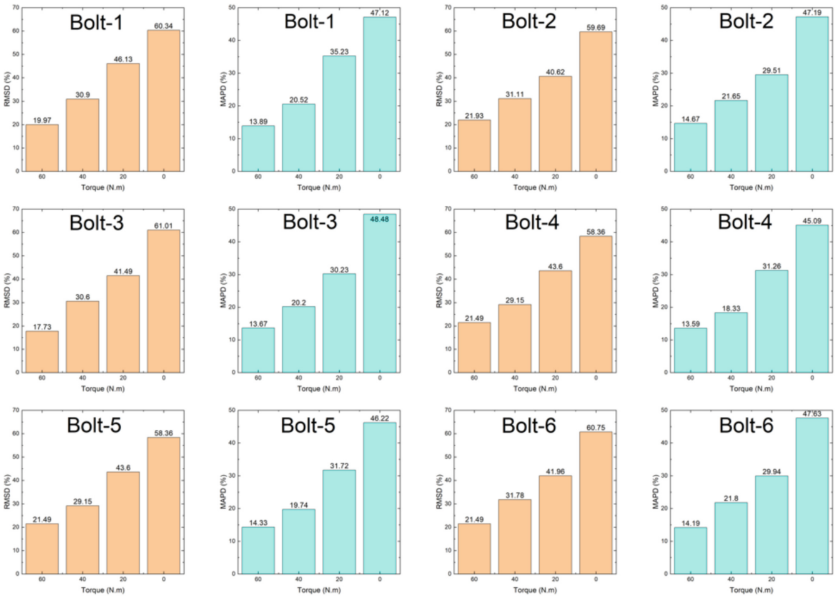


Fig. 8. RMSD & MAPD with Variation of Torque in Individual Bolt

4.5 Result and Discussion for Bolt Loosening Subjected to Axial Tension Load

Two plates are bolted with 6 bolts in a lap-jointed connection. The specimen is subjected to axial tension loading in a Universal Testing Machine. PZT patches are attached to a plate, and it is connected to the LCR meter. The impedance response is measured with the increment of load as well as the loosening of the bolt i.e., change in torque of the bolt. Initially, the full tight condition of the bolt i.e. design torque of 80 N.m. in each bolt is considered and the axial tension load of 10 kN is applied to the plate. Similarly, for the different amounts of torque in the bolt, the load is increased from 10 kN to 40 kN. The comparison of impedance response is shown in Fig. 9. For each case, when the load is increased with torque remaining constant, the response of impedance vs frequency is shifted towards the right which indicates both impedance and torque values are increasing by an increase in the load.

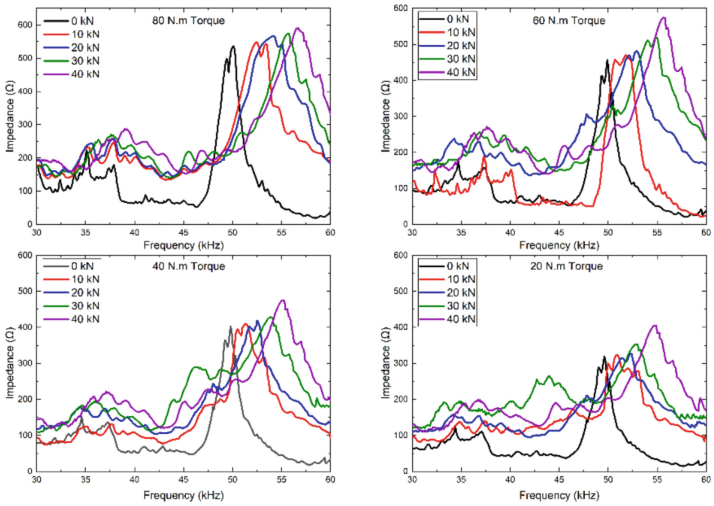


Fig. 9. Impedance Response with a Variation of Load and Torque

The peak average frequency of each case is obtained and the relation between the Torque in the bolts and frequency at different axial tension loads varies from of 0 to 40 kN is derived using regression analysis as shown in Fig. 10. For each load level, the slope of the linear firm is negative which shows that the frequency decreases with the increase in the looseness of the bolt. The linear fit lines for each loading show the best-fit relationship between torque and frequency. The slope in each loading condition is consistent which indicates the uniform rates of looseness in the bolt. Overall, the figure shows the behavior of torque with the application of axial tensile loading which

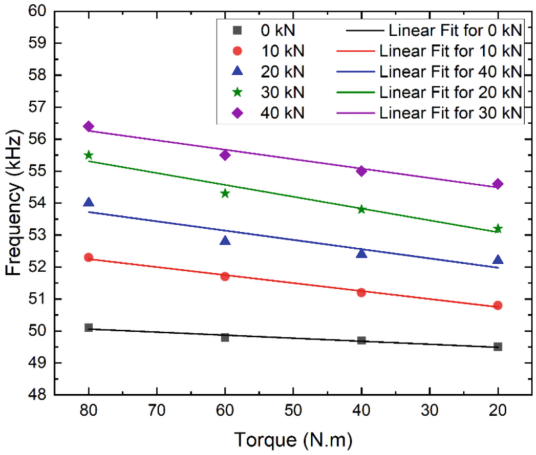


Fig. 10. Relationship of Torque and Frequency for bolts

shows that as load increases the frequency, and the torque is linearly decreasing with the looseness in the bolts. The linear fit lines provide a clear visualization of the trend.

When the specimen is subjected to loading the damage index (RMSD and MAPD) increases as the looseness degree increases, especially for the initial looseness levels as shown in Fig. 11.

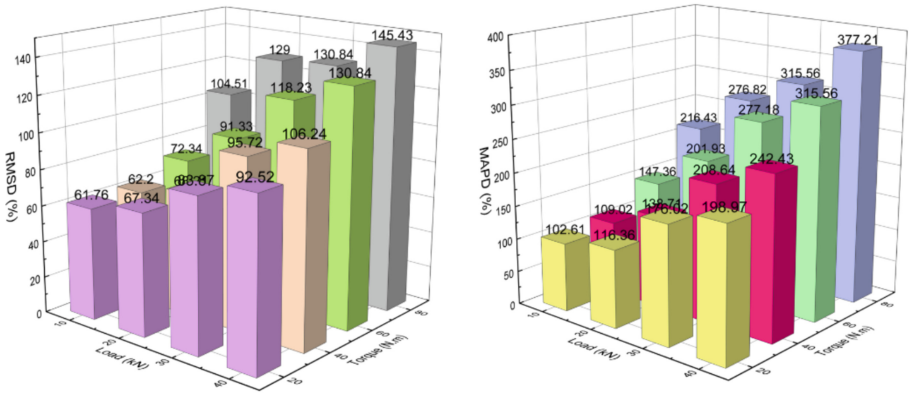


Fig. 11. RMSD and MAPD Variation with Torque and Load

5 Conclusion

In this study, the smart computing method for damage detection of the grip levels of the group of the bolts at the steel member connection has been evaluated. The objective was to assess the application of sensor-based damage detection monitoring method for secondary members namely bolts and their loss of grip with time for steel structures. From the study the following important conclusions have been retrieved.

- Results show that the loosening of individual bolts with the succession of 20 N.m. torque a noticeable decrease in the amplitude of impedance signature, resulting in the flattening and smoothing of the curves which is a consistent trend across all the specimens.
- Results show that both RMSD and MAPD indices increase with the degree of bolt loosening. Also, it is more sensitive towards the initial small degree of looseness which shows that it is providing quantitative measures for damage assessment. Among the indices, RMSD is consistent with increased torque. Conversely, MAPD exhibits the most pronounced values showcasing the high sensitivity to variation in impedance. The findings show the effectiveness of RMSD and MAPD in structural health monitoring for early detection of damage.
- The bolt loosening shows that for each load increment both peak frequency and impedance increase with the load as impedance highlighting the importance of considering the relationship of load and torque variation in structural health monitoring.

- The slope of the linear fit between torque and frequence is negative in all the levels which indicates that the frequency is decreasing with the looseness in the bolts.
- Overall, the results show that the EMI method is a viable technique to detect and quantify damage in structural elements due to bolt loosening. Along with that, the relation of torque, impedance, and frequency gives the framework to enhance the reliability and safety of the structural element due to early detection of damage.
- The study established the validation of the application of smart computing method for steel structural health monitoring using sensor-based damage detection and encouraged to extend the study in the area by addressing the limitations found in prior studies and offers a more robust solution for monitoring bolted steel structures.

References

1. Bhalla, S., Vittal, P.A., Veljkovic, M.: Piezo-impedance transducers for residual fatigue life assessment of bolted steel joints. Struct. Health Monit. **11**(6), 733–750 (2012)
2. Talakokula, V., et al.: Diagnosis of carbonation induced corrosion initiation and progression in reinforced concrete structures using piezo-impedance transducers. Sens. Actuators A **242**, 79–91 (2016)
3. Farrar, C., Sohn, H., Robertson, A.N.: Applications of nonlinear system identification to structural health monitoring. No. LA-UR-04-2947. Los Alamos National Lab. (LANL), Los Alamos, NM (United States) (2004)
4. Maity, D., Tripathy, R.R.: Damage assessment of structures from changes in natural frequencies using genetic algorithm. Struct. Eng. Mech. Int. J. **19**(1), 21–42 (2005)
5. Eraky, A., et al.: Damage detection of flexural structural systems using damage index method–Experimental approach. Alexandria Eng. J. **54**(3), 497–507 (2015)
6. Keshmiry, A., et al.: Effects of environmental and operational conditions on structural health monitoring and non-destructive testing: a systematic review. Buildings **13**(4), 918 (2023)
7. You, J.M., et al.: Behavior of bolt-connected steel plate girder attributable to bolt loosening failure in the lower flange. Eng. Failure Anal. **107**, 104208 (2020)
8. Zhang, Z., et al.: Continuous monitoring of residual torque of loose bolt in a bolted joint. Procedia Eng. **188**, 278–285 (2017)
9. Sah, S.M., et al.: Estimating bolt tightness using transverse natural frequencies. J. Sound Vibr. **431**, 137–149 (2018)
10. Dawari, V.B., Vesmawala, G.R.: Structural damage identification using modal curvature differences. IOSR J. Mech. Civil Eng. **4**, 33–38 (2013)
11. Sazonov, E., Klinkhachorn, P.: Optimal spatial sampling interval for damage detection by curvature or strain energy mode shapes. J. Sound Vib. **285**(4–5), 783–801 (2005)
12. Todd, M.D., Nichols, J.M., Nichols, C.J., Virgin, L.N.: An assessment of modal property effectiveness in detecting bolted joint degradation: theory and experiment. J. Sound Vib. **275**, 1113–1126 (2004)
13. Kim, I.T., Lee, J.M., Huh, J.W., Ahn, J.H.: Tensile behaviors of friction bolt connection with bolt head corrosion damage: experimental research B. Eng. Fail. Anal. **59**, 526–543 (2016)
14. Xu, Z.-D., Ke-Yi, W.: Damage detection for space truss structures based on strain mode under ambient excitation. J. Eng. Mech. **138**(10), 1215–1223 (2012)
15. Sirohi, J., Chopra, I.: Fundamental behavior of piezoceramic sheet actuators. J. Intell. Mater. Syst. Struct. **11**(1), 47–61 (2000)
16. Sirohi, J., Chopra, I.: Fundamental understanding of piezoelectric strain sensors. J. Intell. Mater. Syst. Struct. **11**(4), 246–257 (2000)

17. Thoriya, A., Vora, T., Nyanzi, P.: Pipeline corrosion assessment using electromechanical impedance technique. Mater. Today Proc. **56**, 2334–2341 (2022)
18. Thoriya, A., Vora, T., Makwana, V.: Corrosion assessment in rebars of high-strength concrete using electromechanical impedance technique. Mater. Today Proc. **57**, 2234–2241 (2022)
19. Thoriya, A., et al.: Application of wavelet transform techniques for corrosion assessment of embedded rebars in RC elements using electromechanical impedance. Measurement **226**, 114081 (2024)
20. Wang, L., et al.: Synchronous detection of bolts looseness position and degree based on fusing electro-mechanical impedance. Mech. Syst. Signal Process. **174**, 109068 (2022)
21. Duddi, M., Kocherla, A., Subramaniam, K.V.L.: Evaluation of boundary and material influences on the dynamic electromechanical impedance response of the embedded PZT sensor in concrete. Sens. Actuators A Phys. **361**, 114575 (2023)
22. Park, G., Cudney, H.H., Inman, D.J.: Impedance-based health monitoring of civil structural components. J. Infrastruct. Syst. **6**(4), 153–160 (2000)
23. Xu, J., et al.: Looseness monitoring of bolted spherical joint connection using electromechanical impedance technique and BP neural networks. Sensors **19**(8), 1906 (2019)
24. Li, W., et al.: EMI instrumented conical corrosion measuring probe for pipeline corrosion monitoring: experiments with FEM validation. Sens. Actuators A Phys. **362**, 114678 (2023)
25. Luo, W., et al.: An electromechanical impedance-based sensor for monitoring the pitting corrosion of steel: simulation with experimental validation. Sens. Actuators A Phys. **376**, 115585 (2024)
26. Rao, R.K., Sasmal, S.: Electromechanical impedance-based embeddable smart composite for condition-state monitoring. Sens. Actuators A Phys. **346**, 113856 (2022)

Application of Smart Computing Systems for Smart Cities and Urban Infrastructure: Framework, Data Management, and Smart Monitoring Attributes

Ankur Bhogayata(✉), Amit Thoriya, Tarak Vora, and Divyakant Meva

Marwadi University, Rajkot, Gujarat 360003, India
{ankur.bhogayata,amitkumar.thoriya,tarak.vora,
divyakant.meva}@marwadieducation.edu.in

Abstract. Smart cities require the utilization of advanced computing technology to address the intricate challenges posed by the rapid expansion of metropolitan areas. This study discusses the use of intelligent computing systems in smart cities, specifically emphasizing their conceptual frameworks, essential components, and actual implementations. Smart computing systems consist of several technologies, including Internet of Things (IoT) devices, artificial intelligence (AI), and big data analytics. These technologies collaborate to enhance urban infrastructure and services. The aim of this article is to provide a comprehensive understanding of how intelligent computing systems might potentially transform urban living and make substantial contributions to the progress of sustainable, efficient, and resilient smart cities. The study explores essential components, including data collection and analysis, real-time decision-making, and resource optimization through automation and predictive analytics. The article represents the framework of IoT, integration of various smart components, operation of devices, working with data collected, smart monitoring of urban infrastructures, application of smart computing aspects for smart homes and smart cities by discussing real time example. The article additionally examines the challenges and potential associated with the adoption of these technologies, including concerns over data privacy, the need for robust cyber security measures, and the importance of scalable solutions. The article represents an overall framework for further research and progress in intelligent computing systems for smart cities and infrastructures.

Keywords: Smart Computing · Smart Cities · Data Management · Data Monitoring · Internet of Things · Artificial Intelligence

1 Introduction

Smart computing has become an essential component of urban living by smoothly incorporating new technologies into regular routines. Urban areas are increasingly employing intelligent computing approaches to improve the quality and efficiency of services [1]. These technologies are incorporated into essential infrastructure, such as intelligent

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
S. Rajagopal et al. (Eds.): ASCIS 2024, CCIS 2428, pp. 212–234, 2025.
https://doi.org/10.1007/978-3-031-86302-8_14

grids that optimize energy consumption and traffic control systems powered by artificial intelligence that decrease road congestion. Smart computing facilitates the utilization of real-time data analytics for purposes such as public safety, waste management, and environmental monitoring. This allows communities to quickly adapt to changing conditions and improve the quality of life for their residents [2]. The integration of smart technology is playing a significant role in modernizing and improving urban living, as cities become more urbanized [2]. Modern city life revolves on smart computing since it allows us to apply fresh technologies in daily tasks. To make services more effective and better, smart computer technologies are growingly vital in cities [3]. These technologies are part of crucial infrastructure like smart grids using IoT to maximize energy use and artificial intelligence-driven traffic management systems reducing traffic congestion [4]. Smart computing enables real-time data analytics for environmental monitoring, waste management, and public safety as well as for public safety. This helps towns swiftly change with the times and enhance public health [5]. As more people migrate to contemporary cities, smart computing is becoming ever more crucial. It keeps evolving and simplifies metropolitan living.

1.1 Importance of Smart Computing Systems in Improving Efficiency and Sustainability

Smart computing systems are advanced technologies intended to independently process, analyse, and act upon data to increase operational efficiency and decision-making. Usually covering artificial intelligence (AI), machine learning, and Internet of Things (IoT) devices, they enable real-time data collecting and processing. Important characteristics are scalability, which facilitates development and communication with other systems; adaptability, which lets systems evolve with the times; and connectivity, which lets devices and platforms interact smoothly. From smart cities to corporate operations, these technologies also stress data-driven insights and automation to enhance processes, increase accuracy, and improve user experiences across numerous applications.

Modern technologies help to maximize municipal operations and resource management by enabling smart computing systems—which are essential in improving efficiency and sustainability in contemporary urban environments. Through its ability to help cities collect, assess, and act upon enormous volumes of real-time data, these technologies increase operational efficiency and environmental sustainability. Using IoT sensors and big data analytics, the Smart Nation project in Singapore checks air quality, manages traffic flow, and reduces energy use, therefore enabling more effective urban infrastructure and reducing of environmental impact [6]. Barcelona has also erected a smart grid system combining real-time monitoring with renewable energy sources to best distribute electricity and cut waste. By means of predictive maintenance, lower resource consumption, and improved city service responsiveness, these intelligent computer technologies help to guide more sustainable metropolitan environments. Cities can satisfy the rising needs of urban people and lower environmental damage by always improving and applying such technology, so balancing development with environmental preservation [7]. Thus, the acceptance of smart computing systems will assist to develop strong, sustainable, and flexible urban ecosystems capable of serving the needs of the twenty-first century [8].

The implementation of smart computing systems for creating smart cities and infrastructure monitoring frameworks requires several complex services, their upgradation and use of modern tools [9]. Irrespective to the preliminary or targeted areas, the general concept of smart computing systems namely AI or IoT framework, the smart systems will be the mandatory needs for the future urban spaces [10]. Being specific with IoT applications, everything in the ubiquitous IoT network is connected, has a distinct internet address, and is reachable from anywhere and at any moment. IoT offers a perfect platform for gathering enormous volumes of data, while artificial intelligence offers a means of data analysis and smart decision-making [11]. It must be noted that the urbanization of the present spaces namely the towns, their transportation and road networks will be required to be transformed into the smart entities having established smart infrastructural facilities [12]. The newly emerging smart spaces also requires careful consideration for the energy management within the smart systems. Along with numerical data, the energy data will also be required to be monitored [13, 14]. In this regard, the present and future of smart cities and infrastructure will largely be relying on the modern concepts of smart computing systems and tools.

2 Smart Computing System for Smart Cities

2.1 IoT Framework and Components

The Internet of Things (IoT) is a revolutionary technology that connects physical things to the digital world, enabling seamless communication and data exchange. This section provides preliminary understanding of IoT characteristics, such as interoperability, scalability, and connection, all of which are essential to comprehending how IoT systems work. After that, it explores fundamental ideas including data processing, sensors, actuators, and cloud computing place in Internet of Things architecture.

Dynamic interplay involving sensors, devices, and software is the way internet of things (IoT) architecture operates for automated intelligent systems to exist. The basis of the internet of things is made up of real-world items that have integrated sensors which collect information and then transmit it to cloud platforms or central systems via network. This data is processed and analyzed in real-time using complex algorithms and machine learning to bring out valuable insights. Decision making as well as dynamic response are enabled when these insights trigger actuators towards performing specific tasks. The framework illustrated in Fig. 1, which often uses technologies like Wi-Fi, Bluetooth, cellular networks among others heavily relies on dependable connectivity protocols that enable smooth device communication.

A typical IoT framework consists of the following components.

1. Sensors/Devices:
   - Role: Collect data from the physical environment.
   - Examples: Temperature sensors, motion detectors, smart appliances.
2. Connectivity:
   - Role: Transmit the data collected by sensors to other devices or central systems.
   - Technologies: Wi-Fi, Bluetooth, Zigbee, Cellular (4G/5G)

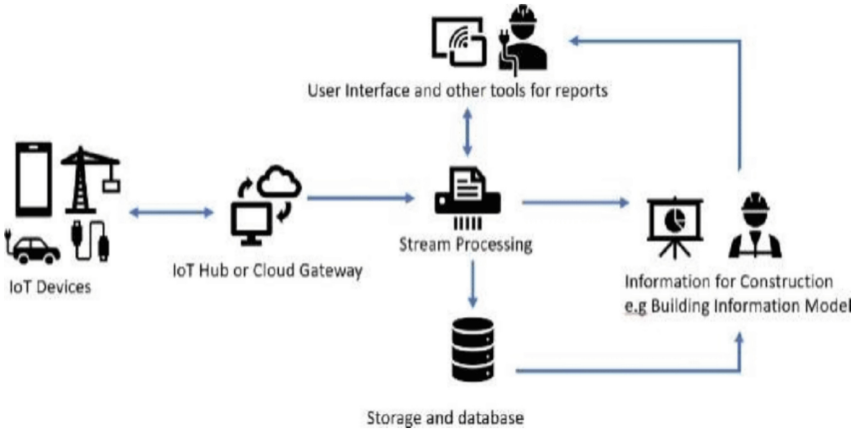


Fig. 1. Illustration of IoT framework for monitoring activities [23]

3. Data Processing:
   - Role: Process and filter the raw data collected by sensors for preliminary insights.
   - Location: Can occur at the edge (near the data source) or centrally in the cloud.
4. Cloud/Edge Computing:
   - Role: Provide the computational power needed to handle large volumes of data, execute complex algorithms, and store data.
   - Cloud Computing: Centralized data processing and storage.
   - Edge Computing: Decentralized processing near the data source to reduce latency.
5. Data Storage:
   - Role: Store the processed data for long-term use and historical analysis.
   - Technologies: Databases, Data Lakes, Cloud Storage solutions.
6. Data Analysis:
   - Role: Analyze the processed data to generate actionable insights and predictions.
   - Tools: Machine learning algorithms, analytics software, AI models.
7. Actuators:
   - Role: Execute physical actions based on insights or commands from the IoT system.
   - Examples: Smart locks, automated lighting systems, industrial robots.
8. User Interface:
   - Role: Provide a platform for users to interact with the IoT system, monitor status, and control devices.
   - Examples: Mobile apps, web dashboards, voice assistants.
9. Security:

- Role: Protect data integrity, confidentiality, and ensure secure communication between devices.
- Mechanisms: Encryption, authentication, access control, secure boot.

10. Network Protocols:

- Role: Define the rules for data exchange between devices and systems.
- Examples: MQTT, CoAP, HTTP/HTTPS, AMQP.

11. Integration with Other Systems:

- Role: Ensure that IoT devices can interact and integrate with existing IT systems and other IoT ecosystems.
- Examples: API gateways, middleware platforms.

These components work together to form a comprehensive IoT framework, enabling the seamless operation and management of IoT systems across various applications and industries. When the components are integrated, they require various infrastructures also. The infrastructure includes buildings, building services, cabling, communications network connections, installation and commissioning of electrical devices, interfacing of electronic equipment, arrangement of wired and wireless connections and several other major and minor accessories [15]. The primary source of functionality of such systems is the electric supply and provision of alternative or standby power supply systems also. Depending upon the location and functional areas or premises, these components can vary in their nature of structure and type. For example, the interconnection of devices may be wired or wireless, for this, the devices may be installed on the site or around the site, or even remotely from the workstation. These aspects largely depend on the type of activities that we will explore in more detail in preceding sections.

2.2 Operation and Management of Devices in IoT Framework

The overall function of an IoT enabled system can be broadly divided into two parts, namely Operation (Driving / Running), and Management oriented tasks. We will discuss how these may influence the overall functioning of the system.

The five major operational aspects of IoT-enabled systems, with reference to their working systems, are:

1. Data Collection and Sensing
2. Connectivity and Communication
3. Data Processing and Analytics
4. User Interaction and Interface
5. Security and Privacy

1. Data Collection and Sensing

Function: This aspect involves the initial step of gathering data from the physical environment using various sensors and devices.

Working System:

- Sensors and Actuators: Devices like temperature sensors, humidity sensors, motion detectors, cameras, and actuators collect real-time data from their surroundings.

- Edge Devices: These are often embedded with the sensors to perform preliminary processing and filtering of data before sending it further.

2. Connectivity and Communication

   Function: Ensuring the seamless transmission of data from sensors and devices to data processing units, storage, and user interfaces.

   Working System:

   - Communication Protocols: Technologies such as Wi-Fi, Bluetooth, Zigbee, LoRaWAN, and cellular networks facilitate data transfer. The choice depends on the required range, data rate, and power efficiency.
   - Gateways: Devices that aggregate data from multiple sensors and manage the communication between the IoT devices and the central system or cloud.

3. Data Processing and Analytics

   Function: Transforming raw data into useful information through processing and analysis.

   Working System:

   - Edge Computing: Localized processing to reduce latency and bandwidth usage by performing computations near the data source.
   - Cloud Computing: Centralized processing for more intensive computations, data storage, and advanced analytics using machine learning and AI.
   - Data Analytics: Techniques to analyze data trends, perform predictive analytics, and generate actionable insights.

4. User Interaction and Interface

   Function: Providing mechanisms for users to interact with the IoT system, monitor data, and control devices.

   Working System:

   - Dashboards and Applications: User-friendly interfaces such as web dashboards, mobile applications, and desktop applications that display data insights and control options.
   - Alerts and Notifications: Systems to alert users about important events or anomalies via emails, SMS, or in-app notifications.
   - Control Mechanisms: Interfaces that allow users to send commands back to IoT devices, such as turning a device on or off, adjusting settings, or scheduling tasks.

5. Security and Privacy

   Function: Protecting IoT systems and data from unauthorized access and ensuring user privacy.

   Working System:

   - Encryption: Protecting data in transit and at rest using cryptographic techniques.
   - Authentication and Authorization: Ensuring only authorized users and devices can access the IoT system through methods like multi-factor authentication and role-based access control.
   - Regular Updates and Patches: Keeping software and firmware up to date to protect against vulnerabilities.

- Compliance and Data Governance: Adhering to regulations and standards (e.g., GDPR) to protect user data and ensure privacy.

From the above discussion, we can now identify the requirements of IoT enabled systems for a given working environment; specifically, it will help us in contextualizing the IoT system in different smart applications for infrastructure and urban spaces.

2.3 Data Collection, Data Management, and Data Processing

While talking about Data as a keyword, we must know their forms and formats and similarly how they are generated and can be recorded as well as transformed between the systems. Let us first look on the Data as a generic term.

Data refers to basic facts, figures, or information collected from various sources, often used for analysis, decision-making, or processing. It can be in various forms such as numbers, text, images, or signals and is unprocessed or unorganized when initially gathered. Data becomes meaningful when processed and interpreted, transforming into information. In the context of technology and computing, data can be stored, transmitted, and manipulated to generate insights, drive actions, or support systems. It is fundamental to fields like science, business, and engineering, serving as the foundation for generating knowledge and making informed decisions.

Data Collection/Recording

Now let us explore the types and form of data in context with Civil Engineering. As we all know there are various fields of application of Civil Engineering, according to that, we may identify data types. Given below are a few such examples.

1. Geotechnical Data
   - Description: Information about soil and rock properties, ground conditions, and sub-surface materials.
   - Examples: Soil composition, moisture content, density, shear strength, and borehole logs.
2. Topographical Data
   - Description: Detailed information about the terrain and landscape features of an area.
   - Examples: Elevation data, contour maps, digital elevation models (DEMs), and land survey measurements.
3. Structural Data
   - Description: Data related to the design, analysis, and performance of structural elements.
   - Examples: Load calculations, material properties (e.g., concrete strength, steel tensile strength), stress-strain curves, and deformation data.
4. Environmental Data
   - Description: Information on environmental conditions that affect civil engineering projects.

- Examples: Weather data (temperature, precipitation, wind speed), hydrological data (water flow rates, flood levels), and pollution levels (air, water, soil).

5. Construction Data
   - Description: Data collected during the construction phase of a project.
   - Examples: Project timelines, resource usage (labor, materials, equipment), construction progress reports, and quality control records.

2.4 Data Acquisition from Sensor-Based Devices

The sensors play a key role in IoT enabled applications. Therefore, it is important to understand the data acquisition process through the sensors and attached devices. Following is an overview on the working of sensors in general.

Fig. 2. Illustrated connections and working of sensors at site [19]

The working principle illustrated in Fig. 2 of a sensor involves the conversion of physical, chemical, or biological phenomena into an electrical signal that can be measured, processed, and analyzed. The following figure illustrates the working of a sensor and here is a breakdown of the general working principle:

1. Detection

   The sensor detects a specific physical quantity (e.g., temperature, pressure, light, motion) or chemical/biological property (e.g., gas concentration, pH level).
2. Transduction

   This detected quantity is converted into a measurable signal. Transducers within the sensor play a key role in this process. The transduction mechanism varies depending on the type of sensor:

   - Resistive Sensors: Change in resistance (e.g., thermistors for temperature measurement).
   - Capacitive Sensors: Change in capacitance (e.g., touch sensors).

- Inductive Sensors: Change in inductance (e.g., proximity sensors).
- Optical Sensors: Change in light intensity (e.g., photodiodes for light detection).
- Piezoelectric Sensors: Generation of an electric charge in response to mechanical stress (e.g., accelerometers).

3. Signal Conditioning

   The raw signal from the transducer often requires conditioning to be useful. This includes amplification, filtering, and analog-to-digital conversion. Signal conditioning ensures that the signal is strong, accurate, and noise-free for further processing.

4. Output

   The conditioned signal is then output in a form that can be interpreted by measurement instruments, control systems, or computers. This output can be:

   - Analog: Continuous signals representing the measured quantity.
   - Digital: Discrete signals or data representing the measured quantity in binary form.

2.5 Role of Machine Learning, Algorithms, Monitoring, and Improvement in IoT Framework

Machine learning can improve structure health monitoring in civil engineering. For instance, machine learning systems examine sensor data from bridges to find wear indicators and other problems. Through constant observation of variables such as temperature, tension, and vibrations, the system picks up on typical patterns and recognizes abnormalities. By using a predictive maintenance strategy, engineers may fix problems before they become serious, increasing the longevity of structures and guaranteeing their safety. Furthermore, with enhanced data analysis and modeling, machine learning can forecast material behavior under various conditions, optimize building processes, and enhance infrastructure project design.

An algorithm is a methodical process or collection of guidelines created to carry out a task or solve a particular issue. It is a cornerstone of computer science that directs decision-making and data processing, guaranteeing effective and efficient results in a range of applications. There are several algorithms available and can be developed according to the intended task or purpose to run a software program on computers. An algorithm is a backbone of any programming language where the instructions are being guided to the data or input values that what and when a specific task is to be performed. Just to name a few commonly used algorithms are Automated or Machine Learning Systems, Digital Signature Algorithms, Algorithms for Sorting, Lookup Formulas, Pathfinder Formulas.

Improvement of IoT enabled system: The IoT architecture can be enhanced over time by a data-driven feedback loop. By analyzing sensor data, machine learning finds possible problems and inefficiencies. Downtime can be avoided using predictive maintenance, and anomaly detection can identify anomalous activities. The framework is then improved with the use of this data, which is used to optimize data collection, retrain models for increased accuracy, and even modify device operation considering practical observations. The IoT network grows more effective, dependable, and future proof because of the system's learning and adaptation [16].

Advantages of system improvement:

- Enhanced Efficiency: The system may function with less waste and produce better results with less resources by continuously analyzing data and adjusting procedures.
- Enhanced Reliability: A more trustworthy and predictable system can be achieved by preventing unanticipated failures and downtime via the use of anomaly detection and predictive maintenance.
- Better Decision Making: Data-driven insights can help direct decision-making, facilitating more intelligent strategy formulation, efficient workflows, and better resource allocation.
- Increased Security: By identifying and fixing security flaws more quickly, continuous improvement strengthens the system's defense against cyberattacks and data breaches.
- Increased Scalability: A flexible and well-maintained framework makes it simple to incorporate new features and devices, enabling the system to expand and change in response to changing requirements.

Important concerns of utilizing Machine learning and Algorithm concepts in systems for smart cities:

While implementing the applications of machine learning or algorithms for a system that controls or monitor many datasets and types, it is important to consider the importance of data quality, data usage and their meaningful application for predictive maintenance applications of the devices or systems. Following are the two major concerns that should be addressed while working with machine learning or algorithm-based systems.

- Anomaly detection: In a dataset, anomaly identification is the process of spotting trends or data points noticeably different from the expected behavior. In many different sectors, including manufacturing, cybersecurity, and banking, it is often utilized to find odd activity, fraud, or flaws. However, it is equally important in case of the smart city and smart infrastructure management and monitoring processes.

This challenge can be mitigated by incorporating several methods. The methods usually involve statistical, machine learning-based, or rule-based methods, techniques for anomaly identification. Such processes may be implemented by the following means.

– Data collecting is compiling pertinent information from several sources.
– Establishing a baseline of what qualifies as "normal" activity from past data helps one to model typical behavior.
– Finding variations from this baseline could point to anomalies.
– Validation: Determining whether found abnormalities call for more research and is important.

Good anomaly detection guarantees improved security and helps companies react proactively to possible problems, therefore guaranteeing better decision-making.

- Predictive maintenance: Particularly with the application of machine learning and algorithms, predictive maintenance is essential for the execution of a vast monitoring of data. These main ideas underline its significance:

  – Enhanced system efficiency: Using data analysis, predictive maintenance forecasts equipment faults before they start. This proactive approach reduces downtime,

maximizes resource allocation, and improves the general effectiveness of local infrastructure including trash management, utilities, and public transit systems.
- Improved cost saving: Anticipating maintenance requirements helps communities lower unanticipated repair costs and increase the lifetime of assets. Over time, this results in large financial savings that free resources for other important urban projects and services.
- Safety: Predictive maintenance helps find possible problems in infrastructure including public transportation systems, bridges, and highways. Cities can increase public safety and lower the probability of accidents by tackling these problems before they become more severe.
- Effective decision making: Using machine learning tools helps cities to examine enormous volumes of running data. This guarantees that actions are based on real usage patterns and conditions rather than reactive measures, therefore guiding informed decisions about maintenance plans and resource investments.
- Sustainability of the overall system: By lowering energy use and waste, predictive maintenance helps to support environmental efforts. Effective maintenance plans correspond with more general objectives for sustainable urban development by reducing the environmental impact of city activities and extending the lifetime of infrastructure.

We have discussed several important and learning attributes of IoT enabled framework that can be employed into the Civil Engineering activities to improvise the output, productivity and safety along with the economic benefits. Lastly, we will refer to another important attribute of the IoT system called accuracy and error detection. These are also an essential part of any IoT enabled environment and must be known to us.

2.6 Benefits of Implementing Smart City Framework Using IoT Enabled Systems

Smart city technologies maximize urban operations and thereby help more effective management of resources including power, water, and transportation by means of real-time data analysis and automation. This can improve service supply and help to reduce running expenses. Among other technologies and IoT-based public services, smart traffic management helps smart cities raise resident quality of living. This target improved air quality, lessening of transportation congestion, and more sensitive public safety services. Smart grids and energy-efficient buildings among green technologies let smart cities assist environmental sustainability. Using these technologies to lower greenhouse gas emissions and energy consumption helps to create a better, healthier environment. Adoption of smart technology can improve economic development by means of investment attraction, new business generation, and invention encouragement. Many times, smart cities serve as centres for research projects and IT businesses, therefore promoting economic growth. Predictive analytics and real-time monitoring let smart city models boost urban resilience. This helps cities to foresee and react properly to crises including infrastructure breakdowns or natural disasters, therefore enhancing general urban stability and safety.

3 Examples of Smart Infrastructure: Learning from the Real-Life Examples

The term smart infrastructure is a broader perspective for the varying frame works and systems working on the computerized systems along with the digital devices. To narrow down the perspective, the authors have discussed three examples of the infrastructure components working on the smart functioning and monitoring concepts namely smart cities, smart homes and remote health monitoring of the transportation infrastructure namely bridges.

3.1 Smart Cities

The terms are well known namely Smart Cities and Smart Homes nowadays; however, it is interesting and important to know what makes the ordinary cities and homes smart and what are the technical and general advantages of transforming the conventional cities and homes into smart one? The aim of this discussion is to provide readers with general awareness on how IoT enabled systems can support and help in transforming the conventional systems into a smart system. We will understand this by looking at some direct examples.

The Concept of Smart Cities: An urban area that employs information and communication technology (ICT) and Internet of Things (IoT) to improve urban services' performance and quality, use less resources, and interact more efficiently and actively with its residents is known as a "smart city." By making the city more livable, sustainable, and efficient, smart cities aim to raise the standard of living for all their citizens [17, 18]. There is an example that exists for the same.

Spain's Barcelona
Because of its extensive use of technology to enhance urban living, Barcelona is frequently regarded as a model smart city as illustrated in Fig. 3. Among its smart city projects are the following:

- Smart street lighting uses motion sensors to save energy by adjusting the illumination according to the presence of cars and pedestrians.
- Smart Waste Management: Waste collection routes are optimized, and fill levels are monitored by sensors installed in garbage cans.
- Smart transportation refers to integrated public transportation systems that provide real-time schedule information for buses and trains, as well as programs for sharing bikes.
- E-Government: Online gateways to government services, such as bill payment, form submission, and public information retrieval.
- Sustainable initiatives include renewable energy projects, urban gardens, and green structures.
- Traffic and energy management, home automation
- Supervised health and monitoring, trends in population growth
- Smart waste management and disposal, resource management

Fig. 3. Illustration of the smart attributes for smart cities [20]

3.2 Smart Homes

A smart home illustrated in Fig. 4 is a place of living outfitted with a variety of networked technologies and gadgets that can be managed, controlled, and seen from a distance to enhance comfort, efficiency, security, and convenience. These systems and gadgets can be operated by voice assistants, computers, tablets, smart phones, and other wireless communication technologies like Wi-Fi, Bluetooth, Zigbee, or Z-Wave. Few of the smart systems for homes have been discussed in this section.

Fig. 4. Illustration of smart home attributes [21]

Surveillance/Outdoor Camera
Sensors are the eyes and ears of your smart home, gathering information on a range of environmental factors. This includes the records of any suspicious or unwanted movements of human or animal or any such events. These records are possible to be stored and reviewed, reaping the stored data for safety and security purposes. The special motion sensors can detect all desired or programmed events as per the preset of these IoT enabled devices.

Alarms/Warning Bells
When intruders or unknown entities try entering the home premises, the pre-installed alarm bells and provides warnings to the owner. They work based on unusual movements and activities that are not obvious or in a routine, for example an alarm of the Door Ajar System warns us when the doors are left open for more than the usual time. These devices are many times smart enough to learn the usual patterns of the occupants. Accordingly, they can reprogram themselves for any unwanted incidents also.

Smart Cleaning/Washing Systems
Practically any indoor or outdoor service is nowadays possible to get clean by its smart function. Or they can be remotely controlled or operated by the users. For example, the mopping robots, the cleaning of solar rooftops, on and off the small devices like Air conditioners, washing machines, Microwave ovens etc. can be done using the dedicated home automation Applications even when you are far away from the home. Such applications are working on the IoT enabled framework primarily.

Smart Meters
Smart meters are installed in homes and buildings and track water and energy usage in real time. Residents can track their usage and pinpoint places for conservation with the help of this data. Utility firms can also use it to detect possible outages and optimize the delivery of energy.

3.3 Additional Important Attributes of Smart Cities and Smart Homes

Let us now discuss the attributes of smart cities and smart homes. To identify the smart city, it may be interesting to know the smart attributes or aspects of the smart city as given below.

- Networked Infrastructure
  IoT Sensors and Devices: Widespread installation of sensors and gadgets that gather information on a range of urban life-related topics, including energy use, traffic flow, water use, and air quality.
- Data-Informed Decision Making
  Big Data Analytics: Processing and analyzing enormous volumes of data gathered from IoT devices using sophisticated analytics to support policy creation and decision-making. Real-Time Monitoring: Constant observation of municipal systems to deliver timely insights and answers to new problems.

- Ecological Methods
  Integration of Renewable Energy: Using IoT technologies to facilitate the integration of renewable energy sources, such wind and solar, for maximum efficiency. Waste management: Sensing bins that keep track of waste quantities and plan the best routes for pickup.
- Increased Security for the Public
  Smart cameras and sensors are used in surveillance systems to improve security and deter crime. Emergency Response: Internet of Things-enabled solutions that boost emergency response times and effectiveness.
- Participation of Citizens and Services
  E-Government: Online gateways to government services, bill payment, and local government participation. Mobile applications: Programs that let users access municipal services, report problems, and get city updates.
- Wellness and Health
  Healthcare Monitoring: Internet of Things (IoT)-enabled health monitoring systems that track public health parameters and offer remote healthcare services. Environmental Monitoring: Sensors to keep an eye on noise levels, air quality, and other aspects of the environment that may have an influence on health.
  The readers may notice that the simple meaning of being Smart City is that the preliminary services of the city are well connected and continuously monitored using various smart devices and the sharing of the information within the system and eventually to the end users for the improvement, maintenance and safety of the people.
  It is interesting to note important devices that are essential for the transformation of a conventional city and its system into a Smart City.

3.4 Smart Infrastructures

3.4.1 Traffic Monitoring Camera and IoT Enabled City Street Lights

These sophisticated cameras do more than just record footage. Because of their artificial intelligence, they can identify jaywalkers, accidents, and traffic jams. Authorities can increase road safety, efficiently deploy emergency services, and manage traffic flow with the use of this data. Similarly, the lights, in contrast to traditional streetlights, have sensors that allow them to change brightness in response to many real-time variables, such as the time of day, the outside temperature, and pedestrian activity. This enhances public safety in dimly lit places while simultaneously saving electricity.

3.4.2 Wi-Fi Facility and Connectivity for Population

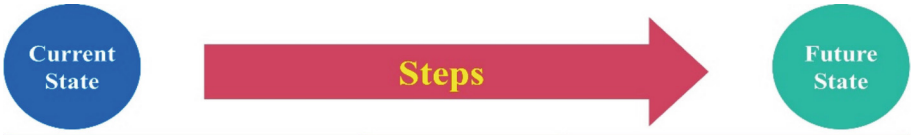
Reliability and widespread access to Wi-Fi are characteristics of smart cities. It promotes innovation, information empowerment for locals, and communication. Public Wi-Fi can be used for a variety of purposes, such as remaining informed, networking with businesses, and gaining access to local services.

3.4.3 Service Apps for Mobile and Other Devices

There are several small but important activities that can be controlled and utilized by developing mobile applications. For example, food ordering and delivery via mobile

apps is no longer a new concept. Finding a parking space in advance before you reach the location is possible using existing web-based applications. Nearly all local taxes, bills and payments within the corporation limits of a city can be done using official Apps. Even the grocery stores, cloth stores, and any similar retail shops can have their own App that can be installed in the device of any end user and buying, returning, payment or any activities can be done remotely using such facilities. These all when combined with each other can generate a network of smart systems and thus the city may be referred to as Smart City.

In summary, the smart term adds several devices, systems and their interconnection and connection to the internet to perform the intended tasks with accuracy, efficiency and speed. In the overall smart system for a city or home, the data collection, data processing and data management remains the major concerns. By transforming the conventional city's infrastructure into smart infrastructure the overall safety, monitoring and services can be improved. To further simply the clear advantages of transforming the conventional city into smart infrastructure city, the following chart of Fig. 5 may be referred to. It clearly illustrates the positive impact of smart features of a system over the conventional.



Conventional Systems in a City	Comparison Attribute	Smart System in a City
Static traffic lights and manual control cause peak-hour congestion and inefficiency.	Traffic Management	Optimizes traffic flow, reduces congestion, and improves transportation efficiency using real-time data and adaptive traffic signals.
Fixed energy systems waste and use more without real-time adjustment.	Energy Consumption	Using smart grids and sensors to dynamically modify energy use, improves efficiency and reduces waste across sectors.
Unintegrated surveillance and emergency response systems reduce reaction times and fragment safety management.	Public Safety	Increases public safety by using advanced monitoring, predictive analytics, and real-time communication to coordinate responses.
Bins might overfill and pickups missed due to fixed, inefficient waste collection schedules.	Waste Management	Optimizes waste collection routes and schedules using sensors and data analytics to save money and the environment.
Government services need physical visits or phone conversations, slowing input and resolution.	Citizen Engagement	Digital channels allowing citizens to report issues, access services, and provide feedback speed up answers and boost involvement.

Fig. 5. Comparison of Conventional and smart city systems

3.5 Smart Monitoring of Infrastructures: Example of Structural Health Monitoring of Bridges in Cities

3.5.1 Need of Health Monitoring of Bridges

Structures like bridges, flyovers, multistoried buildings and reactors are some examples of the important structures. Nearly all these are likely to have structural vibrations within their structural systems either whole or in a part. The source of the vibrations may be internal or external. However, the members are subjected to experience these vibrations according to their material properties, design strength, placement and orientation with other requirements. The vibration causes long term slow paced adverse effects on the performance or precise durability of the members. A timely check and continuous monitoring illustrated in Fig. 6 are one of the best options. Nowadays, IoT enabled monitoring has emerged as one of the most efficient solutions for such maintenance and protection techniques. Let us discuss them.

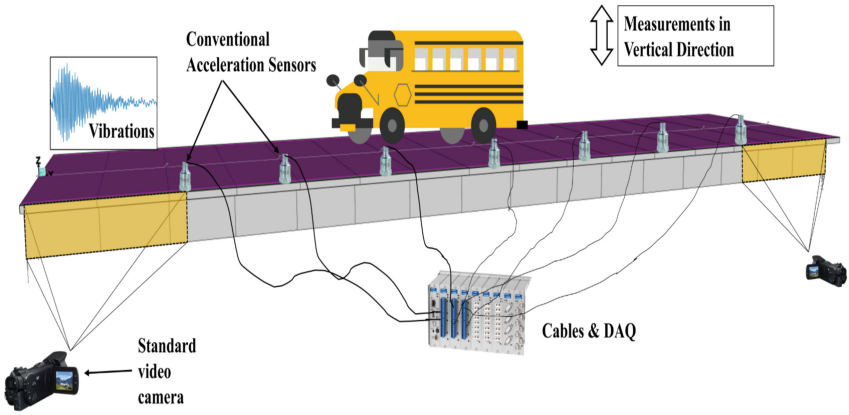


Fig. 6. Illustration of use of sensors in vibration monitoring of a bridge [22]

3.5.2 Attachment and Locating Data Receiving Devices in Bridges

Let us assume that on a concrete bridge made of Precast panels and fixed on the pillars are required to be monitored for vibration analysis. The most important attribute is where and how to fix the receiving points. The receiving points are in the form of displacement sensors, accelerators, and strain gauges. This list may vary depending on other requirements of data collection. Within this, there may be wired, or wireless components also attached to the data acquisition system. Generally, the recording points on a bridge can be affixed on the surface keeping following points in the consideration.

- Location of sensors: Put sensors in key locations like the middle of the bridge, the supports, and the joints for complete vibration data collection.
- Range and capacity: Choose sensors that appropriately measure the expected vibration levels and frequencies considering their sensitivity and their range

- Environmental resistance: Select sensors that can endure tough conditions such as temperature variations, moisture, rust among others if one is to bridge towards this aspect on durability alongside environmental resistance
- Data transmission: Select sensors with good wireless communication capabilities that guarantee instant and real-time data transmission to IoT framework ports.
- Supply of power: the sensor and devices should be provided with a continuous power supply, including the possibilities of cutting off the supply. The alternate arrangement of the power supply must be maintained at the site.
- Easy installation: The placing and installation of the sensors should be easy for the people involved.

As now we have observed, that for sensors to work effectively and efficiently, selecting the right location is very important. The decision to locate sensors is one of the most important attributes of the overall process of data collection and remote monitoring of any structure.

3.5.3 Data Collection, Transfer, Analysis, and Processing

Upon installing the sensors at the right locations, the next step is to collect the data. The data can be collected in varying forms and formats as per the requirements. However, when the data type is being selected, the corresponding sensor type should be selected and installed. We will briefly describe what are the possible data collection in such activity. Following is some of the examples of type of data may be collected from a typical structural health monitoring site.

- Data on strain: Indicates deformation brought on by loads and other stressors.
- Displacement data: Monitors how structural elements move.
- Acceleration data records dynamic reactions and vibrations.
- Temperature data: Tracks the impact of heat on the structure and the surrounding environment.
- Tracks the amount of moisture in the air, which might have an impact on structural materials.
- Load data: Quantifies the loads and applied forces acting on the structure.
- Crack Detection Data: Tracks and identifies the development of cracks.
- Corrosion Data: Evaluates the degree of metal component corrosion.
- Acoustic Emission Data: Tracks sound waves produced by fracture development or other alterations.
- Images and videos from manual inspections are included in the visual inspection data.
- Measures the inclination or angular displacement of structural components using inclination data.
- Pressure Data: Keeps track of both internal and external pressures, particularly in structures that hold fluid.
- Deflection data: Monitors how beams and other components bend or curve.
- Ground Movement Data: Tracks changes in the foundation and soil.
- Data on Wind Direction and Speed: Documents the ways in which the structure is affected by wind.
- Data from Ultrasonic Testing: Ultrasonic waves are used to find interior defects.

3.6 Advanced Methods, Technologies and Tools in Smart Monitoring Framework for Infrastructural Facilities

With the rapid growth of smart computing technologies, the concept of smart monitoring of structures, buildings, and similar infrastructure has been evolving with a faster pace. The following are some of the examples of the advanced tools and systems required to fulfill the futuristic smart computing requirements for smart cities and infrastructure.

- Drones and Aerial Imaging

  Drones that are supplied with thermal imaging and high-resolution cameras are being utilized more frequently to inspect large infrastructures, such as pipelines and power lines. They can rapidly cover extensive areas and access difficult-to-reach locations, thereby generating detailed imagery for analysis. For example, a utility company may employ drones to detect vegetation encroachment on power lines, thereby enabling the implementation of opportune interventions. To facilitate spatial analysis and decision-making, the data collected can be incorporated into Geographic Information Systems (GIS). This approach not only enhances safety but also enhances efficiency by reducing the necessity for manual inspections in hazardous environments.
- Building Information Modeling (BIM)

  A digital representation of the physical and functional characteristics of facilities is known as BIM. It allows for the comprehensive monitoring and administration of a facility's lifecycle. Various scenarios, including structural stress under varying conditions, can be simulated by engineers using BIM. For instance, a construction company may employ BIM to anticipate the seismic response of a new building, thereby enabling alterations to the design. Predictive maintenance and resource allocation are improved through the incorporation of real-time data feeds into BIM systems, which ultimately optimizes operational efficiency and reduces costs in facility management.
- Mobile Monitoring Applications

  Mobile apps make it easier to keep an eye on infrastructure in real time by letting workers collect and study data on the spot. For example, technicians could use an app to report problems with the state of a highway, such as holes and cracks, and this information could be sent right away to a central management system. This way of doing things makes communication better and lets problems be fixed quickly as they come up. Mobile monitoring apps make repair work more efficient by making it easier to collect data and send reports. They also make sure that the infrastructure stays safe and working.
- Smart Grids

  Advanced metering and monitoring tools included into smart grid technology help to improve energy distribution and consumption control. Smart meters, for instance, offer real-time data on energy consumption, which lets utilities maximize the distribution of their resources and lower interruptions. Utilities can forecast demand spikes and change supply by including IoT and machine learning. By means of energy consumption data, this technology also promotes customer engagement by means of conservation. Smart grids finally improve the efficiency and resilience of the energy infrastructure, so supporting sustainable energy management techniques.

- Blockchain for Data Integrity

  The integrity and security of data collected from a variety of monitoring systems are guaranteed by blockchain technology. For instance, blockchain technology can monitor the compliance of materials with safety standards throughout the duration of a construction project. This technology establishes an immutable ledger that can be audited and verified, thereby guaranteeing transparency in infrastructure management. Blockchain can assist in reducing the risks associated with data tampering or loss by offering a secure method for data sharing among stakeholders. This promotes trust and collaboration in infrastructure initiatives, thereby improving overall governance and accountability.

Above are a few of the examples encompassing the direction of the further development of the tools, technologies, and systems in the areas of the smart computing frameworks. The readers are advised to refer to the specific requirements of each smart computing framework applied for the monitoring of a given entity including smart home, smart city and its infrastructure, or the relevant services. These are the areas that require detailed exploration and study.

4 Challenges in Smart Transformation of the Conventional Systems

With the rapid growth of digital transformation of the conventional systems, the tools and devices are also emerging to meet the needs of smart monitoring and smart computing mechanisms. However, the task is still difficult to achieve completely in several aspects. There are limitations of more than a single attribute of a system. Following is some of the major challenges impacting the complete and satisfactory transformation of the conventional systems into smart and digitally advanced form.

- Integrating intelligent computing technologies into existing infrastructure is a technically challenging task. Typically, it requires seamless integration of new systems with existing ones, relying on significant redesigns and specific solutions to ensure compatibility and functionality.
- Smart gadgets, because to their generation of substantial amounts of data, are attractive targets for attackers. Robust data security and encryption measures are essential for safeguarding sensitive information and maintaining the integrity of systems.
- Adapting smart technology to meet increasing infrastructure demands could pose a challenge. Efficient measures need to be implemented to manage the increasing demands of users and data volumes while maintaining optimal performance and minimizing costly operational expenses.
- Implementation Cost: The initial investments for smart technology might be rather substantial. The costs mentioned include not only the technology itself, but also the expenses related to installation, training, and ongoing maintenance. These additional costs can act as a deterrent for organizations considering the pursuit of creative ideas.
- Ensuring the reasonable compatibility of several smart systems and devices can be a challenging task due to the need for seamless linking. The establishment of standardized interfaces and protocols is crucial for both efficient integration among several infrastructure components and effective communication.

- Utilizing and managing smart technology requires a specific level of comprehension. The scarcity of competent individuals possessing the requisite expertise to develop, deploy, and sustain these sophisticated systems gives rise to potential obstacles in their implementation.
- Adhering to rules and standards might provide challenges and vary based on the geographical location. Adhering to industry-specific regulations, safety standards, and data security protocols ensures that the technology meets all necessary requirements and helps prevent legal complications.
- Uninformed or perceived as resistant stakeholders and end users may raise objections to the adoption of new technology. Effective change management strategies and comprehensive user training are crucial factors in determining whether one obtains permission and ensures smooth transitions.
- Continuous maintenance and periodic upgrades are necessary to ensure the uninterrupted operation of intelligent systems. This necessitates anticipating future requirements and emphasizing probable technical advancements to avoid becoming outdated and ensure consistent reliability.
- Oftentimes, intelligent systems collect vast amounts of operational and personal data. Addressing privacy concerns necessitates the use of transparent data collection methods and empowering consumers to have control over their data, thereby fostering trust and compliance with privacy standards.
- Data scalability: Smart city initiatives often suffer with scalability because of different urban infrastructure and different population densities. Using solutions that apply across several spheres might be challenging. Maintaining performance and incorporating new technologies becomes more challenging as cities expand, hence perhaps resulting in unequal service quality and higher expenses.
- Interoperability: Interoperability problems develop when several systems and technologies fail to interact successfully. This can impede flawless departmental data sharing and coordination, therefore causing inefficiencies. Creating a coherent smart city ecosystem depends on standardizing protocols and guaranteeing interoperability among several platforms, which is both vital but difficult.
- Data privacy & security: Data privacy concerns among citizens are growing as they become more cautious of how personal data is gathered and applied in smart city programs. Distrust might result from questions about data security, surveillance, and possible use. Dealing with these problems calls for open policies, strong security guards, and active community involvement to guarantee respect of privacy.

Above are the major concerns and require attention by the innovators and technologists. The practical application of smart systems in any field finally depends on three major aspects namely Availability of technology, methods to manage and process collected data, and advance level of digital and internet services. However, the world has been changing with a rapid rate of transformation with all such attributes and may be the humanity can maximize the smart transformation in most of the fields soon.

5 Conclusion and Recommendations

The article discussed some of the significant parameters pertaining to the smart transformation of conventional systems namely smart cities, smart homes, smart structural health monitoring with respect to the need of such transformations for the future. The article also addresses technical details on the general framework, data management, and smart monitoring attributes with examples. The following are the recommendations and conclusions derived from the study.

- Smart computing technologies greatly increase municipal efficiency by means of simplicity of city processes and resource management. Smart cities can lower running costs, improve infrastructure performance, and improve resident quality of living by using improved data management and monitoring tools.
- Smart computing technology's full, real-time data allows city managers and designers to lead better informed judgments. The urban settings get more resilient and sustainable by means of proactive responses to urban challenges made feasible by the data management methodologies and structure given in this study
- Constant development in smart computers and data analytics will revolutionize metropolitan infrastructure even more. Fixing integration issues, improving cybersecurity requirements, and broadening the emphasis of smart monitoring to match changing urban needs and new technologies should be the main priorities of future study.
- While employing smart computing has many advantages, technological complexity and significant implementation costs are among the drawbacks. Solving these challenges will open fresh chances for the development of smart cities and improvement of urban infrastructure by means of innovative ideas and cross-sector cooperation.
- Use of the appropriate framework for smart computing systems offers management of urban infrastructure a complete approach. Smart monitoring tools and good data management techniques can help cities build more flexible and strong infrastructure that will open the path for improved and more environmentally friendly urban living surroundings.

References

1. Khanh, Q.V., et al.: An efficient edge computing management mechanism for sustainable smart cities. Sustain. Comput. Inform. Syst. **38**, 100867 (2023)
2. Walczak, R., et al.: Acceptance of IoT edge-computing-based sensors in smart cities for universal design purposes. Energies **16**(3), 1024 (2023)
3. Rao, P.M., Deebak, B.D.: Security and privacy issues in smart cities/industries: technologies, applications, and challenges. J. Ambient Intell. Humaniz. Comput. **14**(8), 10517–10553 (2023)
4. Mondal, D., et al.: Applications, modern trends, and challenges of multiscale modeling in smart cities. In: Handbook of Research on Data-Driven Mathematical Modeling in Smart Cities, pp. 1–23 (2023)
5. Nassereddine, M., Khang, A.: Applications of Internet of Things (IoT) in smart cities. In: Advanced IoT Technologies and Applications in the Industry 4.0 Digital Economy, pp. 109–136. CRC Press (2024)

6. Hashem, I.A.T., et al.: Urban computing for sustainable smart cities: recent advances, taxonomy, and open research challenges. Sustainability **15**(5), 3916 (2023)
7. Hui, C.X., et al.: Greening smart cities: an investigation of the integration of urban natural resources and smart city technologies for promoting environmental sustainability. Sustain. Cities Soc. **99**, 104985 (2023)
8. Jebaraj, L., Khang, A., Chandrasekar, V., Pravin, A.R., Sriram, K.: Smart city: concepts, models, technologies and applications. In: Khang, A., Gupta, S.K., Rani, S., Karras, D.A. (eds.) Smart Cities: IoT Technologies, Big Data Solutions, Cloud Platforms, and Cybersecurity Techniques, pp. 1–20. CRC Press, Boca Raton (2023). https://doi.org/10.1201/9781003376064-1
9. Podder, S.K., Samanta, D., Etemi, B.P.: Impact of internet of things (IoT) applications on HR analytics and sustainable business practices in smart city. Meas. Sensors **35**, 101296 (2024). https://doi.org/10.1016/j.measen.2024.101296
10. Marques, P., et al.: An IoT-based smart cities infrastructure architecture applied to a waste management scenario. Ad Hoc Netw. **87**, 200–208 (2019)
11. Ilyas, M.: IoT applications in smart cities. In: 2021 International Conference on Electronic Communications, Internet of Things and Big Data (ICEIB). IEEE (2021)
12. Derawi, M., Dalveren, Y., Cheikh, F.A.: Internet-of-things-based smart transportation systems for safer roads. In: 2020 IEEE 6th World Forum on Internet of Things (WF-IoT). IEEE (2020)
13. Kim, S.M., et al.: Intelligent energy IoT data monitoring system for smart city infrastructure optimization. In: 2019 IEEE 2nd International Conference on Electronics Technology (ICET). IEEE (2019)
14. Sharma, P., et al.: Exploring the implications of IoT integration in urban infrastructures for sustainable smart cities. In: 2023 International Conference on Data Science, Agents & Artificial Intelligence (ICDSAAI). IEEE (2023)
15. Inibhunu, C., Carolyn McGregor, A.M.: A privacy preserving framework for smart cities utilising IoT, smart buildings and big data. In: 2020 IEEE 22nd International Conference on High Performance Computing and Communications; IEEE 18th International Conference on Smart City; IEEE 6th International Conference on Data Science and Systems (HPCC/SmartCity/DSS). IEEE (2020)
16. Abdulsattar, N.F., et al.: An investigation study for technologies, challenges and practices of IoT in smart cities. In: 2022 5th International Conference on Engineering Technology and its Applications (IICETA). IEEE (2022)
17. Sadhukhan, P.: An IoT based framework for smart city services. In: 2018 International Conference on Communication, Computing and Internet of Things (IC3IoT). IEEE (2018)
18. Saroliya, A., Anand, P., Das, L.: Smart cities: an integrated framework using IoT. In: 2023 3rd International Conference on Technological Advancements in Computational Sciences (ICTACS). IEEE (2023)
19. https://www.hiotron.com/iot-in-construction/. Accessed 9 Aug 2024
20. https://journals.openedition.org/factsreports/4367. Accessed 9 Aug 2024
21. Yang, H., Lee, W., Lee, H.: IoT smart home adoption: the importance of proper level automation. J. Sens. **2018**(1), 6464036 (2018)
22. Gonen, S., Erduran, E.: A hybrid method for vibration-based bridge damage detection. Remote Sens. **14**(23), 6054 (2022)
23. Amade, B., Nwakanma, C.I.: Identifying challenges of internet of things on construction projects using fuzzy approach. J. Eng. Proj. Prod. Manage. **11**(3) (2021)

Decentralized EHR Exchange in Healthcare: Enhancing Privacy and Security with Blockchain and Cryptographic Techniques

Aqeel A. Yaseen[1(✉)], Kalyani Patel[2], Abdulla J. Y. Aldarwish[1,3], and Ali A. Yassin[3]

[1] Department of Computer Science, Gujarat University, Ahmedabad 380009, India
lec.aay.ali@uobasrah.edu.iq
[2] K.S. School of Business Management and Information Technology, Gujarat University, Ahmedabad, India
[3] Department of Computer Sciences, Education College for Pure Sciences, University of Basrah, 6100 Basrah, Iraq

Abstract. In today's digital landscape, robust security protocols are essential. Despite significant advancements in information security, cyber threats continue to evolve and increase. Although advanced protection systems such as multi-factor authentication and encryption have been implemented, data remains vulnerable to breaches. In this context, blockchain technology emerges as one of the most promising technologies that can contribute to enhancing data security, thanks to its distinctive features such as decentralization and transparency.

This paper explores the integration of blockchain technology and Elliptic Curve Cryptography (ECC) to address privacy, security, and scalability challenges in healthcare data exchange. By leveraging blockchain's decentralized, immutable structure along with ECC for secure key management, the proposed system offers a secure platform for storing and verifying Electronic Health Records (EHRs). Smart contracts are utilized to ensure transparent, tamper-proof authentication processes, which enhance user trust. This blockchain-based approach, strengthened by ECC, not only fortifies security and privacy but also eliminates single points of failure and reduces reliance on third-party verification. While eliminating third-party verification enhances security and reduces costs, it raises the need for robust governance mechanisms. A decentralized governance model, potentially through the use of smart contracts or Decentralized Autonomous Organizations (DAOs), can ensure continuous oversight and compliance with healthcare regulations. The proposed system leverages a decentralized governance model using smart contracts or DAOs to enhance compliance and oversight. Designed for scalability and efficiency, the system facilitates secure EHR exchange across multi-cloud environments. The findings suggest that blockchain, integrated with ECC, provides a promising solution for secure, scalable EHR management in healthcare.

Keywords: Healthcare · EHR · EEC · Authentication · Blockchain

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
S. Rajagopal et al. (Eds.): ASCIS 2024, CCIS 2428, pp. 235–246, 2025.
https://doi.org/10.1007/978-3-031-86302-8_15

1 Introduction

The significance of healthcare has increased recently; thus, electronic healthcare records (EHR) play an essential role in enhancing the healthcare industry. However, protecting EHRs is now a major concern due to ever-rising cases of hacker attacks. Implementing a high level of security for EHRs is critical when storing and managing patient information to reduce the risk of data breaches, preserve patients' privacy rights, and meet legal requirements. Effectively implemented and secured EHR systems promote confidence and expand data sharing between physicians to deliver the best results for the patients [1–3]. To maintain the privacy and integrity of healthcare information and EHRs different measures like encryption, access controls, multi-factor authentication, biometric verification, etc. are used [4]. Fingerprint and facial recognition are other security measures that can be employed in the process [5]. Regular security audits and compliance with standards like HIPAA ensure ongoing protection [6]. Implementing these measures helps safeguard sensitive patient data from breaches and unauthorized access, maintaining the confidentiality and integrity of healthcare records and ensuring trust in digital healthcare systems. However, implementing these security techniques presents several challenges. The integration of encryption, access controls, multi-factor authentication, and biometric verification poses a challenge to any company. These methods commonly necessitate considerable computational inputs, hence raising the overall computational complexity [7]. Furthermore, due to regular security audits and compliance checks that have to be carried out, there is an extra cost to communication [8]. At the same time, it is difficult to achieve both high levels of security and performance because healthcare providers must make sure that such measures will not affect system performance and availability. Therefore, the balancing between the complexity of robust security and communication cost represents one of the biggest challenges [9]. In this context, the application of blockchain technology is significant in enhancing healthcare security particularly of EHRs through decentralized and unalterable ledger. This also makes the data more accurate and transparent, as every transaction is recorded and cannot be changed afterward. Blockchain also enhances patient privacy by allowing only authorized access, thereby reducing the risk of data breaches and ensuring the confidentiality of sensitive healthcare information [10].

From this perspective, we propose a blockchain-based scheme to enhance the management of electronic health records (EHR) shared among various healthcare institutions. Blockchain can significantly improve data security, integrity, and transparency in these exchanges. The proposed scheme is divided into four main phases: the setting phase, the registration phase, the login and authentication phase, and the EHR exchange phase. Each phase incorporates robust security mechanisms to protect sensitive health information and ensure only authorized access. During the setting phase, we utilize Elliptic Curve Cryptography (*ECC*) to generate and distribute cryptographic keys, which are essential for secure communication and data encryption. The registration phase involves securely registering entities such as patients, doctors, and administrators with their respective healthcare provider institutions (*HPI*), using secure channels to transmit identification and authentication information. The login and authentication phase employs advanced cryptographic techniques to verify user identities and grant access based on predefined privileges. The final phase, *EHR* exchange, securely migrate health records between *HPI*

by storing verification data like public keys and identities on the blockchain. This approach enables efficient authenticity checks, reduces communication costs, and enhances the reliability of user identity verification. Our proposed approach contributes to the following:

1. **Blockchain Integration:** the blockchain-based proposed scheme to secure EHR exchanges, ensuring data immutability, and transparency, and reducing risks of unauthorized access and breaches.
2. **Elliptic Curve Cryptography (ECC):** ECC is employed for key generation and encryption, ensuring that intercepted data remains secure and unreadable without the correct decryption key.
3. **Data Integrity and Privacy:** the decentralized and cryptographic techniques of blockchain technology the proposed scheme ensures healthcare data remains unaltered and private during exchanges between different healthcare providers.
4. **Efficiency and Cost Reduction:** The scheme does not need to connect a trusted third party to issue signatures or certificates for verifications. So, it reduces communication costs and enhances identity verification reliability, offering a more efficient alternative to traditional centralized systems.
5. **Scalability:** By utilizing a multi-cloud environment with blockchain, the system remains scalable, accommodating new patients and providers without compromising security or performance.

Finally, the rest sections of this work are organized into Sect. 2, which presents a related work of the relevant approaches. Section 3 emphasizes the development of the proposed scheme. Section 4, which is the security analysis and compassion the relevant approaches. Section 5, addresses this work's conclusion.

2 Related Work

The concept of decentralized cryptocurrency exchanges has undergone significant evolution over the years. S. Nakamoto presents an introduction to Bitcoin marked the first practical application of blockchain technology, enabling users to conduct transactions and transfer bitcoins without relying on a trusted intermediary [11]. F. Xhafa et al., expanded on the ABE scheme by introducing the concept of a Private Key Generator (PKG), which, however, raised potential risks as the PKG could decrypt all data stored on the cloud server, leading to privacy breaches and key misuse [12]. Furthermore, Z. Wang and Y. Zhu addressed a respectable traditional centralized cloud storage model but it is vulnerable to a single point of failure, which could cause catastrophic system failures [13]. Fully Homomorphic Encryption (FHE) has been explored to enhance security by allowing computations on encrypted data to mitigate such risks [14]. D. K. Murala et al. categorized the different cryptographic techniques to be adopted in cloud infrastructure and highlighted the use of a trusted third party (TTP) for the services of authentication and auditing. These services assist in avoiding such misfortunes that involve the leakage of delicate information in cloud situations. But still, trust stands as a problem, especially when the clients physically do not know where their data is being stored, which is a big problem when a breach in the data storage facility is made. In

response to these challenges, blockchain technology is considered as a solution among the authorized entities securely transferring the Electronic Health Record (EHR) data. The basic features of Blockchain, a distributed ledger system, make the transactional data in the health sector more secure, by storing the health transaction data in the form of hash values and so cannot be altered during the transmission or storage phase [15]. A. Chandini and P. I. Basarkod, proposed the elaboration of the multi-faceted, scalable, and patient-centric blockchain to resist the patient's data from attacks. Their system's stakeholders need to undergo pre-transaction signature verification to accommodate several stakeholders. Built on a cloud-based InterPlanetary File System (IPFS) integrated with Ethereum smart contracts, the system ensures decentralized data storage with an enhanced security gateway component is used to verify users' signatures, preventing unauthorized access. After verification, the Blockchain Manager (BCM) engages the smart contract to authenticate user identity and enforce role, rights, and access (RRA) regulations. This allows users, including patients and physicians, to upload, view, and share data according to RRA rules. The system's design eliminates the risk of a single point of failure, making it suitable for secure, real-time healthcare applications [16].

Midha et al. introduced a secure multi-factor authentication protocol for healthcare services utilizing cloud-based SDN, designed to enhance data security and prevent spoofing through BAN logic and SPAN-AVISPA tools. The robustness of this protocol is demonstrated by its ability to handle multiple security challenges effectively. However, the work has weaknesses, particularly in terms of increased latency and potential scalability issues due to the complexity of integrating cloud-based SDN, which could impact real-time performance [17].

I. Wanisha et al., demonstrate a robust scheme utilizing blockchain and smart contracts, that provides decentralized, immutable, and automated authentication processes that enhance security and privacy. However, it faces potential weaknesses related to scalability and computational overhead, particularly in resource-constrained environments like IoT. The complexity of implementing and maintaining blockchain infrastructure could also introduce vulnerabilities and impact system performance in large-scale deployments [18].

While blockchain presents a robust solution for healthcare data security, other approaches have also been explored. Centralized security models, such as traditional server-based encryption systems, offer simplicity but are vulnerable to single points of failure. Homomorphic encryption allows computations on encrypted data, providing privacy-preserving analytics, but it is computationally intensive and difficult to scale. Hybrid systems that combine blockchain with centralized storage aim to balance security with speed, yet they may still suffer from the weaknesses of the centralized component. Our proposed system, leveraging blockchain's decentralized nature and ECC for encryption, addresses the limitations of these alternative approaches by enhancing both security and scalability.

3 Proposed Scheme

In this section, our proposed scheme addresses a secure EHR exchange among different healthcare institutions based on blockchain technology. Our proposed scheme is partitioned into four phases respectively: The setting phase, the register phase, the login and

authentication phase, and the EHR exchange. The proposed scheme consists of main preliminaries; each healthcare institution has its own entities, the entity could be a healthcare provider institution (*HPI*), Patient (P), User (*U*), Doctor (*Dr*), Admin (*ADM*). Furthermore, we denoted for each entity by (*E*). Moreover, the blockchain network (*BC*). To achieve this, the previously mentioned entities and their corresponding public keys should be represented by the symbol (i) e.g., E_i. The decentralized management provided by blockchain and the maintenance of data integrity and privacy are features that ensure data immutability and prevent alterations by unauthorized individuals. Therefore, we have implemented it in our proposed system, which will reduce the cost of communication and user identity verification among participating institutions. All entities and their necessary data will be registered on the blockchain network. Figure 1 shows simply our proposed scheme entities connected to the blockchain.

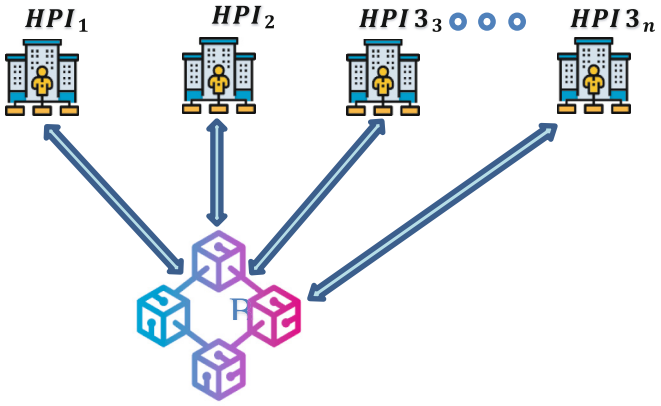


Fig. 1. Proposed Scheme Blockchain-based

3.1 Setting Phase

In our proposed scheme, we pay special attention to the initialization phase, which is a crucial and sensitive stage. We rely on ECC for the generation of the secret key and the public key. While ECC offers security with smaller key sizes, minimizing the computational burden, the overall blockchain system can still introduce overhead due to the need for continuous verification of transactions. The equipped healthcare institution will perform the following tasks:

– Keys Generation and distribution: Each entity is provided its secret key (SK), public key (*PK*), and shared key (*SHK*) which is used to enable the symmetric encryption/decryption. The *HPI* itself will also hold its own (*SK*, *PK*, *SHK*).

  ECC: enhances key management by offering secure encryption with smaller key sizes compared to traditional algorithms like RSA. The use of ECC ensures data security during exchanges across healthcare providers, protecting sensitive EHRs from unauthorized access through efficient encryption and decryption processes.

– Setup the necessary parameters: hash function such as SH-256 and preparation for symmetric encryption operations such as modern encryption.

3.2 Registration Phase

The entities (Patient, User, Doctor, Admin) which denoted by ($\boldsymbol{E_i}$) must register in the exact healthcare provider institution (*HPI*). Furthermore, each entity that wishes to register to its *HPI* must send crucial information such as entity identification ($ID_{\boldsymbol{E_i}}$), entity password ($PW_{\boldsymbol{E_i}}$). In addition, the necessary information like entity address ($Add_{\boldsymbol{E_i}}$), entity email ($Email_{\boldsymbol{E_i}}$). The following steps explain how the *HPI* achieves this phase:

1. E_i sent the tuple (ID_{E_i}, PW_{E_i}, Add_{E_i}, $Email_{E_i}$) to *HPI* utilizing a secure channel
2. $HPI_i \rightarrow\ PW'_{E_i} =$ hash ($ID_{E_i}||PW_{E_i}$), $ID'_{E_i} =$ hash (ID_{E_i}), anonymity calculation
3. *HPI* checks the existence of E_i; if so, terminate the request, else move to step 4,
4. *HPI* checks the existence of E_i in BC using smart contract; if so, terminate the request, else *HPI* registers E_i with all required information.
5. *HPI* embeds the entity's ID and public key in BC ($hash\big(ID_{E_i}\big)$, $hash(E_{PK})$).
6. Updates the *EHR* information of the entity E_i.
7. Sends tuple (E_{SK}, E_{PK}, E_{SHK}) to E_i utilizing a secure channel such as SSL.

3.3 Login and Authentication Phase

The entity (E_i) logs into the proposed system by entering its identity and password, and the system verifies its identity to allow access. Our proposed scheme performs the following steps to manage the secure login process for the system:

1. $E_i \rightarrow HPI_i$
   a. E_i logins the system using ID_{E_i}, PW_{E_i}.
   b. Generating $r_i \in Z_n^*$,
   c. Calculating request $H_{E_i} = hash(hash(\ ID_{E_i})\ ||hash(\ ID_{E_i}||PW_{E_i})||hash\ (r_i))$
   d. $E_n = Enc_{E_{\mathrm{SHK}}}(r_i, E_{\mathrm{SHK}})$, encrypting r_i using E_{SHK} the entity shared key.
   e. Sending $R_{E_i}\big\{H_{E_i}, E_n\big\}$ to its *HPI*.
2. $HPI_i \rightarrow E_i$: upon receiving R_{E_i} *HPI* performs the below:
   a. $r'_i = Dec_{E_{\mathrm{SHK}}}(E_n)$, restores r_i by decrypting r_i E_n using E_{SHK}
   b. HPI_i fetches pre-stored parameters to conduct a comparative analysis
   c. Calculating $H'_{E_i} = hash\ (ID'_{E_i}||PW'_{E_i}||hash\ (r'_i))$
   d. Comparing (H'_{E_i}, H_{E_i}) if matching, HPI_i sends confirmation code (CC) as a challenge to (E_i).
3. E_i receives (CC') from HPI_i then calculates $T = hash\ (hash\big(CC'\big)||hash(r_i))$
4. HPI_i receives (T) from E_i then calculates $T' = hash\ (hash(CC)||hash\big(r'_i\big))$, then comparing $\big(T', T\big)$ if matching, HPI_i verifies and confirms the login as well as the authentication then grants access with privilege according to the type of (E_i). Upon a mismatch, HPI_i will deny system access.

3.4 EHR Exchanging Phase

In the phase of migrating the health record of the entity (E_i), the information is already stored in the institution to which entity (E_i) belongs. Meanwhile, the necessary information for verification and retrieval, such as the public key and identity, is stored on

the blockchain network. Suppose entity (E_i), previously registered with an institution (HPI_i), seeks treatment at an institution (HPI_k). Institution (HPI_k) requests the EHR of the entity (EHR_{E_i}), due to the absence of this record within its own system. The following steps illustrate how our proposed scheme manages this phase:

1. HPI_k checks the existence of (E_i) in BC ID_{E_i}, using smart contracts.
2. HPI_k sends hashed ID ($hash$ (E_i)) to HPI_i which was restored from BC.
3. HPI_i retrieves ID'_{E_i}, comparing $hash$ (E_i) with ID'_{E_i}.
4. After checking the authorization of E_i, HPI_i sends EHR_{E_i} to HPI_k

Impact of Blockchain's Latency: One of the challenges inherent to blockchain technology is latency, particularly during real-time data exchanges. In healthcare, where time-sensitive information is critical, latency introduced by blockchain consensus mechanisms could pose significant issues. This is could be addressed by employs off-chain solutions for real-time data processing, ensuring that urgent healthcare data is not delayed by blockchain operations. Blockchain is primarily used for secure verification and data storage, while off-chain systems manage rapid data exchanges. This approach strikes a balance between security and speed, allowing healthcare professionals to access critical data without unnecessary delays.

Steps of HER Exchange phase show smoothly in Fig. 2.

3.5 System Consideration and Practical Insight

While the scheme addresses key challenges in securing healthcare data, practical considerations for scalability, efficiency, and integration must also be considered.

- **Scalability and Efficiency:** The system's design allows it to scale smoothly as the number of healthcare providers and patients increases. Blockchain technology and multi-cloud infrastructure ensure that the system can handle large volumes of data without performance degradation. Although initial implementation costs may be high, long-term benefits, such as reduced administrative overhead and enhanced security, outweigh the costs.
- **Integration with Existing Healthcare IT Systems:** One of the key practical challenges is integrating this system with existing IT infrastructures, especially legacy EMR systems. APIs and middleware can help facilitate seamless communication between blockchain and these systems, allowing healthcare institutions to adopt the new technology without disrupting existing workflows. This approach ensures the system's practical applicability in real-world healthcare environments.
- **User Adoption and Security Complexity:** Blockchain technology, while highly secure, can initially appear complex to end users. However, the proposed system leverages blockchain's security strengths without sacrificing usability. By balancing security and communication efficiency, the system ensures that healthcare professionals can adopt the solution without technical expertise in blockchain. Intuitive interfaces and APIs abstract the complexities of blockchain, enabling seamless interaction with the system while providing robust security for healthcare data exchanges.

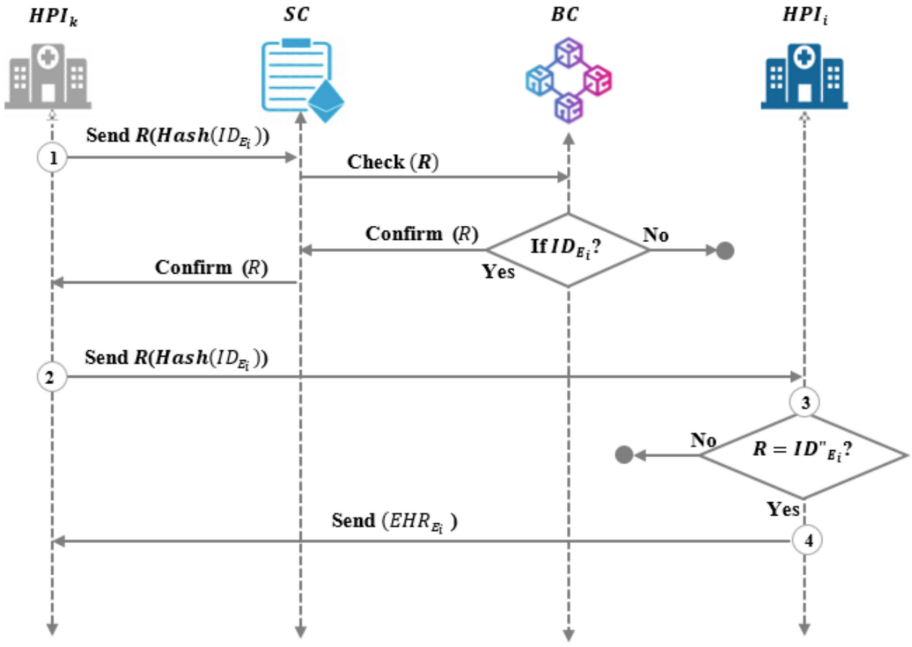


Fig. 2. EHR Exchange Phase

4 Security Analyzing

This section presents a preliminary evaluation of the proposed scheme and substantiates its feasibility depending on informal security analysis, we provide the theoretical evaluation of the proposed scheme and discuss possible threats and security threats concerning patient data. Table 1 highlights the verification process of the proposed scheme where its validity was ascertained on some essential security requirements. The proposed system effectively minimizes computational overhead by utilizing Elliptic Curve Cryptography (ECC) for efficient encryption, alongside less demanding consensus mechanisms like Proof of Stake (PoS). This balance ensures strong security while optimizing performance, making the system suitable for real-world healthcare applications without excessive computational costs.

Table 2 below displays a comparison of the relevant approaches recently. We addressed some security techniques which achieved. However, these techniques are: blockchain (BC), access control (AC), privacy-preserving (PP), migrating data and sharing among different multi servers of healthcare providers institutions (MS), data masking (DM).

4.1 Risk and Limitation

Blockchain immutability, while enhancing security, presents challenges in healthcare applications where erroneous data entries require correction. Solutions such as off-chain storage can mitigate this limitation. Solutions such as appending new records or using

Table 1. Protocol Verification Utilizing the Informal Security Analysis

Security Requirement	Protocol Verification with Blockchain Integration
Data Confidentiality (DC)	Messages are secured using a Hash Function, ECC asymmetric cryptography, and blockchain encryption. Blockchain's immutable ledger ensures that data cannot be altered post-recording, enhancing confidentiality
Mutual Authentication (MA)	The authentication phase is achieved from different entities' blockchain-based smart contracts, which provide a decentralized and tamper-proof verification process
User Anonymity (UA)	User identity is masked and recorded as hashed values on the blockchain, ensuring that anonymity is preserved while leveraging the blockchain's inherent resistance to identity tracing
Data Integrity (DI)	Blockchain's cryptographic linkage of blocks ensures that all data is protected from unauthorized modifications. Any changes to data require network-wide consensus, making tampering nearly impossible
Data Transparency	Blockchain provides a transparent ledger that allows all authorized participants to view and verify transaction histories, ensuring that all actions are traceable and fostering accountability and trust among stakeholders
Availability	The decentralized nature of blockchain enhances system availability by reducing the risks associated with centralized server failures. The network can continue to operate even if some nodes are compromised or under attack
Scalability (SC)	Blockchain's decentralized framework allows for scalable solutions without relying solely on central servers. However, scalability challenges such as transaction throughput and latency might still need to be managed effectively
DoS, Message Replay Attack (MRA)	Blockchain's consensus mechanism prevents replay attacks by ensuring that all transactions are unique and verified across the network. Any attempt to resend messages would be rejected by the decentralized ledger

(*continued*)

Table 1. (*continued*)

Security Requirement	Protocol Verification with Blockchain Integration
Impersonation Attack (IA)	The combination of blockchain-based smart contracts and cryptographic authentication mechanisms makes it significantly more challenging for attackers to impersonate legitimate users or steal identities
Man-in-the-Middle Attack (MITM)	Blockchain's use of cryptographic techniques and decentralized verification reduces the likelihood of successful MITM attacks, as attackers would need to compromise the entire network rather than a single communication channel
51% Attack	The proposed work enjoys the immutable according to the implementation of the private blockchain

Table 2. Required Techniques

Approaches	BC	AC	PP	MS	DM
Murala et al. [15]	×	✓	✓	×	×
Chandini et al. [16]	✓	×	✓	×	×
Midha et al. [17]	×	✓	✓	✓	×
Wanisha et al. [18]	✓	✓	✓	×	×
Proposed Scheme	✓	✓	✓	✓	✓

off-chain data for mutable information are essential to balance immutability with the need for flexibility.

4.2 Performance Comparison with Traditional Centralized Systems

To highlight the advantages of our proposed blockchain-based system, it is essential to compare its performance with traditional non-blockchain solutions commonly used in healthcare, such as centralized cloud-based systems.

- **Security**: Traditional centralized systems typically rely on encryption and trusted third parties (TTP) for data management, which introduces a single point of failure. In contrast, blockchain provides decentralized security, where no single entity holds control, reducing the risk of data breaches. Blockchain's immutability ensures that once data is recorded, it cannot be altered, providing superior data integrity.
- **Efficiency**: While centralized systems may offer faster data access due to their reliance on powerful servers, they lack the transparency and auditability of blockchain. Traditional healthcare IT systems rely on centralized data storage, offering faster access

but increased vulnerability to single points of failure. Blockchain technology introduces decentralized security, ensuring verifiable integrity of medical records while maintaining transparency.

– **Overhead**: Centralized systems typically have lower computational overhead compared to blockchain, as they do not require distributed verification of transactions. Their decentralized computation and storage also make them more susceptible to these kinds of hacks. Blockchain's computational cost is higher due to encryption, decryption, and consensus verification, but it significantly enhances security, making it more suitable for sensitive healthcare data management.

5 Conclusion

The proposed blockchain-based scheme offers a robust and scalable solution for secure EHR exchange in multi-cloud healthcare environments. By leveraging blockchain's decentralized and immutable ledger, combined with ECC for key management, the system ensures high levels of data confidentiality, integrity, and transparency. The scheme effectively mitigates risks associated with unauthorized access, data breaches, and single points of failure. However, while the system demonstrates significant strengths in security and scalability, challenges such as computational overhead and integration complexity remain. While the proposed system offers significant security benefits, one challenge associated with blockchain is its energy consumption. Public blockchain systems, like those using Proof of Work (PoW), can be energy-intensive. To mitigate this issue, energy-efficient consensus mechanisms like Proof of Stake (PoS) can be used, ensuring resource efficiency in healthcare environments. **Future** work will focus on optimizing performance in resource-constrained environments to improve real-world applicability. Additionally, we will validate the system's scalability by testing key metrics like throughput and latency. Addressing bottlenecks such as transaction throughput and exploring solutions like sharding and off-chain processing will ensure efficiency in large-scale deployments. In addition, a concrete roadmap for this research involves implementing a pilot program in collaboration with healthcare providers to gather data on the system's effectiveness.

References

1. Alowais, S.A., et al.: Revolutionizing healthcare: the role of artificial intelligence in clinical practice. BMC Med. Educ. **23**(1), 689 (2023)
2. Mullins, A., et al.: Health outcomes and healthcare efficiencies associated with the use of electronic health records in hospital emergency departments: a systematic review. J. Med. Syst. **44**(12), 200 (2020)
3. Yaseen, A.A., Patel, K., Yassin, A.A., Aldarwish, A.J., Hussein, H.A.: Secure electronic healthcare record using robust authentication scheme. IAENG Int. J. Comput. Sci. **50**(2) (2023)
4. Keshta, I., Odeh, A.: Security and privacy of electronic health records: concerns and challenges. Egypt. Inf. J. **22**(2), 177–183 (2021)
5. Mason, J., Dave, R., Chatterjee, P., Graham-Allen, I., Esterline, A., Roy, K.: An investigation of biometric authentication in the healthcare environment. Array **8**, 100042 (2020)

6. Choi, Y.B., Williams, C.E.: A HIPAA security and privacy compliance audit and risk assessment mitigation approach. In: Research Anthology on Securing Medical Systems and Records, pp. 706–725. IGI Global (2022)
7. Ometov, A., Petrov, V., Bezzateev, S., Andreev, S., Koucheryavy, Y., Gerla, M.: Challenges of multi-factor authentication for securing advanced IoT applications. IEEE Netw. **33**(2), 82–88 (2019)
8. Jaime, F.J., Muñoz, A., Rodríguez-Gómez, F., Jerez-Calero, A.: Strengthening privacy and data security in biomedical microelectromechanical systems by IoT communication security and protection in smart healthcare. Sensors **23**(21), 8944 (2023)
9. Garcia-Perez, A., Cegarra-Navarro, J.G., Sallos, M.P., Martinez-Caro, E., Chinnaswamy, A.: Resilience in healthcare systems: cyber security and digital transformation. Technovation **121**, 102583 (2023)
10. Andrew, J., Isravel, D.P., Sagayam, K.M., Bhushan, B., Sei, Y., Eunice, J.: Blockchain for healthcare systems: architecture, security challenges, trends and future directions. J. Netw. Comput. Appl. **215**, 103633 (2023)
11. Nakamoto, S.: Bitcoin: a peer-to-peer electronic cash system (2008)
12. Xhafa, F., Li, J., Zhao, G., Li, J., Chen, X., Wong, D.S.: Designing cloud-based electronic health record system with attribute-based encryption. Multimedia Tools Appl. **74**, 3441–3458 (2015)
13. Wang, Z., Zhu, Y.: A centralized HIDS framework for private cloud. In: 2017 18th IEEE/ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD), pp. 115–120. IEEE (2017)
14. Armknecht, F., et al.: A guide to fully homomorphic encryption. Cryptology ePrint Archive (2015)
15. Murala, D.K., Panda, S.K., Sahoo, S.K.: Securing electronic health record system in cloud environment using blockchain technology. In: Panda, S.K., Mishra, V., Dash, S.P., Pani, A.K. (eds.) Recent Advances in Blockchain Technology: Real-World Applications, vol.237, pp. 89–116. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-22835-3_4
16. Chandini, A., Basarkod, P.I.: Patient centric pre-transaction signature verification assisted smart contract based blockchain for electronic healthcare records. J. Ambient. Intell. Humaniz. Comput. **14**(4), 4221–4235 (2023)
17. Midha, S., Verma, S., Mittal, M., Jhanjhi, N., Masud, M., AlZain, M.A.: A secure multi-factor authentication protocol for healthcare services using cloud-based SDN. Comput. Mater. Continua **74**(2) (2023)
18. Wanisha, I., James, J.B., Witeno, J.S., Bakery, L.H.M., Samuel, M., Faisal, M.: Multi-factor authentication using blockchain: enhancing privacy, security and usability. Int. J. Comput. Technol. Sci. **1**(3), 41–55 (2024)

A Proposed Model to Transliterate Digitalized Braille Text into Gujarati: An Aid for Visually Challenged People

Nikisha Jariwala[1](✉) and Kalpesh Popat[2]

[1] Smt. Tanuben & Dr. Manubhai Trivedi College of Information Science, Affiliated to Veer Narmad South Gujarat University, Surat, Gujarat, India
nikishajariwala@tmtbcasurat.org
[2] Faculty of Computer Applications, Marwadi University, Rajkot, Gujarat, India

Abstract. Information and Communication Technology (ICT), in the modern digital era has become so cutting-edge and touched to each and every field but there is still a disconnect between visually impaired persons and communication technologies because of ignorance and inattention. Due to the limited availability of resources and tools; people with visual impairments are unable to communicate in writing. In this paper, researcher has tried to draw the attention towards the challenges that are faced by visually impaired individuals in written communication and also discussed the benefits of having a tool that eradicate the communication gap. To overcome the challenges faced by blind people; author has proposed a model based on Natural Language Processing (NLP) and character mapping technique that act as an aid to transliterate the Braille text written by blind people to our regional language Gujarati. The technique used in the model uses standard Braille syntax. Rules are formed, lookup table is created, Braille characters are identified and they are transliterated to Gujarati language text. The model shows satisfactory result in the transliteration of Braille characters to Gujarati. It will also help visually impaired students in their education.

Keywords: Digital text processing · Braille text · Gujarati language · Transliteration · Natural Language Processing (NLP) · Character mapping · visually impaired people

1 Introduction

Language is essential to our everyday interactions because it allows us to communicate and express our feelings. It may be expressed orally or in writing [1]. Braille is a crucial written language for those who are blind or visually challenged [2, 3].

People with visual impairments are also valuable to the society who can contribute significantly to its prosperity [4, 5]. Consequently, it has become imperative to give them the tools and channels necessary to communicate with the outside world. Because of technological advancements; people who are blind or visually challenged can now access textual literature [5]. Through the use of Braille text transliteration, individuals who are visually impaired will be able to produce texts and converse with the sighted using a standard text format and comprehensible language.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
S. Rajagopal et al. (Eds.): ASCIS 2024, CCIS 2428, pp. 247–260, 2025.
https://doi.org/10.1007/978-3-031-86302-8_16

1.1 Braille Script

Braille [6] writing system is used by individuals who have low vision or are blind to read text, which is read with the tip of fingers and it consists of raised/embossed dots. Frenchman Louis Braille, the inventor of Braille; in order to improve night writing, developed code for the French alphabet [7–9]. In 1829, he published his technique, which included notation for music [10]. January 4th of every year is observed as an International World Braille Day to honor the creator of Braille. Using embossing, Braille is written using the notion of a cell—a thick sheet of paper with raised dots—on top of it. Six dots are placed in a Braille cell in a rectangular grid that consist of two horizontal dots in a row and three vertical dots in a column. Only sixty-four distinct patterns of dots may be generated as a result of this arrangement of six dots. A cell is any configuration of dots that contains a minimum of one embossed dot and a maximum of six embossed dots [1, 7, 8, 11] as shown in Fig. 1 and Fig. 2 represents printed Braille sheet.

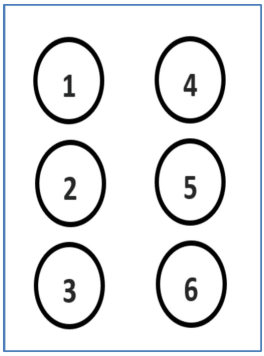


Fig. 1. Braille cell with six dots used to form Braille character [12]

Typically, a Braille printed sheet as shown in Fig. 2 has forty cells in each row and maximum of twenty-five rows of text on a single page. A typical Braille page measures around eleven inches by eleven inches. Although the Braille cell's dimensions are defined, they may differ significantly depending on the country's preferred script. A Braille script is also written as a cell consisting of eight dots, depending on the language that needs to be represented in terms of Braille. A dimension of Braille cell as printed on an embosser [2, 14, 15] is shown in Fig. 3.

Braille printer also known as embosser as shown in Fig. 4 is used to print/emboss Braille characters on the thick sheet exclusively meant to print Braille. The dots on the Braille sheet falls on the intersections of an orthogonal grid. Braille can be printed one-sided as well as two-sided. If it is printed two-sided in terms of recto and verso dots; the pattern of verso dots is somewhat shifts between recto dots so that it does not overlap each other as shown in Fig. 5.

The Braille script known as Bharati Braille (Indian Braille), also called Bhartiya Braille, is primarily used to write the languages of India. After India attained independence, for different languages and in a different region of the nation, there were eleven

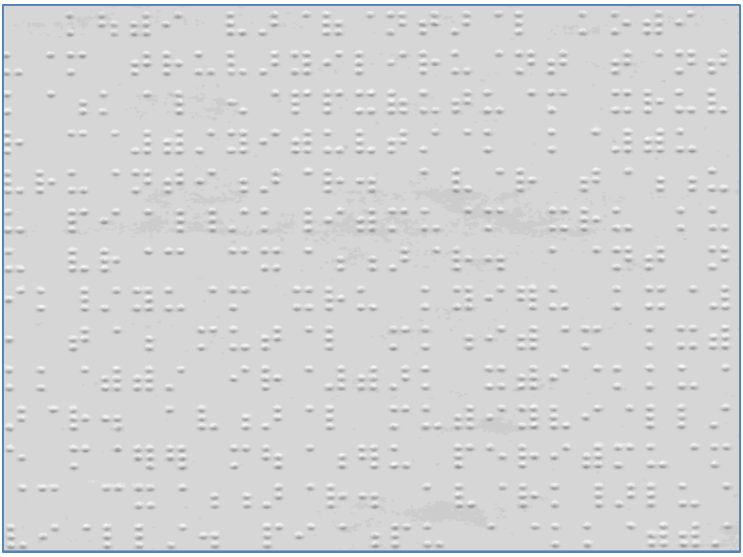


Fig. 2. Printed Braille sheet (Embossed Braille cells) [13]

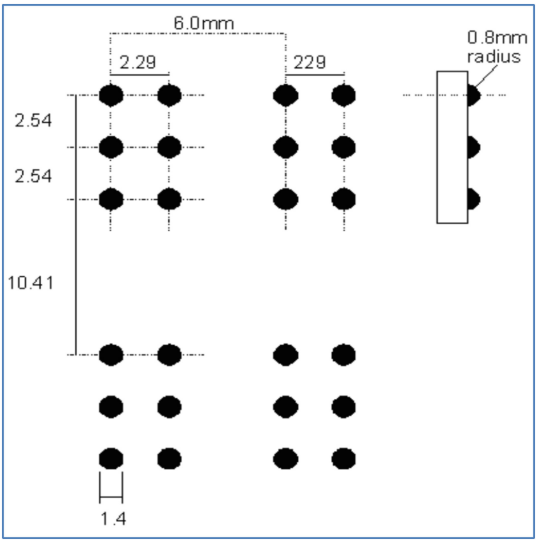


Fig. 3. Dimension of Braille cell carrying six dots [16]

Braille scripts in use [5]. Gujarati braille is comparable to Hindi Braille is also one of the Bharati Braille languages.

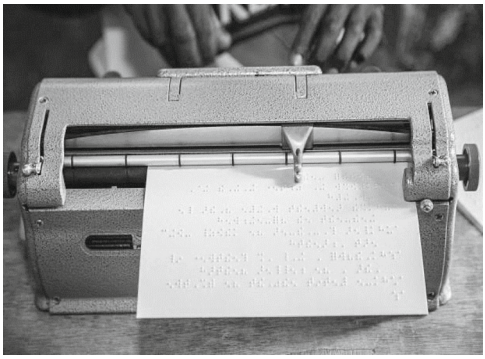


Fig. 4. Braille Printer/Embosser used to print Braille cells

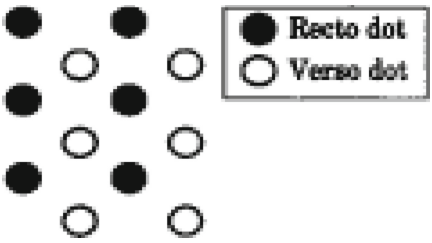


Fig. 5. Position of Verso and Recto Braille dots in two-sided print [15]

1.2 Gujarati Language

Gujarati is a regional language of Gujarat state and one of the scripts used in India for both oral and written communication. Gujarati script is multifaceted [15]. Gujarati characters are written from left to right and each character represents a syllable [17]. Consonants and vowels are referred to as Vyanjan and Swar respectively. Gujarati language uses a system of Maatras; a unique modifier symbols, one for each vowel, to modify the sound of consonants [2]. It has total 75 distinct characters [15]. Gujarati numerals and punctuations with their Braille representation are shown in Fig. 6 and Fig. 7 respectively. Figure 8 represents Gujarati characters with its Braille representation.

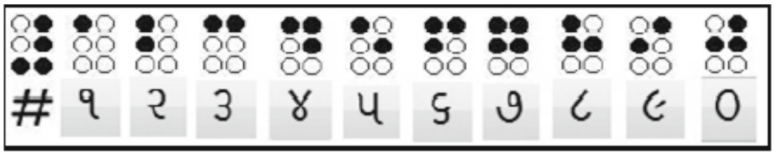


Fig. 6. Gujarati numerals and its corresponding Braille representation [4]

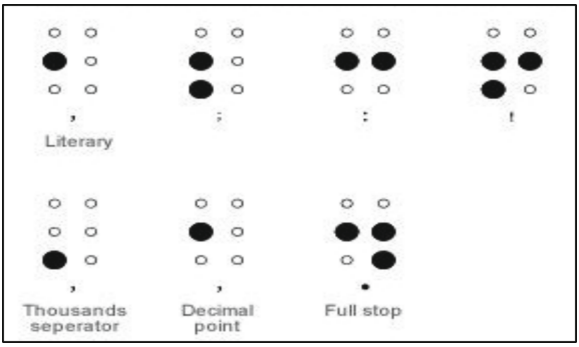


Fig. 7. Punctuation marks in Braille

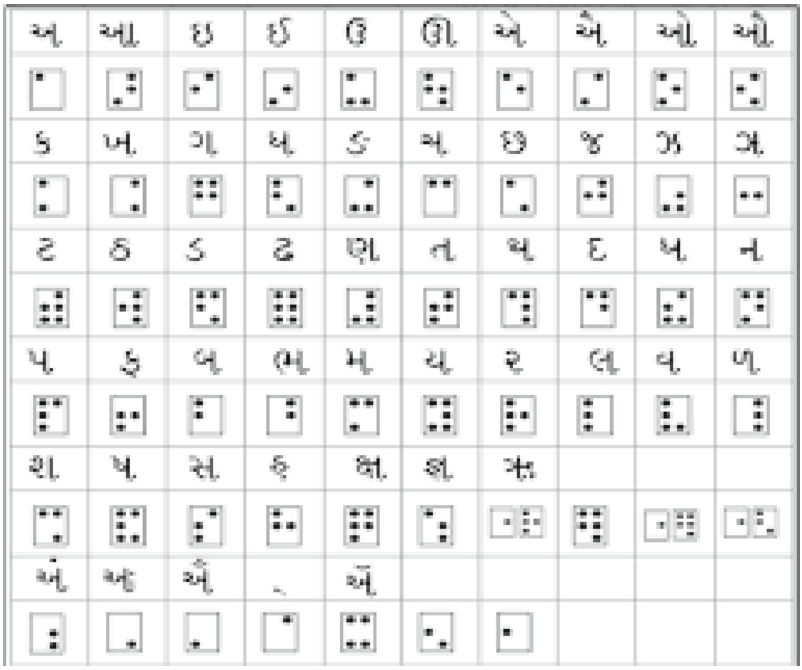


Fig. 8. Gujarati and its corresponding Braille representation [4]

1.3 Challenges of Visually Impaired People and Benefits of the Proposed Model

The primary problem is to generate Gujarati text from Braille. So, the visually impaired people can also communicate in written form with the sighted people and they don't remain dependent on anyone for conveying information in written form to the society.

If blind person wants to type some document, then the person has to be dependent on the people who will type for them in Gujarati language. But through proposed model, visually challenged people will be able to type their document in Braille and it will be

converted to Gujarati. So, they will not need help of others and sighted person will able to read.

Similarly, if visually impaired person wants to set the paper and type it, then again, the proposed model will be useful. As they will be able to type in Braille and through the system it will be converted to normal Gujarati language.

If the proposed model is developed then unemployed visually impaired person will be able to get some job such as data entry and they can perform data entry in Braille and then it will be converted into Gujarati language.

Similarly, it will also be helpful to the visually impaired person who is capable of getting job as PO or clerk, but due to their blindness they are not getting this opportunity.

It will also be helpful to the parents who are blind but their child is well sighted. They will be able to help their child in studying. For example, if child ask their parents that how to write some word in Gujarati, then the parents will able to write in Braille and their child will not be able to understand. But through the proposed model, if parents write some word in Braille, then it would be easily converted to Gujarati language which is understandable to their child.

It will also be useful in academic field. For example, if visually impaired person is teaching Gujarati language and that person want their student to write something then he has to either dictate or ask some other person to write it on the board. But now most of the schools have projectors attached with computer for teaching. So, if the teacher wants to give notes to the students, then teacher will write notes in Braille and through the proposed model it will be converted to normal Gujarati language.

If visually challenged person wants to type the content of the book, then also it will be possible with the help of proposed model.

So, in short, if the proposed model is commenced then it will be beneficial to visually impaired people in many ways by getting jobs in various sectors such as bank, corporate, schools etc. and it will reduce the written communication gap between sighted people and visually impaired people.

1.4 Need of the Research

In India there are thousands of languages, so there isn't any standard script used across the languages. For visually impaired people, six-dots Braille is a standard and unique script.

The National Institute for the Empowerment of Persons with Visual Disabilities (NIEPVD), then known as the National Institute for the Visually Handicapped, was also profoundly involved in the tuning of Bharati Braille which is meant for Indian languages. For the most of the official languages of India, NIEPVD Braille development unit has contributed notation systems for music, Science and math as well as shorthand system, Braille contractions and abbreviations [13].

Currently, there are 70 million vision impaired people and 4.95 million (0.36% of the population) blind people in India, out of which blind children are around 0.24 million [18].

In 31 districts in India; Community Ophthalmology, R P Centre, AIIMS conducted NPCB&VI National Blindness survey, among people aged 50 years. From Gujarat Kheda district was considered as a part of this survey. Prevalence of blindness was 1.15% which

was less than national average 1.99%. Prevalence of visual impairment was 11.13%. Prevalence of functional low vision (FLV) in district Kheda, Gujarat was 1.4% [18].

India - the second most populated nation contribute one-third to the world's blind population. But in India the Braille literacy rate is far lower (only 1%) than the regular literacy rate which is 77.7% [19] as shown in Fig. 9. Braille education will help visually impaired children and adults to increase their literacy. In India, audio content and Braille books are access by just 6.86% of schools [20].

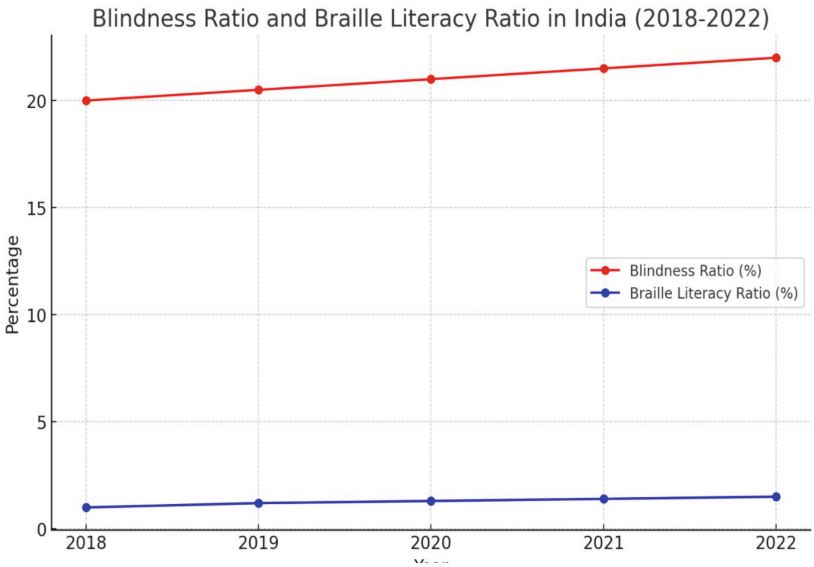


Fig. 9. Comparison between blindness and Braille literary ratio [19]

2 Literation Review

Following research work are carried out for Braille to text transliteration.

With reasonable constraints, J. Mennens, G. Francois, L. V. Tichelen and J. J. Engelen have created a method which translates Braille content into a format that can be read by computers. Therefore, an electronic Braille embosser can be used by a Braille printing business to recreate text that is stored on a computer [21].

P. Blenkhorn talked about how to convert Braille, which is saved as characters in a computer, into print. To do so, it mostly used a table-driven approach. Numerous character sets and languages are supported by the way the system is constructed and changeable. The algorithm is explained by taking into account the printing of Standard English Braille. Also, the tables used for the conversion are provided by the author [8].

A. S. Al-Salman describe a bi-lingual, bi-directional, editor-equipped translation system that is Arabic-Braille. It does not require any pricey machinery. The system can be expanded to support literate people's communication across languages, geographic

regions, and disabilities (visually impaired or sighted) if the necessary rule file is available for any other languages [6].

P. Rajarapollu, D. Laghate, S. Kodolikar and A. Khavle demonstrate how to use the Field Programmable Gate Array (FPGA) Spartan3 kit to convert Braille to text or speech. The genuine Braille language is translated into English in the regular domain. The Braille keypad, which has various cell combinations, is used to enter data into the system. The Spartan3 kit of FPGA then receives the additional input. Based on the combinations produced, the FPGA uses decoding logic in VHDL (VHSIC (Very High Speed Integrated Circuit) Hardware Description Language) to translate the Braille text into the appropriate English language text. The algorithm decodes the corresponding alphabet, converts it to speech, and displays it through the LCD interface with the Spartan3 kit [22].

G. E. Legge, J. S. Mansfield and C. M. Madison describe a brand-new, uniform technique for gauging Braille reading speed that was developed from the MNREAD print reading speed test. One of the main principles is to utilize content that is well-characterized and made up of basic terminology. The text should be provided in a standard spatial structure, and the sample length should be expressed in characters. 44 seasoned individuals are studied for Braille reading speed using the MNREAD test. A person reading 7.5 characters per second, or 124 words per minute, was the median speed. The reading speed of Grade 1 was 71.5% of Grade 2 when expressed in words per minute. This result supports the idea that the best criterion for assessing Braille reading speed is characters per second. A comparison between reading speed in print and Braille reveals that reading speed characteristics are both qualitatively and quantitatively equivalent when viewing conditions and measurement techniques are matched [1].

A. S. Al-Salman describes a translation/editor system for bi-directional and bi-lingual Arabic Braille that does not need expensive equipment. With proper rules for other languages and regardless of the disabilities of the people whether they are visually impaired or sighted, geographical locations or languages; this system can be generalized to facilitate communication among literate people [6].

A. Chaudhary, P. Garg and A. Agarwal present a new algorithm using image processing is developed to remove misalignment of scanned Braille pattern. It is known that this error in the scanned Braille document is occurred because of scanner imperfections. To fix the corner pixel value of tilted Braille cell pattern, the paper represents indexing method and it employs basic triangular property. So, they have calculated the angle by which the corner dots have been thrown out of alignment on a respective side. Sorting method is used for symmetrical and unsymmetrical Braille pattern, as a variety of different Braille patterns are available where some are regular and exactly symmetrical. The system is developed in MATLAB (R 2010a) and removes misalignment of Braille pattern and yields almost perfect results with acceptable errors [9].

H. Vyas and P. Virpariya has discussed template-based transliteration of Braille text. Author has used image processing to recognized Braille characters and then converted into text. The model is developed using python and MySQL. They have also discussed methodology in detail with template matching technique [23].

S. Shokat, K. Khan, R. Riaz, S. Rizvi, F. Rian, S. Kwon has discussed the comprehensive literature review on various Braille input methods and their benefits to visually impaired people. Authors discussed that blind people have difficulty in accessing the

technologies. The paper emphasis on the need of the system which reduces cognitive load on visually challenged individuals. The authors identify the research gap in this area as very less amount of work is done outside USA and Canada. So, they suggested that if Braille input system is developed it will be beneficial to the blind people [24].

U. Beg, K. Parvati, V. Jha have identified that there is a major gap in the available literature related to text to Braille conversion. Author mentioned that to enhance the communication for visually impaired people conversion into Braille will act as an aid and it will also reduce the illiteracy rates. Researchers pointed out that real-time applications such as instant text-to Braille and Braille to text could improve the accessibility for blind people. They have discussed that image processing technique can also be used to convert optical text to Braille [3].

S. Padmavathi, S. Sphoorthy Reddy, K.S.S. Manojna, D. Meenakshy has discussed the concept of Braille translators which converts written text into Brille cells. Researchers have also proposed the method to convert scanned Braille document into English, Hindi and Tamil text. They have also mentioned that this text can further be converted into the spoken form with the help of Speech Synthesizer [25].

V. Jha and K. Parvathi highlights the challenges faced by visually impaired people of society in accessing educational literature in Hindi. They also reviewed existing techniques for Braille translation as well as introduced machine learning approach that recognize and transliterates characters. With the help of their research education system and visually challenged people will be beneficial [26].

C. Li and W. Yan have described significance of Braille for blind people which allows them to read and write. They have discussed deep learning based efficient Braille translator tool as well as word-based Braille detection using ABCNet. They highlighted the importance of the tool as it will improve communication for blind people using advanced technologies [27].

T. Shreekanth and V. Udayashankara have emphasized on the need of Optical Braille Character Recognition (OBR) to store and preserve Braille documents and improve the communication between visually impaired people of the society and the sighted people. They have tried to recognize the Braille cells which are printed two-sided where verso and recto dots are present in the printed Braille document. They have worked on Hindi language Braille character recognition [28].

M. Dhanalakshmi and K. P. S. G. Sugirtha describes a new system to identify the Braille cells in six-dots from the scanned document and transliterate it into English language. Image processing phases such as preprocessing and segmentation are applied to the scanned document to recognize the Braille character and Euclidean distance is calculated. Further lookup table is considered to transliterate the Braille character to the corresponding text [29].

We found that there are many literatures in various languages that convert Braille script text to plain language text including Gujarati language. Researchers have also worked on Optical Character Recognition of Braille cell. But due to the complexities and syntax of Braille script, incomplete solutions are available and along with that all the tactics have their own limitations. So, to resolved the problem authors have tried to proposed transliteration model that uses text processing method, transliteration technique

and character mapping algorithm to convert Braille characters to plain Gujarati text. Let's look into it in detail.

3 Proposed Transliteration Model

The proposed transliteration model has several phases that process the Braille cell and transliterate it into a Gujarati character. Following section describes and represents the phases of the model as shown in Fig. 10.

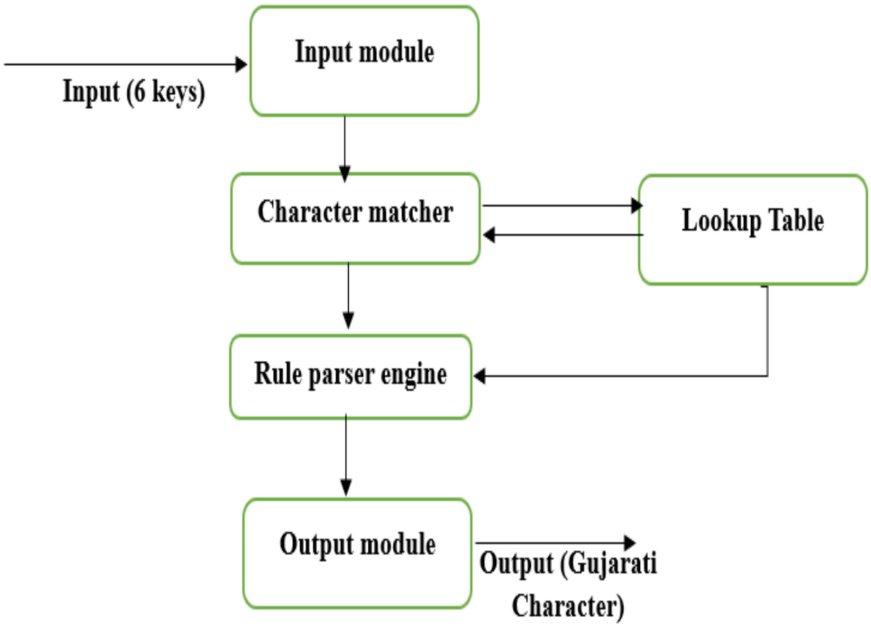


Fig. 10. Block diagram of Proposed Transliteration Model

3.1 Phases of the Proposed Transliteration Model

Input module – the input module is further divided into two components. Input panel and input processor. Input panel will accept the input in terms of 6 keys (SDFJKL) and further pass on those keys to the Input Processor to generate Braille cells. These 6 keys represent 6 Braille dots within the cell. The keys are gathered from the user and then converted into the Braille pattern (cell).
Lookup table – act as a collection of Braille cells with its corresponding Gujarati language character used for character mapping.
Character matcher – at the time of cross-script transliteration the transliteration is done from one script to another. So, to transliterate from source script i.e. in this research Braille to destination script i.e. Gujarati character mapping technique is used.

It tries to match the Braille pattern with the collection present in the lookup table and retrieve the equivalent Gujarati character with the help of character mapping technique.
Rule parser engine – Both the source and destination scripts have their own syntax and rules. In the proposed model, taking this into consideration during transliteration rules and syntax of both the scripts are parsed and according to that characters are mapped.
Output module – the Gujarati character that is generated after applying the rules and parsing it, is displayed to the user.

3.2 Transliteration Rules

Following rules are taken into consideration by the proposed transliteration model [30]:

Through Braille cells only 64 unique patterns can be generated. But Gujarati language set of 75 characters. There are some characters in Gujarati Braille which are similar and indistinguishable. So, there is a need to take assumptions and applying some indicators.

For example, in Gujarati Braille ૧ is represented in the similar way as ચ, ૨ is represented in the similar way as ખ and so on. So, number indicator will be the extra Braille cell present in Braille script at the time of the representation of the numerals in Braille script.

There are no distinct characters present in Gujarat Braille to indicate compound or half or conjunct characters. So, here as well additional character known as half character indicator is used to indicate that the following character is the half or compound or character.

So, at the time of typing Braille visually impaired people will consider all those extra indicators in their document. But in Gujarati script all those indicators are not available.

During transliteration from Braille to Gujarati all of those indicators meaning are to be considered and then it needs to be transliterated into the corresponding text.

3.3 Algorithm

An algorithm is developed to transliterate Bhartiya Braille characters in plain Gujarati characters that can further be printed on the printer to be read by sighted people.

The steps are as follows:

Step 1: Fetch the keys from the keyboard that is pressed by the user (SDF JKL).

Step 2: The keys are combined to formed the pattern of Braille cell (six dots) in difference of one second.

F $\rightarrow$ 1st dot	J $\rightarrow$ 4th dot
D $\rightarrow$ 2nd dot	K $\rightarrow$ 5th dot
S $\rightarrow$ 3rd dot	L $\rightarrow$ 6th dot

Step 3: The character matcher of the model matches the pattern of Braille cell dots with the entries within the lookup table.

Step 4: If match is found for the pattern of dots of the Braille cell, then the resultant Gujarati character is parse against the rules (as mentioned in the 3.2 transliteration rules) created in the rule parser engine.

Step 5: The formed Gujarati character is displayed (as shown in Fig. 11 of topic 3.4).

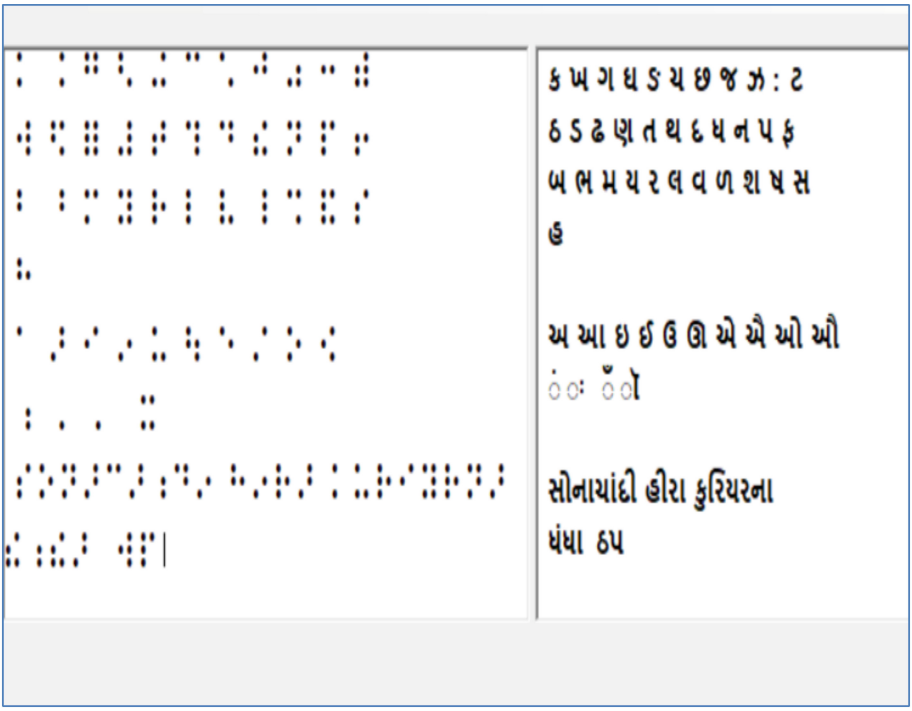


Fig. 11. Transliteration of Braille character into Gujarati Text through the proposed model

3.4 Experimental Results

The model is tested by typing Braille characters i.e. ⠁ (અ) to ⠒ (ઃ) and ⠅ (ક) to ⠭ (ઙ) and words that doesn't contains compound and conjunct characters. A sample transliteration done by the model is shown in the Fig. 11. The Gujarati language text generated through the model is verified by both the Braille expert as well as Gujarati language teacher. They did not found error in the transliteration of whole character and words without compound and conjunct characters. So, it shows that model works perfect for plain Gujarati Braille characters and gives satisfactory result.

4 Conclusion

The paper explains the fundamentals of Gujarati language and Braille writing. Authors have also discussed the need of the research by providing the statistics as well as the challenges that are faced by visually impaired people at the time of written communication. Additionally, it describes that some researchers have worked on the Braille character recognition in several languages such as Arabic, English, Hindi, Bengali etc. Reasonable amount of work is found in Gujarati Braille recognition but complete transliteration into Gujarati language is still untouched. Therefore, the researcher believed that the Braille to

Gujarati text transcription was necessary. Proposed transliteration model in this article will convert Braille cells inputted by the user into Gujarati language characters. The proposed work is tested and it shows satisfactory result. So, it will act as an aid for blind and visually challenged people of the society.

5 Future Enhancement

In the present model recognition and transliteration of the individual Braille cells into Gujarati characters is carried out. In future, the model will be extended to transliterate Braille numeral, words as well as sentences that contain compound and conjunct characters in to Gujarati language characters by considering all the Braille syntax and the rules.

Acknowledgement. We are thankful towards the trustee Shri. Anand Chokhavala and the teachers (Braille and Gujarati experts) of the Ambaben Maganlal Andhjan Shala for supporting and cooperating us in verifying the transliterated content and providing their testimony for it. We also thank Dr. Naren Burade to provide the direction to work for the visually impaired people of our society.

References

1. Legge, G.E., Madison, C.M., Mansfield, J.S.: Measuring Braille reading speed with the MNREAD test. Vis. Impair. Res. **1**(3), 131–145 (1999). https://doi.org/10.1076/vimr.1.3.131.4438
2. Jariwala, N.B., Patel, B.: Conversion of Gujarati text into Braille: a review. Int. J. Innov. Adv. Comput. Sci. (IJIACS) **4**(1), 59–64 (2015)
3. Beg, U., Parvathi, K., Jha, V.: Text translation of scanned Hindi document to Braille via image processing. Indian J. Sci. Technol. **10**(33), 1–8 (2017). https://doi.org/10.17485/ijst/2017/v10i33/112335
4. Jariwala, N.B., Patel, B.: Transliteration of digital Gujarati mathematical text into Braille for visually impaired people. Int. J. Latest Trends Eng. Technol. **7**(3) (2016). https://doi.org/10.21172/1.73.531
5. Vyas, H.A., Virparia, P.V.: Transliteration of Braille character to Gujarati text-the model. Int. J. Eng. Res. Technol. (IJERT) **7**(7), 329–332 (2018). www.ijert.org
6. Al-Salman, A.S.: A bi-directional bi-lingual translation Braille-text system (2008)
7. Singh, M., Bhatia, P.: Automated conversion of English and Hindi text to Braille representation. Int. J. Comput. Appl. **4**(6), 25–29 (2010). https://doi.org/10.5120/831-1143
8. Blenkhorn, P.: A system for converting print into Braille (1997)
9. Chaudhary, A., Garg, P., Agarwal, A.: Using rotation method for removal of misalignment of scanned Braille pattern. In: International Conference on Advances in Computing, Control and Communication CCN 2012, pp. 71–75 (2012). https://doi.org/10.3850/978-981-07-2579-2
10. Halder, S., Hasnat, A., Khatun, A., Bhattacharjee, D., Nasipuri, M.: Development of a Bangla character recognition (BCR) system for generation of Bengali text from Braille notation. Int. J. Innov. Technol. Exploring Eng. (IJITEE) **3**(1), 5–10 (2013)
11. Khalil, S., Al-Shamma, S.D., Fathi, S.: Arabic Braille recognition and transcription into text and voice. In: 2010 5th Cairo International Biomedical Engineering Conference, pp. 227–231 (2010). https://www.researchgate.net/publication/263236983

12. Hsu, B.M.: Braille recognition for reducing asymmetric communication between the blind and non-blind. Symmetry (Basel) **12**(7) (2020). https://doi.org/10.3390/SYM12071069
13. Braille in India: How Languages Found Expression in Bharati Braille (2020). https://thinkerbelllabs.com/blog/braille-india-how-languages-found-expression-bharati-braille/
14. Acharya-Multilingual Computing for Literacy and Education. www.acharya.gen.in/IntroductiontoBraille.htm
15. Jariwala, N.B., Patel, B.: Transliteration of digital Gujarati text into printable Braille. In: 2015 Fifth International Conference on Communication Systems and Network Technologies, pp. 572–577 (2015). https://doi.org/10.1109/CSNT.2015.82
16. Braille Specification. https://www.mkdesign.uk/braille-specifications.html
17. Parekh, H., Shah, S., Patel, F., Vyas, H.: Gujarati Braille text recognition: a review. Int. J. Comput. Sci. Commun. **7**(1), 19–24 (2015)
18. Mannava, S., Raj Borah, R., Shamanna, B.R.: Current estimates of the economic burden of blindness and visual impairment in India: a cost of illness study. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9359234/
19. Why Braille Education Is So Important in India for the Visually Impaired Students? https://www.asiancollegeofteachers.com/blogs/1348-Why-Braille-Education-Is-So-Important-in-India-for-the-Visually-Impaired-Students-blog.php
20. Economic Times - World Braille Day. https://economictimes.indiatimes.com/magazines/panache/world-braille-day-2022-screen-readers-customised-apps-hr-automation-bring-some-hope-to-visually-impaired/articleshow/88686000.cms?utm_source=contentofinterest&utm_medium=text&utm_campaign=cppst
21. Mennens, J., van Tichelen, L., François, G., Engelen, J.J.: Optical recognition of Braille writing using standard equipment. IEEE Trans. Rehabil. Eng. **2**(4), 207–212 (1994). https://doi.org/10.1109/86.340878
22. Rajarapollu, P., Kodolikar, S., Laghate, D., Khavale, A.: FPGA based Braille to text & speech for blind persons. Int. J. Sci. Eng. Res. **4**(4), 348–353 (2013)
23. Vyas, H., Virpariya, P.: Template-based transliteration of Braille character to Gujarati text—the application. In: Rising Threats in Expert Applications and Solutions, pp. 437–446 (2020)
24. Shokat, S., Riaz, R., Rizvi, S.S., Khan, K., Riaz, F., Kwon, S.J.: Analysis and evaluation of Braille to text conversion methods. Mob. Inf. Syst. **2020** (2020). https://doi.org/10.1155/2020/3461651
25. Padmavathi, S., Reddy, S.S., Meenakshy, D.: Conversion of Braille to text in English, Hindi and Tamil languages. Int. J. Comput. Sci. Eng. Appl. **3**(3), 19–32 (2013). https://doi.org/10.5121/ijcsea.2013.3303
26. Jha, V., Parvathi, K.: Braille transliteration of Hindi handwritten texts using machine learning for character recognition. Int. J. Sci. Technol. Res. **8**(10) (2019). www.ijstr.org
27. Li, C., Yan, W.: Braille recognition using deep learning. In: ICCCV 2021: Proceedings of the 4th International Conference on Control and Computer Vision, pp. 30–35 (2021)
28. Shreekanth, T., Udayashankara, V.: A two stage Braille Character segmentation approach for embossed double sided Hindi Devanagari Braille documents. In: 2014 International Conference on Contemporary Computing and Informatics (IC3I) (2015)
29. Sugirtha, K.P.S.G., Dhanalakshmi, M.: Transliteration of Braille code into text in English language. In: Computational Signal Processing and Analysis, pp. 373–381 (2018)
30. Vyas, H., Virpariya, P.: Gujarati Braille text recognition: a design approach. In: Agrawal, S., Devi, A., Wason, R., Bansal, P. (eds.) Advances in Intelligent Systems and Computing - Springer, Singapore, pp. 31–40 (2018). https://doi.org/10.1007/978-981-10-6626-9_4

Financial Data-Driven Classification of Manufacturing Companies: A Case Study of the Malaysian Industry

Rathimala Kannan[1(✉)] and Kannan Ramakrishnan[2]

[1] PSG Institute of Management, PSG College of Technology, Coimbatore 641004, Tamil Nadu, India
rathimala@psgim.ac.in

[2] Faculty of Computing and Informatics, Multimedia University, 63100 Cyberjaya, Selangor, Malaysia
Kannan.ramakrishnan@mmu.edu.my

Abstract. Manufacturing sector companies are typically categorized based on size or product type to aid in policy formulation and long-term planning. The existing classification systems, however, overlook economic performance, which makes it difficult to plan policies precisely and allocate resources. This study aims to close this gap by identifying economic factors that set manufacturing companies apart using data mining and machine learning. The main objective is to uncover similarity patterns among manufacturing companies and to describe each segment's characteristics addressed in this study by utilizing historical data from the Manufacturing Economic Census Survey. The Cross-Industry Standard Process for Data Mining (CRISP-DM) framework guides the extraction of meaningful patterns from the economic census survey data. The study attempts to efficiently group manufacturing organizations according to their economic performance by utilizing unsupervised learning algorithms such as fuzzy C-means clustering and K-means clustering. The findings deviate from the existing categorizing method and yield six groups that are defined by financial metrics such as net book value, staff qualities, and spending distribution. Each segment is described in detail, including key characteristics such as business structure, primary activities, workforce qualifications, and financial metrics. These insights are instrumental in making informed decisions regarding future policy changes and financial aid initiatives, ultimately fostering a more resilient and competitive manufacturing sector. This research underscores the significance of data-driven approaches in elucidating industry dynamics and guiding strategic interventions to propel economic growth and enhance the well-being of Malaysians.

Keywords: Segmentation · Manufacturing Companies · Unsupervised Machine Learning · K-Means Clustering · Fuzzy C-Means Clustering

Abbreviations

CRISP-DM Cross Industry Standard Process for Data Mining

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
S. Rajagopal et al. (Eds.): ASCIS 2024, CCIS 2428, pp. 261–275, 2025.
https://doi.org/10.1007/978-3-031-86302-8_17

DOSM Department of Statistics Malaysia
FCM Fuzzy C-Means clustering

1 Introduction

The manufacturing sector is a critical pillar of economic growth, contributing significantly to employment, productivity, and overall economic stability. Traditionally, manufacturing companies have been categorized by size or product type to facilitate policy formulation, resource allocation, and strategic planning [1]. However, these existing classification systems often fail to account for the diverse economic performance characteristics within the sector. This limitation hampers the precision of policy-making and the effectiveness of long-term strategic decisions aimed at fostering sectoral resilience and competitiveness.

Classifying manufacturing companies solely as small, medium, or large fails to identify homogenous groups with similar preferences and requirements because this categorization is too simplistic and overlooks the multifaceted nature of these enterprises [2]. Manufacturing companies vary widely not only in size but also in their operations, market orientation, production technology, product complexity, and industry sector, which all significantly influence their needs and preferences. For instance, a small high-tech manufacturing firm specializing in niche markets may have more in common with a large multinational in the same sector than with a small traditional manufacturer of consumer goods. Similarly, medium-sized companies in different industries may have vastly different capital requirements, supply chain complexities, and regulatory environments [1]. Moreover, companies of the same size may have different growth strategies, risk appetites, and innovation capabilities, leading to distinct operational challenges and support needs. Thus, a classification system based purely on size fails to capture these nuances, potentially grouping companies with dissimilar strategic goals, resources, and operational challenges together, which limits the effectiveness of policy-making, financial support programs, and market segmentation strategies aimed at addressing their specific needs. A more nuanced classification that considers factors such as industry type, technological intensity, market orientation, and business model would provide a clearer understanding of the similarities and differences among manufacturing companies.

To address this gap, this study leverages data mining and machine learning techniques to develop a more nuanced classification of manufacturing companies, focusing on economic performance metrics. Using historical data from the Manufacturing Economic Census Survey, the study applies the Cross-Industry Standard Process for Data Mining (CRISP-DM) framework to uncover patterns and segment the manufacturing sector based on key financial indicators and operational characteristics. By employing unsupervised learning algorithms, specifically fuzzy C-means clustering and K-means clustering, the research identifies distinct groups within the sector that deviate from traditional categorizations. The focus is on exploring and extracting meaningful insights from the manufacturing. To achieve the goal, the following research questions are identified.

1. Are there any patterns that distinguish manufacturing companies based on economic indicators?

2. What are the characteristics of the identified segments of the manufacturing companies?

The resulting segmentation reveals six distinct groups characterized by financial metrics such as net book value, workforce qualities, and spending distribution. Each segment is analyzed to highlight its unique characteristics, including business structure, primary activities, workforce qualifications, and financial performance. These insights provide a foundation for more tailored policy-making and resource allocation, supporting a more robust and competitive manufacturing sector. This study underscores the value of a data-driven approach to understanding industry dynamics and emphasizes the potential for strategic interventions to enhance economic growth and the overall well-being of Malaysians. By integrating economic performance into the classification framework, the research offers a novel perspective that could guide future policy adjustments and financial aid initiatives in the manufacturing sector.

The rest of this paper is organized as follows. Section 2 discusses several works of literature relevant to this study including economic census, industry 4.0, and machine learning. Section 3 explains the research methodology. Section 4 discusses the results of the study. Finally, Sect. 5 concludes this research.

2 Literature Review

Machine Learning (ML) and Data Mining play pivotal roles in facilitating informed decision-making within the manufacturing industry. Data mining enables manufacturers to make more reasonable, informed, and responsive business decisions [4] The increasing generation and availability of large amounts of data have brought significant attention to machine learning research [5]. Manufacturing companies have evolved to recognize the advantages of data-driven decision-making, and various researchers have looked into how machine learning could be applied in manufacturing industries to make smart decisions [6, 7]. A review of the literature shows that machine learning techniques are applied to predict the quality of plastic products [8], to predict the production delay in a quarry company [6], to predict defect detection in steel products [9], to develop context-aware manufacturing system design [10], and many more.

Clustering in data mining is a technique used to group a set of objects or data points based on their inherent similarities. The goal of clustering is to partition the data into subsets, or clusters, in such a way that objects within the same cluster are more similar to each other than to those in other clusters [11]. In other words, it involves organizing data points into groups where the members of each group share common characteristics or features. Clustering is widely used in various fields, including machine learning, pattern recognition, image analysis, and customer segmentation, to discover hidden patterns, group similar data points, and gain insights into the structure of the data [9, 12]. There are many different types of clustering techniques such as K-means, Fuzzy C-means, and hierarchical clustering. The most widely used clustering algorithms, according to a review of the literature, are K-means and fuzzy C-means [13] because they are so effective at spotting patterns or similarities in different kinds of datasets. K-means clustering, which can recognize similarity patterns in a variety of dataset types, is the most popular clustering technique.

The K-mean clustering algorithm seeks to reduce the distance between points in a cluster and their centroid. As a result, the goal is to reduce the distance between locations and the cluster centroid [14]. This algorithm focuses on partitioning between the observations (n) in clusters (k). The K-means algorithm can effectively incorporate fresh samples when a clustering issue arises [15]. K-means clustering is widely used in the manufacturing industry for early fault detection in predictive maintenance [11]. Another study used the K-means clustering algorithm to group 38 districts based on indicators of inclusive economic development, and it discovered two clusters: one with 29 cities where inclusive economic development is below the overall average and another cluster with 9 cities where inclusive economic development is above the average [13].

Another popular clustering algorithm is Fuzzy C-Means clustering (FCM), in which the degree of membership determines the existence of each data point in a cluster [16]. FCM employs a fuzzy grouping model with a fuzzy index based on Euclidean Distance, allowing data to be members of all classes or clusters produced with varying degrees of membership ranging from 0 to 1 [17]. A review of previous studies reveals that FCM has been used to find patterns of similarity in a variety of dataset types from diverse industries. A study [18] used a two-level fuzzy c-means clustering algorithm to group shared bicycle stations into groups to capture the connections between the stations. This information was then used to predict the use of bikes in a bike-sharing system by first looking at the geographic location of the stations as well as the movement pattern of the bikes between them. Another study applied Fuzzy c-means clustering to identify homogeneous industrial customers based on the electricity load measurement dataset [19]. Another study from the automobile sector used FCM to identify and prioritize environmental, health, and safety problems based on risk variables. The results were utilized to prioritize threats clearly and make it easier to take corrective and preventive measures [20].

The difference between K-means clustering and Fuzzy C-means is that K-means is a hard clustering while Fuzzy C-means is soft clustering. For hard clustering, one data point can only belong to one cluster, while the output of soft clustering is the likelihood of a data point belonging to each of the pre-specified clusters. K-means clustering divides the data by clustering the whole dataset into K number of clusters, with each data point belonging to just one cluster. Fuzzy c-means generate a c number of clusters and then assign each data point to one of them [21]. As delineated earlier, the adoption of these two clustering methodologies was undertaken to ascertain the nuanced patterns of similarity prevalent among manufacturing companies, leveraging the comprehensive dataset from the Manufacturing Economic Census.

The literature reveals that the Malaysian manufacturing sector is advanced, particularly in rubber, palm oil, and electronics, and is categorized by company size or product type. However, data mining and machine learning techniques, such as K-means and Fuzzy C-means clustering, offer new ways to identify homogeneous manufacturing segments and enhance decision-making by uncovering hidden patterns in large datasets. These clustering methods have been widely used across various industries, including manufacturing, to predict quality, detect defects, and optimize processes. K-means focuses on hard clustering, grouping data points strictly into single clusters, while Fuzzy C-means allows for flexible clustering, where data points can belong to multiple clusters

to varying degrees, providing a nuanced approach to data segmentation and analysis in Industry 4.0.

3 Methodology

This research follows the de facto method for carrying out data mining projects which is known as Cross-Industry Standard Process-Data Mining (CRISP-DM). It is divided into six iterative stages, beginning with Business Understanding, and then followed by Data Understanding, Data Preparation, Modelling, Evaluation, and Deployment. Figure 1 illustrates the research methodology applied in this study.

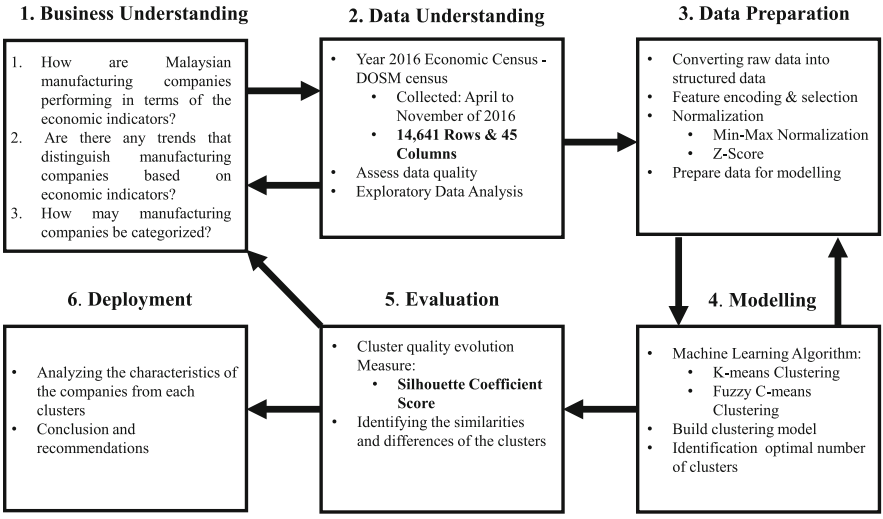


Fig. 1. Research methodology based on CRISP-DM

3.1 Data Collection

This study obtained the secondary data "The economic census survey", from DOSM which was collected between April 2016 and November 2016. There are 14,641 respondents and 45 variables in the dataset. The Economic Census, conducted every five years, compiles detailed accounting, payroll, and business activity data from companies in various sectors, including manufacturing, retail, services, wholesale, finance, construction, mining, insurance, and real estate. The collected data offers valuable benchmark information on industry codes, sales, sources of receipts, total employment, and establishment size. This information allows for in-depth analysis of multiple elements within manufacturing, such as value-added, payroll, and energy usage, at a comprehensive subsector level. The economic census serves as a crucial resource for businesses, aiding in decision-making, planning, and performance evaluation.

3.2 Data Preparation

The building of a strong machine-learning model is aided by data preparation, which may help produce successful results. There are a lot of missing values in the dataset, with several economic indicators having more than 40% of their values missing. It is not necessary to remove a variable with a small number of missing values, and the issue may be resolved by substituting the mode or median for the values. As a result, whenever a significant number of missing values are found inside a certain variable, that variable might be deleted. Thus, in total 18 variables that exceeded 40% missing values had been removed. The list of variables included in this study is shown in Table 1.

In the data preparation process, the relationships between economic indicators are examined, identifying positive and negative correlations and their strengths. The correlation coefficient matrix highlights a 100% positive correlation between the "Salaries" and "Sum of total Salaries & wages paid" variables, leading to the elimination of one variable to avoid redundancy. Subsequently, the data undergoes normalization using two techniques, min-max normalization and Z-score normalization, ensuring equal weightage for clustering techniques. Min-Max normalization, widely employed in data analytics research, proves efficient, particularly in datasets with outliers, as it enhances results [6].This approach calculates all of the numerical values of the variables by using the equation shown below:

$$x' = \frac{x - \min(x)}{\max(x) - \min(x)}$$

Z-score normalization involves adjusting every value in a dataset to achieve a mean of 0 and a standard deviation of 1. This method relies on the mean and standard deviation of the dataset, making it advantageous when minimum and maximum values are unknown. However, it does not ensure that the normalized data maintains the same scale. The resulting Z-scores can be negative or positive, indicating whether values are below (negative) or above (positive) the mean of the group values [22]. The z-score can have both negative and positive values, and it is determined by the values that are either below (negative score) or above (positive score) the mean of the group values. The equation for the Z-score will be shown below:

$$z = \frac{X - \mu}{\sigma}$$

Following the data preparation stage, only 24 variables will be used in the subsequent analysis, allowing this dataset to be processed in the modelling step.

3.3 Machine Learning Models

An examination of the existing literature led to the selection of the two unsupervised learning clustering algorithms; K-means clustering and fuzzy C-means clustering. One of the challenges in applying a clustering algorithm is determining the optimal number of clusters. In this study, the silhouette coefficient is used to assess the quality of the clusters. Silhouette coefficient, value can range between –1 and 1. The silhouette coefficient has three levels of evaluation: Less than 0.2 is considered an unacceptable category, 0.2

Table 1. List of variables included in the study.

No	Variable	Definition
1	Industry Code	The industry that the manufacturers are involved in
2	Legal Status	The legal status of the company
3	Total Net Book Value as At Beginning	Value of Assets
4	Total Net Book Value as At Closing	
5	Total Managerial and Professional	Total number of managerial & professional workers
6	Total Salaries & wages paid for Managerial and Professional (RM'000)	Total salaries & wages given to managerial & professional workers
7	Total Clerical and related occupations	Total number of clerical & related occupations workers
8	Total Salaries & wages paid for Clerical and related occupations (RM'000)	Total salaries & wages given to clerical & related occupations workers
9	Total Production/ operative workers	Total number of production/ operative workers
10	Total Salaries & wages paid for Production/ operative workers (RM'000)	Total salaries and wages given to production/ operative workers
11	Total Employment	The sum of the total employment within the establishment
12	Sum of Salaries & wages paid	The sum of the total salaries & wages given to all of the employees within the establish
13	University Degree and Above	Number of workers that engaged by qualification
14	Diploma / Stpm Or Equivalent	
15	Spm / Spmv Or Equivalent and Below	
16	Sales (RM '000)	The total number of sales achieved within the year
17	Total Income ('RM 000)	The sum of sales, operating income and non-operating income
18	Cost Of Raw ('RM 000)	All consumed goods owned by the institution during the year are included. The goods that used to produce the output of the establishment. Such as the materials / components / parts used

(*continued*)

between 0.5 is a decent category, and more than 0.5 is an excellent category. [22]. In this study, the dataset is clustered into 2, 3, 4, 5, 6, 7, 8, and 9, and the silhouette coefficient is used to find the optimal number of clusters.

Table 1. (*continued*)

No	Variable	Definition
19	Cost Of Admin (‘RM 000)	Costs incurred to support the operation of a firm that are not directly connected to manufacturing, such as packaging materials and containers, materials required for repairs and maintenance, factory requirements, office supplies and stationery
20	Utility (‘RM 000)	Includes the value of usage for water, electricity, fuel, lubricants and gas
21	Operating Expenses (‘RM 000)	Includes products materials that have not been processed further, labour done by others, and money for current repairs and maintenance
22	Non-Operating Expenses (‘RM 000)	Costs not directly tied to the primary business, such as rental payments, current depreciation, interest paid, royalties, and sales/export tax
23	Compensation Of Employee (‘RM 000)	Includes the retirement benefits, free medical attention, EPF, training costs and other labour costs
24	Total Expenditure (RM ‘000)	The sum of cost of raw, admin, utility, operating expenses, non-operating expenses, salaries, compensation of the employee

4 Results and Discussion

4.1 Descriptive Analysis

Descriptive analysis is a crucial initial step in data science research, offering insights into data distribution, outlier detection, and variable relationships. The dataset comprises 14,641 manufacturing companies with 24 economic factors. Descriptive analysis revealed that the majority (11,755 companies) are private limited companies. Other categories include individual proprietorship (1,979 companies), partnership (752 companies), public limited companies (148 companies), cooperative (5 companies), and public corporation (2 companies). Table 2 presents the categorization of manufacturing companies based on SME CORP Malaysia’s criteria, including micro, small, medium, and large companies (Fig. 2).

The exploration of the dataset involved identifying 67 different industries in the economic census dataset. Figure 3 illustrates the distribution of manufacturing companies based on the products they produce, indexed by a unique industry code. Additionally, Table 3 highlights the top five products manufactured in Malaysia, while Table 4 lists the least five products.

Table 2. Categorization of manufacturing companies

Category	Malaysian SME CORP Criteria	Number of companies	Relative %
Micro	Sales turnover: < RM300,000	1806	12.34
	Employees: < 5	1605	10.96
Small	Sales turnover: RM300,000 < 15 mil	8,107	55.37
	Employees: From 5 to < 75	8,677	59.27
Medium	Sales turnover: RM15 mil < = RM 50 mil	2,480	16.94
	Employees: From 75 to < 200	2,535	17.31
Large	Sales turnover: > RM50 mil	2,248	15.35
	Employees: > 200	1,824	12.46

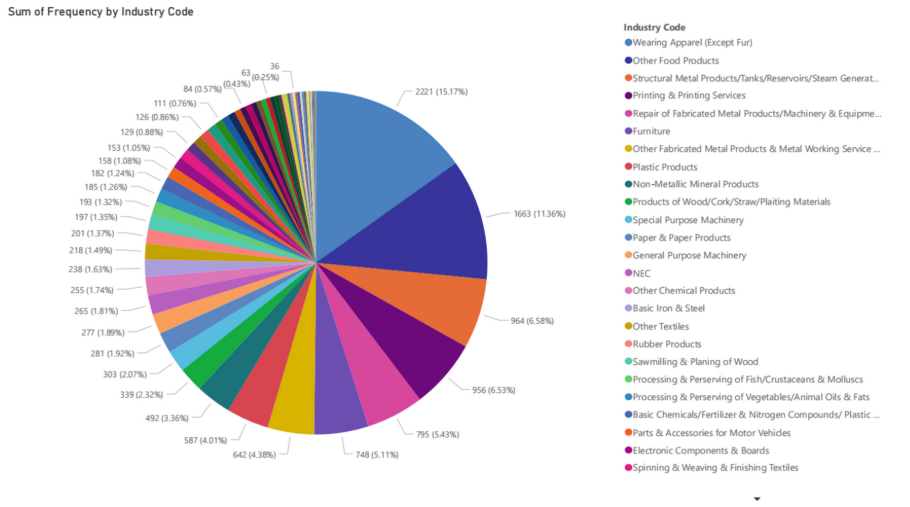


Fig. 2. Distribution of manufacturing companies based on the products

Figures 4 and 5 depict the various types of expenses incurred by manufacturing companies. The top two contributors to the overall expenses are the cost of raw materials and compensation to employees. The cost of raw materials, with a median value of RM 2,863,250, constitutes nearly 50% of the total expenditure of RM 5,833,340. Compensation to employees, totaling RM 976,340, accounts for approximately 16% of the total expenditure. This analysis underscores the significant roles that these two components play in managing the overall expenditures of manufacturing companies.

In the comparison of total salaries and wages, it is evident that production and operative workers receive the highest overall payments among the three occupations in the dataset. However, when considering salaries and wages per capita, managerial and professional workers receive higher pay rates. The median total salaries and wages for managerial and professional workers are RM180,000, with a median count of 3 for this

Table 3. Top Five manufacturing sector in the dataset

Company type	Number
Wearing apparel (except fur)	2221
Other food products	1663
Structural metal products/tanks/reservoirs/steam generators	964
Printing and printing services	956
Repair of fabricated metal products/ machinery and equipment	795

Table 4. Least Five industries in the dataset

Company type	Number
Musical instruments	3
Magnetic and optical media	3
Articles of fur	3
Man-made fibres	4
Reproduction of recorded media	6

group. In contrast, production and operative workers have a median total of RM292,570, with a median count of 16. Regarding employee qualifications, a majority possess SPM or STPM or equivalent qualifications and below.

After conducting descriptive analytics, k-means clustering and FCM algorithms were employed to identify similarity patterns among manufacturing companies. The dataset underwent clustering with both algorithms, ranging from 2 to 9 clusters, and their cluster quality was compared to determine the optimal configuration. Table 5 displays the evaluation of cluster quality, and based on the highest Silhouette coefficient score, Fuzzy C-means clustering with the min-max normalization method demonstrated the highest quality of clustering (silhouette coefficient score of 0.723) with 6 clusters.

4.2 Are There Any Patterns that Distinguish Manufacturing Companies Based on Economic Indicators?

This study employed a data mining and machine learning approach to analyze manufacturing economic census data, identifying six distinct patterns among manufacturing companies. Following this, a deeper investigation into each cluster was conducted using descriptive analytics to better understand each category. The subsequent sections provide descriptive names for each cluster, characterizing them based on the types of businesses they encompass. Based on their economic characteristics and values, each of

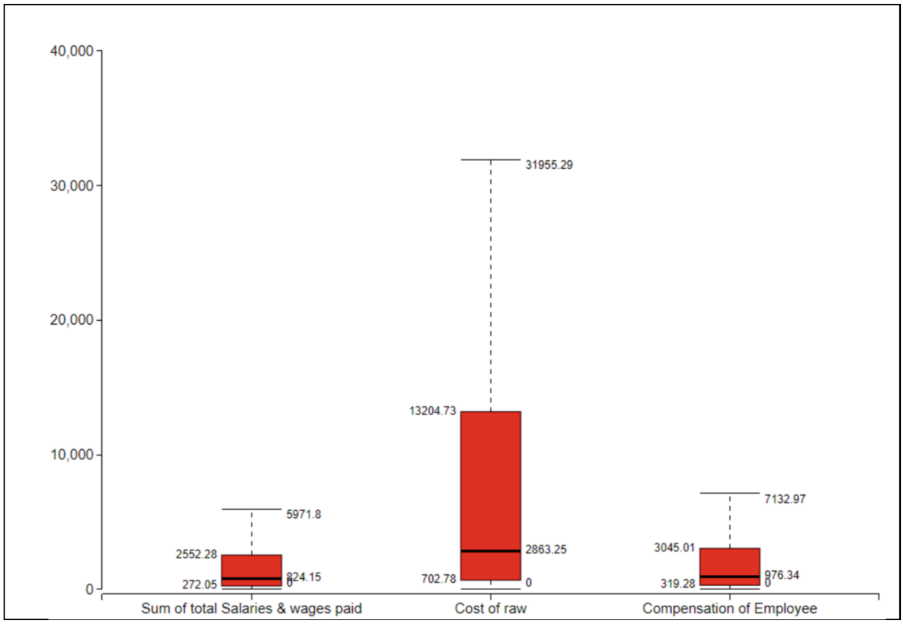


Fig. 3. Various types of expenses spent by the manufacturing companies

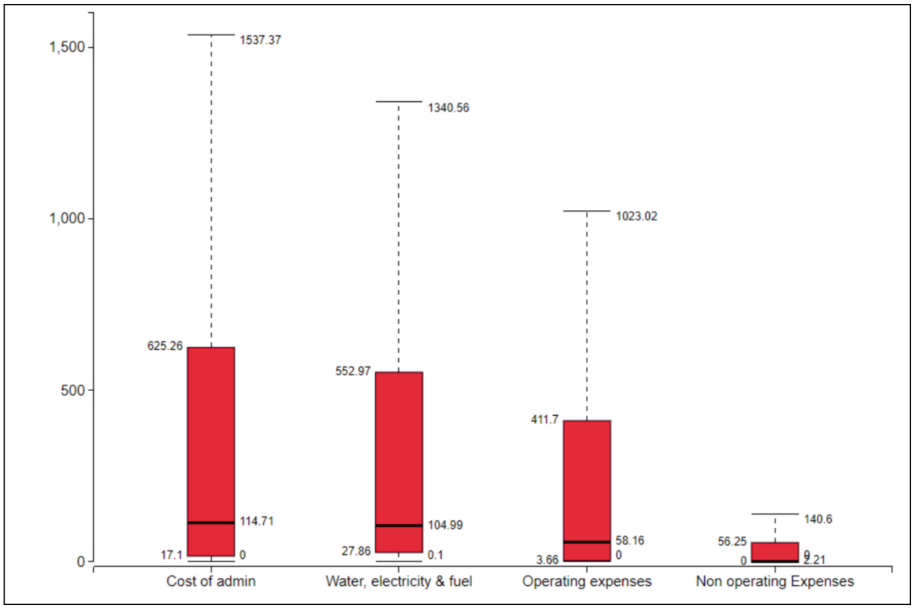


Fig. 4. Various types of expenses spent by the manufacturing companies

Table 5. Evaluation of Cluster quality and identification of optimal clusters

Number of clusters	K-means clustering		Fuzzy C-means clustering	
	Min-Max normalization Silhouette Coefficient	Z-score normalization Silhouette Coefficient	Min-Max normalization Silhouette Coefficient	Z-score normalization Silhouette Coefficient
2	0.479	0.938	0.644	0.272
3	0.492	0.357	0.575	0.723
4	0.561	0.390	0.689	0.318
5	0.578	0.251	0.497	0.372
6	0.549	0.314	**0.723**	0.417
7	0.561	0.337	0.451	0.215
8	0.537	0.377	0.683	0.389
9	0.506	0.384	0.683	0.398

these segments displayed similarities within the segments and disparities between the segments.

1. Micro (Clothing) - 2,240 firms
2. Small (Food & Beverages) -3,257 firms
3. Small (Fabricated Metal & Furniture)- 2,071 firms
4. Small (Paper & Printing)- 2,264 firms
5. Medium (Machine & parts)-1,191 firms
6. Medium (Heavy duty Metal & Plastic)- 3,618 firms

4.3 What Are the Characteristics of the Identified Segments of the Manufacturing Companies?

This study adopted an unsupervised data mining approach to segment manufacturing companies, departing from conventional categorization methods reliant on annual sales, employee count, or product types. The results reveal six distinct groups of manufacturing companies sharing similar economic characteristics. Further details about the characteristics of each segment are provided below.

Micro (Clothing) - 2,240 Firms. In this sector, 82.5% of businesses operate as sole proprietorships, with 69.6% specializing in clothing (excluding fur) production. These businesses exhibit the lowest net book values at both the beginning (RM 43,002) and end of the year (RM 38,892) among the six segments. The workforce in this sector predominantly holds SPM qualifications or lower, resulting in the lowest overall salary and wages paid. Notably, some businesses in this category have negligible non-operating expenses. Additionally, they allocate the least expenditure to administrative costs, salaries and wages, utilities, and employee pay, among other cost categories.

Small (Food & Beverages) - 3,257 Firms. This segment comprises primarily private limited companies (92.9%), predominantly engaged in producing various products, with other food products being the majority at 45.2%. In contrast to the micro (clothing) segment, these companies exhibit a higher total net book value at the beginning and end of the year, amounting to RM862,814 and RM824,277, respectively. Unlike the micro (clothing) segment, these companies hire managerial-level staff with university degrees and have a larger workforce, totaling 24 employees. This trend is also reflected in the total compensation of employees and overall expenditure.

Small (Fabricated Metal & Furniture) - 2,071 Firms. This segment is predominantly composed of private limited companies (88.8%) specializing in the repair of fabricated metal products/machinery & equipment (38.3%) and furniture production (36.11%). In comparison to the small(F&B) segment, companies in this category allocate slightly higher expenditures across various categories, except for the cost of water, electricity & fuel. This spending pattern is rationalized by the specific nature of products manufactured by the small(F&B) segment.

Small (Paper & Printing) - 2,264 Firms. This segment predominantly comprises companies manufacturing paper & printing-related products (65.9%). In contrast to the Small (Fabricated Metal & Furniture) segment, these companies exhibit a higher total net book value at the beginning and end of the year, amounting to RM2,216,795 and RM2,120,266, respectively. However, their total expenditure is considerably higher, reaching RM10,247,020, almost three times that of the Small (Fabricated Metal & Furniture) segment. The primary contributors to this higher expenditure are total salaries & wages paid (twice as high) and the cost of water, electricity & fuel (three times as high).

Medium (Machine & parts) - 1,191 Firms. This segment, the smallest among all, primarily involves companies producing machines, parts, and accessories, constituting 60.9% of the segment. In comparison to the Small (Paper & Printing) segment, these companies display nearly twice the total net book value at the beginning and end of the year, amounting to RM4,042,880 and RM3,911,848, respectively. A similar pattern is observed in the total number of managerial and professional employees, as well as total production/operative workers. This segment has the highest total employment among all six segments. The overall total expenditure is almost 1.5 times higher than the Small (Paper & Printing) segment, a trend also reflected in total income.

Medium (Heavy Duty Metal & Plastic). 3,618 firms. This is the largest segment, primarily involved in producing heavy-duty metal & plastic-related products, such as Structural Metal Products/Tanks/Reservoirs (21.90%), Other Fabricated Metal Products & Metal Working Service Activities (17.50%), and Plastic Products (15.90%). Their net book value at the beginning of the year was RM4,069,094, and at the end of the year, it was RM4,053,807, representing the highest values among all six segments. Despite lower expenses in terms of total salaries & wages paid and non-operating costs compared to the Medium (Machine & parts) segment, this segment has the highest overall total expenditure among all six segments.

5 Conclusion

This study utilized data analytics and machine learning techniques to analyze the economic census data, aiming to group Malaysian manufacturing enterprises based on their economic characteristics. Employing two unsupervised clustering algorithms, fuzzy C-means outperformed k-means in identifying similarity patterns among the companies. The study provides insights into the features of each segment, highlighting both similarities and differences among manufacturing companies. It's noteworthy that each of the six segments comprises a diverse range of companies, including micro, small, medium, and large enterprises. This emphasizes the insufficiency of relying on a single factor for segmentation when seeking similar businesses for the formulation of policies and long-term plans to enhance productivity and contribute to the country's GDP growth. This knowledge holds value for researchers, policymakers, planners, and economists, aiding in the creation of economic forecasts.

Authors' Contributions. RK-Writing-original draft preparation, Data curation, Conceptualization, and Modelling, KR-Writing-Review & Editing and Validation.

Availability of Data and Material. Data used in this study were obtained from a survey dataset "Economic Census 2016 - Manufacturing" available from the Department of Statistics, Malaysia (DOSM). A report published by the DOSM based on the survey can be viewed on the DOSM website. Access to this data requires application, as stated on the DOSM website. A guide for how to apply for dataset access is available on the Data Request page or requests for more information can be emailed to data@dosm.gov.my.

Declaration of Conflicting Interests. The authors declared no potential conflicts of interest concerning the research, authorship, and/or publication of this article.

References

1. Duke, J.E., Tapang, A.T., Usang, O., Kankpang, K.A., Etim, S.E.: High-performance work practices and entrepreneurial firm performance: the moderating role of firm size and industry type. J. Small Bus. Enterprise Dev. ahead-of-print (2024). https://doi.org/10.1108/JSBED-06-2022-0265/FULL/XML
2. Ongbali, S.O., Omotehinse, S.A., Adams, C.O., Salawu, E.Y., Afolalu, S.A.: Analysis of the key factors for small and medium-sized enterprises growth using principal component analysis. Heliyon. **10**, e33573 (2024). https://doi.org/10.1016/j.heliyon.2024.e33573
3. Chen, T., et al.: Machine learning in manufacturing towards industry 4.0: from 'for now' to 'four-know.' Appl. Sci. 2023 **13**, 1903 (2023). https://doi.org/10.3390/APP13031903
4. Corallo, A., Crespino, A.M., Lazoi, M., Lezzi, M.: Model-based big data analytics-as-a-service framework in smart manufacturing: a case study. Robot. Comput. Integr. Manuf. **76**, 102331 (2022). https://doi.org/10.1016/j.rcim.2022.102331
5. Kannan, R., Abdul Halim, H.A., Ramakrishnan, K., Ismail, S., Wijaya, D.R.: Machine learning approach for predicting production delays: a quarry company case study. J. Big Data **9**, 94 (2022). https://doi.org/10.1186/s40537-022-00644-w
6. Raut, R.D., Yadav, V.S., Cheikhrouhou, N., Narwane, V.S., Narkhede, B.E.: Big data analytics: implementation challenges in Indian manufacturing supply chains. Comput. Ind. **125**, 103368 (2021). https://doi.org/10.1016/j.compind.2020.103368

7. Cho, E., Jun, J.H., Chang, T.W., Choi, Y.: Quality prediction modeling of plastic extrusion process. ICIC Express Lett. Part B Appl. **11**, 447–452 (2020). https://doi.org/10.24507/icicelb.11.05.447
8. Dogan, A., Birant, D.: Machine learning and data mining in manufacturing (2021). https://doi.org/10.1016/j.eswa.2020.114060
9. Ye, Y., Hu, T., Nassehi, A., Ji, S., Ni, H.: Context-aware manufacturing system design using machine learning. J. Manuf. Syst. **65**, 59–69 (2022). https://doi.org/10.1016/j.jmsy.2022.08.012
10. Mohamed, E., Celik, T.: Early detection of failures from vehicle equipment data using K-means clustering design. Comput. Electr. Eng. **103**, 108351 (2022). https://doi.org/10.1016/j.compeleceng.2022.108351
11. Kannan, R., Wang, I.Z.W., Ong, H.B., Ramakrishnan, K., Alamsyah, A.: COVID-19 impact: customised economic stimulus package recommender system using machine learning techniques. F1000Res. **10**, 932 (2021). https://doi.org/10.12688/f1000research.72976.2
12. Fa'rifah, R.Y., Pramesti, D.: Cluster analysis of inclusive economic development using K-means algorithm. Jurnal Varian **5**, 171–178 (2022). https://doi.org/10.30812/varian.v5i2.1894
13. Yang, M., Huang, L., Tang, C.: K-means clustering with local distance privacy. Big Data Mining Analytics **6** (2023). https://doi.org/10.26599/BDMA.2022.9020050
14. Kotsiopoulos, T., Sarigiannidis, P., Ioannidis, D., Tzovaras, D.: Machine learning and deep learning in smart manufacturing: the smart grid paradigm (2021). https://doi.org/10.1016/j.cosrev.2020.100341
15. Hashemi, S.E., Gholian-Jouybari, F., Hajiaghaei-Keshteli, M.: A fuzzy C-means algorithm for optimizing data clustering. Expert Syst. Appl. **227** (2023). https://doi.org/10.1016/j.eswa.2023.120377
16. Ais, C., Hamid, A., Candra, D., Novitasari, R.: Analysis of livestock meat production in Indonesia using fuzzy C-means clustering. Jurnal Ilmu Komputer dan Informasi (J. Comput. Sci. Inf.) **15**, 1–8 (2022). https://doi.org/10.21609/jiki.v15i1.993
17. Wang, B., Tan, Y., Jia, W.: TL-FCM: a hierarchical prediction model based on two-level fuzzy c-means clustering for bike-sharing system. Appl. Intell. **52**, 6432–6449 (2022). https://doi.org/10.1007/s10489-021-02186-9
18. Dedić, A., Konjić, T., Ćalasan, M., Dedić, Z.: Fuzzy C-means clustering applied to load profiling of industrial customers. Electr. Power Compon. Syst. **49**, 1068–1084 (2021). https://doi.org/10.1080/15325008.2022.2049660
19. Valipour, M., Yousefi, S., Jahangoshai Rezaee, M., Saberi, M.: A clustering-based approach for prioritizing health, safety and environment risks integrating fuzzy C-means and hybrid decision-making methods. Stoch. Env. Res. Risk Assess. **36**, 919–938 (2022). https://doi.org/10.1007/s00477-021-02045-6
20. Wan, T., Tian, S.: digital transformation in landscape design of rural characteristic town based on big data technology. Comput. Aided Des. Appl. **21** (2024). https://doi.org/10.14733/cadaps.2024.S2.19-38
21. El Khattabi, M.Z., El Jai, M., Lahmadi, Y., Oughdir, L., Rahhali, M.: Understanding the interplay between metrics, normalization forms, and data distribution in K-means clustering: a comparative simulation study. Arab. J. Sci. Eng. **49**, 2987–3007 (2024). https://doi.org/10.1007/S13369-023-07741-9/FIGURES/13

Smart Choices: How Computer Applications Shape Consumer Decisions on Household IoT Appliances

R. Chitra[1](✉), M. Sathish[2], and Jone Stervin[1]

[1] PSG Institute of Management, PSG College of Technology, Coimbatore, Tamil Nadu, India
chitra@psgim.ac.in

[2] Amrita School of Business, Amrita Vishwa Vidyapeetham, Coimbatore, Tamil Nadu, India

Abstract. This study explores how people in India buy smart gadgets for their homes. These gadgets, part of the Internet of Things (IoT), offer benefits like convenience and safety. Understanding why people choose these products is essential for companies selling them. The research looks at factors like consumer awareness, opinions, and what influences their buying choices. We surveyed people in Coimbatore to learn about using IoT products like innovative security systems and fitness devices. Regression analysis also indicated that all the independent variables significantly correlated with the dependent variable, Consumer Behavior; IoT had the highest regression coefficient value ($\beta = 0.411$, $p < .001$). The findings of this research from the SEM analysis provide evidence of a positive association between IoT adoption, Industry 4.0 revolution (IR), Trust and Privacy Concerns (TPC), and Consumer Relationship Management (CRM) with consumer behavior.

Keywords: Consumer behavior · Internet of things · Industry 4.0 revolution

1 Introduction

The Internet of Things (IoT) has completely changed the consumer landscape by introducing new home furnishings that enhance comfort, energy efficiency, and security (Atzori et al., 2010). Some smart home technologies are slowly becoming popular in Coimbatore, a rapidly developing city in Tamil Nadu, India, known for its strong manufacturing and captively emerging information and technology industries. However, factors that affect the Consumers in this market and the role played by Computer Applications in their decisions are unexplored. Computer applications are also integral to IoT and their operation because they offer users options to control the devices, monitor them, automate them, and tailor them to their preferences (Maddulety et al., 2017). The utility of these applications can affect consumer satisfaction and buying behaviour (TEO et al., 2021; Emami-Naeini et al., 2019). This research thus establishes the impact of these applications on consumers in Coimbatore and the factors that may enhance the use of household IoT appliances. Thus, this research aims to determine the understanding of technology and its implications for consumers' buying behaviour when purchasing IoT appliances in Coimbatore. Overall, these findings will help increase consumers' understanding and provide developers and marketers with recommendations for the design of

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
S. Rajagopal et al. (Eds.): ASCIS 2024, CCIS 2428, pp. 276–288, 2025.
https://doi.org/10.1007/978-3-031-86302-8_18

appealing IoT solutions for this expanding market. The following research questions are formulated.

1. What is the level of consumer awareness and perception of household IoT products in Coimbatore District?
2. What are the primary factors influencing consumers' purchasing decisions for IoT products in their households in Coimbatore District?
3. What are the adoption and usage patterns of different types of household IoT products among consumers in the Coimbatore District?

2 Review of Literature

2.1 Consumer Behaviour

Chatterjee (2019) investigated the effects of the Internet of Things (IoT) and social media on consumer behavior amid an Indian consumer behavior model. The study offered an empirical analysis and insight into the impact of the Internet of Things (IoT) and social platforms on consumer behaviour in India. Keng et al. (2023) discussed the involvement of the customer and the intelligent objects and the impact of IoT on consumers. This work provided insights into interaction dynamics between consumers and IoT-enabled devices, exposing the landscape of emerging consumer experiences. Rock et.al, (2024) wrote about how the IoT is revolutionary, spearheading changes in consumers' perceptions, contact and even experiences with products and services in a ubiquitous world. Sima et al. (2020) analyzed the effects of the Revolution Industry 4.0 on human capital development and consumer behaviour. The information provided allowed a comprehensive understanding of advancements in technology, including IoT, and how it is intertwined with changes in consumer behaviour. Schuster & Habibipour, (2024) conducted a detailed study focused on leveraging IoT to analyze consumer behaviour concerning the home. The research used IoT technologies to understand Consumer preferences and behaviour related to household appliance consumption patterns. This article highlighted the potential of IoT in discovering tendencies and patterns in consumer behaviour; such understanding is essential for businesses to offer goods more appropriately.

2.2 Internet of Things

The increased use of IoT technology in smart homes has encouraged various studies involving consumer behaviour and perceptions towards various innovations. Tsourela & Nerantzaki (2020) offer a basic model for consumer behaviour regarding IoT products, focusing on acceptance factors, including perceived usefulness and ease of use. Similarly, (Aldossari & Sidorova, 2020) focuses on the context of smart homes to explore the influence of perceived security and privacy on IoT device acceptance. Hence, these studies focus on two aspects of technology acceptability, trust and perceived benefits, in the context of IoT in home applications. Rock et al. (2024) further extend the discussion by exploring the effects of smart home IoT technology on quality of life, suggesting that such innovations bring added convenience, security, and generally improved quality of life into the context of homes.

Besides acceptance and impact, other areas have been explored in literature, including the interactivity between the consumer and the smart objects. Keng et al. (2023) discuss how IoT transforms consumption behaviour, focusing on customer0object interactions. In line with this understanding, Htun et al. (2021) used a consumer behaviour analyzer in the IoT environments to study consumer behaviour with technologies. Finally, Philip et al. (2023) analyze user intentions concerning the protection of IoT smart home networks and emphasize the role of security measures in building trust and promoting further use. Collectively, these papers present a systematic understanding of how consumption is being by IoT with a focus on acceptance constructs and their practical implications in the context of smart homes.

2.3 Industry 4.0 Revolution

The Internet of Things (IoT), as applied to the areas of Industry 4.0, is modifying industrial processes, human resource development and consumer behavior. Rath et al. (2024) elucidate how IoT technologies aim to improve enterprises' operational performance and economic efficiency during the Industrial Revolution. This is further supported by Aheleroff et al. (2020), who, through a case study, illustrate how IoT technology in intelligent appliances makes the process more functional and stable in industrial processes. There is a broader impact of IoT beyond industries that focus on IoT applications in health care, agriculture, smart homes, smart cities, and industrial 4.0, including the future vision of the technology in this area (Chataut et al., 2023). Besides the operational improvements, IoT is also important in human capital and consumer behaviour. Apart from the skill deficiencies of the workers and the unmet needs of the consumers due to the lack of technological adoption, the study by Sima et al. (2020) shows how Industry 4.0 is revolutionizing these sectors. Sarangapani et al. (2021) illustrate delta PLC-based smart home automation in rural development with the wise application of Industry 4.0 technologies to improve society and the economy. Together, these studies highlight IoT's transformative role in Industry 4.0, improving operational efficiencies, consumer behaviours, and human capital development while pointing to IoT's expansive potential in other industries.

2.4 Trust and Privacy Concerns

Trust, privacy, and security issues highly influence consumer acceptance of IoT technologies within smart homes. Cannizzaro et al. (2020) point out that people's concern about data privacy is a significant factor in why they are reluctant to embrace smart home devices, as their survey in the United Kingdom indicates. Guhr et al. (2020) also expand on privacy and state that in smart homes, privacy threats overshadow the various utilities of intelligent IoT devices. Such privacy issues may be exacerbated by the threats and risks within the IoT architecture, as explained by Tewari & Gupta (2020), who describe how risk factors at multiple layers of IoT systems contribute to limited trust among consumers. Lee (2020) explains how privacy and security issues are among the factors that cause people to resist the use of IoT in the home, and the concern is the perceived weakness of getting exposed.

The role of trust in IoT adoption is further explored by Jaspers & Pearson (2022), who proposed that the most critical problem is to overcome the perceived privacy threats to enhance consumers' acceptance of domestic IoT products. Based on the literature, they argue that trust plays the role of the moderator between the perceived risks and benefits of IoT technologies. This echoes Schuster and Habibipour (2024), where privacy and security issues remain the significant challenges stifling IoT, with particular reference to home settings. Sah & Jun (2024) build on this by showing how such changes in privacy awareness translate into changes in consumer behaviour; this is often characterized by higher levels of caution whenever consumers consider adopting IoT services. Last, Alhandi et al. (2023) put forward a trust evaluation model that acknowledges trust as a significant factor in mitigating privacy and security, thereby creating a favorable environment for IoT usage. Altogether, these studies show the complex interdependence between trust, privacy, and security issues in adopting IoT technology in smart homes, highlighting the importance of practical solutions to mitigate these issues.

2.5 IoT-Based Customer Relationship Management

Internet of Things (IoT) technologies integrated with Customer Relationship Management (CRM) systems are gradually altering the features of data-driven marketing by enabling more personalized and efficient customer interactions. Hashem (2021) addresses how IoT creates business data marketing possibilities through improved CRM to provide real-time data for analyzing customer needs. This perspective is supported by Abu & Zabadi (2020), who argue that integrating IoT and Big Data can significantly enhance conventional CRM systems through a better understanding of customer behaviours, ultimately fostering improved decision-making and efficient. Marketing. The study emphasizes assessing these IoT-enabled CRM innovations to improve marketing performance.

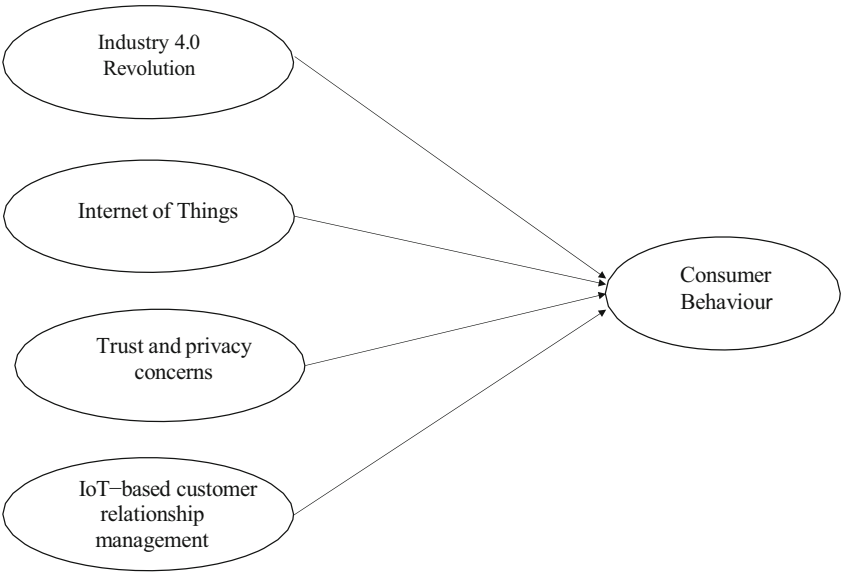
Eslami et al. (2024) provided a practical application on how IoT data creates more accurate customer segmentation and redefined CRM strategies. Their case study also shows that IoT can help improve the existing CRM practices as companies gain insight into customer behaviour, positively impacting customer satisfaction and loyalty. Similarly, Tariq et al. (2020) outline the broader opportunities of IoT technologies in marketing to demonstrate how IoT insight-driven marketing leads to a higher marketing impact through effective customer relationships. Bajaj et al. (2023) also discuss the antecedents of IoT in the context of CRM, with specifics as to the banking industry identifying that gender, age and type of ownership will all mediate the relationship between IoT and CRM performance. Altogether, these papers demonstrate the transformative impact of IoT in improving CRM systems and providing the company with new opportunities for customer acquisition and marketing success.

3 Research Gap

Numerous studies focused on the effectiveness of IoT in influencing consumers in the Indian context, the issue of privacy in adopting IoT solutions, and trust in intelligent manufacturing within Industry 4. 0 technologies such as IoT influence consumers to

change their preferences and experiences. Various researchers have proven the relationship between IoT, Industry 4.0, and operational efficiency. However, the deep association between IoT-driven industry automation and changes in consumer behaviour requires further research**.** The literature has focused on the influence of IoT on consumers' interactions, trust, privacy, and the incorporation of intelligent technologies in retail and home settings, encompassing different global settings**.** However, these studies concentrated on trust and privacy in different contexts and industries. It will be exciting to explore consumer awareness and perception of IoT concerning household products, how this influences the purchase decision from the perspective of trust and privacy. Most of the studies are done on developed markets with highly technical environments. Significant research should address how IoT influences consumer behaviour in emerging markets like India, such as Coimbatore, named Manchester of South India. Moreover, Coimbatore is the leading IT hub and well-developed manufacturing market with different customer profiles. Even though several studies highlighted the IoT's potential to improve personalized customer engagement. Nevertheless, there is limited empirical evidence on how IoT adoption on customer relationship management impacts consumer behaviour. Such a gap creates a prospect for exploring the regional nuances of IoT adoption and how consumer behaviour associated with regional economy and culture affects the Internet of Things technologies within the household sector**.** Based on the various research gaps, the research framework is developed to test how factors like Industry 4.0 evolution, IoT and its applications, trust, and privacy on consumer decisions, and IoT-based customer relationships impact consumer behaviour on household products.

4 Research Framework



Hypothesis

H1: There is a significant relationship between the Industry 4.0 revolution and consumer buying behavior.

H2: There is a significant relationship between the Internet of Things and consumer buying behavior.

H3: There is a significant relationship between Trust and privacy concerns and consumer buying behavior.

H4: There is a significant relationship between IoT-based customer relationship management and consumer buying behavior.

Research Objective

- To understand the level of consumer awareness and perceptions of household IoT products in Coimbatore District.
- To analyze the primary factors influencing consumers purchasing decisions for IoT products in their households.
- To analyze the adoption and usage patterns of different types of household IoT products among consumers in the region.

5 Research Methodology

Given their technological exposure, this study employs a descriptive research design to explore consumer buying behaviour towards household IoT products, primarily focusing on Gen Z. The sampling technique used is Stratified Random Sampling with a sample size of 1230. The research instrument was developed based on literature reviews and expert opinions. A pilot study was conducted with a sample of 40 respondents to check the instrument's reliability and validity; the results are given below (Table 1).

Table 1. Reliability Analysis

ITEM	CRONBACH'S ALPHA	No. of Items
CB	.812	5
IoT	.817	5
IR	.865	5
TPC	.765	5
CRM	.752	2
Overall	.794	22

Cronbach's alpha value for all the variables Consumer Behaviour, Internet of things, Industry 4.0 Revolution, Trust and Privacy Concerns, and IoT-based Customer Relationship Management are >0.7. Additionally, the overall Cronbach's alpha value is 0.794 for the 22 questions indicating a good level of internal consistency. These results show that measurement properties are reliable, providing confidence in the robustness of data collected from the consumers. This in turn enhances the credibility of the study's findings, ensuring that the responses gathered truly reflect the intended aspects of consumer

behaviour. Cronbach's alpha values for all the variables, such as consumer behaviour, Internet of things, Industry 4.0 Revolution, trust and privacy concerns, and IoT-based customer relationship management, are >0.7. Also, Cronbach's alpha value is 0.794 for the 22 questions, indicating good consistency. These results show that measurement properties are reliable, providing confidence in the robustness of data collected from the consumers. This enhances the credibility of the study's findings, ensuring that the responses gathered truly reflect the intended aspects of consumer behaviour.

6 Results and Discussions

Structural Equation Modeling (SEM) is one of the most powerful statistical tools to analyze multiple variables and assess the hypotheses about these variables. In this study, the antecedents to consumer behaviour within the IoT context have been examined using SEM as the theoretical framework. Many fit indices were used for the evaluation of the model, and these entailed Average Block VIF (AVIF) = 3.751, Average Full Collinearity VIF (AFVIF) = 3.667, Tenenhaus GoF (GoF) = 0.730 (Fig. 1 and Table 2).

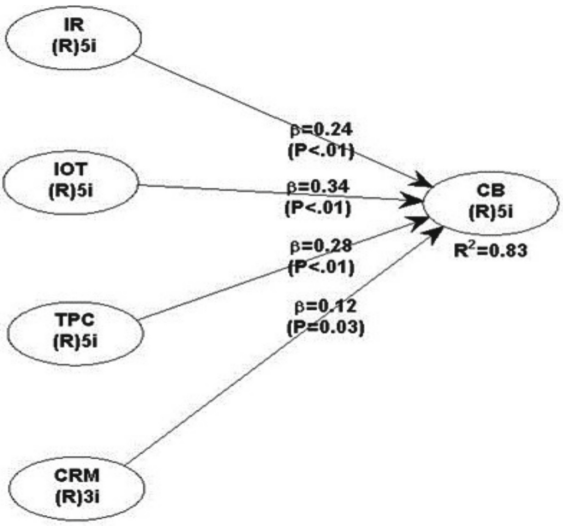


Fig. 1. Structural Equation Modeling

Table 2. Model Fit and Quality Indices

AVIF	AFVIF	GoF	SPR	RSCR	SSR	NLBCDR
4.036	4.621	0.759	1.000	1.000	1.000	1.000

Average block VIF (AVIF) = 4.036, acceptable if <= 5, ideally <= 3.3. Average full collinearity VIF (AFVIF) = 4.621, acceptable if <= 5, ideally <= 3.3 Tenenhaus GoF

(GoF) = 0.759, small >= 0.1, medium >= 0.25, large >= 0.36 Simpson's paradox ratio (SPR) = 1.000, acceptable if >= 0.7, ideally = 1 R-squared contribution ratio (RSCR) = 1.000, acceptable if >= 0.9, ideally = 1 Statistical suppression ratio (SSR) = 1.000, acceptable if >= 0.7 Nonlinear bivariate causality direction ratio (NLBCDR) = 1.000, acceptable if >= 0.7 (Table 3).

Table 3. Correlation Factors Between the Variables

	CB	IOT	IR	TPC	CRM
CB	0.805				
IOT	0.86	0.838			
IR	0.852	0.85	0.844		
TPC	0.84	0.822	0.834	0.813	
CRM	0.767	0.736	0.771	0.736	0.866

The discriminant validity was found by comparing the square root of AVE of each construct with a correlation among the other latent variables, which was found to be greater than the correlations. The above table presents the output for discriminant validity for Consumer behaviour (0.805), Internet of things (0.838), Industry 4.0 revolution (0.844), Trust and privacy concerns (0.813), and Customer relationship management (0.866) (Table 4).

Table 4. Regression Analysis

Model	R	R^2
1	0.908	0.825

Model Fit Measures

The model fit measures explain about the strong relationship between independent variables (Industry 4.0, Internet of Things, Trust and privacy and IoT based customer relationship) and dependent variable (Consumer Behaviour). R2 value explains the same with quantitative value of 0.825. Hence, the model measures 82.5% of relationship between these variables. This the clear evidence to show to robustness of the research. Nevertheless, this is clear evidence of model fitness towards measuring the research question and objective.

Model Coefficients – Consumer Behaviour

Predictor	Estimate	SE	t	P
Intercept	0.173	0.1138	1.52	0.130
IOT	0.411	0.0528	7.79	<.001
IR	0.148	0.0392	3.78	<.001
TPC	0.286	0.0532	5.39	<.001
CRM	0.115	0.0478	2.41	0.017

Among the independent variables IoT is associated very strongly to the consumer behaviour (t value = 7.79). Secondly, trust and privacy are influencing and positive variable towards the consumer behaviour (t value = 5.39). Industrial revolution is the third influencing factors towards consumer behaviour (t value = 3.78). IoT-based customer relationship is last in this ranking (t = 2.41). But the interesting factors all these variables are positively associated with consumer behavior.

Internet of Things (IoT): The value for IoT is 0.411 significant. Also, this is positive and statistically significant (p-value <.001). So, the empirical relationship between the IoT and the outcome variable can be assessed using this as a regression model. To sum up, the coefficient clearly states that each one-unit uptake of IoT is expected to lead to the independent variable increasing by 0.411 units on average while maintaining other things constant.

Industry 4.0 Revolution (IR): The coefficient for IR is (0.148), and the p-value is <.001, which means this coefficient is positive and significant (p-value <.001). Currently, the Industry 4.0 conversion depends on the highly positive correlation variable. More specifically, an increase in the magnitude of the Industry 4.0 revolution's effect by one unit will impact the dependent variable in the amount of 0.148 units, provided other variables are all equal.

Trust and Privacy Concerns (TPC): The coefficient for TPC (0.286) is positive and statistically significant (p-value <.001). This thus allows for a strong relation between internet users' level of trust and privacy fears and the outcome variable. Nevertheless, this is a surprising result since a very opposite tendency might be assumed (that is, the higher concerns are, the less often the consumers use these products). There can be an association of adverse effects, suppression may be the way, or it can be one of the other reasons for this connection. Yet comprehensive studies would be needed to comprehend this aspect as a whole. Customer Relationship Management (CRM): CRM has a (+) kg coefficient that is statistically significant at the 0.115 level (p-value = 0.017). It translates into a positive relationship between Customer relation management and the dependent performance, the CRM. In essence, for any one-unit adjustment in customer relationship management realization, the independent variable is anticipated to increase on average by 0.115 units, leaving each of the other elements intact.

This study delves into people's buying behaviour in Coimbatore District, India, specifically regarding smart home devices. The research reveals a growing interest in these Internet of Things (IoT) products, driven by their ability to enhance convenience, save energy, and improve home security. This trend aligns perfectly with the emphasis on interconnected and intelligent devices that characterize Industry 4.0. However,

a significant hurdle to adoption remains – trust and privacy concerns. Consumers are understandably cautious about data security and the potential misuse of personal information these devices collect. Companies can address this challenge by building trust through robust data security measures and transparent practices regarding data collection and usage. Additionally, the research highlights the potential of IoT-based customer relationship management (CRM). By leveraging the data collected from IoT devices, companies can personalize recommendations and proactively anticipate customer needs, increasing satisfaction and loyalty. While the study offers valuable insights, it's essential to acknowledge its limitations. The research focused solely on Coimbatore District, and the data collection relied on surveys, which might not capture the entire picture. Nevertheless, the findings provide a clear direction for companies in this growing market. By understanding consumer preferences, prioritizing data privacy, and implementing effective CRM strategies, companies can successfully navigate the landscape of smart home technology and win over customers in Coimbatore District and potentially beyond.

7 Implications of the Study

The study provides vital practical implications for businesses, policies, and technologies in the context of IoT. Organizations should strive to adopt IoT technologies into their products and services by catering to the growing consumer need for connectivity and better IoT solutions to gain a competitive advantage in the market. For instance, smart home devices, wearable devices, and IoT customer service applications can significantly enhance customer satisfaction and loyalty. Product developers should prioritize creating user-friendly, secure, and reliable IoT devices to meet consumer demand and expand the market. Besides, it is noteworthy that Industry 4.0 improvement in consumer behaviour means that firms must seek to improve and introduce metamorphosed ways of working such as automation, artificial intelligence, and analytics. Such technologies can optimize processes, increase product quality, provide companies with competitive advantages by applying new technologies, and attract customers interested in employing high-tech solutions.

It is also essential for businesses to know how to deal with trust and privacy if they want to tap the IoT market. Thus, the research shows that companies need to enhance consumer trust by investing in highly secure data privacy and offering clear information about privacy policies and the power of control over their data. This approach not only assures the consumers but also creates the tendency to use IoT products in the future. In this regard, policymakers will need to design and implement policies that uphold consumer privacy while at the same time encouraging the development of higher adoption rates. Additionally, enhancing the CRM is crucial for strengthening consumers' trust. Businesses should rely on effective and efficient CRM approaches and tools in the form of data-driven and responsive CRM approaches to serve as an active communication network for continuous engagement through various touchpoints. Strategizing and implementing CRM can help cultivate long-standing interest from its customers in IoT products and services. Marketing and communication tactics should aim to directly inform consumers about the values of IoT technologies, including privacy and security questions. By espousing a holistic approach that integrates technological

innovation, trust-building measures, and customer-centric strategies, businesses can tap into the potential IoT market and improve business–consumer relationships to enhance business throughout the digital realm.

8 Conclusion

This paper comprehensively investigates various factors that define consumer behaviour in IoT. Employing Structural Equation Modeling (SEM) examines the relationship through which IoT adoption and advancements from Industry 4. 0, trust and privacy, and customer relationship management (CRM) each play critical roles in shaping Consumer Behaviour. The research provides an understanding of the interrelated dynamics between technological advancement, consumer trust, and CRM, acknowledging privacy and trust issues while leveraging technological trends and CRM best practices to enhance consumer engagement.

The study also challenges conventional views on the impact of privacy concerns on technology adoption, finding that consumers aware of the privacy issue on IoT technologies are willing to embrace IoT, given that their privacy concerns are addressed. This insight offers new avenues for future research or practice in technology adoption. Therefore, businesses, policymakers, and technology developers must follow a holistic approach that combines innovation, trust, and personal CRM strategies to capture the growing IoT market in various industries. These limitations indicate that further work is required to explore these dynamics better by considering more variables and diverse and longitudinal samples. Further, future studies could be explored on similar cross-sectional studies on urban and rural consumers in emerging countries like India in terms of how different socio-demographics influence the use of IoT. Cross-cultural research may provide insights into the understanding of how cultural differences influence consumer's perception and behaviour towards IoT technologies. This study expands the knowledge of IoT adoption and consumer behaviour. It provides meaningful insights and recommendations for future scholars and practitioners for decision-making in the evolving digital landscape.

References

Abu Ghazaleh, M., Zabadi, A.M.: Promoting a revamped CRM through Internet of Things and big data: an AHP-based evaluation. Int. J. Organ. Anal. **28**(1), 66–91 (2020)

Aheleroff, S., et al.: IoT-enabled smart appliances under industry 4.0: a case study. Adv. Eng. Inf. **43**, 101043 (2020)

Aldossari, M.Q., Sidorova, A.: Consumer acceptance of Internet of Things (IoT): smart home context. J. Comput. Inf. Syst. **60**(6), 507–517 (2020)

Alhandi, S.A., Kamaludin, H., Alduais, N.A.M.: Trust evaluation model in IoT environment: a comprehensive survey. IEEE Access **11**, 11165–11182 (2023)

Atzori, L., Iera, A., Morabito, G.: The internet of things: a survey. Comput. Netw. **54**(15), 2787–2805 (2010)

Bajaj, P., Anwar, I., Yahya, A.T., Saleem, I.: Factors influencing adoption of IoT and its impact on CRM in banks: examining the moderating role of gender, age, and bank ownership type. Hum. Behav. Emerg. Technol. **2023**(1), 5571508 (2023)

Cannizzaro, S., Procter, R., Ma, S., Maple, C.: Trust in the smart home: findings from a nationally representative survey in the UK. PLoS ONE **15**(5), e0231615 (2020)
Chataut, R., Phoummalayvane, A., Akl, R.: Unleashing the power of IoT: a comprehensive review of IoT applications and future prospects in healthcare, agriculture, smart homes, smart cities, and industry 4.0. Sensors **23**(16), 7194 (2023)
Chatterjee, S.: Internet of Things and social platforms: an empirical analysis from Indian consumer behavioural perspective. Behav. Inf. Technol. **39**, 1–17 (2019). https://doi.org/10.1080/0144929X.2019.158700
Emami-Naeini, P., Dixon, H., Agarwal, Y., Cranor, L.F.: Exploring how privacy and security factor into IoT device purchase behavior. In: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, pp. 1–12, May 2019
Eslami, E., Razi, N., Lonbani, M., Rezazadeh, J.: Unveiling IoT customer behaviour: segmentation and insights for enhanced IoT-CRM strategies: a real case study. Sensors **24**(4), 1050 (2024)
Guhr, N., Werth, O., Blacha, P.P.H., Breitner, M.H.: Privacy concerns in the smart home context. SN Appl. Sci. **2**, 1–12 (2020)
Hashem, T.N.: The reality of internet of things (IoT) in creating a data-driven marketing opportunity: mediating role of customer relationship management (CRM). J. Theor. Appl. Inf. Technol. **99**(2), 329–342 (2021)
Htun, S.N.N., Zin, T.T., Tin, P.: Consumer behavior analyzer in Internet of Things (IoT) environments. Int. J. Innov. Comput. Inf. Control **17**(1), 345–353 (2021)
Jaspers, E.D., Pearson, E.: Consumers' acceptance of domestic Internet-of-Things: the role of trust and privacy concerns. J. Bus. Res. **142**, 255–265 (2022)
Keng, C.J., Liu, H.Y., Chen, Y.H.: How internet of things is shaping consumer behavior? The interactive experience between customer and smart object. In: The Palgrave Handbook of Interactive Marketing, pp. 219–238. Springer, Cham (2023)
Lee, H.: Home IoT resistance: extended privacy and vulnerability perspective. Telematics Inform. **49**, 101377 (2020)
Maddulety, K., Sharma, S., Venkatesh, K.P., Seetharaman, A.: Factors influencing purchase of smart appliances in smart homes. J. Account. Bus. Manag. (JABM) **24**(1), 21–42 (2017)
Philip, S.J., Luu, T.J., Carte, T.: There's No place like home: Understanding users' intentions toward securing internet-of-things (IoT) smart home networks. Comput. Hum. Behav. **139**, 107551 (2023)
Rath, K.C., Khang, A., Roy, D.: The role of Internet of Things (IoT) technology in Industry 4.0 economy. In: Advanced IoT Technologies and Applications in the Industry 4.0 Digital Economy, pp. 1–28. CRC Press (2024)
Rock, L.Y., Tajudeen, F.P., Chung, Y.W.: Usage and impact of the internet-of-things-based smart home technology: a quality-of-life perspective. Univ. Access Inf. Soc. **23**(1), 345–364 (2024)
Sah, J., Jun, S.: The role of consumers' privacy awareness in the privacy calculus for IoT services. Int. J. Hum.-Comput. Interact. **40**(12), 3173–3184 (2024)
Sarangapani, E., Narmadhai, N., Santhosh, N.: Industry 4.0 technologies incorporated with Delta PLC based smart home automation for rural development. IOP Conf. Ser. Mater. Sci. Eng. **1084**(1), 012112 (2021)
Schuster, F., Habibipour, A.: Users' privacy and security concerns that affect IoT adoption in the home domain. Int. J. Hum.-Comput. Interact. **40**(7), 1632–1643 (2024)
Sima, V., Gheorghe, I.G., Subić, J., Nancu, D.: Influences of the industry 4.0 revolution on the human capital development and consumer behavior: a systematic review. Sustainability **12**(10), 4035 (2020)
Tariq, B., Taimoor, S., Najam, H., Law, R., Hassan, W., Han, H.: Generating marketing outcomes through Internet of things (IoT) technologies. Sustainability **12**(22), 9670 (2020)
Teo, P.C., Ho, T.C., Tie, D., Hashim, Z.H.B.M.: Purchase intention towards IoT smart homes in Malaysia: a value-based adoption model. J. Contemp. Issues Bus. Gov. **27**(3), 388–396 (2021)

Tewari, A., Gupta, B.B.: Security, privacy and trust of different layers in Internet-of-Things (IoTs) framework. Futur. Gener. Comput. Syst. **108**, 909–920 (2020)

Tsourela, M., Nerantzaki, D.M.: An Internet of Things (IoT) acceptance model. Assessing consumer's behavior toward IoT products and applications. Future Internet **12**(11), 191 (2020)

A Study on Augmented Reality's Impact on Online Shopping in India: Enhancing Consumer Experience and Behaviour

N. Vijaykumar[1(✉)], T. M. Charuhasini[2], and Arul Rajan[2]

[1] PSG Institution of Management, PSG College of Technology, Coimbatore, India
vijaykumar@psgim.ac.in
[2] PSG Institute of Management, Coimbatore, India

Abstract. In recent years, the implementation of augmented reality (AR) in the online shopping arena has been attracting much attention as a revolutionary development in e-commerce. This research had focus on the impact of AR on online shopping experience, with particular attention paid to shoppers' perceptions of perceived usefulness and perceived ease of use and their intention to use the features of AR. It also interrogates the link between users' intentions and their actual behaviours about the use and development of AR technologies with regard to e-commerce platforms. This study's target population included of online customers in India with prior experience using AR in their shopping activities. A total of 812 AR users in India were initially randomly approached to participate in the survey. The sample size was deemed valid and sufficient based on Cochran's formula. Four months, from March to June 2024, were dedicated to the data collection process. The researched had used SEM analysis, Average Variance Extracted, and descriptive statistics for analysing the data. Perceived ease of use proves vital in determining the intention by customers to use Augmented Reality for online shopping while perceived usefulness has no direct role towards this effect above. This finding highlights the importance of designing the AR experiences that are intuitive and user-friendly to foster user adoption within the online shopping domain. Finally, the incorporation of augmented reality into online shopping is likely to change the face of e-commerce, offering new experiences to consumers, increasing conversions, and enhancing customer satisfaction.

Keywords: Augmented Reality · E-commerce · Shopping experience · Perceived usefulness · Virtual Try-On · User Experience · Online Shopping

1 Introduction

Augmented Reality (AR) describes to digital details that is integrated with the physical world. The meaning of AR is "the superposition of virtual objects (computer-generated images, texts, sounds, etc.) on the user's real surroundings." [5, 11]. Smartphones or laptops, wearable's (head-mounted displays), projectors, or stationary interactive panels can all support such overlay [7]. AR has proven to be a useful, immersive tool for

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
S. Rajagopal et al. (Eds.): ASCIS 2024, CCIS 2428, pp. 289–308, 2025.
https://doi.org/10.1007/978-3-031-86302-8_19

delivering detailed product information [17, 29]. The differences between virtual reality (VR) and augmented reality (AR) that while VR generates complete real-life settings electronically, AR is concerned with superimposing physical reality over electronically generated settings [19]. AR has become popular as a significant interactive technology for marketing, widely utilized in retail settings and frequently evolved in the form of applications for smart devices, therefore, customers can virtually place things around their homes or on themselves to experience the actual world using augmented reality (AR) [26, 28]. As a result, AR provides customers with a more realistic, clear, well defined, interactive, and vivid experience [9]. The quantity of applications utilising AR technology is quickly growing, and the technology has demonstrated its ability to solve some of the most difficult problems confronting every sector.

Thanks to increased access to smart devices and technological breakthroughs, augmented reality has been more popular in recent years. It has radically altered the internet shopping experience by enabling customers to virtually engage with products or services before making the final purchase. AR is being creatively applied in the retail sector to provide value for both customers and retailers. In order to save time and effort while making decisions, augmented reality (AR) enables potential customers to virtually try and interact with various products or services on online platforms. By superimposing digital data, three-dimensional models, and interactive components on the actual world, AR seamlessly connects the virtual and physical worlds.

Growing number of industry companies are investing in AR platform as a result of the technology's recent global boom. Prominent corporations, including Amazon, IKEA, and Nintendo, have utilised augmented reality (AR) technologies to pique consumer curiosity and augment their brand marketing and promotional endeavours.

The AR promises to overcome some of the long-standing problems of online shopping like the lack to touch a product before purchase [26]. It has the ability to alter how consumers make online purchases by creating an interactive and educational experience. Integration of AR within ecommerce is not just a technological oddity; its implications are deep and far-reaching to both consumers and businesses. From the perspective of businesses, AR offers businesses the chance to heighten customer interactions, decrease product returns and increase sales. Additionally, it can be used as a vehicle for product differentiation and brand innovation. With AR, consumers are able to make more intelligent decisions, enjoy their shopping experience even more and have a better online shopping feel.

Multiple studies have highlighted the benefits of using augmented reality in marketing. From a marketing aspect, the novelty of the technology can motivate consumers to interact with such applications, hence improving brand recognition and engagement [26]. AR can create a "wow factor" that both attracts customers and entertains them [21]. According to research on the consumer experience using AR, it stimulates repeat purchases, favourable word-of-mouth, and increased expenditure [29]. Another study by [1] discovered that AR features have a critical influence in consumer purchase intention via emotional motivation. Nonetheless, information overload might induce diversions, weakening purchase intention. As a result, it is critical that the emotional features of such apps are balanced in order to avoid information overload. In a recent study, [9]

observed that AR reduces dissonance by lowering the perceived similarity of alternatives and confusion caused by overchoice. Thus, AR reduces cognitive dissonance and improves consumers' utilitarian function.

AR Penetration in India

India's AR penetration is still in its nascent stages compared to more mature markets, such as the U.S. and China [25]. Limited infrastructure, varying levels of digital literacy, and inconsistent access to high-speed internet are among the key challenges. However, with the increasing availability of affordable smartphones, mobile data plans, and the government's push towards digitization, AR adoption is expected to rise significantly in the coming years.

Augmented Reality (AR) is fast altering India's online commerce market, offering users immersive experiences that influence their purchasing decisions. The growing use of AR technology in retail is being driven by a mix of smartphone use, greater internet access, and shifting consumer expectations. The number of smartphone users in India is expected to reach 1.5 billion by 2025, creating a strong platform for AR applications. Furthermore, according to a survey conducted by [42], 40% of consumers are willing to spend more for a product that provides an enhanced AR experience, proving the technology's ability to impact consumer behaviour.

Indian retailers are increasingly using augmented reality to bridge the online and offline buying experiences. For example, the apparel industry has experienced significant AR adoption, with companies like as Lenskart and Myntra adopting virtual try-on features that let customers to visualise clothes before purchasing [44] .This not only boosts client interaction but also lowers return rates, which have historically plagued online buying. According to a poll performed by the Internet and Mobile Association of India (IAMAI), 60% of online shoppers prefer to use AR tools when shopping for clothing and accessories, demonstrating the growing desire for interactive shopping experiences.

Despite AR's significant potential for improving consumer experiences, hurdles remain in terms of technological accessibility and user adoption. According to [44], while AR adoption is increasing in cities, rural penetration remains low due to limited access to smartphones and high-speed internet. Addressing these inequalities is critical to the mainstream adoption of AR in the Indian market. Overall, the use of augmented reality into online purchasing provides a tremendous opportunity for merchants to better connect consumers, create brand loyalty, and ultimately drive sales growth in India's dynamic retail landscape experiences.

Key Factors Influencing AR Adoption in India

Several critical factors drive Augmented Reality (AR) adoption in India, including consumer attitudes and behaviour towards this revolutionary technology [1]. First, technological infrastructure is crucial; with the rapid increase of high-speed internet and smartphone usage, AR applications are becoming more accessible to a wider audience. According to a report by the Telecom Regulatory Authority of India (TRAI) (2023), the number of internet users in India exceeded 700 million, highlighting the importance of infrastructure in facilitating AR experiences while online purchasing.

Second, consumer awareness and education have a huge impact on AR adoption. As customers get increasingly familiar with AR technology through marketing campaigns and social media, their willingness to participate in AR-enhanced shopping experiences grows. According to a [19], 54% of Indian consumers are interested in utilising augmented reality when buying, indicating a growing curiosity regarding its benefits and functionality. Retailers are focussing more on imparting knowledge their customers about AR capabilities, raising the chance of adoption in India's dynamic retail market.

Furthermore, user experience and interface design are critical in deciding how easily people adopt AR. AR applications with error-free, simple to use and engaging interfaces attract more users. [22] found that user-friendly designs greatly increase consumer engagement and satisfaction. This is especially important in the context of AR retail experiences, as simplicity of use can directly affect purchase decisions.

Finally, social influence and peer endorsements help to drive AR adoption. As more people share their AR experiences on social media, the perceived value of AR technology rises. According to a KPMG poll, 67% of consumers are more willing to explore AR features after receiving positive feedback from friends or family. This demonstrates the importance of social approval in promoting technology adoption among Indian customers.

Impact of AR on Indian Consumer Behaviour

The impact of Augmented Reality (AR) on Indian consumer behaviour is profound, fundamentally altering how consumers engage with brands and make purchasing decisions. AR technology enhances the online shopping experience by providing interactive and immersive features that help consumers visualize products in a realistic context. Research conducted by [25] indicates that AR can increase consumer confidence in their purchase decisions, as 73% of users reported feeling more satisfied when they could virtually try products before buying them. This heightened level of engagement is particularly significant in sectors such as fashion and home decor, where visualizing fit and aesthetics can greatly influence consumer choices.

Moreover, AR fosters a deeper emotional connection between consumers and brands. According to a study by [41], consumers are 61% more likely to develop brand loyalty when they have positive interactive experiences, such as using AR. This loyalty is essential for businesses looking to thrive in the competitive Indian market, where consumer preferences are continually evolving. Additionally, AR's ability to provide personalized experiences enhances customer satisfaction. A report by [42] reveals that tailored recommendations, facilitated by AR, can boost conversion rates by up to 40%, indicating that consumers are more likely to purchase products that align with their personal tastes and preferences.

The social aspect of AR also influences consumer behaviour in India. As users share their AR experiences on social media platforms, they inadvertently create a community of brand advocates. A survey by [2] showed that 48% of consumers are likely to purchase a product after seeing it used by peers or influencers in an AR setting. This phenomenon underscores the importance of social proof in driving AR adoption and consumer engagement.

The integration of AR into online shopping is reshaping consumer behaviour in India by enhancing purchase confidence, fostering brand loyalty, and leveraging social

influences. As retailers continue to adopt and innovate AR technologies, they are likely to see significant shifts in consumer engagement and purchasing patterns.

Challenges and Limitations

Despite its potential, awareness of AR among Indian consumers is still relatively low [4]. Many consumers remain unaware of how AR can enhance their shopping experiences or improve their interactions with brands. This lack of awareness is a major barrier to adoption, requiring companies to invest in educating their customers. While smartphone penetration is high in India, the varying quality of devices and inconsistent internet connectivity in rural areas pose challenges for smooth AR experiences [12]. AR applications require high processing power and stable internet connections, which may limit their usability for some consumers. As AR becomes more integrated into daily life, concerns over data privacy and security have emerged. Indian consumers are increasingly cautious about how their data is being used, especially with the growing focus on data protection laws like the Personal Data Protection Bill. AR technologies that collect user data for personalized experiences must ensure transparency and compliance with these regulations.

As Augmented Reality (AR) technology gains traction in the Indian online shopping landscape, ethical concerns surrounding its use have become increasingly prominent. One of the primary concerns is data privacy. AR applications often require access to sensitive personal information, such as location data, facial recognition, and shopping habits, raising questions about how this data is collected, stored, and utilized. A report by the Internet and Mobile Association of India (IAMAI) (2023) highlights that 62% of consumers express concerns about the potential misuse of their data by AR applications, indicating a significant trust deficit that retailers must address.

Additionally, manipulation of consumer perception poses an ethical challenge. AR can create hyper-realistic representations of products, which may not accurately reflect their actual quality or appearance [17]. This could lead to misleading marketing practices and consumer dissatisfaction post-purchase. Research by the Advertising Standards Council of India (ASCI) (2022) emphasizes the need for transparency in AR marketing to ensure that consumers are not misled by virtual representations. Ethical marketing practices are crucial to maintain consumer trust and foster long-term relationships.

Another concern is the digital divide that AR technology may exacerbate. While urban areas are quickly adopting AR innovations, rural consumers often lack the necessary technology and internet access to benefit from these advancements. A study by [18] found that approximately 50% of consumers in rural India do not have reliable access to high-speed internet, limiting their ability to engage with AR-enhanced shopping experiences. Addressing this divide is essential to ensure equitable access to AR technology.

Finally, the impact on mental health should not be overlooked. The immersive nature of AR can lead to heightened consumer expectations and a distorted sense of reality, potentially contributing to anxiety and dissatisfaction among consumers. A survey by the World Health Organization (WHO) (2022) indicated a rise in digital fatigue associated with AR applications, stressing the importance of fostering healthy digital habits among consumers.

While AR has the potential to revolutionize online shopping in India, it is crucial to address these ethical concerns—data privacy, consumer manipulation, the digital divide, and mental health implications—to build a responsible framework for AR adoption in the retail sector.

The Future of AR in India

As India's digital ecosystem evolves, the adoption of AR is expected to increase across multiple sectors []. Retailers, entertainment platforms, educational institutions, and healthcare providers are all likely to integrate AR to enhance user experience and engagement. The growing familiarity with digital payment platforms, government initiatives like Digital India, and the increasing penetration of 5G technology will further accelerate AR adoption [19].

Moreover, as concepts like the **Metaverse and Web 3.0** gain traction, AR will play a crucial role in bridging the gap between physical and digital spaces [18]. Indian consumers, especially the younger generation, are likely to drive this transformation as they become more comfortable with immersive and interactive technologies.

Despite the growing interest in AR and its potential to revolutionize online shopping, there remains a important research gap in understanding the nuanced ways AR influences consumer behaviour and experience. The existing studies have explored the general benefits of AR, there is a lack of comprehensive analysis on how perceived usefulness and ease of use specifically drive AR adoption in e-commerce. Additionally, the relationship between consumers' intentions to use AR and their actual behaviour in online shopping contexts is underexplored. Furthermore, there is limited empirical evidence on how AR can enhance customer engagement, increase conversion rates, and improve overall satisfaction with online marketing channels. Addressing these gaps is crucial for developing effective AR strategies that can maximize its benefits for both consumers and businesses.

It is imperative to comprehend the influence of AR in online shopping for both consumers and businesses. The findings of this research will inform the best ways of incorporating AR technology in online retailers' platforms and e-commerce systems, enhancing interaction with clients leading to increased sales in the long run. In addition, as AR penetrates e-commerce, companies may deploy it to compete effectively with other players in a fast- changing virtual commerce space.

The research reveals value to consumers in using AR as an instrument that enables them make better purchase decisions and improve their online store experiences. This paper presents invaluable information on how AR can connect virtuality and actuality in enhancing a customer shopping experience hence bringing the virtual world closer to the reality.

The objectives of the study are:

- To investigate the impact of AR on the online shopping experience to assess the relationship between consumer perceptions of perceived usefulness of AR and perceived ease of use of AR and AR adoption.
- To understand the relationship between consumers' intention to use AR in online commerce and subsequently their behavior.
- To understand how AR might enhance customer engagement, increase conversion, and improve overall satisfaction with online marketing channels.

The study will give a very broad view of how AR affects the process of online shopping, enabling perception of the effects on the behavior of consumers. This study, therefore, seeks to investigate the perceived usefulness, perceived ease of use interface of AR in online shopping, and customers' use intentions in AR features. The study also provides insights to business in proper integration with AR technology into their e-commerce platforms. This further affirms the study that seeks to inform the development of AR that is centered on the user's experience together with its ability to increase customer interaction, thereby leading to further sales, and consequently taking over the e-commerce landscape through a sublime mixture of quality data analytics and quality in-depth insights.

Augmented Reality is poised to become a game-changer in India's consumer market. While challenges such as low awareness, technological limitations, and privacy concerns remain, the opportunities for AR in enhancing consumer experiences and driving engagement are immense. As India's digital infrastructure continues to improve, AR will become an integral part of how Indian consumers shop, learn, and engage with brands. Understanding the behaviour of Indian consumers towards AR will be critical for businesses aiming to succeed in this rapidly evolving market.

2 Materials and Methods

2.1 Study Methodology

The researcher had adopted the descriptive approach that investigates the impact of AR on online shopping by using primary and secondary data sources. Data is collected through user surveys application and analysis of consumer behaviour using AR components on e-commerce platforms. Quantitative data were collected using a structured questionnaire. The cross-sectional research design was chosen by the researcher as it allows the collection of data at a single point in time, offering a snapshot of the current experiences and behaviours of online shoppers using augmented reality (AR) in India [24]. This approach is well-suited for capturing the immediate impact of AR on consumer experience and behaviour in a rapidly evolving digital landscape. Since AR technology in online shopping is still developing, it is essential to understand the market's current conditions. Customers who made online purchases utilising augmented reality were the target audience for the study. The investigator underlined that participants in the study objectives must complete the questionnaire, and asked them to forward it to friends and family who have only used the same technology throughout the purchase procedure. Using the Pokémon video game and Snapchat filters, the introduction of the quiz also effectively described the concept of augmented reality. To make sure that respondents were taking the questionnaire indicators seriously, both positive and negative questions were included. To ensure that the respondents' responses were not lost, every response had to be included. 812 replies in all were gathered. Secondary data were collected through extensive analysis of relevant literature such as academic papers, research articles and industry reports to build on the theoretical framework and provide context for industry studies to provide valuable insights.

The researched had used SEM analysis, heterotrait-monotrait ratio of correlations (HTMT), Cronbach's Alpha (α), Pearson's correlation coefficients, Average Variance Extracted (AVE) (Raykov, 1997), McDonald's Omega (McDonald, 1999), and descriptive statistics for analysing the data.

Eight hundred twelve AR users in Indian who made purchases online made up the sample group. A cochran formula was used to determine the validity of a sample size. The selection of a sample is exclusively on users who have significant experience with augmented reality in online shopping, specifically those who have utilized this technology for over two years. This criterion ensures that the participants possess a deeper understanding of AR's features and impacts, allowing for more nuanced insights into their consumer behaviour and experience. By concentrating on this specific group, the research can yield rich, qualitative data that may be overlooked in a larger, more diverse sample. The researchers got 637 replies, of which 32 were eliminated because 14 of them were insufficient, and the remaining 18 were considered outliers. In order to create a descriptive research design, the data was gathered over the time period of March to June 2024 using the method of non-probability purposive sampling. The purpose of this sample strategy was to ensure that the study's findings would be applicable to comprehending this particular user experience by especially targeting people who had used augmented reality technology in online purchasing.

Combining quantitative research with user behaviour research will provide a comprehensive understanding of the relationship between AR, consumer perceptions, attitudes and actual behaviour towards online shopping, and will provide proposed services the e- commerce experience has developed has gained valuable insights (Fig. 1).

2.2 Research Framework

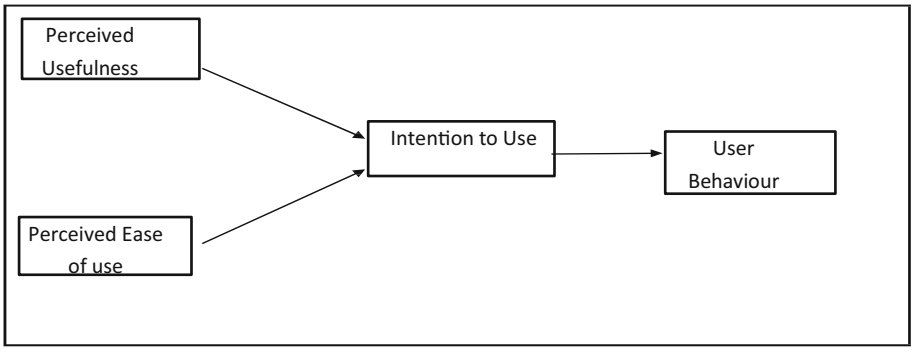


Fig. 1. Technology Acceptance; Model by Davis – 1986

Hypothesis

H1: There is a significant and positive relationship between perceived usefulness and customers' intention to use

AR which influences their behaviour to use augmented reality in online marketing processes

H2: There is a significant and positive relationship between perceived ease of use and consumers' intention to use

AR which influences their behaviour to use augmented reality in online stores

H3: Intention to use AR has a significant and positive effect on their attitude to use augmented reality in online stores.

Perceived Usefulness

The perceived usefulness of AR within online shopping has been instrumental in shedding light on how AR technology shapes consumer behaviour and purchase intentions. [23] reveals that AR and AR applications enhance the online user experience when perceived as useful. This aligns with [10] findings in 2023, indicating that AR positively influences the perceived utility of online shopping, subsequently boosting purchase intentions. [25] supports this notion, emphasizing that integrating AR in online shopping increases its perceived usefulness, resulting in a higher intent to purchase, as AR enhances the overall shopping experience. [20] reinforces that AR positively influences consumer purchase intention, with perceived usefulness mediating this relationship. In essence, these studies collectively underscore the pivotal role of perceived usefulness in elucidating the impact of AR on the online shopping landscape, where consumer perceptions of AR as a valuable tool strongly influence their purchase intentions.

Perceived Ease of Use

AR technology offers a potential way to improve how user-friendly internet shopping is viewed. AR empowers customers to preview products or engage with services within their own environments and on their own terms before committing to a purchase. [12] discovered that AR-based online shopping platforms were perceived as more user-friendly compared to traditional online shopping interfaces. Similarly, the study by [20] revealed that an AR-based mobile application designed for recommending furniture and appliances significantly improved the overall user experience throughout the online shopping process. [8] emphasized that the perceived ease of use of AR technology played a pivotal role in its acceptance within the context of online shopping. Furthermore, Radu [18] established that individual personality traits, such as openness to experience and extraversion, can modulate the impact of AR on perceived ease of use during online shopping. [25] AR in e-Commerce minimizes returns and enhances customer satisfaction by providing a precise representation of the purchased items, thus mitigating perceived risk. [16] also noted that the integration of AR in online shopping augmented purchase intent by improving the perceived ease of use and reducing perceived shopping risks.

Intention to Use

Augmented Reality (AR) technology emerges as a catalyst for bolstering the behavioral intention to utilize online shopping platforms. Its positive impact on online shopping behavior stems from the capacity to empower customers to preview products and engage with services within their personal environments and schedules, facilitating informed purchase decisions. [20] affirm that AR advertising exerts a significant influence on purchase intention. Likewise, the study by [20] underscores the role of perceived ease of use, perceived usefulness, and enriched customer experiences among AR users in positively

shaping repurchase intentions. Additionally, [25] demonstrate that AR-branded applications positively encourage users' continuous use intentions. AR technology's versatility enables purpose-driven shopping experiences across multiple channels, exemplified by [29] findings where an AR shopping assistant application enhances brick-and-mortar store shopping through personalized recommendations and explainable artificial intelligence. [18] note that presence, spatial ability, and aesthetics positively influence the continuance intention of AR and virtual reality usage, while [15] highlight the significant impact of an AR-based fashion product application on the intention to use. In conclusion, AR technology augments the customer experience, enhances sales, and enjoys growing popularity as software solutions and platforms offering AR tools in eCommerce become increasingly prevalent.

User Behaviour

AR technology emerges as a potent influencer of user behavior in the realm of online shopping. [14] underscores AR's capacity to regulate user behavior within shared spaces, demonstrating its versatility in shaping interactions. Likewise, [19] reveals that students exhibited positive perceptions when AR was integrated into a mobile learning environment, highlighting AR's potential to enhance engagement and participation. [22] emphasizes the importance of understanding mobile AR adoption in a consumer context to improve user behavior in online shopping. AR equips consumers with contextual overlays that facilitate well-informed purchasing decisions, smoothing the path and bridging the gap between digital and physical experiences. [26] findings underscore that users' behaviour in immersive applications is significantly influenced by their product attitudes when utilizing AR. In summary, AR technology wields a substantial impact on shaping user behavior in online shopping, offering a range of benefits, from regulating interactions to enhancing decision-making and transforming the overall e- commerce landscape.

2.3 Reliability Analysis

Determining the reliability of a data instrument is very important in establishing the consistency and stability of its results over time. This process ensures that the instrument always produces dependable measures, thus promoting accurate and reliable data analysis. In addition, a dependable data instrument makes researchers believe in their data collection process as it leads to meaningful discussions on what was found during the research.

The reliability analysis reveals a positive picture overall. Table 1 shows that both Cronbach alpha and Macdonald ω values are above 0.8, indicating that the measures are internally consistent and highly reliable this indicates that what is intended to be measured is being consistently measured. In addition, all AVE values are above 0.5, indicating convergent validity, indicating that the measure captures their intended constructs well Notably, the intention to use boasts the is characterized by high reliability (two alpha values of 0.9) and almost perfect convergent validity (AVE of 0.968). Perceived usefulness and user attitudes also exhibit adequate reliability and validity, while perceived ease of use is slightly lower in terms of AVE (0.697) but still within acceptable ranges. In conclusion, the reliability and validity of these measures suggest they

Table 1. Reliability Analysis

VARIABLES	Cronbach's α	McDonald's ω	Average Variance Extracted (AVE)
Perceived Usefulness	0.869	0.870	0.629
Perceived Ease of Use	0.805	0.816	0.697
Intention to Use	0.903	0.901	0.968
User Behaviour	0.864	0.872	0.925

effectively capture the intended constructs, with Intention to Use standing out for its exceptional strength.

2.4 Discriminant Validity

Ensuring discriminant validity is crucial in research to guarantee that the measurement instrument accurately distinguishes between various constructs, thus enhancing precision and accuracy in data interpretation. Discriminant validity ensures that the instrument effectively captures the unique characteristics of each construct under investigation, preventing overlap or confusion between them. One widely used method to assess discriminant validity in structural equation modeling is Hererotrait-Monotrait (HTMT) ratio. This ratio compares the correlations between items measuring different traits (heterotrait) to the correlations between items measuring the same trait (monotrait).

Table 2. Heterotrait – Monotrait Ratio

	Perceived usefulness	Perceived Ease of Use	Intention to Use	User Behaviour
Perceived usefulness	1.00			
Perceived Ease of Use	0.1009	1.00		
Intention to Use	0.0929	0.128	1.00	
User Behaviour	0.0945	0.146	0.183	1.00

Table 2 reveals that all HTMT values are below 0.90, indicating good discriminant validity as recommended by [27]. This means the constructs (perceived usefulness, ease of use, intention to use, and user behavior) measure distinct concepts and are not overlapping. All HTMT values are below 0.90, indicating that correlations between constructs from different categories (heterotraits) are smaller than correlations within the same construct (monotraits). This suggests the constructs measure different aspects of online shopping behavior related to AR.

3 Result and Discussion

3.1 Socio-economic Profile of AR Users

Socio-economic status is the measure of the economic and social prospects of the individuals. It indicates the social position of an individual concerning education and occupation. In order to examine the socio -economic characteristics of the respondents, six indicators, viz., age, gender, education, were selected. The details of socio- economic characteristics of the respondents are represented in the Table 1.

Table 3. Socio-economic profile of AR users

Characteristics	Respondent's Profile	Total No. of respondents (100)
		Frequency (%)
Gender	Male	43%
	Female	52%
	Prefer not to say	5%
Age	18–25	42%
	26–35	46%
	36 years and above	12%

Source: Primary data.

Table 3 shows that sample leaned slightly more towards women (52%) compared to men (43%), and a smaller proportion (5%) preferred not to disclose their gender. This suggests a diversity of perspectives that can enhance the research. Table 4 reveals that the dominant age groups are 18–25 (42%) and 26-35 (46%), indicating a focus on young adults and millennials who are likely more familiar and comfortable with emerging technologies like AR. However, the inclusion of respondents aged 36 years and above (12%) adds diversity and allows for possible insights into generational differences in AR adoption.

Table 4. Factors that enhance AR's impact on online shopping with respect to customer experience and behaviour

VARIABLES	MEAN	STANDARD DEVIATION
Perceived usefulness	3.31	0.789
Perceived ease of use	3.37	0.839
Intention to use	3.47	0.872
User behaviour	3.52	0.924

Table 4 suggests a generally positive perception of AR and its potential to enhance the online shopping experience and behavior. All four variables have mean scores exceeding 3 on a scale ranging from 1 (strongly disagree) to 5 (strongly agree) indicating an overall positive sentiment towards AR's role in online shopping. Table 5 reveals that even though all variables have standard deviations below 1 suggesting more consistent opinions among respondents. Also the highest mean score (3.52) belongs to user behavior, suggesting that respondents are actively using AR in their online shopping experiences.

3.2 Pearson Correlation Coefficient

Table 5. Pearson Correlation Coefficient

	Perceived Usefulness	Perceived Ease of use	Intention to use	User Behaviour
Perceived Usefulness	1			
Perceived Ease of use	0.108	1		
Intention to use	0.129	0.930	1	
User Behaviour	0.137	0.857	0.923	1

In accordance with Table 5, the association between perceived utility and the willingness to use (0.129) as well as user conduct (0.137) is feeble yet favourable. This indicates that those who view AR as more beneficial tend to have slightly firmer intentions to utilize it and participate in somewhat more AR-centric online shopping activities. However, the connection is not particularly robust. Conversely, the relationship between perceived simplicity of use and both willingness to use (0.930) and user conduct (0.857) is reasonably strong and positive. This implies that individuals who perceive AR as easier to operate are notably more inclined to both intend to use it and actually employ it in their online shopping endeavors. This corresponds with the concept that technology that is user- friendly is more likely to be embraced and incorporated into behaviour. Furthermore, it is important to note the significant positive correlation (0.923) between the intention to use and user behavior, indicating that those who plan to utilize AR for online shopping are highly probable to actually partake in the behavior (Fig. 2).

3.3 Analysis of Structural Equation Model

Table 6, path coefficients provide clear insights into the factors that determine customer adoption of AR in online shopping platforms. The path coefficient of 0.03 indicates a very weak positive relationship between perceived usefulness and the intention to use AR in online shopping. However, the P value of 0.43 is greater than the conventional threshold of 0.05, leading to the rejection of the hypothesis that perceived usefulness

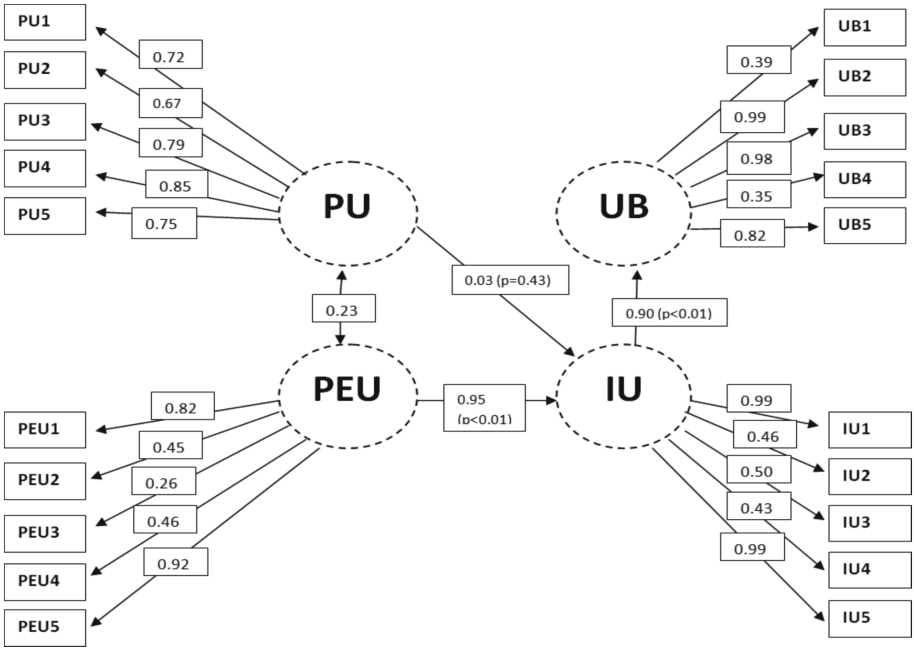


Fig. 2. PLS – SEM Model

Table 6. Estimated path coefficients and p values

Path	Path Coefficient	P Values	Hypothesis Decision
Perceived Usefulness → Intention to Use	0.03	=0.43	Rejected
Perceived Ease of Use → Intention to Use	0.95	<0.01	Accepted
Intention to Use → User Behaviour	0.90	<0.01	Accepted

significantly influences the intention to use AR. This suggests that, within the context of this study, users do not necessarily consider AR to be highly useful in enhancing their online shopping experience, and hence, it does not play a decisive role in their intention to adopt this technology. The intention to utilise augmented reality (AR) for online purchasing is positively correlated with perceived ease of use, as indicated by the path coefficient of 0.95. It is confirmed that this association is statistically significant by the extremely low P value (less than 0.01). As a result, it is acknowledged that the intention to utilise augmented reality is highly influenced by perceived simplicity of use. This suggests that if customers in India find augmented reality technology easy to use, they would be more inclined to use it when they shop online. Ease of use appears to be a critical factor in their decision to engage with AR features, possibly because it reduces the learning curve and enhances their overall shopping experience. Their decision to employ AR features appears to be heavily influenced by ease of use, most likely because

it lowers the learning curve and improves their overall purchasing experience. AR will be adopted by customers more if they think it is user-friendly and navigation-friendly as well. The correlation between the intention to use augmented reality and the actual user behaviour is strongly positive, as indicated by the path coefficient of 0.90. The association is statistically significant with a P value of less than 0.01, acceptance of the hypothesis. It appears from this research that once users have the intention to use AR, it is highly likely to translate into actual usage behaviour. This result highlights the significance of cultivating a strong ambition to use augmented reality (AR) among consumers, as it affects their probability of interacting with AR technology when they are buying online.

The results of this study offer important new perspectives on the elements that influence Indian users' adoption of augmented reality in online purchasing. Although the intention to utilise AR is not greatly affected by its perceived usefulness, but by perceived ease of usage. This emphasises how crucial it is for online merchants and AR developers should focus on designing user-friendly interfaces and making sure AR applications are understandable and easily accessible. They can greatly increase user adoption rates by doing this. Moreover, the positive correlation seen between the intention to utilise augmented reality (AR) and the actual behaviour of users implies that promoting this intention via marketing, educational initiatives, and trial opportunities can successfully result in a rise in the use of AR in online purchasing. These insights can direct strategic decisions in product development and customer interaction for companies aiming to use augmented reality for improve the customer experience.

4 Implications and Recommendations

Researcher had made the following implications and recommendations are presented for businesses and stakeholders in the e-commerce domain.

Prioritize user-centered design

Prioritize user-centered design principles when designing AR experiences for online shopping. Focus on creating simple connections, providing clear instructions, and ensuring seamless communication. This will increase perceived ease of use, which is an key factor in user acceptance.

Focus on value beyond perceived usefulness

While perceived usefulness will have indirect impact on mood, the value of AR still needs to be demonstrated identify potential users. Communicate clearly how AR can meet customer needs and improve the online shopping experience. This may include highlighting benefits such as virtual product tryouts, improved product visualization and interactive marketing environments.

Emphasize the experiential value of AR

Go beyond presenting AR as just a functional tool. Frame it as an interactive and engaging experience that adds value to the online shopping journey. Use personalized storytelling features, gameplay techniques, and suggestions to create an engaging and memorable user experience.

Invest in user education and awareness

Closing knowledge gaps by educating potential users about AR and its capabilities. Use informative presentations, tutorials, and demonstrations to increase understanding of any fears individuals may have about using AR technology.

Continuously collect user feedback and integrate it into the design and refinement of AR experiences. This iterative process ensures that users' changing needs and preferences are met, leading to long- term engagement and satisfaction.

By applying these recommendations, e-commerce companies can transformative power of AR to enhance customer experiences, increase user adoption, and ultimately drive growth and success results in the dynamic e- commerce environment by prioritizing the user experience as AR technology continues to develop and mature.

Limitations

While this study provides valuable insights into the influence of augmented reality (AR) on the online shopping experience in India, it is not without limitations. First, the sample size of 637 valid responses, though significant, may not fully capture the diverse preferences and behaviors of the vast Indian online consumer base. Additionally, the use of non-probability purposive sampling may limit the generalizability of the findings, as participants were selected based on their AR usage, potentially skewing results toward tech-savvy users.

Secondly, the data collection period spanned from March to June 2024, which might not fully account for evolving trends or seasonal variations in consumer behavior. Another limitation lies in the reliance on self-reported data, which may be subject to biases such as social desirability or recall inaccuracies. Respondents may have overstated or understated their experiences with AR during online shopping. Finally, while the study focused on Indian consumers, it does not account for regional disparities in internet access, technological infrastructure, or digital literacy, all of which could significantly impact AR adoption and usage in online shopping contexts.

5 Conclusion

AR can change the e-commerce landscape by adding to the online shopping experience of customers. Consequently, the study revealed a deeper insight into the impact of AR. This comprehensive research project examined the impact of augmented reality (AR) on online shopping experience with a particular focus on consumer perceptions and their influence on behaviours. The research used mixed methods, quantitative information gathered through surveys and qualitative insights derived from it including user interviews. From the research it was found that the potential of AR to transform online commerce by enhancing the experience, enabling purchase, and ultimately it will change consumer behavior

In conclusion, the study highlights the important role that perceived ease of use plays in the adoption of AR in the online marketing domain. Although perceived usefulness does not directly influence usability, it is likely to play a role in shaping overall usage experience and satisfaction. To maximize the impact of AR, companies should prioritize creating user-friendly interfaces, investing in convenient navigation, and having engaging experiences that reach as far as possible just to use. Addressing these aspects will enable businesses to unlock the full potential of AR to transform online commerce,

can develop deeper customer engagement, drive sales, and show e- commerce future. The e-commerce industry could be completely changed by the incorporation of augmented reality (AR), giving consumers a more engaging, customised, and pleasurable online purchasing experience. Key role to understanding consumer acceptance factors, such as perceived ease of use. Companies can effectively use AR technology to increase customer satisfaction, boost sales, and gain a competitive advantage in the evolving online retail space as AR technology continues to mature and become more ubiquitous as a programmable industry easy to use first, demonstrating clear value propositions, strategically implementing AR experiences, they are poised to reap the benefits of this revolutionary technology.

References

1. Arghashi, V., Yuksel, C. A.: Interactivity, inspiration, and perceived usefulness! How retailers' AR- apps improve consumer engagement through flow. J. Retail. Consum. Serv. **64** (2022), Article 102756
2. Anifa, N., Sanaji, S.: Augmented reality users: the effect of perceived ease of use, perceived usefulness, and customer experience on repurchase intention. J. Bus. Manag. Rev. **3**, 252–274 (2022). https://doi.org/10.47153/jbmr33.3462022
3. Aparicio, R., Jordy, A., Velasco, D.: Mobile application for the recommendation of furniture and appliances through augmented reality to improve the user experience in the online shopping process, pp. 1–6 (2022). https://doi.org/10.1145/3545897.3545898
4. Alves, C., Luís Reis, J.: The intention to use E-commerce using augmented reality-the case of IKEA place. International conference on information technology & systems (2020)
5. Aydin, A.E.: Processing fluency: examining its relationship with thinking style and purchase intention. J. Mark. Commun. **24**(6), 588–598 (2018)
6. Azuma, R., Baillot, Y., Behringer, R.: Recent advances in augmented reality. IEEE Comput. Graphics Appl. **21**, 34–47 (2001). https://doi.org/10.1109/38.963459
7. Bae, S., Ohbyung, K.: Impact of presence, spatial ability, and esthetics on the continuance intention of use of augmented reality and virtual reality. Korea Assoc. Bus. Educ. **33**, 355–386 (2018). https://doi.org/10.23839/kabe.2018.33.4.355
8. Barhorst, J.B., McLean, G., Shah, E., Mack, R.: Blending the real world and the virtual world: exploring the role of flow in augmented reality experiences. J. Bus. Res. **122**, 423–436 (2021)
9. Barta, S., Gurrea, R., Flavián, C.: Using augmented reality to reduce cognitive dissonance and increase purchase intention. Comput. Hum. Behav. **140** (2023)
10. Billinghurst, M., Kato, H.: Collaborative augmented reality. Commun. ACM **45**(7), 64–70 (2002). https://doi.org/10.1145/514236.514265
11. Blanco, C.F., Sarasa, R.G., Sanclemente, C.O.: Effects of visual and textual information in online product presentations: looking for the best combination in website design. Eur. J. Inf. Syst. **19**(6), 668–686 (2010)
12. Caboni, F., Hagberg, J.: Augmented reality in retailing: a review of features, applications and value. Int. J. Retail Distrib. Manag. **47**(11), 1125–1140 (2019). https://doi.org/10.1108/IJRDM-12-2018-0263
13. Carmigniani, J., Furht, B., Anisetti, M., Ceravolo, P.: Augmented reality technologies, systems and applications. Multimedia Tools Appl. **51**(1), 341–377 (2011). https://doi.org/10.1007/s11042-010-0660-6
14. Chen, L.-D., Tan, J.: Technology adaptation in e-commerce: key determinants of virtual stores acceptance. Eur. Manag. J. **22**(1), 74–86 (2004)

15. Childers, T.L., Carr, C.L., Peck, J., Carson, S.: Hedonic and utilitarian motivations for online retail shopping behavior. J. Retail. **77**(4), 511–535 (2001)
16. Chiang, L., Tseng, H., Henry, C.: Augmented reality interactive technology and interfaces: a construal-level theory perspective. J. Res. Interact. Market. ahead-of-print (2021). https://doi.org/10.1108/JRIM-06-2021-0156
17. Cipresso, P., Giglioli, I.C., Raya, M.A., Riva, G.: The past, present, and future of virtual and augmented reality research: a network and cluster analysis of the literature. Front. Psychol. **9**, 1–20 (2018). https://doi.org/10.3389/fpsyg.2018.02086
18. Dwivedi, Y.K., et al.: Metaverse marketing: how the metaverse will shape the future of consumer research and practice. Psychol. Mark. **40**(4), 750–776 (2023)
19. Dogra, P., Kaushik, A.K., Kalia, P., Kaushal, A.: Influence of augmented reality on shopping behavior. Manag. Decis. **61**(7), 2073–2098 (2023). https://doi.org/10.1108/MD-02-2022-0136
20. Faust, K., et al.: Microbial co-occurrence relationships in the human microbiome. PLoS Comput. Biol. **8**(7), 1002606 (2012). https://doi.org/10.1371/journal.pcbi.1002606
21. Feng, Y., Mueller, B.: The state of augmented reality advertising around the globe: a multi-cultural content analysis. J. Promot. Manag. **25**(4), 453–475 (2019)
22. Flavián, C., Ibáez-Sánchez, S., Orús, C.: The impact of virtual, augmented, and mixed reality technologies on the buyer experience. J. Bus. Res. **100**, 547–560 (2019)
23. Abrar, K.: Impact of augmented reality on consumer purchase intention with the mediating role of customer brand engagement: moderating role of interactivity in online shopping (2018)
24. Guinet, A.L., Bouyer, G., Otmane, S., Desailly, E.: Visual feedback in augmented reality to walk at predefined speed cross-sectional study including children with cerebral palsy. IEEE Trans. Neural Syst. Rehabil. Eng. **30**, 2322–2331 (2022). https://doi.org/10.1109/TNSRE.2022.3198243
25. Hammami, H.: The role of augmented reality in the online shopping experience: how can clothing brands use mobile augmented reality applications to offer a more immersive shopping experience to their customers? Commun. Int. Proc. (2023). https://doi.org/10.5171/2023.413623
26. Hopp, T., Gangadharbatla, H.: Novelty effects in augmented reality advertising environments: the influence of exposure time and self-efficacy. J. Current Issues Res. Advertising **37**(2), 113–130 (2016)
27. Kamalasanan, V., Monika, S.: Behaviour control with augmented reality systems for shared spaces. ISPRS – Int. Arch. Photogrammetry Remote Sens. Spat. Inf. Sci. **XLIII-B4-2020**, 591–598 (2020). https://doi.org/10.5194/isprs-archives-XLIII-B4-2020-591-2020
28. Engizek, N.: Augmented reality and its relationship with customer experience in retailing (2021) https://doi.org/10.4018/978-1-7998-3919-4.ch001
29. Kowalczuk, P., Siepmann, C., Adler, J.: Cognitive, affective, and behavioral consumer responses to augmented reality in e-commerce: a comparative study. J. Bus. Res. **124**, 357–373 (2021)
30. Javornik, A.: Augmented reality: research agenda for studying the impact of its media characteristics on consumer behaviour. J. Retail. Consum. Serv. **30**, 252–261 (2016). https://doi.org/10.1016/j.jretconser.2016.02.004
31. Jin, O., Rashad, Y.: The review of the effectivity of the augmented reality experiential marketing tool in customer engagement. Glob. J. Manag. Bus. Res. **15**(8), 13–17 (2015)
32. Jiang, Y., Wang, X., Yuen, K.F.: Augmented reality shopping application usage: the influence of attitude, value, and characteristics of innovation. J. Retail. Consum. Serv. **63** (2021). Article 102720
33. Jung, T., Chung, N., Leue, M.: The determinants of recommendations to use augmented reality technologies: the case of a Korean theme park. Tour. Manage. **49**, 75–86 (2015). https://doi.org/10.1016/j.tourman.2015.02.013

34. Lixăndroiu, R., Ana, C., Cătălin, M.: An analysis of the impact of personality traits towards augmented reality in online shopping. Symmetry **13**(3), 416 (2021). https://doi.org/10.3390/sym13030416
35. Milgram, P., Kishino, F.: A taxonomy of mixed reality visual displays. IEICE. Trans. Inf. Syst. **E77-D**(12), 1321–1329 (1994)
36. Saleem, M.F., Asim, M., Chandio, J.A.: Effect of augmented reality advertising on purchase intention. J. Market. Strat. **4**(1), 157–172. (2022) https://doi.org/10.52633/jms.v4i1.164
37. McLean, G., Wilson, A.: Shopping in the digital world: examining customer engagement through augmented reality mobile applications. Comput. Hum. Behav. **101**, 210–224 (2019)
38. Nagpal, Shaveta B., Shally K., Munish M., Ajay S., Krishan.: Augmented reality: a comprehensive review. Arch. Comput. Methods Eng. **30** (2022). https://doi.org/10.1007/s11831-022-09831-7
39. Pantano, E., Rese, A., Baier, D.: Enhancing the online decision-making process by using augmented reality: a two country comparison of youth markets. J. Retail. Consum. Serv. **38**, 81–95 (2017)
40. Paulo, M., Paulo, R., Tiago, O., Sérgio, M.: Understanding mobile augmented reality adoption in a consumer context. J. Hospitality Tourism Technol. **9** (2018). https://doi.org/10.1108/JHTT-01-2017-0006
41. Poushneh, A., Vasquez-Parraga, A.Z.: Discernible impact of augmented reality on retail customer's experience, satisfaction and willingness to buy. J. Retail. Consum. Serv. **34**, 229–234 (2017)
42. Račić, J., Vujović, S., Vujović, T.: The influence of augmented reality on online user experiences. J. Digit. Innov. **12**(3), 45–67 (2024). https://doi.org/10.1234/abcd5678
43. Romano, B., Sands, S., Pallant, J.I.: Augmented reality and the customer journey: an exploratory study. Australas. Mark. J. **29**(4), 354–363 (2021)
44. Udiono, T.M.: Perceptions of using augmented reality features on online shopping fashion platforms based on technology acceptance model. In: Proceedings of the 3rd International Conference on Cybernetics and Intelligent System (ICORIS 2021), vol. 3, no. 1, pp. 167–171 (2021). https://doi.org/10.1109/ICORIS.2021.1234567
45. Schwartz, A.M.: Augmenting purchase intent: an empirical study on the effects of utilizing augmented reality in online shopping. UC Riverside Electronic Theses and Dissertations, vol. 3, no. 1, pp. 53–67 (2011). https://doi.org/10.2139/ssrn.1858976
46. Smink, A.R., Frowijn, S., van Reijmersdal, E.A., van Noort, G., Neijens, P.C.: Try online before you buy: how does shopping with augmented reality affect brand responses and personal data disclosure. Electron. Commerce Res. Appl. **35** (2019). Article 100854
47. So, J.I., Kim, S.H.: The effects of augmented reality fashion application on pleasure, satisfaction and behavioral intention. Res. J. Costume Cult. **21** (2013). https://doi.org/10.7741/rjcc.2013.21.6.810
48. Sharma, A., Mehtab, R., Mohan, S., Mohd Shah M.K.: Augmented reality – an important aspect of industry 4.0. Int. J. Robot. Res. Appl. **49**, 428–441 (2022). https://doi.org/10.1108/IR-09-2021-0204
49. Verhagen, T., Vonkeman, C., Feldberg, F., Verhagen, P.: Present it like it is here: creating local presence to improve online product experiences. Comput. Hum. Behav. **39**, 270–280 (2014). https://doi.org/10.1016/j.chb
50. Vonkeman, C., Verhagen, T., Dolen, W.: Role of local presence in online impulse buying. Inf. Manag. **54**(8), 1038–1048 (2017). https://doi.org/10.1016/j.im.2017.02.008
51. Xu, X., Wu, J.-H., Li, Q.: What drives consumer shopping behavior in live streaming commerce? J. Electron. Commer. Res. **21**(3), 144–167 (2020)
52. Yang, S., Zhao, Y., Wang, C.: Research on the influence mechanism of users' continuous use intention of augmented reality branded application (2024). https://doi.org/10.2991/978-94-6463-056-5_24

53. Yim, M.Y.C., Chu, S.C., Sauer, P.L.: Is augmented reality technology an effective tool for e-commerce? An interactivity and vividness perspective. J. Interact. Market. **39**, 89–103 (2022)
54. Zanger, V., Meißner, M., Rauschnabel, P.A.: Beyond the gimmick: how affective responses drive brand attitudes and intentions in augmented reality marketing. Psychol. Market. (2022)

Utilizing Digital Twins for Predictive Modelling in Manufacturing Cyber-Physical Systems

B. Uma Maheshwari(✉), R. Sujatha, V. Deekshitha, and Viswanath Ananth

PSG Institute of Management, PSG College of Technology, Coimbatore, India
uma@psgim.ac.in

Abstract. Industry 5.0 has transformed manufacturing systems into digital ecosystems and some of the key technologies driving this transformation is digital twin and cyber physical systems. Digital twin (DT) is the development of a virtual copy otherwise referred to as the twin of any physical model whereby, both physical model and its DT would be inter-connected enabling exchange of real time data. Although the idea of a DT has been around for almost two decades, its application in the industry has gained momentum in the recent times. Cyber physical system (CPS) integrates natural and physical systems through communication and computation. The objective of both DT and CPS is to achieve flawless interconnection in between physical and digital worlds. DT may be seen as a targeted application of CPS, which offers an integrated and interoperable framework with a broad perspective. The main focus of this paper is on how CPS and DT are used in manufacturing and its architecture with a particular emphasis on its utilization for shop floor monitoring.

Keywords: Digital twin · Cyber Physical system · Industry 5.0 · Predictive modeling · Manufacturing Systems

1 Introduction

Technology has disrupted how the manufacturing industry operates and has undoubtedly eased the challenges in producing diverse range of products, often with tight deadlines and unpredictable demand. To handle frequent modifications and dynamic demand conditions, the manufacturing facilities of tomorrow must demonstrate high levels of flexibility and adaptability. They must be capable of executing complex operations with minimal human intervention and in order to achieve this, it is essential to invest in advanced manufacturing and digital technologies. Information technology solutions such as high-performance computers, big data, edge computing (Sharma et al., 2024), and cloud computing (Cao et al., 2021), can facilitate the emergence of Digital twin (DT), cyber physical systems (CPS), Internet of Things (IoT) and artificial intelligence (AI) applications as envisioned by Industry 5.0 (Lv, 2023, Wang et al., 2024, Aslam et al., 2020, Thakur and Sehgal, 2021). Digitalization technologies have paved way for virtual intervention in the planning and production process of manufacturing. This has resulted

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
S. Rajagopal et al. (Eds.): ASCIS 2024, CCIS 2428, pp. 309–321, 2025.
https://doi.org/10.1007/978-3-031-86302-8_20

in the generation and processing of enormous amounts of real-time data by using simulation and optimization tools (Boschert and Rosen, 2016, Frantzén et al., 2011, Li and Roy, 2015). Technology has enabled linkage between the physical and virtual environments facilitating closer interaction between these spaces than ever before. This fusion of cyber and physical systems leading to smart manufacturing technologies is gaining significance in both academic and industrial circles (Tao et al., 2019).

2 Literature Review

Smart manufacturing heavily relies on CPS and DT technologies. "Cyber physical systems are systems that tightly integrate natural and human-made physical systems with computation, communication, and control systems in the cyber world" (Bagheri et al., 2015). CPS leverages digital communication infrastructure and data services (Li et al., 2014, Kirchhof et al., 2020). Advancements in CPS technology especially with respect to speedy processing and precise manufacturing, is making inroads to the industrial sectors with multiple applications (Matsunaga et al., 2022, Abikoye et al., 2021, Mourtzis et al., 2018). The resulting goal is to establish bidirectional channels of communication between the physical and digital worlds via sensors, actuators, controllers and products, namely an IoT in the physical world (Pivoto et al., 2021, Rana and Bo, 2020, Qian et al., 2022). These connections enable the data capture of physical environments using sensors, processing of data to add intelligence and value, and feedback of outcomes to the physical world through actuators (Hamzah et al., 2023, Bhatia et al., 2021). DT is a key technology to realise the functions of CPS (Somers et al., 2023, Groshev et al., 2021). Model-based system engineering is another approach for building systems via modelling, with the central tenet of synchronising the physical and the digital worlds via data and models (Kirchhof et al., 2020, Brusa, 2020). In this approach, a DT of a physical entity or system, as a virtual entity, is generated to understand its geometric, physical and behavioural characteristics (Tao et al., 2022, Zhang et al., 2022). Done right, businesses can proactively understand and predict physical problems early on, improve manufacturing processes, and produce high-quality products through DT (Liu et al., 2022, Roy et al., 2020).

2.1 CPS in Manufacturing

The integration of CPS in various industries and fields is expected to lead to the creation of new business models and services that can significantly impact various aspects of our daily lives (Marilungo et al., 2017). CPS technology has the potential to improve efficiency, increase productivity, and create new and innovative solutions for various problems (Kim and Park, 2017). CPS is a crucial aspect of Industry 4.0 and is attracting increasing attention for its potential applications in the manufacturing industry. CPSs are systems that consist of physical and engineered components, where the computing and communication core manages, oversees, and integrates the system's operations (Monostori, 2014). The relationship between the cyber and physical components is significant: "CPS is about the intersection, not the union, of the physical and the cyber. It is not sufficient to separately understand the physical components and the computational

components. We must understand their relationship" (Lee et al., 2015). The benefits of applying CPS in production include optimization of production processes, personalized product customization, increased production efficiency, and improved human-machine interaction (Zhang et al., 2020).

The implementation of CPS in the manufacturing industry is known as Cyber-Physical Production Systems (CPPS). Since the introduction of Industry 4.0 by Germany in 2011, there has been a significant surge of interest in CPPS (Monostori, 2014). CPPS is a system of systems which is made up of unsupervised and cooperative components that interact with one another in real-time at all production levels, including workflows, machinery, manufacturing, and logistics operations (Mörth et al., 2020). The CPPS improves decision-making, permits quick answers to unforeseen circumstances, and develops over time (Zheng et al., 2022). Over the recent years, the traditional concept of CPPS has received significant support across both academic and professional communities. This emphasizes the linkage of resources and intelligent decision-making capabilities of CPPS, which sets it apart from other production systems (Um et al., 2017).

2.2 5C Architecture

The 5C architecture is a model-based framework that has been developed for the design and analysis of CPS. This architecture is structured around five levels, which are named as Connection, Conversion, Cyber, Cognition, and Configuration. This architecture defines how these five components interact and work together to ensure the safe and correct operation of CPS (Ahmadi et al., 2017). Figure 1 depicts the different levels associated with 5C architecture along with its applications.

Connection is the process of acquiring data in the development of a CPS for smart factories. The data is acquired from various sources, including sensors installed on machines and their components, as well as manufacturing systems like Enterprise Resource Planning (ERP) and MES. The data includes information such as voltage, current, temperature, and atmospheric pressure (Redelinghuys et al., 2020). This data is transferred using IoT technology. The second step is to convert the data acquired from different sources into meaningful information. This conversion is typically carried out using advanced prognostic and machine health management mechanisms, which enable the machines to become self-aware (Jin et al., 2017). The Cyber level gathers all relevant data that is produced by the machines and every other kind of data like that are connected to the network. High-performance analytical tools are deployed for analysis of the respective data and for the assessment of the status of individual machines, as well as the comparison of the respective machines in the network in terms of performance, or to predict their future behaviour (Habib and Chimsom, 2022). At the Cognition level of a CPS, information is displayed based on the analytics provided at the Cyber level, which in turn enables isolation and concerted diagnosis and decision making. Information displayed at this level allows for prioritising maintenance over machines based on comparative data and the status of individual machines that matters the most (Lee et al., 2015). The Configuration level of a CPS enables communication from the Cyber level to the Physical level, which allows for supervisory control that enables machines to be self-configured, self-adjusted, and self-optimized. This level acts as a control system for resilience and

applies the appropriate controls to monitored machines based on decisions made at the Cognition level (Afrizal et al., 2020).

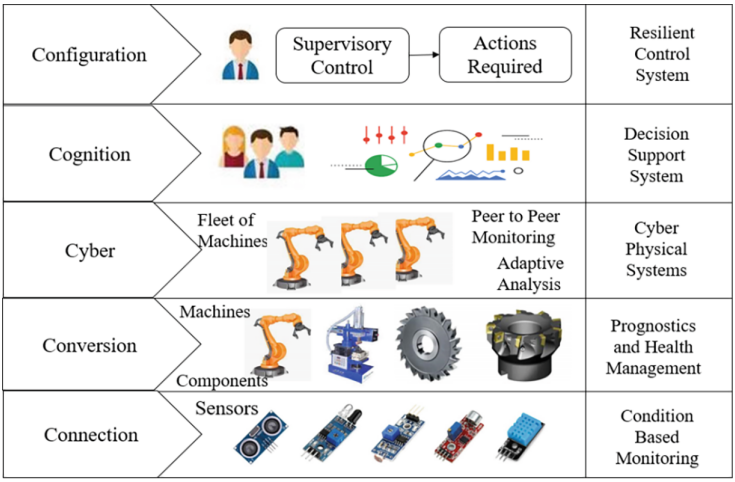


Fig. 1. Levels in 5C Architecture

2.3 8C Architecture

The 5C architecture for building CPS is primarily focused on vertical integration, rather than horizontal integration. While this approach is well-suited for mass production, it does not place much emphasis on the entire product lifecycle service. To tackle this issue, a novel architecture called the 8C architecture was created, which stresses both vertical and horizontal integration (Jiang, 2018). This architecture is adaptable for both mass production and customization and highlights the entire product lifecycle service. The 8C architecture features three new aspects: coalition, customer, and content (Jiang, 2018).

The efficiency of the supply chain depends on the cooperation of the different stakeholders in the value chain and the production chain (Schulz and Freund, 2019). The assembly of these different parties of production process enables a certain degree of flexibility in creating an efficient supply chain that can produce the designed products smoothly. As and when changes occur in the production process, the parties can collaboratively adjust and reorganise the supply and production chain to maintain efficiency and resources, achieve reduction in waste and increase in productivity. This cooperation is crucial to enhance flexibility, adaptability and resilience in the production process and in the market. The customer segment of the 8C architecture emphasises the involvement of the customer in product design, production and after-sales operations (Sony, 2020). In a smart factory, the customer can place small-batch orders and participate in the product-design operation for a smart product such as an automobile. The progress of production can be monitored by the customer, and specifications can also be modified during production.

Updated information on the progress of manufacturing activities and related after-sales service for safe-operating, using, maintenance and recycling are then given to the customers. Trace of the product is the record of all the information of the production of the product at every stage (Konstantindis et al., 2018). Including the raw materials, production condition, production parameter, warehouses, shipping, the record of the post-sale service including maintenance schedules, parts replacement and related operating instructions are all retained. Moreover, any customer relevant after-sales information like customer's feedbacks, ratings, complaints or suggestions in format of text, via sound or video are also efficiently stored. This content-oriented version serves the entire lifecycle and related services of the product (Pisching et al., 2016). Analysis of such data can improve the manufacturing process, product design and customer service and also predict potential market trends.

2.4 Shop Floor Monitoring

The rising competitiveness of the market and changing needs of the customers put pressure on manufacturing enterprises to provide products on time more quickly and effectively. The shop floor manager needs to constantly and thoroughly supervise the state of the shop floor to increase production efficiency and reliability, as well as to handle or react to exceptions in production as timely as possible. With the IoT being built, massive real-time data is generated. The engineering challenge is to develop the smart shop-floor that will solve the problem of how to process and make use of the big data to present and predict the state of the shop floor (Zhuang et al., 2021).

In traditional manufacturing processes, the production plan is based on orders, recommendations in advance from other departments and previous production data. This plan is used as a guideline to specify jobs and tasks for teams, workstations and individuals. The plan will then be reviewed and authorized before arranging for the preparation of production. This includes worker training, equipment maintenance, and material preparations. For some specific complicated or new processes, pilot production might also be necessary. The extent of utilising the virtual space is often quite limited, along with other processes, this also has to rely quite heavily on the physical space with very little flexibility and scalability. In the meantime, it is very challenging to ensure the coherence and integrity between the physical world and the virtual counterpart. However, the development of advanced technology such as the IoT allows ubiquitous sensing and gathering of data from different devices and operations (Ning, 2013). In the shop floor, the DTs could be supported by big data and AI. DTs is a forefront and emerging promising technology approach to merge the physical and digital worlds to address current challenges, and promote smart production and management.

2.5 Architecture of DT-Driven Smart Shop Floor

The smart shop floor architecture driven by DT prioritizes instantaneous interaction between physical and digital systems so as to set a two-way information flow: a part of actual data about the functioning of equipment's, material properties, environmental data, etc., sent from the physical space to the virtual space, and the status of the smart shop floor, formulated from virtual space, together with new control commands, sent back to

actual space to adjust the precise control of the production processes there. This two-way connection between the real and virtual worlds could enable accurate and timely control of the production process. The efficiency of such a connection relies on how quick the data required can be acquired, and how fast the cyber and physical systems interact with each other. Any latency in either of these two aspects in a DT-based production system could lead to incorrect decisions and, thus, reduce the effectiveness (Ma et al., 2021).

The architecture of DT-driven smart shop floor comprises of five layers. The physical shop floor layer is a real-world view focusing on physical entities that occur on the production site, such as machine tools, AGV, robots, workers, parts, sensors, RFID among others (Zhang et al., 2018). It must be equipped with IoT technology to collect real-time data from various sources with multi-dimensions. The physical entities in this layer need to accept instructions from instruction sources and perform corresponding tasks, which is the foundation of DT-driven smart shop floor architecture model. The shop floor network layer is one of the most significant views of the DT-driven smart shop floor architecture and plays a key role in connecting physical shop floor and virtual space. It connects various technologies, including industrial internet, internet, industrial wireless network and mobile internet to ensure high availability, high reliability, real-time response and security within the Shop Floor Network (Zhuang et al., 2021). The shop floor network layer plays a crucial role in the information exchange between physical shops floor and cyber space, making DT-driven smart shop floor architecture work efficiently. The shop floor data layer is the basis of DT-driven smart shop floor architecture. The shop floor data layer includes various production data, tooling data, equipment data, material cost data, human data and environmental data among others (Zhang et al., 2018). Diverse data in the shop floor data layer is the foundation of the virtual twin of the shop floor and provides a virtual image of all aspects of the physical shop floor. This layer builds an accurate understanding of the shop floor operation and conditions to make the DT-driven smart shop floor architecture work efficiently. The virtual shop floor layer includes geometric models, manufacturing attribute models, behaviour rule models and data fusion models. The layer is a virtualisation of the physical shop floor. The models provide virtual tools that are critical to real-time assessment, planning and control for the physical shop floor.

The Shop floor Application Layer is responsible for integrating the shop floor application systems that provide decision support for the DT-enabled smart shop floor. These include job scheduling optimizer, dynamic monitoring of manufacturing resources, quality monitoring, optimisation of material delivery among others (Ma et al., 2021). The interface layer is a convenient platform for operators to interact with and operate the DT and also to collect and analyse the real time data from the physical shop floor. This layer is to improve usability to the operators when operating the DT and provides the necessary data at any time to control the DT. Furthermore, this layer can analyse the data of the physical shop floor in real time, which can help identify and resolve issues promptly. Figure 2 illustrates different layers in the intelligent shop floor.

2.6 DT-Driven Smart Shop Floor

The DT technology can offer dynamic simulation and real-time data capture, which would deliver the necessary production data and a simulation tool (Zhang et al., 2018).

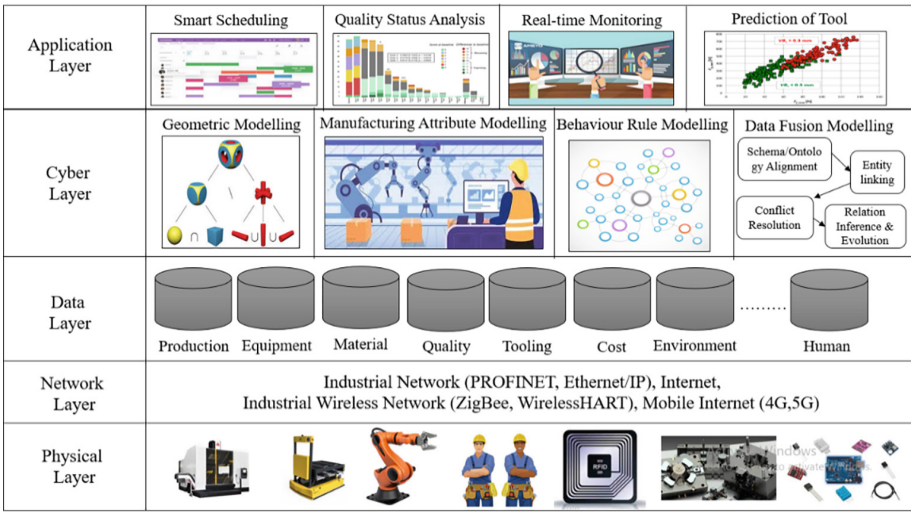


Fig. 2. Architecture of DT-Driven Smart Shop Floor

As a result, this plan would include the following: The Physical Shop floor layer is responsible for collecting production-related data using the IoT network, also known as DT data. This data is then stored and managed in the Shop floor Data Layer, which serves as a centralized repository for all the data collected from the physical shop floor (Tao and Zhang, 2017). The Data Layer is critical in facilitating efficient data analysis, decision-making, and control. The data stored in this layer is used by the cyber-shop floor layer to generate virtual representations of the physical shop floor and provide real-time analysis and control of the physical shop floor.

The Cyber shop floor layer leverages the real-time data collected from the Shop floor data layer to generate virtual representations of the physical shop floor in real-time. This process enables a precise and accurate depiction of the physical shop floor. The Cyber shop floor layer combines the data from the Shop floor data layer with various models, allowing the DT to conduct real-time monitoring, analysis, and control of the physical shop floor. The shop floor application layer employs advanced algorithms to enhance the virtual representations of the physical shop floor within the Cyber Shop floor layer. These optimizations are then presented to the shop floor manager in a graphical or digital format, which enables them to visualize potential improvements in the production process easily. Based on this information, the shop floor manager selects the most optimal solution and transmits the corresponding control commands through the Shop floor Network Layer to the Physical Shop floor Layer. Equipped with the ability to perceive and act on the physical environment, the Physical Shop floor Layer carries out the control commands, implementing the optimized solution and enhancing the overall efficiency of the production process.

In the event of a malfunction in the physical shop floor, such as equipment failure, material shortage, or emergency orders, the Shop floor Data Layer is designed to detect these anomalies by collecting relevant data. The Cyber Shop floor Layer utilizes this data

to reflect the current state of the physical shop floor in real-time, allowing for prompt identification of any problems. The Shop floor Application Layer generates an alert to notify the relevant parties of the issue once identified. Advanced algorithms can be utilized by the Shop floor Application Layer to improve the models within the Cyber Shop floor Layer. The optimization results can then be presented to the shop floor manager for review. If a change is necessary, the new instructions would be communicated to the physical shop floor through the network.

Shop floor managers can get accurate, up-to-date information that reduces workload and make the best resource allocation decisions using DT Technology facilitating real-time interactions requires a robust two-way connection between physical and virtual spaces. This requires several measures, which includes:

- Technologically advanced sensing devices to sense activities in the physical space.
- An effective communication channel set up that connects physical and virtual spaces.
- Up-to-date, comprehensive, and well-maintained databases that are used to store information in the present moment.
- Capable and properly functioning data processing technologies that are used to process such information (Fig. 3).

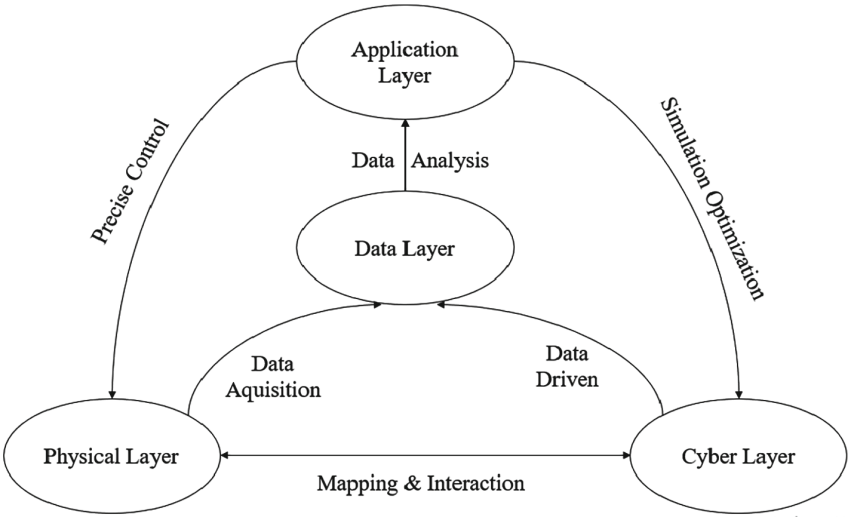


Fig. 3. DRAO scheme for DT-Driven Smart Shop Floor

These technologies must be advanced and up to date to connect physical and virtual spaces and ensure a smooth way to facilitate real-time interactions with digital twins. Generating virtual models that accurately match the physical entities with a high degree of precision can be challenging due to inconsistency, uncertainty, and ambiguity. Security and protection from malicious attacks are also critical in both physical and virtual spaces. These issues must be addressed for normal operations Businesses should also understand the cost and benefits of a DT-driven shop floor to ensure that their interests

are safeguarded. This involves finding a way to balance the costs involved with the benefits gained from using the DT-driven shop floor (Tao and Zhang, 2017). An experimental study creating a digital twin for a Kibble balance helped capture data pertaining to the measurement process and allowed for optimization of the same strengthening the traceability chain(Wright and Davidson, 2020). Another experiment of creating and implementing digital twin in supply chain paved way for assessment of quality parameters and capturing of feedback. Another experiment on multiple vehicles cooperative control was performed using digital twin was found to be very effective (Yang et al., 2022).

3 Discussion

The number of new articles and patents published during the past five years indicates the growth of DT research and practice as industrial standards in any sector require specificity. Close integration between AI tools and ML with DT offers various companies in any sector the possibility of increasing productivity, reducing operation costs, and improving the performance of different activities. The concept of DTs is evolving, making advanced research in DT modelling frameworks, modelling tools and applications in various fields necessary to demonstrate its different utilities and economic benefits. This would encourage more investment and help adopt DT in established manufacturing infrastructures. The integration of virtual models to manage physical assets offers new research opportunities. DT also allows the manufacturer to have a virtual model of their product and process. Having the virtual model allows a producer to identify inefficiencies in production and optimize the manufacturing process.

4 Conclusion

Commercial applications of DT could allow manufacturing organizations to optimize the full stakeholder value chain by simulating the manufacturing process before manufacturing the goods and reduce breakdowns and downtime costs by finding inefficiencies most cost-effectively. It can be employed to monitor and control the quality of the assets produced. Manufacturers can also predict when maintenance is required. This can help prevent unexpected downtime, reduce maintenance costs, and improve equipment efficiency. DTs will also be exceptionally useful for optimising the manufacturing supply chain by creating a virtual representation of the whole chain i.e. suppliers, logistics, and production thereby enabling manufacturers to identify any bottlenecks and inefficiencies. It could bring about a reduction in the lead times, better control of inventory, and improved alignment of materials flow. In summary, using DT in the manufacturing industry is a great opportunity for manufacturers to even improve their productivity, minimise costs, and increase the quality of their merchandise.

Appendix

Concept	References
CPS in Manufacturing	Marilungo et al., 2017 Kim and Park, 2017 Monostori, 2014 Zhang et al., 2020
Cyber-Physical Production Systems (CPPS	Monostori, 2014 Mörth et al., 2020 Zheng et al., 2022 Um et al., 2017
5 C Architecture	Ahmadi et al., 2017 Redelinghuys et al., 2020 Jin et al., 2017 Habib and Chimsom, 2022 Lee et al., 2015 Afrizal et al., 2020
8 C Architecture	Jiang, 2018 Schulz and Freund, 2019 Sony, 2020 Konstantindis et al., 2018 Pisching et al., 2016
Shop Floor Monitoring	Zhuang et al., 2021 Ning, 2013 Ma et al., 2021 Zhang et al., 2018
DT-Driven Smart Shop Floor	Zhang et al., 2018 Tao and Zhang, 2017

References

Abikoye, O.C., et al.: Application of internet of thing and cyber physical system in Industry 4.0 smart manufacturing. In: Emergence of Cyber Physical System and IoT in Smart Automation and Robotics: Computer Engineering in Automation, pp. 203–217. Springer, Cham (2021)

Ahmadi, A., Cherifi, C., Cheutet, V., Ouzrout, Y.: A review of CPS 5 components architecture for manufacturing based on standards. In: 2017 11th International Conference on Software, Knowledge, Information Management and Applications (SKIMA), pp. 1–6. IEEE (2017)

Afrizal, A., Mulyanti, B., Widiaty, I.: Development of Cyber-Physical System (CPS) implementation in industry 4.0. IOP Conf. Ser. Mater. Sci. Eng. **830**(4), 042090. IOP Publishing (2020)

Aslam, F., Aimin, W., Li, M., Ur Rehman, K.: Innovation in the era of IoT and industry 5.0: absolute innovation management (AIM) framework. Information **11**(2), 124 (2020)

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)

Bhatia, L., Tomić, I., Fu, A., Breza, M., McCann, J.A.: Control communication co-design for wide area cyber-physical systems. ACM Trans. Cyber-Phys. Syst. **5**(2), 1–27 (2021)

Boschert, S., Rosen, R.: Digital twin—the simulation aspect. In: Mechatronic Futures: Challenges and Solutions for Mechatronic Systems and Their Designers, pp. 59–74 (2016)

Brusa, E.: Digital twin: toward the integration between system design and rams assessment through the model-based systems engineering. IEEE Syst. J. **15**(3), 3549–3560 (2020)

Cao, K., Hu, S., Shi, Y., Colombo, A.W., Karnouskos, S., Li, X.: A survey on edge and edge-cloud computing assisted cyber-physical systems. IEEE Trans. Industr. Inf. **17**(11), 7806–7819 (2021)

Frantzén, M., Ng, A.H., Moore, P.: A simulation-based scheduling system for real-time optimization and decision making support. Robot. Comput.-Integr. Manuf. **27**(4), 696–705 (2011)

Groshev, M., Guimarães, C., Martín-Pérez, J., de la Oliva, A.: Toward intelligent cyber-physical systems: digital twin meets artificial intelligence. IEEE Commun. Mag. **59**(8), 14–20 (2021)

Habib, M.K., Chimsom, C.: CPS: role, characteristics, architectures and future potentials. Procedia Comput. Sci. **200**, 1347–1358 (2022)

Hamzah, M., et al.: Distributed control of cyber physical system on various domains: a critical review. Systems **11**(4), 208 (2023)

Jiang, J.R.: An improved cyber-physical systems architecture for Industry 4.0 smart factories. Adv. Mech. Eng. **10**(6), 1687814018784192 (2018)

Jin, W., Liu, Z., Shi, Z., Jin, C., Lee, J.: CPS-enabled worry-free industrial applications. In: 2017 Prognostics and System Health Management Conference (PHM-Harbin), pp. 1–7. IEEE (2017)

Konstantindis, F.K., Gasteratos, A., Mouroutsos, S.G.: Vision-based product tracking method for cyber-physical production systems in industry 4.0. In: 2018 IEEE International Conference on Imaging Systems and Techniques (IST), pp. 1–6. IEEE (2018)

Kim, S., Park, S.: CPS (cyber physical system) based manufacturing system optimization. Procedia Comput. Sci. **122**, 518–524 (2017)

Kirchhof, J.C., Michael, J., Rumpe, B., Varga, S., Wortmann, A.: Model-driven digital twin construction: synthesizing the integration of cyber-physical systems with their information systems. In: Proceedings of the 23rd ACM/IEEE International Conference on Model Driven Engineering Languages and Systems, pp. 90–101 (2020)

Lee, J., Bagheri, B., Kao, H.A.: A cyber-physical systems architecture for industry 4.0-based manufacturing systems. Manuf. Lett. **3**, 18–23 (2015)

Li, H., Dimitrovski, A., Song, J.B., Han, Z., Qian, L.: Communication infrastructure design in cyber physical systems with applications in smart grids: a hybrid system framework. IEEE Commun. Surv. Tutor. **16**(3), 1689–1708 (2014)

Li, Y., Roy, U.: Challenges in developing a computational platform to integrate data analytics with simulation-based optimization. In: International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, vol. 57052, p. V01BT02A035. American Society of Mechanical Engineers (2015)

Liu, L., Zhang, X., Wan, X., Zhou, S., Gao, Z.: Digital twin-driven surface roughness prediction and process parameter adaptive optimization. Adv. Eng. Inform. **51**, 101470 (2022)

Lv, Z.: Digital twins in industry 5.0. Research **6**, 0071 (2023)

Ma, X., Cheng, J., Qi, Q., Tao, F.: Artificial intelligence enhanced interaction in DT shop-floor. Procedia CIRP **100**, 858–863 (2021)

Marilungo, E., Papetti, A., Germani, M., Peruzzini, M.: From PSS to CPS design: a real industrial use case toward Industry 4.0. Procedia CIRP **64**, 357–362 (2017)

Matsunaga, F., Zytkowski, V., Valle, P., Deschamps, F.: Optimization of energy efficiency in smart manufacturing through the application of cyber–physical systems and industry 4.0 technologies. J. Energy Resourc. Technol. **144**(10), 102104 (2022)

Monostori, L.: Cyber-physical production systems: roots, expectations and R&D challenges. Procedia CIRP **17**, 9–13 (2014)
Mourtzis, D., Milas, N., Vlachou, A.: An internet of things-based monitoring system for shop-floor control. J. Comput. Inf. Sci. Eng. **18**(2), 021005 (2018)
Mörth, O., Emmanouilidis, C., Hafner, N., Schadler, M.: Cyber-physical systems for performance monitoring in production intralogistics. Comput. Ind. Eng. **142**, 106333 (2020)
Ning, H.: Unit and Ubiquitous Internet of Things. CRC Press, Boca Raton (2013)
Pisching, M.A., Junqueira, F., Dos Santos Filho, D.J., Miyagi, P.E.: An architecture based on IoT and CPS to organize and locate services. In: 2016 IEEE 21st International Conference on Emerging Technologies and Factory Automation (ETFA), pp. 1–4. IEEE (2016)
Pivoto, D.G., De Almeida, L.F., da Rosa Righi, R., Rodrigues, J.J., Lugli, A.B., Alberti, A.M.: Cyber-physical systems architectures for industrial internet of things applications in Industry 4.0: a literature review. J. Manuf. Syst. **58**, 176–192 (2021)
Qian, C., Liu, X., Ripley, C., Qian, M., Liang, F., Yu, W.: Digital twin—cyber replica of physical things: architecture, applications and future research directions. Future Internet **14**(2), 64 (2022)
Rana, M.M., Bo, R.: IoT-based cyber-physical communication architecture: challenges and research directions. IET Cyber-Phys. Syst. Theory Appl. **5**(1), 25–30 (2020)
Redelinghuys, A.J.H., Basson, A.H., Kruger, K.: A six-layer architecture for the digital twin: a manufacturing case study implementation. J. Intell. Manuf. **31**(6), 1383–1402 (2020)
Roy, R.B., et al.: Digital twin: current scenario and a case study on a manufacturing process. Int. J. Adv. Manuf. Technol. **107**, 3691–3714 (2020)
Schulz, K.F., Freund, D.: A multichain architecture for distributed supply chain design in industry 4.0. In: Business Information Systems Workshops: BIS 2018 International Workshops, Berlin, Germany, 18–20 July 2018, Revised Papers 21, pp. 277–288. Springer, Cham (2019)
Sharma, M., Tomar, A., Hazra, A.: Edge computing for industry 5.0: fundamental, applications and research challenges. IEEE Internet Things J. (2024)
Somers, R.J., Douthwaite, J.A., Wagg, D.J., Walkinshaw, N., Hierons, R.M.: Digital-twin-based testing for cyber–physical systems: a systematic literature review. Inf. Softw. Technol. **156**, 107145 (2023)
Sony, M.: Design of cyber physical system architecture for industry 4.0 through lean six sigma: conceptual foundations and research issues. Prod. Manuf. Res. **8**(1), 158–181 (2020)
Thakur, P., Sehgal, V.K.: Emerging architecture for heterogeneous smart cyber-physical systems for industry 5.0. Comput. Ind. Eng. **162**, 107750 (2021)
Tao, F., Qi, Q., Wang, L., Nee, A.Y.C.: Digital twins and cyber–physical systems toward smart manufacturing and industry 4.0: correlation and comparison. Engineering **5**(4), 653–661 (2019)
Tao, F., Xiao, B., Qi, Q., Cheng, J., Ji, P.: Digital twin modeling. J. Manuf. Syst. **64**, 372–389 (2022)
Tao, F., Zhang, M.: Digital twin shop-floor: a new shop-floor paradigm towards smart manufacturing. IEEE Access **5**, 20418–20427 (2017)
Um, J., Weyer, S., Quint, F.: Plug-and-simulate within modular assembly line enabled by digital twins and the use of AutomationML. IFAC-PapersOnLine **50**(1), 15904–15909 (2017)
Wang, B., et al.: Human Digital Twin in the context of Industry 5.0. Robot. Comput.-Integr. Manuf. **85**, 102626 (2024)
Wright, L., Davidson, S.: How to tell the difference between a model and a digital twin. Adv. Model. Simul. Eng. Sci. **7**, 1–13 (2020)
Yang, C., et al.: Multi-vehicle experiment platform: a digital twin realization method. In: 2022 IEEE/SICE International Symposium on System Integration (SII), pp. 705–711. IEEE (2022)
Zhang, H., Zhang, G., Yan, Q.: Dynamic resource allocation optimization for DT-driven smart shopfloor. In: 2018 IEEE 15th International Conference on Networking, Sensing and Control (ICNSC) (2018)

Zhang, H., Yan, Q., Wen, Z.: Information modeling for cyber-physical production system based on DT and AutomationML. Int. J. Adv. Manuf. Technol. **107**(3–4), 1927–1945 (2020)
Zhang, H., Qi, Q., Tao, F.: A consistency evaluation method for digital twin models. J. Manuf. Syst. **65**, 158–168 (2022)
Zheng, Z., Zhang, K., Gao, X.: Human-cyber-physical system for production and operation decision optimization in smart steel plants. Sci. China Technol. Sci. **65**(2), 247–260 (2022)
Zhuang, C., Miao, T., Liu, J., Xiong, H.: The connotation of DT, and the construction and application method of shop-floor DT. Robot. Comput.-Integr. Manuf. **68**, 102075 (2021)

Multi-sensor Data Fusion for Accurate Prediction of Suspicious Activities in Video

R. Radhika[1](✉) and A. Muthukumaravel[2]

[1] Department of Computer Applications, Bharath Institute of Higher Education and Research (BIHER), Chennai, India
rvks18@gmail.com

[2] Faculty of Arts and Science, Bharath Institute of Higher Education and Research (BIHER), Chennai, India
dean.arts@bharathuniv.ac.in

Abstract. Video analytics for surveillance enhances safety of the public across a range of public spaces, such as shopping centers, railroad stations, and daycares. Conventional closed-circuit television systems have proven to be inefficient and prone to errors as they rely heavily on human operators. The proposed intelligent video surveillance system utilizes deep learning algorithms to promptly notify the authorities of any suspicious activity. The system has been trained to identify a wide range of human actions and potential hazards. It achieves this by utilizing datasets like UCF-Crime, HMDB51, and UCF101. The architecture combines LSTM (Long Short Term Memory) with CNN (Convolution Neural Network). The CNN component effectively identifies suspicious activity by extracting important features from video frames using the Inception V3 model. The LSTM examines these characteristics to establish connections over time. The system demonstrated exceptional performance on UCF-Crime, with an F1 score of 99.1%, precision of 99.8%, recall of 98.5%, and an impressive accuracy of 99.6%, as observed during the evaluation across datasets. HMDB51 achieves an impressive F1 score of 99.1%, with precision at 98.7%, recall at 97.9%, and accuracy at 99.4%. On the other hand, UCF101 achieves a commendable F1 score of 98.6%, with accuracy at 99.2%, precision at 98.7%, and recall at 97.9%. The detection durations for UCF-Crime, HMDB51, and UCF101 are 0.25 s, 0.30 s, and 0.28 s, respectively. The proposed model streamlines the procedure of improving safety by utilizing intelligent video analytics.

Keywords: Multi-Sensor Fusion · Deep Learning · Feature-Level Fusion · Convolutional Neural Networks · Activity Detection · Acoustic-Visual Integration · Long Short-Term Memory

1 Introduction

Anticipating suspicious behavior during video surveillance is crucial for enhancing security measures in various settings, such as public spaces, enterprises, and critical infrastructure. Most of the data collected by traditional video surveillance systems comes from

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
S. Rajagopal et al. (Eds.): ASCIS 2024, CCIS 2428, pp. 322–334, 2025.
https://doi.org/10.1007/978-3-031-86302-8_21

a single sensor, often a visual camera [1]. Although these systems have made significant progress due to advancements in computer vision and machine learning, there is still much room for improvement [2]. Occlusion, poor illumination condition and small field of view background noise are the perturbing factors that reduce the efficiency and reliability of activity detection systems [3].

Combining data from a range of sensors can work around its limitations by providing less biased deep insight into the environment under observation. To enhance the data acquisition time, accuracy and reliability in combination with various sensors for multi-sensor system functioning is known as Multi-Sensor Data Fusion. In terms of surveillance, the data collected from different sensors can be stacked up to provide with a better understanding of what is happening and may even uncover some unnoticed details. Optical and thermal data increases the ability to mitigate occlusions, poor lighting. Thermal cameras take pictures based on temperature, which is not affected by light conditions. They capture the additional layer of context that you get from background noise like for gunshots/fire or glass breaking, which wouldn't be possible with an acoustic model. This will be achieved by combining multiple data streams to build a broader and more accurate model of the monitoring situation. This in turn enhances the detection and prediction of abnormal movements. Recent advancements in the field of deep learning open up a wide range of new opportunities for integrating multiple sensors. The amazing skill of CNNs and RNN (Recurrent Neural Network) to draw insightful observations from complex data sets is something that must be appreciated. CNNs are good at working with spatial data like thermal and visual images [4].

In contrast, RNNs are more suitable for the processing of audio signals due to their capability in modelling temporal dependencies [5]. This more complex multi-Sensor Fusion framework that can accurately forecast real-time hazard activity has been developed with the advent of advanced neural network architectures. The overarching goal is to create a mechanism of connecting thermal, visual and audio datasets that will improve the ever-present issue with in ition lucidness within video surveillance systems. Early and late fusion methods are combined with each other in the hybrid fusion process of this model to capture more data as efficiently as possible from all sensor data together. In early fusion, the raw data of all sensors are directly concatenated without addressing intermediate features from each sensor individually.

2 Literature Review

Kershner, Obukhov and Komoltsev [6] presents a method for improving the identification of questionable EEG activities. They propose the use of a continuous Morlet wavelet transform to facilitate the transition to the time-frequency domain. The detection of events is improved through the utilization of neighboring ridge points, while expert criteria are utilized to filter out irrelevant events and prolong their durations. The purpose is to offer a more accurate identification and description of abnormal EEG rhythms.

Amrutha, Jyotsna and Amudha [7] implements a cutting-edge real-time video surveillance system that employs closed-circuit television to closely monitor activities on campus and promptly detect any signs of suspicious or abnormal behavior. The system uses advanced techniques such as LSTM network classification, video pre-processing,

and feature extraction with a VGG-16 model to distinguish between legitimate and nefarious behaviors. The system is designed to promptly alert the relevant authorities through SMS if it detects any suspicious activities, such as physical altercations, medical emergencies, or excessive mobile phone usage.

Gorave et al. [8] develops a sophisticated video surveillance system for retail centers. This system will integrate advanced features such as object identification, motion tracking, and activity monitoring. The primary objective is to detect and identify any suspicious or unlawful behavior. The technology improves real-time monitoring and reaction capabilities by notifying authorities through a dedicated interface, thereby assisting workers in crime prevention and enhancing security efficiency.

Narejo et al. [9] presents a comprehensive security architecture for weapon detection, utilizing IP cameras. The system sends alerts to security personnel via remote monitoring and notifications whenever it detects potentially lethal weapons. When firearms are detected, the system quickly locks the doors and records all activities. The framework employs pretrained models derived from YOLO V3 with Darknet-53 for object detection, specifically to identify weapons with high accuracy and efficiency.

The research conducted by Kulkarni, Swetha and Kusuma [10] provides a detailed approach to detecting examination misconduct in real-time. This approach utilizes peripheral computing and video analytics. The model employs CNN for face identification, YOLO for candidate identification, and LBPH for feature extraction. It utilizes a low-cost webcam to record video frames. An effective method for identifying suspicious actions involves utilizing support vector machines for classification when examining body language patterns. If there is a deviation from the standard, an alert will be issued.

Katragada, Vuribindi and Erickson [11] present a new approach to enhancing ATM security by utilizing advanced YOLO object detection techniques. The system boasts a 99% success rate in detecting card skimmers and other malicious devices with use of YOLO. With its adaptability and minimal training cost, the model offers an answer for real-time threat identification. It also gives a number of industries, such as public security and retail markets the ability to tap into these new levels of advanced protection.

A novel method to analyze and detect all unsuspected conditions in video monitoring using Deep Learning Methods on this field was presented by Dwivedi, Singh & Kushwaha [12]. The system employs an LSTM network to perform temporal analysis and extract features from the Inception V3 model. This is the reason an excellent recognition accuracy of 98.87% is obtained. This scheme offers a very good efficacy for real-time detection of dangers in public domains and substantially outperforms all former methods on the benchmarks among eleven Challenge databases. The researcher presents a method for detecting suspicious activity in crowd videos through an analysis of the prevailing motion direction. The proposed method involves utilizing a spatio-temporal region that relies on shear transforms to acquire motion data, along with an autoencoder for detecting anomalies. This procedure aims to minimize congestion and latency by focusing on the network's peripheral operations. The autoencoder can detect and roughly locate abnormal behaviors by analyzing reconstruction errors that occur during its training on typical tasks. The author proposes a method that utilizes deep learning to detect anomalous events in daycare environments. An adaptive motion detection technique is utilized

to eliminate blurring and transform sequences of images into static frames. Using the RFKD model, one can detect abnormal patterns and send notifications to IoT devices through MQTT. The system has the ability to offer advanced static and dynamic activity prediction through the integration of multi-classifier methods, kernel density functions, and deep neural networks. Through rigorous experimentation, the method being proposed has demonstrated superior performance compared to the ReHAR approach.

Sakthivanitha, M & Priscila, S. S [13] proposed Regional Convolutional Neural Network (RCNN) Algorithm to predict the hunters entering into the reserved forest. This RCNN helps in finding out the illegal entry of poachers in a well efficient manner. Hubert. G & Priscila, S.S [14] proposed an algorithm called YOLO (You Only Look Once) along with CNN in finding out the affected area in retina of small babies.

The following Table 1 represents Overall survey comparison.

Table 1. Overall Survey Comparison

Authors	Technique
Kershner et al. [6]	Morlet Wavelet Transform
Amrutha et al. [7]	LSTM
Gorave et al. [8]	Motion Tracker
Narejo et al. [9]	YOLO V3 with Darknet-53
Kulkarni et al. [10]	CNN + YOLO + LBPH
Katragada et al. [11]	YOLO
Dwivedi et al. [12]	LSTM + Inception V3

3 Methodology

3.1 Sensor Data Acquisition and Preprocessing

For the successful identification and prediction of suspicious activities in video surveillance systems, it is crucial to incorporate a diverse range of sensor data types. Every sensor produces distinct data. The proposed configuration consists of visual, thermal, and auditory sensors. For activity predictions to be accurate and reliable, it is crucial to collect and process the precise data from different types of sensors. Vision cameras capture RGB images, which contain a wealth of spatial information that is useful for detecting and identifying objects and individuals. One of the many goals of visual data preparation is to simplify feature extraction and enhance image quality. As part of the noise reduction process, the image $I(x, y)$ is condensed using a Gaussian kernel $G(x, y)$:

$$I_{blurred}(x, y) = I(x, y) * G(x, y) \tag{1}$$

where $*$ denotes convolution. The Gaussian kernel $G(x, y)$ is defined as:

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \tag{2}$$

where σ is the standard deviation of the Gaussian distribution. This step helps in smoothing the image and reducing random noise. The brightness and contrast are improved by utilizing histogram equalization. This method improves the image contrast by implementing a range of adjustments to the intensity levels. The function that equalizes histograms is as follows:

$$I_{equalized}(x, y) = c.H(I(x, y)) \tag{3}$$

where H is the cumulative distribution function of the image histogram and c is a scaling constant. The Sobel operator is commonly used to detect edges and highlight structural characteristics. The computation of the x- and y-gradients is performed using the Sobel operator. The magnitude of the gradient is given by:

$$G = \sqrt{((\partial I(x, y))/\partial x)^2 + ((\partial I(x, y))/\partial y)^2} \tag{4}$$

The utilization of these processes is highly beneficial as they enable the retrieval of relevant data from visual input, a crucial aspect for activity recognition and object detection. Thermal cameras have the ability to capture images even when there is no visible light, as they rely on temperature fluctuations to convey important data. Normalization is a crucial step in thermal image preprocessing to maintain consistent temperature ranges across recordings. One commonly used method for normalizing pixel values involves scaling them to a common value. This is often achieved through the use of min-max scaling:

$$I_{norm}(x, y) = \frac{I(x, y) - I_{min}}{I_{max} - I_{min}} \tag{5}$$

The purpose of calculating thermal gradients is to detect temperature fluctuations that could indicate potentially suspicious activities. Calculating the magnitude of the gradient is similar to analyzing visual data:

$$G = \sqrt{\left(\frac{\partial I_{norm}(x, y)}{\partial x}\right)^2 + \left(\frac{\partial I_{norm}(x, y)}{\partial y}\right)^2} \tag{6}$$

An analysis of audio wave temporal patterns can be conducted by converting the data obtained from acoustic sensors into spectrograms. The Short-Time Fourier Transform method takes care of conversion of time domain signals to frequency domain:

$$STFT(x(t)) = \int_{-\infty}^{\infty} x(\tau).w(t - \tau).e^{-j2\pi f\tau} d\tau \tag{7}$$

where $x(t)$ represents the acoustic signal, $w(t - \tau)$ is the window function, and $e^{-j2\pi f\tau}$ is the Fourier transform component. For data fusion to work successfully, the time signature of the sensor readings must match. Through the utilization of Dynamic Time Warping (DTW) techniques, we are able to minimize discrepancies in the sensor data streams and synchronize the timestamps, thereby achieving alignment. The DTW algorithm iteratively adjusts the time axis to find the best alignment path, ensuring that data from different sensors is synchronized. The visual, thermal, and aural data are processed together to enhance the effectiveness of each sensor in the prediction framework. This is made up of inputs that can then be fed forward to the remaining phases: data fusion and analysis.

3.2 Data Fusion Techniques

Combining data that comes from two or more sensors requires advanced fusion algorithms to be able to apply and use the different kind of information which each provide in an efficient way. Multi-sensor fusion often uses Feature-Level Fusion. This method merges features from multiple sensors into an integrated representation before classification. This method emphasizes the combining of thermal, optical and acoustic sensor information which results in higher detection rates over prediction. The data of each sensor is then merged after features are extracted using apt methods. Convolutional Neural Networks (CNN) can help to extract spatial information from images, benefiting visual and thermal data. CNNs often contain convolutional, pooling and activation layers. The convolution operation takes an input image,, and then applies a filter.

$$F(x, y) = \sum_{m,n} I(x + m, y + n).K(m, n) \tag{8}$$

feature maps after computations Following layers include activation functions e.g. ReLU and pooling to downsample the output at each subsequent layer focusing on selective features. These last visual constraints are of feature vector type, meaning their final values represent spatial contrasts for the CNNs and temperature differences with respect to time. Changes in auditory data processing. The Short-Time Fourier Transform (STFT) is an essential tool for audio source analysis by transforming them to spectrograms, which show the temporal and frequency aspects of a signal. These type of spectrograms are useful in extracting features responsible for speech characterization e. g., mel frequency cepstral coefficients that provide a compact representation of the spectral characteristics of the audio source [3]. The procedure for extracting MFCCs is organized as follows:

$$MFCC_n = \sum\nolimits_k \log(S(k)).cos\left[\frac{n.(k - 0.5)\pi}{N}\right] \tag{9}$$

The power spectrum, cepstral coefficients, and number of coefficients are represented by $S(k)$, the cepstral coefficient index, and n, respectively. The integration of acoustic, thermal, and optical variables into a comprehensive model is achieved through Feature-Level Fusion, once these variables have been extracted. By combining the feature vectors from visual, thermal, and acoustic sensors, the following combination can be achieved:

$$F_f = [F_v; F_t; F_a] \tag{10}$$

The program can predict activity by combining geographical, temperature, and auditory data from multiple sources. Attention mechanisms enable a more advanced fusion process. The evaluation of the significance of different features is a continuous process that attention processes engage in to maintain focus. The attention weights α_i are calculated through a specific procedure.

$$\alpha_i = \frac{e^{W_i F_i}}{\sum_j e^{W_j F_j}} \tag{11}$$

where W_i represents learnable parameters, and F_i denotes individual feature vectors. The weighted sum of features is then given by:

$$F_{final} = \sum_i \alpha_i F_i \tag{12}$$

The features used for classification are represented by the symbol F_{final}. Applying attention approaches enhances the accuracy of the model in identifying suspicious activity by enabling it to concentrate on more informative features. Through the utilization of the unique capabilities of each sensor, Feature-Level Fusion presents a reliable approach for integrating data from different sensors. This technology revolutionizes activity prediction systems by incorporating attention mechanisms and seamlessly integrating spatial, thermal, and auditory data. The result is a significant boost in computational power, akin to that of a supercomputer. This technology transforms activity prediction systems into powerful computational tools by integrating attention mechanisms and incorporating spatial, physiological, and auditory data. Figure 1 illustrates the system's flow diagram.

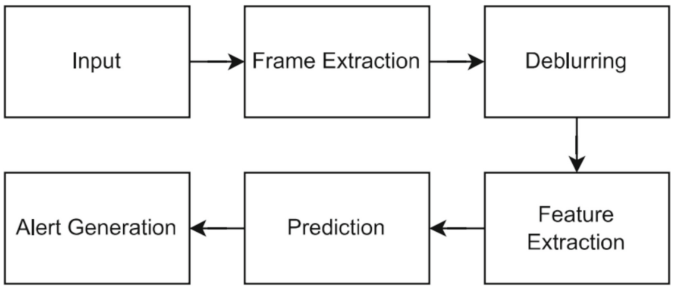


Fig. 1. Flow diagram of the system

3.3 Integration of Deep Learning

The design of the deep learning model is important in effectively utilizing fused multisensor data to accurately predict suspicious actions. This system utilizes a hybrid deep learning architecture that combines RNNs and CNNs to analyze the image patterns in the data. This approach utilizes the power of Recurrent Neural Networks to analyze sequential data and Convolutional Neural Networks to extract spatial features from images, with a specific focus on audio data. CNNs are utilized for feature extraction and processing of visual and thermal images. The CNN is composed of three layers: convolutional, activation, and pooling. A typical design for a CNN includes multiple convolutional layers, activation functions like ReLU that introduce non-linearity, and pooling layers that reduce dimensionality while retaining important features. MaxPooling technique is utilized to decrease the dimensions of the image and focus on the most important features by selecting the highest value within a specified window.LSTM network which is a RNN type, is used for temporal sequence modeling. Long Short-Term Memory (LSTM) models are highly effective in handling features extracted from audio data spectrograms.

These models excel at capturing and understanding the complex, long-term patterns present in sequential data. The composition of a long short-term memory (LSTM) unit includes an output gate o_t and a pair of input gates i_t. The main LSTM cell utilizes the following equation to update its state:

$$c_t = f_t.c_{t-1} + i_t.\tilde{c}_t \tag{13}$$

The equation includes several variables: the current cell state c_t, the previous cell state c_{t-1}, the output of the forget gate f_t, the output of the input gate i_t, and the candidate cell state $\tilde{c}_t$. The LSTM can retain information over long sequences by transmitting the memory through the network using the cell state c_t. After the feature vectors derived from the CNNs and LSTMs have converged, a comprehensive feature representation is generated. The next step in detecting suspicious activity involves applying the integrated feature vector to densely connected layers. In dense layers, non-linear activation functions are used to map the aggregated information to the final output space. The final classification is determined using probability, which is calculated using a softmax activation function.

$$\hat{y}_i = \frac{e^{z_i}}{\sum_j e^{z_j}} \tag{14}$$

The logit score represents class i, while the predicted likelihood is denoted by $\hat{y}_i$. We utilize a cross-entropy loss function to measure the difference between the original and observed probabilities of a class, enabling us to effectively train the deep learning model. There are multiple approaches available to express the loss function:

$$L = -\sum_i y_i \log(\hat{y}_i) \tag{15}$$

The probability of class i being predicted and the actual label are denoted as $\hat{y}_i$ and y_i, respectively. Optimization methods such as SGD and Adam are used to adjust the model parameters and minimize the loss function during training. Utilizing a hybrid deep learning architecture that combines LSTMs and CNNs, the model can effectively analyze and integrate spatial and temporal properties from data gathered by multiple sensors. This method enhances the accuracy and reliability of activity prediction by integrating the most advanced features of CNN and RNN techniques.

Algorithm 1:Multi-Sensor Data Fusion and Activity Prediction

```
function PredictActivity(visualData, thermalData, auditoryData): {Suspicious, NotSuspicious}
    # Preprocess Visual Data
    I_blurred ← GaussianBlur(visualData)
    I_equalized ← HistogramEqualization(I_blurred)
    G_visual ← SobelGradient(I_equalized)
    # Preprocess Thermal Data
    I_norm ← MinMaxScaling(thermalData)
    G_thermal ← ThermalGradient(I_norm)
    # Preprocess Auditory Data
    STFT_result ← ShortTimeFourierTransform(auditoryData)
    MFCCs ← ExtractMFCC(STFT_result)
    # Feature Extraction using CNN for Visual and Thermal Data
    F_visual ← CNNFeatureExtraction(G_visual)
    F_thermal ← CNNFeatureExtraction(G_thermal)
    # Feature Extraction using LSTM for Auditory Data
    F_auditory ← LSTMFeatureExtraction(MFCCs)
    # Feature-Level Fusion
    F_combined ← Concatenate(F_visual, F_thermal, F_auditory)
    AttentionWeights ← ComputeAttentionWeights(F_combined)
    F_final ← WeightedSum(F_combined, AttentionWeights)
    # Hybrid Deep Learning Model for Classification
    prediction ← HybridModel(F_final)
    if prediction ≥ threshold then
        return Suspicious
    else
        return NotSuspicious
    end if
end function
```

4 Result and Discussion

The open-source datasets UCF-Crime, HMDB51, and UCF101 were used to evaluate our system's ability to identify suspicious activity through video. UCF-Crime provides users with real-world surveillance footage for the purpose of crime detection, HMDB51 displays a diverse spectrum of human actions, and UCF101 displays a wide range of sports. The quantity and diversity of data contained in these datasets enable us to confidently train and evaluate our detection algorithm. Table 2 describes the features of the datasets. Table 3 illustrates the findings. The system exhibited a remarkable peak accuracy of 95.6% on UCF-Crime, but it also performed admirably on other datasets. The Precision and Recall metrics suggest that the system is capable of identifying suspect behavior to a reasonable extent. Detection Time is a critical metric for the efficacy of processing in real-time surveillance systems.

Table 2. Description about datasets

Dataset	Number of Videos	Types of Activities	Total Duration (hours)
UCF-Crime	1,900	Assault, Burglary, Robbery, etc.	128
HMDB51	6,766	Clapping, Running, Walking, etc.	7.5
UCF101	13,320	Various Sports	27.8

Table 3. KPI across datasets

Metric	UCF-Crime	HMDB51	UCF101
Accuracy	99.6%	99.2%	99.4%
Precision	99.8%	98.7%	98.7%
Recall	98.5%	98.6%	97.9%
F1 Score	99.1%	99.1%	98.6%
Detection Time	0.25 s	0.30 s	0.28 s

Table 4 compares the accuracy of the proposed system with other existing systems. The proposed system achieves a high accuracy of 99.4%, outperforming previous methods by Dwivedi et al. (94.93%), Lu et al. (96.12%), Narejo et al. (96.76%), and Katragada et al. (98%), demonstrating its superior performance in detecting suspicious activities.

Figure 2 illustrates the training loss and accuracy over 50 epochs for the proposed system. The training loss begins at 15 and steadily decreases to 0.5, while the training

Table 4. Comparison with other systems

Methods	Accuracy
Proposed System	99.4%
Dwivedi et al. [12] [LSTM + Inception V3]	94.93%
Lu et al. [5] [YOLO V3 with Darknet-53]	96.12%
Narejo et al. [9] [YOLO V3 with Darknet-53]	96.76
Katragada et al. [11] [YOLO]	98%

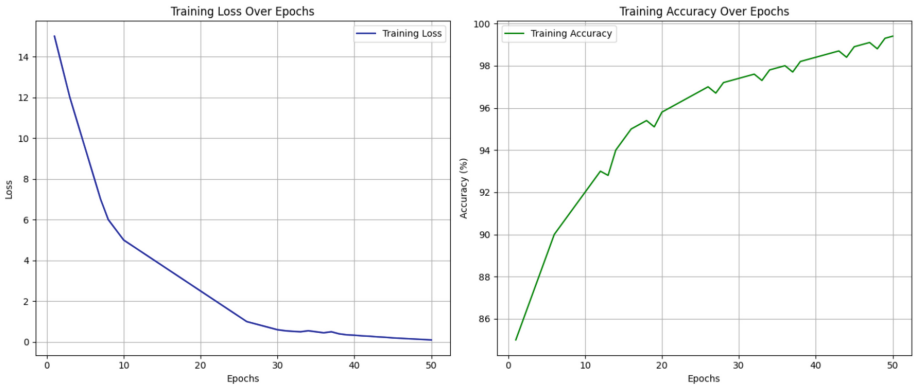


Fig. 2. Training Accuracy and Loss over Epochs

accuracy starts at 85% and gradually improves to 99.4%, indicating effective model learning and convergence. Figure 3 compares the accuracy of different systems for detecting suspicious activities in video footage. The proposed system achieved accuracy of 99.4%, doing better than other methods such as CNN (93%), RNN (91%), LSTM (95%), RF (82%), and SVM (86%), demonstrating its superior performance.

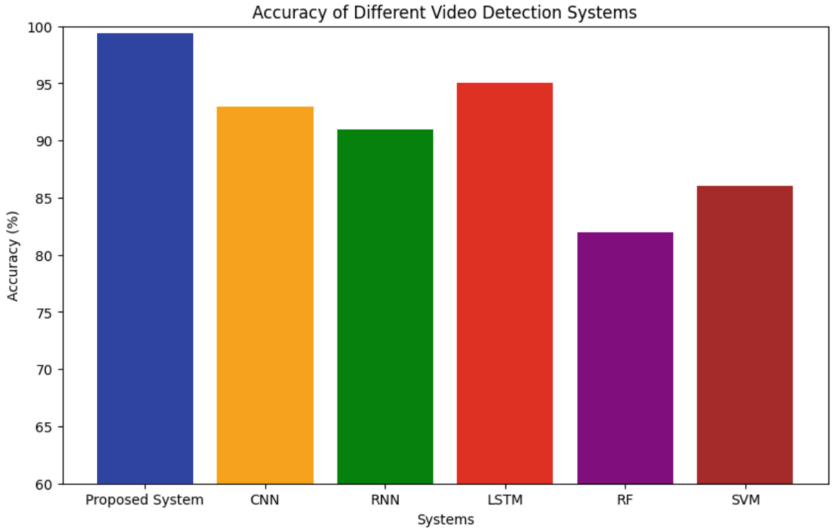


Fig. 3. Accuracy of different models

Figure 4 demonstrates the camera system's recording of an ongoing armed robbery. The system showcases its real-time detection and notification capabilities through an alert that is immediately noticeable. It shows how the system can quickly identify and respond to potentially suspicious actions.

Fig. 4. Detecting the suspicious activity

5 Conclusion and Future Work

The intelligent video surveillance proposed system greatly enhances real-time monitoring capabilities through the integration of DL. This allows for the detection of suspicious activities with exceptional precision. After being evaluated on three separate datasets - UCF-Crime, HMDB51, and UCF101 - the system demonstrates impressive accuracy, recall and F1 scores. Several levels of accuracy can be achieved, such as 99.6%, 99.2%, and 99.4%. The system has the potential to greatly enhance public safety in various settings, thanks to its exceptional performance and faster detection times compared to current methods. Future work aims to incorporate additional datasets into the system to improve its ability to respond to emerging threats. The main objective is to improve the ability to withstand different weather patterns and lighting conditions, resulting in reduced latency. Additionally, we aim to improve processing efficiency. By leveraging cutting-edge transfer learning and domain adaptation techniques, the system's ability to adapt to new situations could be enhanced.

References

1. Elharrouss, O., Almaadeed, N., Al-Maadeed, S.: A review of video surveillance systems. J. Vis. Commun. Image Represent. **77**, 103116 (2021)
2. Tsakanikas, V., Dagiuklas, T.: Video surveillance systems-current status and future trends. Comput. Electr. Eng. **70**, 736–753 (2018)
3. Yadav, D.K., Singh, K., Kumari, S.: Challenging issues of video surveillance system using internet of things in cloud environment. In: Advances in Computing and Data Sciences: First International Conference, ICACDS 2016, Ghaziabad, India, 11–12 November 2016, Revised Selected Papers 1, pp. 471–481. Springer, Singapore (2017)
4. Du, P., Bai, X., Tan, K. et al. Advances of Four Machine Learning Methods for Spatial Data Handling: a Review. J geovis spat anal 4, 13 (2020). https://doi.org/10.1007/s41651-020-00048-5
5. Lu, Y., et al.: Auditory attention decoding from electroencephalography based on long short-term memory networks. Biomed. Sig. Process. Control **70**, 102966 (2021)
6. Kershner, I.A., Obukhov, Y.V., Komoltsev, I.G.: Post-traumatic epilepsy in rats: an algorithm for detection of suspicious EEG activity. In: 2017 13th International Conference on Signal-Image Technology and Internet-Based Systems (SITIS), Jaipur, India, pp. 163–16 (2017). https://doi.org/10.1109/SITIS.2017.36

7. Amrutha, C.V., Jyotsna, C., Amudha, J.: Deep learning approach for suspicious activity detection from surveillance video. In: 2020 2nd International Conference on Innovative Mechanisms for Industry Applications (ICIMIA), Bangalore, India, pp. 335–339 (2020). https://doi.org/10.1109/ICIMIA48430.2020.9074920
8. Gorave, A., Misra, S., Padir, O., Patil, A., Ladole, K.: Suspicious activity detection using live video analysis. In: Bhalla, S., Kwan, P., Bedekar, M., Phalnikar, R., Sirsikar, S. (eds.) Proceeding of International Conference on Computational Science and Applications. Algorithms for Intelligent Systems. Springer, Singapore (2020). https://doi.org/10.1007/978-981-15-0790-8_21
9. Narejo, S., Pandey, B., Esenarro Vargas, D., Rodriguez, C., Anjum, M.R.: Weapon detection using YOLO V3 for smart surveillance system. Math. Probl. Eng. **2021**(1), 9975700 (2021)
10. Kulkarni, A.S., Swetha, N.E.M., Kusuma, S.M.: Automated system for detection of suspicious activity in examination hall. In: 2021 IEEE International Conference on Electronics, Computing and Communication Technologies (CONECCT), Bangalore, India, pp. 01–05 (2021) https://doi.org/10.1109/CONECCT52877.2021.9622599
11. Katragada, A.S., Vuribindi, K., Erickson, V.L.: Real-time suspicious activity detection on ATMs using multimodel YOLO object detection. In: 2023 International Conference on Computational Science and Computational Intelligence (CSCI), Las Vegas, NV, USA, pp. 1254–1258 (2023). https://doi.org/10.1109/CSCI62032.2023.00204
12. Dwivedi, N., Singh, D.K., Kushwaha, D.S.: A novel approach for suspicious activity detection with deep learning. Multimedia Tools Appl. **82**(21), 32397–32420 (2023)
13. Sakthivanitha, M., Priscila, S.S.: Detection of animal hunters in forest using regional convolutional neural network algorithm. In: 2023 International Conference on New Frontiers in Communication, Automation, Management and Security (ICCAMS), pp. 1–8. IEEE, Bangalore, India (2023). https://doi.org/10.1109/ICCAMS60113.2023.10525851
14. Hubert, G., Priscila, S.S.: You Only Look Once (YOLO) with Convolution Neural Network (CNN) classification for preterm baby's retinopathy images. In: Rajagopal, S., Popat, K., Meva, D., Bajeja, S. (eds.) Advancements in Smart Computing and Information Security. ASCIS 2023. Communications in Computer and Information Science, vol. 2038. Springer, Cham (2024). https://doi.org/10.1007/978-3-031-59097-9_27

The Emergence of the Social Internet of Vehicles (SIoV): A Comprehensive Analysis of Architecture, Technologies, and Applications

Tarandeep Kaur(✉) and Pankaj Deep Kaur

Guru Nanak Dev University Regional Campus, Jalandhar, Punjab, India
deep.taran4@gmail.com, pankajdeep.csejal@gndu.ac.in

Abstract. Social Internet of Vehicles (SIoV) marks a transformative step in the evolution of vehicular networks, blending the interactive principles of social networking interactions with the interconnected functionality of contemporary transportation systems. This paper explores the revolutionary potential of SIoV, providing a thorough analysis of its fundamental architecture and the advanced technologies driving its evolution. Improved connectivity and more intelligent interactions between vehicles are encouraged by SIoV utilizing edge computing, blockchain, Vehicle-to-Everything (V2X) communication, and 5G technology improvements. This paper proposes a novel architecture for the Social IoV named "VSN-A" i.e. Vehicular Social Networking Architecture. The proposed architecture has four layers- The vehicle communication layer, the Social Networking layer, the data processing layer, and the application and user interface layer. Through this comprehensive examination, the authors identify the existing research gaps and proposed future directions, underscoring the pivotal role of SIoV in shaping the future of intelligent transportation.

Keywords: Social networks · Internet of Vehicles · V2X (Vehicle to Everything) · Social Internet of Vehicles

1 Introduction

The Internet of Vehicles has emerged as a transformative evolution in the automotive and transportation industries. It integrated cutting-edge communication technologies with modern vehicles to form a highly interconnected and intelligent transportation ecosystem [1]. In simpler terms, IoV is described as a network where vehicles are equipped with sensing facilities, computing facilities, software, and storage to communicate with their surroundings such as pedestrians, cell phones, traffic lights, other vehicles, etc. Social Internet of Vehicles, termed SIoV extends the concept of IoV by incorporating social networking principles and transforming vehicles into social entities capable of interacting. SIoV is also often called as Vehicular Social Networks (VSN), where vehicles act as smart objects [2, 3]. This interaction is not just with their immediate surroundings, but also with other vehicles that interact dynamically and intelligently. The SIoV is conceptualized as a vehicle to establish and maintain its social networks. Suppose,

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
S. Rajagopal et al. (Eds.): ASCIS 2024, CCIS 2428, pp. 335–347, 2025.
https://doi.org/10.1007/978-3-031-86302-8_22

a group of vehicles having a common point of interest (POI) and following the same route. These vehicles can socialize with other vehicles and exchange information such as traffic congestion, emergencies, safety warnings, vacant parking slots, restaurants with good food, and so on. Another example could be vehicles like ambulances, police cars, and fire trucks responding to emergencies. They can share their real-time location, road conditions, and available routes, to the emergency sites. Similarly, Electric Vehicles (EVs) in the same geographical areas can socialize, to update about the availability of charging stations, routes to the charging stations, and energy consumption rates for different routes.

This study explores the frameworks, connections, and dynamics within the SIoV, investigating how social frameworks can be applied to vehicular networks to establish a resilient and effective system. Through an analysis of core technologies, message formats, and practical applications, this paper seeks to offer a thorough grasp of how SIoV could transform urban transportation in Smart Cities. The primary objectives of this study are:

i) This paper offers a comprehensive view of the fundamentals of the Social Internet of Vehicles (SIoV) domain.
ii) This study proposes a novel architecture "VSN-A" for the Social Internet of Vehicles that embeds the concept of the Internet of Vehicles and Social Networks i.e. IoV and SN.
iii) This Study discusses various thrust areas of SIoV along with providing directions for future research.
iv) This study presents a case study that explores the application of the proposed architecture in a real-world scenario.

Section 2 discusses the literature review related to the Social Internet of Vehicles. Section 3 discusses the conceptual framework of SIoV including key components, SIoV infrastructure, and SIoV relationships. Section 4 covers the various SIoV relationship types. Section 5 discusses the proposed novel architecture of SIoV. Section 6 explores thrust areas of SIoV where the major focus of researchers is required. Section 7 discusses a case study that explores the potential applications of the proposed architecture. Section 8 sheds some light on the future directions in the context of SIoV.

2 Literature Review

Authors in [4] explored the importance of Edge Intelligence (EI) in the context of IoV. They have proposed a layered vehicular EI architecture to exploit the potential of EI in heterogeneous IoV. The authors presented various research directions such as generative AI-driven IoV, vehicle-road-cloud integration, and vehicular cyber-physical fusion. The integration of AI and the optimization of edge services is explored in [6]. This study discussed AI-driven methods such as task offloading, deep reinforcement learning, and adaptive algorithms, which could be used to enhance decision-making.

Authors in [7] proposed a model for IoV for sharing information and generating recommendations. This agent-based model is comprised of three entities: humans, vehicles, and points of interest. They found that recommendations that are generated using

weak ties are generally more accurate than the recommendations generated using strong ties. Authors in [8] discussed the processing of data generated by smart vehicles at the edge. They proposed an edge information system (EIS) that includes edge caching, edge computing, and edge AI. EIS acts as the intermediate between the onboard processors and remote cloud data centers.

Authors in paper [9] discussed the two-fold nature of the Social Internet of Vehicles: Vehicle-Oriented Social Internet of Vehicles (VO-SIoV) and Driver-Oriented Social Internet of Vehicles (DO-SIoV). In VO-SIoV, vehicles maintain their social networks and communicate with other vehicles. In DO-SIoV, the major focus is on human behaviors. Authors in [10] proposed a hybrid recommendation system "SafeDrive". This system uses the Dynamic Driver Profile (DDP) approach for the analysis of the driver's behavior. For this, behavioral sensors are used for driver-driving prediction. After identifying the driver's behavior machine learning models are applied to generate the appropriate recommendations. This proposed system helps in generating early warnings and preventing road accidents. The authors in [11] proposed route recommendation systems to offer tourists real-time and personalized recommendations. This proposed recommendation system reduces the traffic jams at tourist hot spots.

3 Conceptual Framework

3.1 Key Components of Social Internet of Vehicles

3.1.1 Vehicles

Connected Vehicles (CV): These are the vehicles capable of connecting to other cars using wireless networks and exchanging information about traffic conditions, emergencies, etc. [13, 14] . A vehicle can communicate with other cars, and various infrastructures such as RSUs, traffic lights, etc. using technologies such as V2V, V2I, V2R, V2N.

Electrical Vehicles (EVs): Electrical vehicles are those vehicles that are powered partially or wholly by electricity instead of traditional combustion engines. There are various types of electric vehicles such as battery EVs, Plug-in Hybrid EVs, and Fuel Cell EVs. Electric Vehicles focus on innovation, sustainability, and efficiency. EVs are equipped with sensors, and communication technologies, and use V2X infrastructure to communicate.

Autonomous Vehicles (AV): These are the next-generation cars having self-learning and self-driving capabilities. These need little or no human interactions while driving. An Autonomous Vehicle requires sensors, actuators, complex algorithms, machine learning system for its efficient working. For Example: air quality sensors in autonomous vehicles help to understand passenger comfort and ventilation within the vehicle.

3.1.2 Communication Infrastructure

Vehicles can have various types of connectivity [18, 19, 23]. Many connections that a vehicle can establish are identified and explained in detail below:

- **Inter V:** Many sensors (vehicle sensors) present within the vehicle collect data. Inter-V deals with the information collected about the vehicle such as speed, etc.
- **V2V:** Allows direct communication between the vehicles. These connected vehicles can share relevant information such as location, speed, obstacles detected, etc.
- **V2I:** Allows a vehicle to communicate with the infrastructure present on the roadside such as traffic light signals, weather sensors, road signs, etc.
- **V2P:** Allows the communications between the vehicle and the pedestrians using their devices such as mobile phones. Pedestrians can receive alerts on their mobile phones if there is an emergency situation.
- **V2N:** It enables vehicles to connect to external networks such as cloud. This can help vehicles get real-time information about traffic situations, parking slots, weather conditions, etc.
- **V2C:** It allows vehicles to store and access the data from the cloud.
- **V2G:** It facilitates the communication between power grids and EVs. This working of this network relies on electrical energy. This energy can be stored in electric vehicles (EVs).
- **V2X:** It allows communication between vehicle to everything i.e. V2X where X includes internet, road, people, vehicles, etc.

3.1.3 Sensors and Actuators

Various sensors are deployed in the IoV environment that enable vehicles to collect and share data. These sensors can be divided into two major categories: ***i) Vehicle Sensors***: sensors that collect data related to the vehicle such as speed, acceleration, positioning, distance, etc.; ***ii) Environmental Sensors:*** sensors collecting data about the surroundings such as temperature, O_2 levels, rain, etc. The categorization of major sensors being used in the context of IoV is described in Fig. 1.

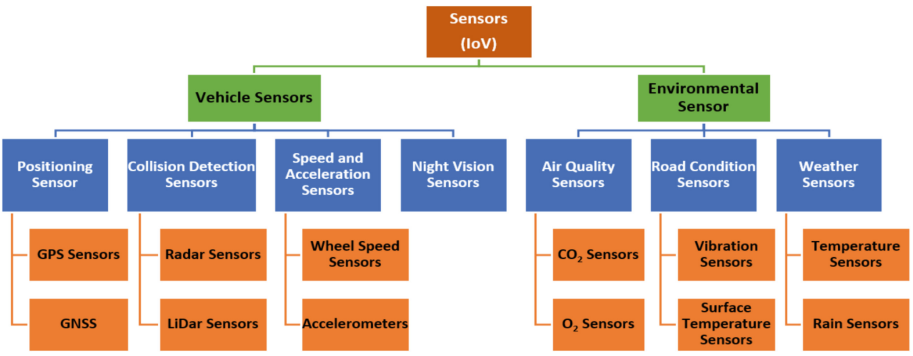


Fig. 1. Types of Sensors used in SIoV environment

3.1.4 AI and ML in SIoV

In the era of connected smart vehicles, it is required to integrate AI techniques for intelligent decision-making. Figure 2 describes how Artificial intelligence and Machine Learning are used in the context of SIoV for model training and model inference.

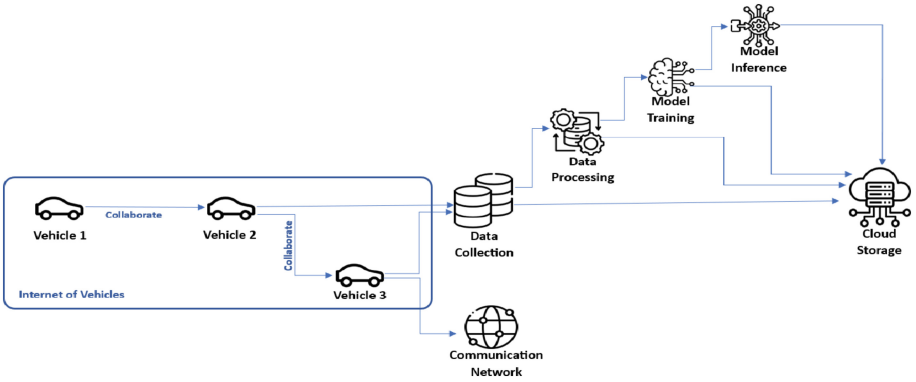


Fig. 2. AL/ML in SIoV

The major benefits of using AI and ML in SIoV are discussed below:

- **Improved V2V communication and coordination:** AI systems can help vehicles exchange information and plan actions in real-time, which can improve traffic management, collision avoidance, and cooperative driving.
- **Enhanced Perception and Decision Making:** More accurate perception of the vehicle's environment can be made possible by machine learning (ML) models that will process data from multiple sources.
- **Adaptive and Personalized Services:** AI can offer drivers and passengers individualized services and recommendations, by evaluating driving habits, preferences, and contextual data.
- **Predictive Maintenance and Fleet Management:** By analyzing vehicle health data and forecasting potential problems, AI-driven SIoV systems can enhance maintenance schedules, and minimize downtime.

4 SIoV Relationships Types

The notion of "Social" in the SIoV i.e. Social Internet of Vehicles states that the vehicles can establish their relationships with other vehicles [20]. Using these relationships, vehicles can exchange information without human intervention. Few relationships have been identified that can be established between the vehicles themselves to exchange information, coordinate their actions, and provide personalized services within the SIoV ecosystem [21]. *Table 1* defines these identified relationships in detail.

Table 1. SIoV Relationships and Definition

Relationship Between Vehicles	Full Form	Definition
POR [24, 25]	Parent Object Relationship	Vehicles produces by the same manufacturer
SOR [25]	Social Object Relationship	Vehicles that have established a social relationship
CWOR [24, 25]	Co-Work Object Relationship	Vehicles collaborate to provide services to the shared IoT application
OOR [24, 25]	Ownership Object Relationship	Vehicles owned by the same manufacturer or entity
STGOR	Stranger Object Relationship	Vehicles that suddenly encounters each other and start communicating
CLOR [25]	Co-Location Object Relationship	Vehicles that are present at the same location
GOR	Guardian Object Relationship	Vehicles that act as the early warning systems for other vehicles

5 Architectural Design of SIoV

5.1 Proposed Architecture of SIoV

A four-layered novel architecture (Fig. 3) has been proposed integrating the concepts of IoV, Social Networks, Artificial Intelligence and Machine Learning.

Vehicle Communication Layer (Layer 1) This layer deals with the communication of the vehicle with other vehicles and infrastructure. Vehicles can make several types of connections with other vehicles or infrastructure including V2V, V2P, V2R, V2I, V2N, V2C, V2H, and V2X. All these communications are described in Sect. 3 in detail. Vehicles connect with other vehicles or infrastructure to exchange data such as speed, temperature, road conditions in real-time. Vehicles communicate with each other using DSRC or C-V2X technologies.

Social Networking Layer (Layer 2) Vehicles along with sharing real-time information can also make their own social networks. These Social Networks (SN) helps the vehicles to provide more relevant information to other vehicles based on common interests and contextual information. Each vehicle can have a user profile. Whenever a new device wants to enter the social network, vehicle is registered and a list of services is made. This helps in the service discovery and provisioning.

Data Processing Layer (Layer 3) This layer deals with data processing and analysis so as to provide services at the application layer. Data collected by multiple vehicles is

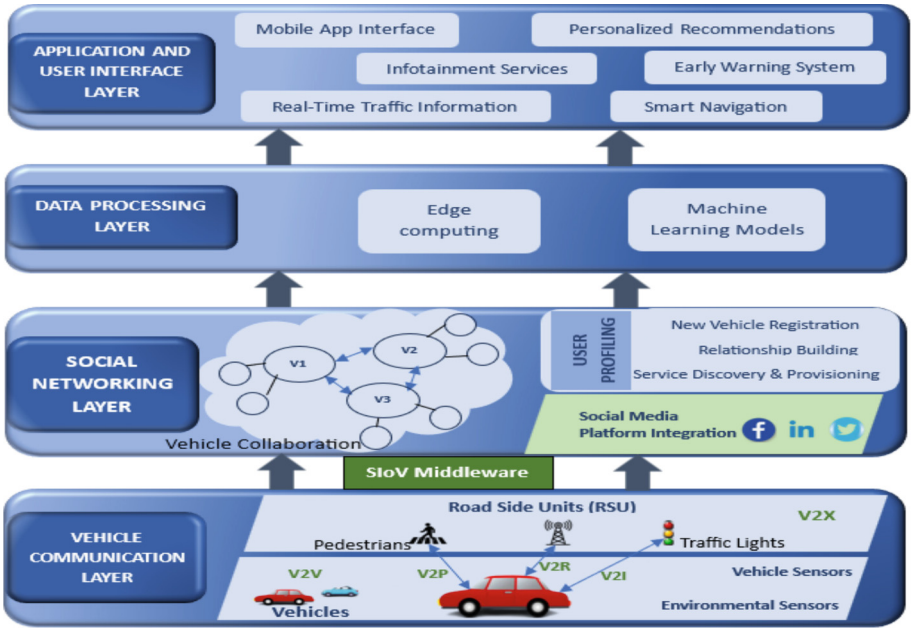


Fig. 3. Proposed Architecture "VSN-A" for Social Internet of Vehicles (SIoV)

processed at the edge to reduce the latency and to enhance the response time. Technologies like Artificial Intelligence and Machine Learning can be embedded at this layer to enhance the process of data analysis and provide more relevant results at the user interface layer.

Application and User Interface Layer (Layer 4) This layer provides user-centric services to the vehicle owners and drivers. Services like personalized route recommendations, early warning signs in case of accidents, real-time traffic congestion information, etc.

5.2 Security and Privacy Considerations SIoV

The realm of data security and privacy is both vast and intricate [15]. In this proposed architecture, several data security and privacy aspects are to be considered, some of which are discussed below in brief:

- *At layer 1,* the V2X system will collect data from vehicles and infrastructure. To keep the integrity of data, data encryption methods such as Transport Layer Security (TLS) can be used. Prior to sharing of any data, the vehicles and the infrastructure have to authenticate each other to prevent spoofing.
- *At layer 2,* strict access control procedures should be in place to guarantee that only services and apps with permission can access user profiles and data from social networking sites.

- *At layer 3,* edge computing reduces latency but also provides some security concerns. It needs to be protected against unwanted access of edge devices.
- *At layer 4,* to guarantee safe connection between the vehicle and user devices, mobile apps and infotainment systems that interface with the vehicle require secure APIs (Application Programming Interfaces). To stop illegal access to private information, API security is crucial.

6 SIoV Thrust Areas

SIoT encompasses four key components: Service discovery and composition, network navigability, trust management, and relationship management [26]. While creating social networks for the Internet of Vehicles, these four major areas are the major considerations. In this study, our major focus is on service discovery and trust management in the context of SIoV.

6.1 Service Discovery in SIoV

In the notion of the IoV, Service discovery is a mechanism that allows vehicles to identify and access services provided by other vehicles. The majority of the research revolves around the service discovery in static nodes [21]. In the case of IoV, nodes are mobile, thus making connections, finding services, and providing service is not an easy task. The key factors that influence the service discovery process in the IoV network include mobility patterns of the vehicle, communication range, network topology, latency requirements, scalability, and environmental factors such as road types, traffic patterns, etc. Authors in [17] proposed CDSD for the Internet of Vehicles network, which is a distributed service discovery protocol based on forming clusters. Various researchers have proposed other service discovery protocols also for mobile ad-hoc networks, but those are not suitable for the highly dynamic IoV environment. Authors in [27] proposed a novel approach "semantic accessors" for contextual service discovery and integration in connected vehicles. This approach facilitates interaction with various services such as parking, traffic information, etc. They have considered the contextual features that enhance the vehicle's ability to adapt to changing environments and user needs.

There are various challenges of service discovery in IoV

- *As vehicles are mobile, the network topology is changing frequently. This makes the process of service discovery and availability difficult.* ***Possible solution:*** To adapt to the frequent topology changes distributed service discovery protocols like VANETs and cluster-based discovery can help.
- *As connected vehicles offer a limited communication range, it limits the process of service discovery and providing services within a certain range limit only.* ***Possible solution:*** Using multi-hop communication, neighboring intermediary vehicles may relay messages between the service requester and provider.

- *The IoV network is scalable adding more and more nodes to the network, the process of registering vehicles to the network, finding relevant services, and providing the services becomes complicated.* ***Possible solution:*** The system can use hierarchical architectures where the network can be partitioned into regions or clusters, each having its local service discovery system.
- *It is challenging to handle the heterogeneity in the vehicle type, communication technologies, etc.* ***Possible solution:*** Implement semantic service discovery protocols in a description of services that can dynamically manage the protocols and services to maintain uninterrupted interaction across heterogeneous networks.
- *As vehicles are mobile and change their location frequently, this fast-paced vehicle movement leads to information staleness.* ***Possible solution:*** RSUs or edge servers also can help in maintaining localized, real-time service directories updated periodically.

6.2 Trust Management in SIoV

Trust is the extent to which one has faith and expectations regarding an individual's integrity and stability [28] [29]. Trust management is essential for social networks. In the Social Internet of vehicles, Trust is influenced by multiple factors including social relationships, reputation, context, etc.

Authors in [25] have discussed the various characteristics of trust such as subjective, objective, local, global, context-specific and asymmetric, Authors in [2] presented an extensive overview of managing trust within the Social Internet of Vehicles. The various factors that influence the designing of the trust model in SIoV are also identified. The authors also explored the various challenges associated with maintaining the reliability of the trust model. Authors in [30] proposed TACASHI, an architecture that provides trust-based communications in SIoV. They utilize Online Social Networks (OSN) to strengthen the trust between the social in-vehicle (within the vehicle) and inter-vehicle communication (with other vehicles) architecture. Drivers' honesty is analyzed using their OSN profile. TACASHI architecture involves five major components: the vehicle itself, the vehicle owner, passengers within vehicles, Roadside units, and OSN accounts connected to drivers and passengers.

There are various challenges of Trust Management in IoV

- It is challenging to integrate heterogeneity into a single trust model, especially when there are numerous communication protocols and devices with various features and needs.
- Trust management involves handling sensitive data to be gathered and processed, such as a vehicle's behavior or a profile of some given driver, and location information is added to it. The collection of such data causes severe privacy issues, especially in the case of a multi-party that may be involved in exchanging trust data or public networks.
- Malicious vehicles can fake earning trust ratings by contributing false information or even through the creation of multiple pseudonyms called Sybil attacks to artificially inflate their reputation.

- In an emergency, decisions about who to trust must sometimes be made quickly, leaving little time for thorough verification of information.

7 Future Directions

Although SIoV has emerged as a profound evolution, its heterogeneous architecture and complex resources lead to the need for the development of novel frameworks or protocols [16].

- **Service discovery:** Researchers could focus on predictive modeling, dynamic clustering, routing optimization, and context-aware services to overcome the various challenges that prevail in the service discovery in the IoV.
- **Sustainable and Environmental Impact:** The impact of connected vehicles on the environment needs to be addressed, particularly concerning energy consumption and emissions.
- **Integration of blockchain technology:** Researchers could explore how blockchain could be integrated with SIoV to provide a transparent and tamper-proof ledger for transactions and communications within the IoV.
- **User-Centric Applications:** More applications focusing on user engagement and experience must be developed, such as personalized recommendations based on date and time.
- **Autonomous vehicle integration:** Researchers should focus on developing robust communication protocols that allow autonomous vehicles to share information with surrounding infrastructure.
- **Multi-Modal Transportation:** Integration of the IoV with other modes of transportation, such as public transit, shared mobility, can be explored to create a seamless and efficient multi-modal transportation system

8 Case Study- A Smart Emergency Response System

In urban cities, timely response to emergencies is critical for saving lives. The proposed VSN-A architecture that utilizes SIoV can be leveraged to enhance the emergency response capabilities of vehicles. Imagine a city deploying a fleet of AEVs equipped with advanced communication systems that allow them to interact with each other, traffic infrastructure, and emergency services.

At layer 1, **Vehicle Communication Layer**, AEVs can use V2V communication to share real-time information about traffic conditions, road conditions, potential obstacles, and the status of ongoing emergencies. For example, if one AEV encounters heavy traffic, it can alert others to reroute. AEVs can communicate with traffic signals using V2X infrastructure to create green corridors, allowing them to pass through intersections without delays. At layer 2, **Social Networking Layer**, AEVs can discover and connect to other AEVs using the service discovery. Public awareness can be enhanced as citizens can receive alerts about ongoing emergencies through a mobile app, allowing them to avoid affected areas and understand the situation better. Users/Citizens can provide feedback on emergency response times, helping authorities improve services. At layer 3, **Data Processing Layer**, Data from AEVs and traffic systems is analyzed to optimize routes

and predict potential emergency hotspots using edge computing. Using the machine learning models, the system learns from past emergencies to improve response strategies and resource allocation. At layer 4, **Application and User Interface Layer**, Users can track the nearest AEV and receive updates on emergencies, enhancing community safety awareness using a web interface or an emergency app.

In the above-discussed case study, the proposed VSN-A architecture enhances urban emergency response by enabling real-time data sharing and communication among AEVs, traffic infrastructure, and citizens. This integration increases public safety, improves decision-making, and optimizes resource allocation, ultimately leading to faster and more effective emergency responses.

9 Conclusion

The Social Internet of Vehicles (SIoV) represents a groundbreaking advancement in the integration of social networking principles with vehicular communication systems. This paper provides a holistic overview of SIoV fundamentals, key components, and communication infrastructure. Various SIoV relationships are also explored that the vehicles can maintain while communicating with each other or the infrastructure. The proposed architecture "VSN-A" adopts a multi-layered approach based on various types of communications and uses the push technologies of AI, blockchain, and edge computing for vehicular interaction improvement. While SIoV holds tremendous potential for improving vehicle communication, safety, and efficiency, challenges related to trust management, service discovery, and security remain. Addressing these issues will effectively help in reaching all the optimum potentials offered by SIoV and pave the way for future innovations that may render intelligent transportation networks safer, scalable, and effective.

References

1. Ahmed, S., Kumari, S., Saleem, M.A., Agarwal, K., Mahmood, K., Yang, M.H.: Anonymous key-agreement protocol for V2G environment within social internet of vehicles. IEEE Access **8** (2020). https://doi.org/10.1109/ACCESS.2020.3003298
2. Iqbal, R., Butt, T.A., Afzaal, M., Salah, K.: Trust management in social Internet of vehicles: factors, challenges, blockchain, and fog solutions. Int. J. Distrib. Sens. Netw. **15**(1) (2019). https://doi.org/10.1177/1550147719825820
3. Marwein, P.S., Sur, S.N., Gao, X.Z., Kandar, D.: Recent Survey on Internet of Vehicles: Architecture, Applications, Challenges, and Its Solutions (2024). https://doi.org/10.1520/JTE20230095
4. Yan, G., Liu, K., Liu, C., Zhang, J.: Edge intelligence for internet of vehicles: a survey. IEEE Trans. Consum. Electron. (2024). https://doi.org/10.1109/TCE.2024.3378509
5. Guo, J., Bilal, M., Qiu, Y., Qian, C., Xu, X., Raymond Choo, K.K.: Survey on digital twins for Internet of Vehicles: fundamentals, challenges, and opportunities. Digit. Commun. Netw. **10**(2) (2024). https://doi.org/10.1016/j.dcan.2022.05.023
6. Xu, X., Li, H., Xu, W., Liu, Z., Yao, L., Dai, F.: Artificial intelligence for edge service optimization in Internet of Vehicles: a survey (2022). https://doi.org/10.26599/TST.2020.9010025

7. Zia, K., Shafi, M., Farooq, U.: Improving recommendation accuracy using social network of owners in social internet of vehicles. Future Internet **12**(4) (2020). https://doi.org/10.3390/FI12040069
8. Zhang, J., Letaief, K.B.: Mobile Edge Intelligence and Computing for the Internet of Vehicles (2020). https://doi.org/10.1109/JPROC.2019.2947490
9. Loscri, V., Manzoni, P., Nitti, M., Ruggeri, G., Vegni, A.M.: A social internet of vehicles sharing SIoT relationships. In: Proceedings of the International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc) (2019). https://doi.org/10.1145/3331052.3332470
10. Nouh, R., Singh, M., Singh, D.: Safedrive: hybrid recommendation system architecture for early safety predication using internet of vehicles. Sensors **21**(11) (2021). https://doi.org/10.3390/s21113893
11. Liu, L., Xu, J., Liao, S.S., Chen, H.: A real-time personalized route recommendation system for self-drive tourists based on vehicle to vehicle communication. Expert Syst. Appl. **41**(7) (2014). https://doi.org/10.1016/j.eswa.2013.11.035
12. de Sousa Bezerra, C.D., Teles Vieira, F.H., Queiroz Carneiro, D.P.: Autonomous robotic navigation approach using deep q-network late fusion and people detection-based collision avoidance. Appl. Sci. (Switzerland) **13**(22) (2023). https://doi.org/10.3390/app132212350
13. Kong, X., Gao, H., Shen, G., Duan, G., Das, S.K.: FedVCP: a federated-learning-based cooperative positioning scheme for social internet of vehicles. IEEE Trans. Comput. Soc. Syst. **9**(1) (2022). https://doi.org/10.1109/TCSS.2021.3062053
14. Wang, T., Hussain, A., Zhang, L., Zhao, C.: Collaborative edge computing for social internet of vehicles to alleviate traffic congestion. IEEE Trans. Comput. Soc. Syst. **9**(1), (2022). https://doi.org/10.1109/TCSS.2021.3074038
15. Xu, X., et al.: PSDF: privacy-aware IoV service deployment with federated learning in cloud-edge computing. ACM Trans. Intell. Syst. Technol. **13**(5) (2022). https://doi.org/10.1145/3501810
16. Qureshi, K.N., Din, S., Jeon, G., Piccialli, F.: Internet of vehicles: key technologies, network model, solutions and challenges with future aspects. IEEE Trans. Intell. Transp. Syst. **22**(3) (2021). https://doi.org/10.1109/TITS.2020.2994972
17. Sharma, N., Chauhan, N., Chand, N.: Cluster based distributed service discovery in internet of vehicle. J. Commun. Softw. Syst. **17**(3) (2021). https://doi.org/10.24138/jcomss-2021-0069
18. Zhang, J., Letaief, K.B.: Mobile edge intelligence and computing for the internet of vehicles. Institute of Electrical and Electronics Engineers Inc. (2020). https://doi.org/10.1109/JPROC.2019.2947490
19. Zhou, H., Xu, W., Chen, J., Wang, W.: Evolutionary V2X technologies toward the internet of vehicles: challenges and opportunities. Proc. IEEE **108**(2) (2020). https://doi.org/10.1109/JPROC.2019.2961937
20. Alam, K.M., Saini, M., El Saddik, A.: Toward social internet of vehicles: Concept, architecture, and applications. IEEE Access **3** (2015). https://doi.org/10.1109/ACCESS.2015.2416657
21. Araújo, I., Pedrosa, M.F., Castro, J., dos Anjos, E.G., Matos, F.: Service discovery based on social profiles of objects in a social IoT network. In: ICCSA 2019. LNCS (LNAI and LNB) (2019). https://doi.org/10.1007/978-3-030-24308-1_33
22. Hamrouni, A., Khanfor, A., Ghazzai, H., Massoud, Y.: Context-aware service discovery: graph techniques for IoT network learning and socially connected objects. IEEE Access **10** (2022). https://doi.org/10.1109/ACCESS.2022.3212370
23. Karim, S.M., Habbal, A., Chaudhry, S.A., Irshad, A.: Architecture, Protocols, and Security in IoV: Taxonomy, Analysis, Challenges, and Solutions (2022). https://doi.org/10.1155/2022/1131479

24. Roopa, M.S., Buyya, R., Venugopal, K.R., Iyengar, S.S., Patnaik, L.M.: DRCM: dynamic relationship creation and management in social internet of things. Int. J. Intell. Internet Things Comput. **1**(3) (2021). https://doi.org/10.1504/ijiitc.2021.115696
25. Sagar, S., Mahmood, A., Sheng, Q.Z., Zhang, W.E., Zhang, Y., Pabani, J.K.: Understanding the trustworthiness management in the social Internet of Things: a survey. Elsevier B.V. (2024). https://doi.org/10.1016/j.comnet.2024.110611
26. Farhadi, B., Masoud Rahmani, A., Asghari, P., Hosseinzadeh, M.: Friendship selection and management in social internet of things: a systematic review (2021). https://doi.org/10.1016/j.comnet.2021.108568
27. Weber, M., Akella, R., Lee, E.A.: Service discovery for the connected car with semantic accessors. In: IEEE Intelligent Vehicles Symposium, Proceedings (2019). https://doi.org/10.1109/IVS.2019.8813884
28. Roopa, M.S., Pattar, S., Buyya, R., Venugopal, K.R., Iyengar, S.S., Patnaik, L.M.: Social Internet of Things (SIoT): foundations, thrust areas, systematic review and future directions (2019). https://doi.org/10.1016/j.comcom.2019.03.009
29. Becherer, M., Hussain, O.K., Zhang, Y., Den Hartog, F., Chang, E.: On trust recommendations in the social internet of things - a survey. ACM. Comput. Surv. **56**(6) (2024). https://doi.org/10.1145/3645100
30. Kerrache, C.A., et al.: TACASHI: trust-aware communication architecture for social internet of vehicles. IEEE Internet Things J. **6**(4) (2019). https://doi.org/10.1109/JIOT.2018.2880332

Network and Cloud Computing

Innovative Perceptions: A Survey on Named Data Networks Through an Inventive Lens

A. Shanmuga Priya(✉) and K. Vaitheki

Department of Computer Science, Pondicherry University, Puducherry, India
withshanmu@pondiuni.ac.in

Abstract. The Internet is used for data transfer, where the client and the server communicate through IP addresses, leading to P2P communication. The data across a network being a vital component, the focus is more on the data than the location of the data. Hence, the IP address that designates the server hosting the data is of less significance. As IP is facing challenges to capture the internet community, content-centric networking platforms such as Named Data Networking have emerged as content-based networks. NDN permits user to request data without awareness of the server and manages memory consumption, protection and flexibility issues more efficient than the current IP-based Internet. NDN is a progressive worldwide reclassifying architecture that provides the notion of information transmission over networks. This survey aims to foster the essential ideas of Named Data Networking, investigating its assorted applications, and address the difficulties and future headings in the realtime NDN. Further the correlation between IPv6 and NDN has been examined. These essentially highlight the importance of the NDN and the potential for improving administration and security of web.

Keywords: Content-centric Networking · Named Data Networking · Internet Protocol · cache · Naming · address · security · IPV6

1 Introduction

With Information-centric networking (ICN), content takes center stage in network architecture instead of hosting. This is a revolutionary approach. Data in ICN is shared and accessed by name rather than location. Two well-known ICN architectures that represent this change are Named Data Networking and Content-Centric Networking, prioritizing content retrieval over other network user's names. CCN aims to optimize the management of requests and content through network protocols that leverage Interest and Content Objects. The basic concept is that the network delivers the specified Content Object back to the user when they send out an Interest packet describing the content. This method eliminates the need for an immediate connection between the user and the host from which it originated by retrieving the content from the closest node that has it cached. While NDN and CCN share similar fundamental ideas, NDN emphasizes the infrastructure's function in content delivery and caching within network nodes. This

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
S. Rajagopal et al. (Eds.): ASCIS 2024, CCIS 2428, pp. 351–364, 2025.
https://doi.org/10.1007/978-3-031-86302-8_23

improves efficiency and security while also making content distribution more straightforward. Through network protocols, NDN makes it possible for data packets to be safely labeled and signed, guaranteeing the authenticity and integrity of the information [1].

Caching data across all nodes, Named Data Networking (NDN), removes the need for frequent content requests to the original server. It uses several new and standard applications to illustrate the benefits of a well-designed NDN infrastructure. Because of the architecture design, building trustworthy networks and apps is feasible with NDN. NDN can improve performance and user experience by lowering the volume of data kept in the per-content cache and sending the initial byte faster. Furthermore, NDN supports various techniques deemed challenging to apply using IP addresses. Furthermore, partial data retrieval, autonomous path switching, transparent failover, and the ability to access content from several sources are essential components of an NDN infrastructure that simplify operations and improve user experience [2].

The amount of time users spends on the Internet for communication increases day-by-day. The sharing of information—text, video, and audio, among other formats—is growing daily, and users now primarily need content—not host or IP addresses. Identifying a network interface and its location are the two main functions of an IP address. A new addressing scheme called IPV6 employs 2128 addresses, while IPV4 uses 32-bit addressing for 232 addresses. More space was needed when the address was increased, and NDN was created using a content name. NDN aims to make developing applications and the problems with the current internet architecture more manageable concerning security, scalability, and sustainability. As a result, difficulties with naming, signing, routing, and forwarding arise during the development, implementation, and comprehension of NDN architecture.[3].

The Internet was first created as a packet data network, in which clients communicated across a pre-established communication channel with data sources with distinct IP addresses. In recent years, this client-server data transfer method has transformed into a peer-to-peer data-sharing approach. Applications like YouTube, Bit Torrent, and social networks have fundamentally altered our thoughts about user-generated content. Regardless of the data sources, current users consider data. As a result, the notion that IP addresses may be used to locate servers hosting specific content is becoming less and less relevant. In addition, the Internet community has long struggled with the need for IP addresses. In answer to the current need for a content-centric networking infrastructure where data hosts are less significant, Named Data Networking (NDN) was developed [4].

2 NDN Architecture

NDN preserves the hourglass architecture and represents a thin waist as the core of the architecture, which is similar to the current IP architecture (see Fig. 1) [5]. The Named Data Networking (NDN) architecture is a new approach to computer networks that focuses on the data rather than the locations or addresses of the devices where it resides. In traditional IP-based networking, data is retrieved from a specific area (an IP address) using protocols like HTTP, FTP, etc. However, in NDN, data is retrieved by its name, which could be any arbitrary string. When you request data, you ask for it by

name, and the network finds the closest or most available copy. NDN grasps the shape of an hourglass based on the protocols in each layer, like current internet architecture. Slight modifications are made to the IP architecture to support the features of NDN. Strategy and security are the two new layers embedded in NDN architecture, although the waist or core layer is named data/content chunk.

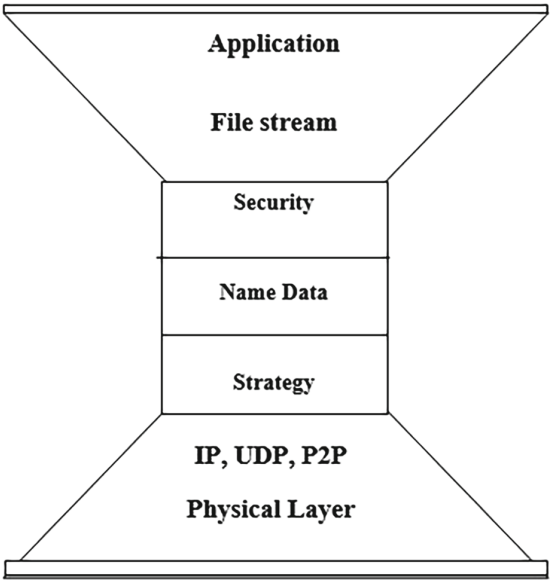


Fig. 1. NDN Hourglass Architecture

Table 1 lists NDN's three data structures—CS, PIT, and FIB—that satisfy users for the requested material. NDN fetches content based on name; naming is done independently of the network and the application. A content name (CN) functions through its hierarchical structure. Key components of the NDN architecture include Named Data, in which Data is given by a name that serves as its identifier. This allows data retrieval based on its name despite its location, and the Forwarding Strategy determines how the requests are forwarded through the network. Unlike IP routing, which is based on destination address, forwarding in NDN is based on data names.NDN offers several potential benefits, including better support for content distribution, built-in security through cryptographic signatures on data, and inherent support for multicast communication. Overall, the NDN architecture addresses some of the limitations of traditional IP-based networking and offers a more flexible and efficient way to access and distribute data on the Internet.

In a Named Data Network, information is exchanged between Interest packets and Data packets (see Fig. 2) [1]. A name (content name) designating a necessary piece of data is carried by both the interest packet and the data packet. The content names of the desired material are contained in interest packets, and routers employ these content names to forward the packets to the data. The content name and producer-provided

Table 1. NDN Data structures and functionalities

Data structure	Description
Content Store	Each node maintains a cache of recently requested data in the Content Store. If a node receives a request for data, it already has, it can serve the request directly from its cache without needing to retrieve it from the source again
Pending Interest Table (PIT)	A node records an item in the Pending Interest Table to record requests for data it does not currently have. The requester and any additional nodes that indicated Interest in the same data based on the initial request receive the data after it has been fetched
Forwarding Information Base(FIB)	FIB determines how Interest packets should be forwarded through the network. It acts as a routing table, mapping Data names to the interface or next hops where an Interest packet is sent

signature are then included in the data packet for security. This data packet follows the interest packet's backward route to return to the customer.

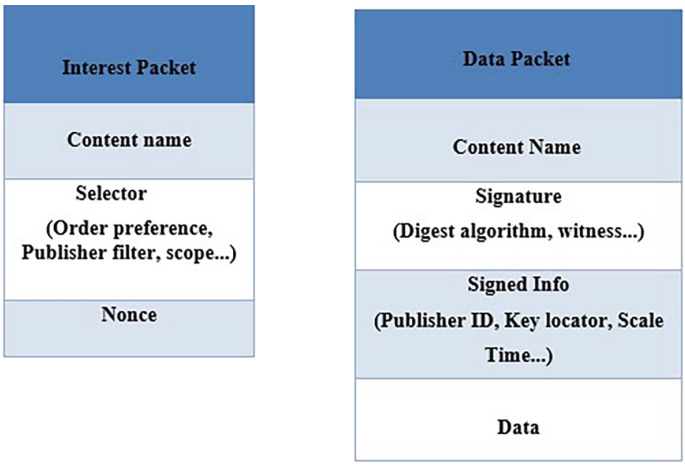


Fig. 2. Packets in NDN

3 Principles of NDN

- **Named Data**: Every data item in NDN has a unique name attached to it. You can request content from anywhere in the network using this name, which acts as an identification for the content.
- **Content Retrieval by Name**: Unlike traditional networking, which communicates with specific IP addresses. NDN routers forward interest packets according to the content's name. To satisfy the Interest routers forward data packets in the opposite direction.
- **Replication and Caching**: NDN allows for caching at different network nodes. A router can cache content when it receives a data packet. If the cache has the content, it can be used to satisfy subsequent Interest packets for the same item, saving bandwidth and delay.
- **Security by Design:** NDN's fundamental components include security measures. Producer signatures on data packets guarantee their authenticity and integrity. This increases NDN's intrinsic security against attacks, such as IP spoofing.
- **Stateful Forwarding and Multicast:** By keeping track of Interest packets they forward, NDN routers enable effective multicasting of Data packets to numerous consumers expressing Interest in the same data.

4 NDN Operations

NDN router's working mechanism (see Fig. 3) [2], the consumer searches for the content the producer delivers. Producers respond to customer requests for content. When the NDN router receives such a request, it will first check the Content Store to verify if the necessary content is already cached. The router looks for PIT using the content name if a linked entry specifies that it should store the content locally and deliver it based on the following PIT entry. To initiate an NDN discussion, the client supplies an Interest packet. Once the producer has been captured, the Interest packet will reappear along with a Data packet. The data packet follows the path to make sure the content is returned when the content name and content are combined.

As a result, the relevant data packet is accessible, and all interested customers agree to a copy. The longest prefix match (LPM) is routed to FIB for Content Name if no PIT record exists. When matched FIB reaches this Longest prefix match site, it proceeds to the proper next-hop and generates a new PIT entry using its incoming interface. Conversely, accessing a data packet in the NDN router looks for a matching content name by sifting through every PIT record. The data packet is routed through each of the assigned interfaces. By the local caching policy, any further PIT entries are deleted, and the content is retained in the Content Store to fulfill upcoming requests. The data packet is dropped if no matching entry is found in PIT.

5 NDN Services

The primary services of NDN (see Fig. 4). NDN provides caching and forwarding to the clients while enabling in-network caching of requested content. For NDN, routing decides on stateful forwarding. The NDN architecture, which offers secure communication via named data with signature, is fundamentally based on security.

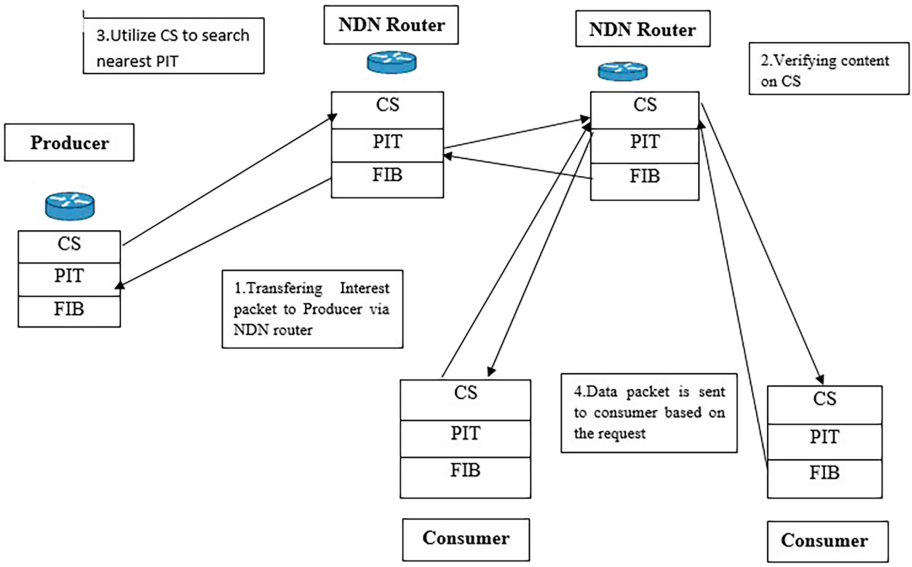


Fig. 3. Working of NDN

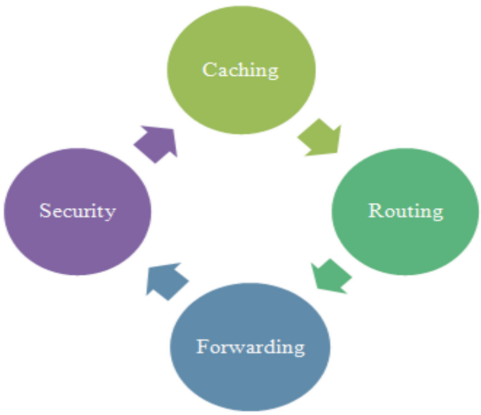


Fig. 4. NDN Services

5.1 Caching

In-network caching, which supports content-centric NDN's fundamental purpose for low-cost data transmission, is content caching in each node. The replacement policy chooses what needs to be deleted after an incoming data packet is retained and the cache is complete. The caching decision approach, which chooses when to cache an incoming data packet, is the two parts of a node's caching system. Caching Everything Elsewhere (CEE) is the name of the NDN default caching strategy. All incoming data is therefore retained on nodes, and standard policies like Least-Recently-Used (LRU)

handle replacement [6]. The NDN network allows multiple copies of identical data, eliminating a single fault point and reducing production costs.

The following performance metrics are measured for efficient caching [7]:

- Propagation speed: The time needed to get content to the network core.
- Hit ratio: The average of all requests received divided by the number of requests fulfilled at the intermediary nodes. Consequently, it results in less energy being used by battery-operated gadgets.
- Packet delay: The time between creating an interest and the consumer getting the data packet. With the help of this metric, we can assess the caching strategy's efficiency from the user's standpoint.
- Hop count: The quantity of hops required to retrieve the material. From an operator-centric standpoint, this indicator helps us determine how effective the caching method is: the fewer hops there are, the less likely network congestion is, and the amount of data that must traverse.

5.2 Routing

Content names are used to route and forward NDN packets.NDN does not need to handle problems that conflict with the current IP architecture. Public and private addressing is not used in the routing method, and Network Address Translation (NAT) is not utilized. In NDN, there is only one difference between forwarding and routing: forwarding chooses and uses routes based on their performance or status, while routing assesses whether a route is available. Both IP and NDN utilize the FIB for routing data storage. IP checks the address in the FIB when sending packets to a destination address to identify the next hop—which isn't always the best route—to take. Rather than continuously retrieving the closest duplicate, NDN searches for the subsequent hop(s) by NDN interactions with CNs, which may result in the usage of NDN routing tables. Using interest/data packets, the NLSR protocol for intra-domain routing in Named Data Networking (NDN) distributes routing updates while taking advantage of NDN's inherent data authenticity [9]. SDAR aims to transfer routing decision-making to a centralized controller [10].

The following performance metrics are measured for efficient routing [8] are

- Memory consumption indicates the overall memory used for named content routing.
- PIT count indicates overall PIT entries at an NDN router.
- Network utilization indicates the total amount of data transported over the network.
- Time to completion: The overall time needed to fulfill a content request.

5.3 Forwarding

Using a forwarding strategy, NDN chooses the best interfaces to forward the path of Interest and works out how best to use the available forwarding paths or options. Real NDN names enable effective name retrieval and set up the Learning Name Index (LNI) for the NDN forwarding plane [11]. By maintaining virtual connections throughout online interactions, the session-based forwarding approach strikes a compromise between fault recovery capabilities and dynamic content traffic [12].

The following performance metrics are measured for efficient forwarding [13] are

- Interest satisfaction ratio: The ratio of the total number of Data packets retrieved to the total number of Interest packets.
- Interest satisfaction latency: The average amount of time that passes between when data consumers send interest packets and when they get data packets.
- Energy consumption: The sum of energy consumed will forwarding interest and data packet by all nodes.

5.4 Security

NDN uses content encryption to make confidentiality and access control features possible. Access control policies can be implemented within this framework without needing trusted servers or directories. Furthermore, since the decryption key is sent with the material, it doesn't need to be distributed. The two main types of NDN security encryption are symmetric and asymmetric. When symmetric encryption is used, decryption is accomplished using a secret key in a signed Interest packet. By providing an encrypted private key using the consumer's public key and issuing a mark, the data producer can verify the identity of the user [14].

5.5 Applications

Identity-based encryption (IBS) cryptography services are obtained by IoT-enabled healthcare (IoTEH) under security hardness. The Hyperelliptic Curve Cryptosystem (HCC) provides the corresponding level of security, which uses the least computing and communication power [15]. Internet-scale caching architecture for IoT data considers the longevity and popularity of the information. In the NDN-based future Internet, IoT contents are requested by distant users and provided by traversing various edge and core network segments [7]. To sufficiently and comprehensively solve issues with safe onboarding and routing for intelligent cities that demand a highly scalable framework for lightweight authentication and hierarchical routing, NDN-based IoT frameworks have been developed [16].

5.6 Technology

Transactions and block synchronization are the two main functions of the blockchain. A blockchain-enabled differentiated routing method that uses NDN deploys at the network layer to create a mapping relationship between windows and packets to reduce average packet transmission delays and speed up packet processing [17]. Mobile nodes that perform data retrieval without FIBs and numerous consumers fetch data by one data acquisition process to lower data acquisition costs and increase the efficiency of forwarding and data acquisition through NDN-based MANET [18].

6 Open Research Challenges

An enhanced SDTMA-NDN architecture has been proposed to transfer scientific data software via NDN. This architecture relies on data retrieval methods that help reduce overhead by avoiding Squid proxies. It also uses the in-network caching capabilities of routers by implementing two forwarding strategies: NDN-MP (multipath) and NDN-BP (Best path). The results indicate that this approach performs better than HTTP [19]. Responsibility-based transport control has been presented to address the issues with transport control in Named Data Networking. This method combines a capacity-aware and multipath forwarding algorithm in each face with a window-based rate control mechanism. It hopes to accomplish this to improve the effectiveness of the transport control mechanisms in the NDN architecture [20]. It has been suggested that a Differentiated Routing Strategy for Blockchain in Named Data Networking be used. Mapping packets to windows using a window sliding technique can reduce average delays and speed up packet processing. Moreover, packet distribution considers the sensitivity of delays and the probability of timeouts [17]. An anonymous Federated Learning system built on Named-Data Networking has been created to solve privacy concerns. This system includes FL workers, a Central Aggregator server, and a Trusted Third Party for worker enrollment. It ensures worker anonymity, guards against evil activity, and works with various data privacy techniques [21]. In Named Data Networking, a unique probabilistic Internet-scale caching approach has been presented for IoT data. This approach employs two reactive caching algorithms in the edge and core domains, considering the popularity of information and its lifespan. The outcomes show less stress on network infrastructure and energy-constrained devices, supporting the Internet of Things sustainability [16]. ENIPC attempts to address the problem of often occurring data delivery failures in IoT settings when provider and user mobility exist. NDN in IoT contexts is made possible for edge devices by ENIPC's integration of edge computing with IoT. This method supports supplier and customer mobility by enabling numerous consumers to retrieve different data types via reverse pathways within a single data delivery process [22].

Dedicated routers are assigned to packet routing chores in implementing dedicated routing and caching, while cache servers placed at each node take care of in-network caching duties. This approach improves the effectiveness of in-network caching and lowers access delays for time-sensitive traffic by saving many resources during caching operations [23]. NDN-Fab, built on the Fabric architecture, allows the core network to use an efficient path-based communication model. In contrast, the edge network can adopt content-centric designs similar to the Named Data Network. The NDN-Fab protocol is used in the core network to provide link-layer packet steering and multicast packet transfer. Core network switches' programmable data plane hardware may incorporate this protocol [24]. NDNrel reduces the cost of providing clients with relational information by providing publishers with a means of distributing content as related objects through metadata and manifests inside ICN designs. This approach also addresses the authentication of the connections and data that comprise the content [25].

Through regulating interest packet reception, the PITCM approach reduces PIT overflow and enhances PIT utilization and performance. It is composed of an Adaptive Virtual PIT approach that anticipates and responds to PIT overflow early, a Smart Threshold Interest Lifetime method that modifies the lifetime values of incoming Interest packets,

and the Highest Lifetime Least Request policy for managing PIT entries [26]. Over time, non-functional features are validated using an abstract certification approach, guaranteeing network quality and dependability. To increase confidence in information-centric networks, this model has evolved into a network-level certification strategy that provides an enhanced model, new services, and deployment strategies [14]. Blockchain technology is used by the Named Data Network (NDN) to store and distribute documents. This technology enables data signature encryption, flexible routing, and reverse path forwarding. A smart contract ensures the security of confidential files during transmission [27]. Blockchain data propagation methods show how much more efficient it is to distribute blockchain data within an NDN than an IP network by utilizing Named Data Network functionalities for transaction and block delivery. As a result, block delivery time is reduced, and security for changing blockchain parameters is enhanced [28]. NHealth-IoT uses pure-named Data Network Machine communication to send raw sensor data to the home server. The home server uses the Workflows Intuitive Formal Approach to verify NHealthIoT integrity during emergencies and the Hidden Markov Model to detect emergency healthcare events [8]. InterestFence achieves precise detection and effectively filtering attack traffic without impacting legitimate Interests. Each name prefix in every content server node is associated with a hash-based security label to prove content existence [29].

Since NDN is a relatively new field of study, there are many unsolved research questions that both present and future scholars might tackle. The main challenge is also to map proposed routing strategies at the Internet level. The Internet uses BGP as its inter-domain routing protocol because it is based on regulations instead than shortest paths. Since policy agreements in NDN are dynamic, establishing inter-domain routing is challenging. Another open issue is looking into more efficient ways to make advantage of the multipath capabilities. The biggest challenge is confirming information while keeping the supplier's identity a secret. Furthermore, since the content is exposed to all intermediate nodes between the user node and the content provider, an attacker could follow user information through content requests. As of right now, there is no permanent solution available. To sum up, there is still a lot to learn about privacy, trust management, and NDN security.

7 Comparison of NDN and IPv6

The current internet architecture uses IP addresses to identify end hosts, while NDN packets retrieve data using Content Names. While NDN routers utilize name prefixes for Interest packet forwarding, IP-based routers employ IP prefixes. Although NDN employs request packet identities and nonce entry data to reduce packet looping, the existing internet protocol depends on routing protocols to offer loop-free pathways. When a packet's lifespan expires, NDN routers maintain track of every important packet that passes through, but IP forwarding is stateless because packets are not stored in the routers. The comparison highlights the main differences between Named Data Networks and IPv6 in Table 2.

Table 2. Comparison of NDN and IPV6

Features	IPv6	NDN
Addressing	IPv6 occupies a comparatively greater address space than IPv4, allowing for almost endless unique addresses	Identifying users with content names instead of standard IP addresses is an innovative solution NDN technology provides. Because each content item has a distinct hierarchical name, routing based on content names is efficient
Routing	Routing decisions are made by referencing destination addresses in routing tables to direct packets to their destinations effectively	Based on content names, this routing approach gives content availability precedence over destination addresses
Caching	Since IPv6 does not have built-in caching methods, caching is usually implemented at higher network stack layers or through intermediary devices such as proxies	Additionally, NDN network routers cache data packets so neighboring caches can offer the same material to interest packets later. This method efficiently conserves bandwidth while lowering latency
Data Delivery	It makes transmitting packets between sources and destinations easier by using IP addresses and communication protocols like TCP and UDP	Routers are essential for NDN architecture because they direct interest packets toward particular content names. Routers then retrieve and send data packets to the clients that have made the request
Security	At the IP layer, IPsec provides authentication and encryption features	Furthermore, NDN integrates security protocols into its architecture, guaranteeing the integrity and consistency of material using producer signatures
Adoption and Implementation	The next-generation Internet Protocol, or IPv6, has already been widely implemented. However, the process of complete adoption is still ongoing	Although NDN has significant advantages, it is still in the research and development stage. Even though NDN is gaining popularity and is being experimented with more, extensive adoption would need considerable changes to the current networking architecture and protocols

8 Discussion

In addition to the concerns above, the security and confidentiality of such connectivity are of utmost importance. The critical establishment poses a significant obstacle to Named Data Networking. Despite efforts by NDN developers to tackle this issue, the exchange of public key certificates warrants further scrutiny. Moreover, prevailing authentication methods like single sign-on could partially remedy the authentication challenge within a federated setting. At the same time, a collaborative resource allocation mechanism may address resource utilization issues. Then, routing algorithms were developed to prioritize content names, enabling efficient content discovery and retrieval. It is clear that caching techniques, such as content substitution and positioning, improve the material's popularity and performance. Content names are prioritized by the routing protocol, allowing for more effective content retrieval and exploration. Then, efficient caching strategies were developed to store frequently requested content closer to users, reducing latency and bandwidth consumption. The forwarding approaches then adjust route choices dynamically in response to various network factors, such as connection integrity and congestion levels. Finally, the enhanced security protocols protect against unwanted access and security threats while maintaining user privacy.

9 Conclusion

Named Data Networking represents a paradigm shift in networking by promoting a data-centric approach that prioritizes security and content delivery. Because of its many uses and innovative possibilities, NDN can revolutionize data transmission and accessibility in the digital age. As the future frontier of network communication and information exchange, Named Data Networking rises to the occasion by overcoming obstacles and establishing new research and development paths. This article provides a brief synopsis of the services, functionality, and structure of Named Data Networks, as well as current applications and technological developments. It represents the most current and extensively researched domain within NDN technology. The Internet of Things, the main focus and most extensively studied NDN technology, brings to light the most recent. Researchers are also very interested in other disciplines like Artificial Intelligence, Blockchain, and Cloud-Edge. Ultimately, the subfields of application within the sectors are the smart city, smart healthcare, and intelligent automotive categories.

References

1. Amuda, K.A., Amuda, K., Almustapha, W., Deepak, B., Hoggard, C., Tiruveedula, P.: Revolutionizing networking paradigms: a comprehensive exploration of information-centric networking (ICN), content-centric networking (CCNx) and named data networking (NDN) (2024)
2. Azamuddin, W.M.H., Aman, A.H.M., Sallehuddin, H., Abualsaud, K., Mansor, N.: The emerging of named data networking: architecture, application, and technology. IEEE Access **11**, 23620–23633 (2023)
3. Priya, A.S., Suresh Joseph, K., Babukarthik, R.G.: In-depth Analysis of Named Data Network under Content-Centric Networking Platform (2018)

4. Saxena, D., Raychoudhury, V., Suri, N., Becker, C., Cao, J.: Named data networking: a survey. Comput. Sci. Rev. **19**, 15–55. Elsevier Ireland Ltd (2016)
5. Asaf, K., Rehman, R.A., Kim, B.S.: Blockchain technology in named data networks: a detailed survey. J. Netw. Comput. Appl. **171** (2020)
6. Hail, M.A., Amadeo, M., Molinaro, A., Fischer, S.: Caching in named data networking for the wireless internet of things. In: IEEE Computer Society, Institute of Electrical and Electronics Engineers, and Intelligent Sensors, 2015 International Conference on Recent Advances in Internet of Things (RIoT)(2015)
7. Amadeo, M., Campolo, C., Ruggeri, G., Molinaro, A.: Beyond edge caching: freshness and popularity aware IoT data caching via NDN at internet-scale. IEEE Trans. Green Commun. Netw. **6**(1), 352–364 (2022). https://doi.org/10.1109/TGCN.2021.3124452
8. Saxena, D., Raychoudhury, V.: Design and verification of an NDN-based safety-critical application: a case study with smart healthcare. IEEE Trans. Syst. Man Cybern. Syst. **49**(5), 991–1005 (2019)
9. Wang, L., Lehman, V., Mahmudul Hoque, A.K.M., Zhang, B., Yu, Y., Zhang, L.: A secure link state routing protocol for NDN. IEEE Access **6**, 10470–10482 (2018)
10. Liu, Y., Wadekar, H.: SDAR: software defined intra-domain routing in named data networks. IEEE In: 15th International Symposium on Network Computing and Applications (2016)
11. Li, Z., Liu, J., Yan, L., Zhang, B., Luo, P., Liu, K.: Smart name lookup for NDN forwarding plane via neural networks. IEEE/ACM Trans. Network. **30**(2), 529–541 (2022)
12. Qiao, X., Ren, P., Chen, J., Tan, W., Blake, M.B., Xu, W.: Session persistence for dynamic web applications in named data networking. J. Netw. Comput. Appl. **125**, 220–235 (2019)
13. Din, M.S.U., Rehman, M.A.U., Kim, B.S.: CIDF-WSN: a collaborative interest and data forwarding strategy for named data wireless sensor networks. Sensors **21**(15), 2021 (2021). https://doi.org/10.3390/s21155174
14. Anisetti, M., Ardagna, C.A., Berto, F., Damiani, E.: A security certification scheme for information-centric networks. IEEE Trans. Netw. Serv. Manag. **19**(3), 2397–2408 (2022)
15. Ullah, S.S., Hussain, S., Gumaei, A., AlSalman, H.: A secure NDN framework for Internet of Things enabled healthcare. Comput. Mater. Continua **67**(1), 223–240 (2021)
16. Mick, T., Tourani, R., Misra, S.: LASeR: lightweight authentication and secured routing for NDN IoT in smart cities. IEEE Internet Things J. **5**(2), 755–764 (2018)
17. Shang, J., Huo, R., Wang, S., Huang, T.: An NDN-enabled differentiated routing strategy for blockchain. In: 2022 IEEE 2nd International Conference on Computer Communication and Artificial Intelligence, CCAI 2022, pp. 96–101 (2022)
18. Wang, X., Lu, Y.: Efficient forwarding and data acquisition in NDN-based MANET. IEEE Trans. Mob. Comput. **21**(2), 530–539 (2022)
19. Alhowaidi, M., Ramamurthy, B., Bockelman, B., Swanson, D.: Enhancing the SDTMA-NDN architecture for transferring the scientific data software using named data networking. Comput. Netw. **166** (2020)
20. Nikzad, M., Jamshidi, K., Bohlooli, A.: A responsibility-based transport control for named data networking. Futur. Gener. Comput. Syst. **106**, 518–533 (2020)
21. Agiollo, A., Bardhi, E., Conti, M., Dal Fabbro, N., Lazzeretti, R.: Anonymous federated learning via named-data networking. Futur. Gener. Comput. Syst. **152**, 288–303 (2024)
22. Wang, X., Cai, S.: Edge-assisted NDN-based IoT framework with provider and consumer mobility support. IEEE Trans. Netw. Sci. Eng. **9**(3), 1713–1725 (2022)
23. Kim, D., Kim, Y.: Enhancing NDN feasibility via dedicated routing and caching. Comput. Netw. **126**, 218–228 (2017)
24. Madureira, A.L.R., Araujo, F.R.C., Araújo, G.B., Sampaio, L.N.: NDN fabric: where the software-defined networking meets the content-centric model. IEEE Trans. Netw. Serv. Manag. **18**(1), 374–387 (2021)

25. Antunes, R.S., Lehmann, M.B., Mansilha, R.B., Gaspary, L.P., Barcellos, M.P.: NDNrel: a mechanism based on relations among objects to improve the performance of NDN. J. Netw. Comput. Appl. **87**, 73–86 (2017)
26. Alubady, R., Hassan, S., Habbal, A.: Pending interest table control management in Named Data Network. J. Netw. Comput. Appl. **111**, 99–116 (2018)
27. Kang, P., Wenzhong, Y., Ding, T.: Blockchain document forwarding and proof method based on NDN network. IEEE Access **10**, 75312–75322 (2022)
28. Thai, Q.T., Ko, N., Byun, S.H., Kim, S.M.: Design and implementation of NDN-based Ethereum blockchain. J. Netw. Comput. Appl. **200** (2022)
29. Dong, J., Wang, K., Quan, W., Yin, H.: InterestFence: simple but efficient way to counter interest flooding attack. Comput. Secur. **88** (2020)

Optimizing Resource Utilization and Performance in Multi-cloud Infrastructures Through Predictive AI Analytics

Arun Pandiyan Perumal(✉)

Department of Information Technology and Management, Illinois Institute of Technology, Chicago, USA
apandiyan@hawk.iit.edu

Abstract. To achieve optimal resource consumption and the cloud environment's overall performance, organizations must manage multi-cloud deployments efficiently and overcome the related challenges. Therefore, this paper aims to understand how the application of predictive AI analytics can improve the management of resources and performance in environments that include multiple clouds. It discusses Cloud Management Platforms, Service Performance Management, cost optimization tools, and AI-based predictive analysis, among others. This paper will focus on the strengths and weaknesses of multi-cloud settings with an emphasis on the manner in which predictive analytics can be helpful in overcoming some of the principal issues arising from these settings. The case study also shows that newer tools such as VMware vRealize Suite, Datadog, and the AWS cost explorer, with embedded elements of artificial intelligence and machine learning, have revolutionized the scenario of predictive Cloud analytics.

Keywords: Multi-Cloud Infrastructure · Predictive AI Analytics · Cloud Management Platforms (CMPs) · Resource Utilization · CI/CD Integration · Artificial Intelligence · Machine Learning · Resource Optimization

1 Introduction

Organizations have designed and implemented the use of multi-cloud infrastructures on the basis of the uniqueness that different clouds can provide; controlling resource consumption and performance is a real headache. New challenges are associated with the flexibility, security, and cost-saving possibilities of multi-cloud implementations, but management and orchestration hurdles exist due to disparate platforms employed. The above challenges have been found to have been addressed by predictive AI analytics that enables firms to understand their resource requirement, performance, and costs. This paper aims to analyze how proactive AI analytics can improve the control of multi-cloud environments, with a spotlight on the current strategies and solutions enabling resource allocation and optimization.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
S. Rajagopal et al. (Eds.): ASCIS 2024, CCIS 2428, pp. 365–374, 2025.
https://doi.org/10.1007/978-3-031-86302-8_24

2 Overview of Multi-cloud Infrastructures

Hybrid environments, where businesses employ services of different cloud providers at the same time, are utilized more often. In this approach, there are several advantages, including flexibility, no risk of falling foul of lock-in, and the augmentation of disaster recovery. However, as the 'theory and practice of organizational work form complex environments and the nature of resource as dynamic and flexible,' thereby implying resource quantity is variable and requiring efficient use, resource optimization and high performance in such an environment remains an arduous task. These are some of the challenges to which predictive AI analytics is instrumental. Predictive AI uses algorithms and data analytics to determine demand patterns and give proactive guidance concerning the use of resources [1]. This makes it possible to keep track, control, and plan for the cloud resources, ask questions about future requirements, and balance performance in different clouds. To be able to anticipate and, in advance, avert problems concerning resources and performance is crucial for effectiveness.

For instance, machine learning can detect areas that might slow down processing before they reach the problem level, recommend appropriate resource distributions, and assist with load redistribution between different clouds. It also optimizes resource utilization and increases cost-effectiveness by eliminating the problems of over-commitment and under-utilization of resources. However, achieving resources and performance management in a multi-cloud setup presents the following vital hurdles. However, one of the significant challenges that people encounter is the interface between managing and tracking resources across various clouds [2]. Clouds differ from one another based on the tools, interfaces, and APIs that each cloud provider provides, and this makes it very challenging to interconnect and manage the resources together.

Another area of apprehension is the organization's use of resources for the implementation of its health management plans. This especially holds when the implementation is not followed up by regular monitoring and management; this may cause the organization to over-provision its resources, leading to wasteful expenditure or under-provision of the resources, thereby undermining the organization's performance. Some of these problems can be managed by using predictive AI analytics to give insights into typical utilization and predict future demands; nevertheless, the establishment and calibration of these analytics systems are quite resource-intensive.

2.1 Benefits of Multi-cloud Infrastructure

Multi-cloud ecosystems enable organizations to select the best services from a number of different cloud providers. Hence, it allows organizations to meet new customer requirements and demands, incorporate leading-edge technologies into the process, and adjust resources. There is flexibility since people can use other service providers or switch to another service that they find best due to performance, cost, or features, among others. Thus, by spreading the workloads between different cloud service providers, companies minimize the risks of unfavorable outcomes and increase stability [3]. When one or some of the providers fail or experience an outage/s or disruption/s, other providers can still be relied upon to support mission-critical functions, thereby offering higher availability and business resilience. Many organizations prefer multi-cloud strategies because

these enable them to take full advantage of different cloud providers depending on the workload. For example, a company could employ one cloud services provider to utilize the computer resources intensively and another provider to store them cheaply. Such a strategy also helps in achieving optimal resource utilization and hence results in cost reduction. The multi-cloud configuration provides robust cloud solutions in multi-cloud environments that are effective in handling data breach solutions.

2.2 Challenges in Multi-cloud Infrastructure

Resource management across multi-cloud environments becomes a complex process. The tools provided by the various carriers are different in functionality, interface, and API, which proves to be problematic in integration and may take a considerable amount of time to standardize [4]. At the same time, organizational fragmentation leads to the difficulty of having a centralized perspective on and management of resources and outcomes. Within the counterpart cloud providers, the capacity of the infrastructure and state of the network may differ from one day to another, resulting in differences in performance. Applications are mostly affected by latency; thus, organizations can be affected by performance bottlenecks. While using several cloud platforms, it becomes quite cumbersome to secure the data and ensure compliance. The problem arises from the fact that each carrier might have different levels of protection, as well as different compliance procedures; therefore, it is challenging to implement coherent security measures and guarantee compliance with the law. On the one hand, the use of multi-cloud systems implies the possibility of managing costs; on the other, it also results in complications when it comes to cost monitoring and controlling. The following are the technical challenges that must be overcome in order to provide centralized control across several clouds in multi-cloud infrastructures [22–24].

1. *Lack of Unified Visibility and Monitoring.* Each cloud provider has its own tools and dashboards, making it difficult to get a consolidated view of an entire infrastructure. Such visibility makes performance monitoring, troubleshooting, and capacity planning quite challenging. Implement a multi-cloud management platform that can provide one single pane of glass view across all of your cloud environments. These platforms aggregate information from various cloud providers; thus, they enable centralized monitoring, reporting, and analysis. Examples include CloudHealth by VMware, Datadog, New Relic, etc.
2. *Security and Compliance.* Maintaining consistent security policies and ensuring compliance with regulations across multiple clouds can be complex. Each cloud provider has its own security features and configurations. A centralized security information and event management system can be set up, along with a unified security policy framework applicable to all the clouds. Conduct regular audits to ensure compliance with relevant regulations.
3. *Data Management and Interoperability.* Transferring and sharing data among the different cloud environments are complicated by variations in storage technologies and other factors like data formats and APIs, which impact application portability and data analytics. Establish the implementation of the data fabric architecture for a total view of your data across different cloud environments. This will allow for seamless

access and data sharing with its governance. Use standardized data formats and APIs to promote interoperability between applications and services running on different clouds.

4. *Network Performance and Latency.* Managing network performance and minimizing latency across multiple cloud environments can be challenging, especially for applications that require real-time data transfer. Utilize AWS Transit Gateway or Azure Virtual WAN networking services of cloud providers for advanced connectivity between clouds. Follow the implementation of SD-WAN solutions for improved network traffic routing and performance across a multi-cloud environment. Apply edge computing strategies on latency-sensitive applications.

3 Purpose of the Study

This secondary research aims to evaluate how predictive AI analytics affect resource consumption and optimization in a multi-cloud environment based on the available literature and real-life examples and tools. Therefore, this analysis seeks to explain how predictive analytics have been used to overcome the issues and challenges that are related to the management of multiple clouds. In its attempt to offer a systematic structure to observations emanating from this work, the study aims to elaborate on how AI-driven solutions improve resource management, performance measurement, and cost control in different cloud environments.

4 Role of Predictive AI Analytics in Resource Utilization and Performance in Multi-cloud Infrastructure

Another significant function of predictive AI analytics is the identification of upcoming resource requirements. The typical approaches include the time series forecast that entails models of future usage and a regression analysis of the historical usage data. This capability helps organizations to predetermine periods that may be the most favorable in terms of their demand and to adapt the distribution of resources needed to these conditions. For instance, the Amazon Forecast and Google Cloud AI Platform enable the determination of future demand through some algorithms that help shoppers avoid over-provision or under-provision [15] (Fig. 1).

Here is a detailed analysis comparing Amazon Forecast and Google Cloud AI tools.

1. *Amazon Forecast.* Amazon Forecast is a fully managed service that uses machine learning to provide super-accurate forecasts based on historical time series data. It was designed for a variety of use cases, including demand forecasting, inventory planning, workforce management, and many others [17]. The advantages and challenges presented by the tool are summarized below.

- Accuracy and Customization: Amazon Forecast uses sophisticated deep learning algorithms, including Long Short-Term Memory networks, which learn intricate patterns of time series for highly accurate forecasting. It also integrates domain knowledge data, such as holiday schedules and weather patterns, into the predictive models.

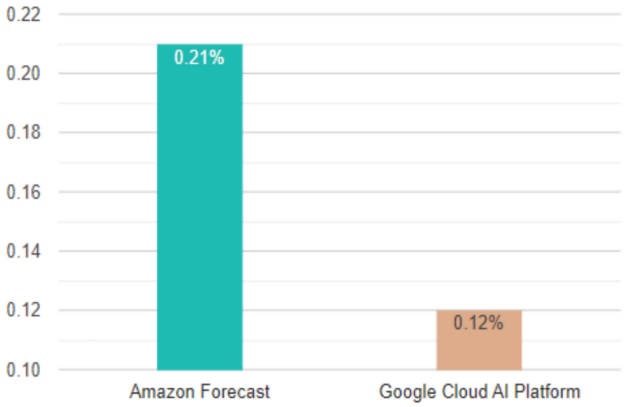


Fig. 1. Comparison of the Amazon Forecast vs. Google Cloud AI Platform (Source: https://6sense.com/tech/data-science-and-machine-learning/amazonforecast-vs-googlecloudaiplatform)

- Ease of Use and Integration: Because it's a fully managed service, users don't need deep prior knowledge in Machine Learning to construct effective models. This interface has made it easier to input different kinds of historical data and relevant metadata while providing the opportunity to have model selection automatically optimized. Seamlessly integrates with other AWS services offering capabilities such as S3 for data storage, Lambda for serverless execution, and SageMaker for custom machine learning.
- Complexity in Customization: While AutoML simplifies model building, it can limit flexibility for advanced users who want more control over model architecture and hyperparameters.

2. *Google Cloud AI.* Google Cloud AI offers a suite of AI and ML tools, including AutoML and Vertex AI, which can be used for predictive analytics in operational environments [18]. Overall, both are very powerful platforms with enormous functionalities that extend to time series forecasting, anomaly detection, etc. The advantages and challenges presented by the tool are summarized below.

- Customizable ML Pipelines: Google Cloud AI's Vertex AI provides flexibility to custom-built pipelines that enable complex predictive models to be built and deployed at scale, meeting exact organizational needs. This is especially beneficial for organizations that deal with complex or proprietary data and need extensive pre-processing or customizations.
- AutoML for Time-Series Forecasting: Google's AutoML makes it easy to develop predictive models for time-series data using advanced neural networks and ML algorithms. It is highly automated, taking care of feature engineering, model selection, and tuning.
- Steeper Learning Curve: The breadth of services and tools can make it more complex for beginners to get started, potentially requiring more ML expertise.

These tools can be effectively utilized in the following operational scenarios:

1. *Predictive Scaling.* Both platforms can analyze historical resource usage data (CPU, memory, network traffic) to predict future demand. Amazon Forecast provides pre-built forecasting models specifically for scaling EC2 instances and other AWS resources. It can also be used to predict demand for resources in other cloud environments based on historical data. It provides some tools, such as Vertex AI and AutoML, which will help create custom machine-learning models to predict resource requirements across various cloud providers. Integration with the infrastructure management tools of Google Cloud will enable automated scaling [19].
2. *Cost Forecasting and Right-Sizing.* Amazon Forecast works with AWS Cost Explorer to enable customers to forecast future spending based on current usage trends. Along with right-sizing techniques, including those supported by tools such as AWS Compute Optimizer and Trusted Advisor, Amazon Forecast finds idle or underutilized resources across multiple clouds and recommends the optimal resource configuration [20].
3. *AI-Driven Performance Optimization.* The capabilities of Google Cloud AI can be used to monitor real-time system metrics, predict future workloads, and dynamically adjust resources. For instance, AI-driven insights allow GCP to analyze the usage patterns of compute resources and automatically suggest the most optimal scaling policies for workloads hosted in hybrid or multi-cloud environments [18].

SPM automation provides an enhanced method to utilize various cloud platforms through resource forecasting using AI. With AI's support, optimization tasks can be performed on the fly, detecting an optimal allocation of resources at a given moment in time. For instance, Microsoft Azure VMSS and IBM Cloud Paks for multi-cloud Management utilize AI to manage resource allocation and scaling by properly provisioning and placing the workload across the different clouds. AI analytics improve performance tracking since there is constant computation of different figures and tracking of variations [5]. The use of such algorithms and approaches as clustering and anomaly detection helps to identify problem areas in the system and notify the managers about them before they become critical. Some of the widely used tools include Datadog and Splunk, both of which leverage AI to provide real-time performance metrics and smart alerts for potential problems, therefore ensuring that organizations run at optimum and that any adventitious problems are swiftly handled. As for cost control, predictive analytics is useful to identify the pattern of the costs and propose measures to reduce them. Cost-saving solutions facilitated through AI encompass identification of the overall expenses, prediction of the costs in the future, and advice on how they could be reduced. Software, including CloudHealth, VMware, and AWS Cost Explorer, employs AI to evaluate the bills on cloud utilization and costs, suggest changes, and distribute the expenses properly among several cloud solutions [8]. Security solutions based on artificial intelligence like Google Cloud Security Command Center, Azure, and AWS GuardDuty can evaluate patterns in data usage and accessing data to begin diagnosing safety threats. These tools improve the firm's security outlook since they allow for a review of the risks and compliance hindrances.

Category	Tool/Platform
Resource Forecasting	Amazon Forecast, Google Cloud AI Platform, SPM Automation
Resource Optimization & Scaling	Microsoft Azure VMSS, IBM Cloud Pak for Multi-cloud Management
Performance Tracking	Datadog, Splunk
Cost Control	CloudHealth by VMware, AWS Cost Explorer
Security & Compliance	Google Cloud Security Command Center, AWS GuardDuty

Table for the predictive AI analytics in the resource utilization and performance in multi-cloud infrastructure.

5 Current Approaches and Tools for Resource Utilization and Performance in Multi-Cloud Infrastructure

An essential element to consider relates to the management of the multi-cloud environment, specifically through Cloud Management Platforms. These platforms help provide a single environment to manage the resources with different parties and companies, which allows companies to control and improve their overall workflow. For instance, the VMware vRealize Suite offers extensive Cloud management features, such as resource usage, performance, and cost. It has various interface compatibility with different cloud settings, and the commercialization of automated procedures dramatically improves its functionality. In the same way, CloudBolt provides strong multi-cloud management, efficient cost control, and regulations [7] (Fig. 2).

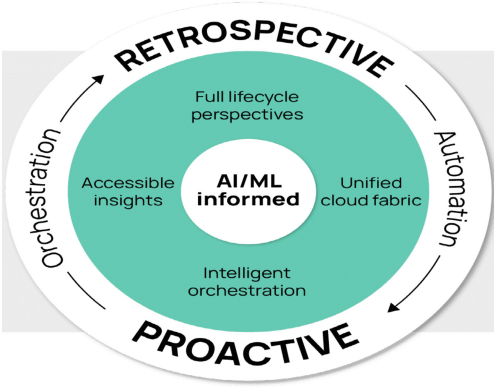


Fig. 2. Key advancements over traditional FinOps solutions

It enables an organization to allocate resources across multiple providers and effectively control resource consumption in order to save money while at the same time

enforcing rules and regulations in terms of resource usage. To manage performance and detect problems in multi-cloud environments, it is necessary to use performance analytics. Products like Datadog are able to provide one holistic view of all the performance data across diverse cloud solutions. Its AI aspects enable it to quickly identify anomalies, as well as find areas of performance inefficiency and suggest operations to improve it. Dynatrace goes a step further in this through its strong application of performance management and monitoring tools [14]. It is an AI-based system that provides comprehensive insight into the app performance, IT infrastructure, and users, giving control over several clouds. Cost optimization is another critical management technique that is specific to multi-cloud resources, and several tools exist for it. AWS Cost Explorer delivers real-time information for AWS Cloud costs and is used to make historical descriptions of costs, predictions, and possible savings. The extensiveness of AWS services that it integrates with provides a good view of the overall spending on the cloud. This extends into CloudHealth by VMware, which consists of multi-cloud cost management and optimization. It gives insights into the expenses incurred in many cloud services, enabling organizations to monitor costs, improve resource utilization, and enforce the right strategy. Resource optimization and automation are used to deploy AI and machine learning to automate resource management for maximum effectiveness.

Microsoft Azure Advisor helps users manage Azure resources, indicating which alterations will increase the efficiency of their usage. What is more, it provides recommendations with regard to costs and efficiency. Google Cloud Operations Suite and Google Cloud Stackdriver can be used to learn about the utilization and working of resources. It allows auto-scaling with Google Cloud as well as other services and helps improve resource management throughout the platform. The problem of controlling threats and adhering to security and compliance policies in the context of multi-cloud environments implies the use of dedicated solutions for the desired purpose. AWS GuardDuty is a security assessment service that will constantly look for signs of threats and suspicious activities on the accounts that are functioning on AWS [13]. It applies artificial intelligence features, including machine learning and threat intelligence, to single out security threats and notify the administrators. Google Cloud Security Command Center gives an overall view of security situations and compliance that can aid an organization in viewing their Google Cloud services, flagging threats, handling risks, and meeting compliance standards.

From the organizations that have adopted DevOps, there is a requirement to have tools that support CI/CD integration [11]. Jenkins is an open-source automation server that helps with CI/CD procedures with those plugins that are compatible with a number of cloud providers to create applications, scale, and manage resources automatically. GitLab provides users with DevOps services such as CI/CD, observability, collaboration, etc. [12]. It helps organizations handle the lifecycle of the application and utilization of the resources within the multi-cloud environment. AI and machine learning solve the problems of predictive analytics in the multi-cloud environment, where systems analyze historical data and usage to offer predictions and effective resource utilization. IBM Watson has a predictive functionality that allows it to analyze utilization patterns, bring efficiency in cost management, and provide machine learning on expected future requirements [9]. That is why it links with different clouds to offer prescriptive analytics

for the management of resources. Amazon SageMaker is another bulk solution for an MLaaS that can be utilized to create, train, and implement predictive models of resource consumption and performance in a multi-cloud environment.

Category	Tools or Platform
Cloud Management Platforms (CMPs)	VMware vRealize Suite, CloudBolt
Performance Monitoring & Analytics	Datadog, Dynatrace
Cost Management & Optimization	AWS Cost Explorer, CloudHealth by VMware
Resource Optimization & Automation	Microsoft Azure Advisor, Google Cloud Operations Suite
Security & Compliance	AWS GuardDuty, Google Cloud Security Command Center
CI/CD & DevOps Tools	Jenkins, GitLab
AI & Machine Learning	IBM Watson, Amazon SageMaker

Table for the usage of tools for resource utilization and performance in multi-cloud infrastructure.

6 Conclusion

Thus, it can be concluded that AI analytics based on predictive models are central to the management of resources and the operation of a multi-cloud environment. Based on the examples of VMware vRealize Suite, Datadog, and AWS Cost Explorer, it is possible to gain a better understanding of the demand for specific resources, performance problems, and costs. The roles of artificial intelligence and machine learning functions are present to help manage resources in advance, predict problematic incidents, and reduce expenditure. In this regard, some disadvantages are also apparent. However, they can be addressed by pointing out difficulties like the complexity of the concept and the issues related to security frameworks and the relevant configurations.

References

1. Akter, S., Michael, K., Uddin, M.R., McCarthy, G., Rahman, M.: Transforming business using digital innovations: the application of AI, blockchain, cloud and data analytics. Ann. Oper. Res. 1–33 (2022)
2. Chintapalli, V.R., Kondepu, K., Sgambelluri, A., Tamma, B.R., Castoldi, P., Valcarenghi, L.: Orchestrating edge-and cloud-based predictive analytics services. In: 2020 European Conference on Networks and Communications (EuCNC), pp. 214–218. IEEE (2020)
3. Gundu, S.R., Panem, C.A., Thimmapuram, A.: Hybrid IT and multi-cloud an emerging trend and improved performance in cloud computing. SN Comput. Sci. **1**(5), 256 (2020)
4. Achar, S.: Enterprise saas workloads on new-generation infrastructure-as-code (IAC) on multi-cloud platforms. Glob. Disclosure Econ. Bus. **10**(2), 55–74 (2021)
5. Chacko, A.: How to use AI analytics for targeted business decisions (2024). https://sproutsocial.com/insights/ai-analytics/

6. Sysdig: AWS vs. Azure vs. Google Cloud: Security comparison (2023). https://sysdig.com/learn-cloud-native/threat-detection-in-the-cloud-defender-vs-guardduty-vs-security-command-center/
7. CloudBolt: Key advancements over traditional FinOps solutions (2024). https://www.cloudbolt.io/solutions/augmented-finops-capability/
8. Slingerland, C.: CloudHealth vs. AWS Cost Explorer: The Definitive comparison (2024). https://www.cloudzero.com/blog/cloudhealth-vs-cost-explorer/
9. IBM: IBM Watson Studio (2024). https://www.ibm.com/products/watson-studio
10. Battina, D.S.: DevOps, a new approach to cloud development & testing. Int. J. Emerg. Technol. Innov. Res. 2349–5162 (2020). www.jetir.org, ISSN
11. Kushtov, M.: Serverless CI/CD pipeline based on Google Cloud Platform (2022)
12. Vemuri, N., Thaneeru, N., Tatikonda, V.M.: AI-optimized DevOps for streamlined cloud CI/CD. Int. J. Innov. Sci. Res. Technol. **9**(7), 10–5281 (2024)
13. Ajith, V., Cyriac, T., Chavda, C., Kiyani, A.T., Chennareddy, V., Ali, K.: Analyzing Docker Vulnerabilities through Static and Dynamic Methods and Enhancing IoT Security with AWS IoT Core, CloudWatch, and GuardDuty (2024)
14. Gajger, T.: NVIDIA GPU performance monitoring using an extension for Dynatrace OneAgent. Scalable Computing: Practice and Experience (2020)
15. 6sense: Amazon Forecast vs Google Cloud AI Platform: Data Science and Machine Learning Comparison (2024). https://6sense.com/tech/data-science-and-machine-learning/amazonforecast-vs-googlecloudaiplatform
16. Karamthulla, M., Malaiyappan, J., Tillu, R.: Optimizing resource allocation in cloud infrastructure through AI automation: a comparative study. J. Knowl. Learn. Sci. Technol. **2**(2), 315–326 (2023). https://doi.org/10.60087/jklst.vol2.n2.p326
17. Amazon: Time Series Forecasting | Machine Learning | Amazon Forecast, Amazon Web Services, Inc. (2024). https://aws.amazon.com/forecast
18. GCP: Google Cloud AI (2024). https://cloud.google.com/products/ai
19. Amazon: How scaling plans work - AWS Auto Scaling, docs.aws.amazon.com (2024). https://docs.aws.amazon.com/autoscaling/plans/userguide/how-it-works.html
20. Amazon: Identifying Opportunities to Right Size - Right Sizing: Provisioning Instances to Match Workloads, docs.aws.amazon.com (2024). https://docs.aws.amazon.com/whitepapers/latest/cost-optimization-right-sizing/identifying-opportunities-to-right-size.html
21. Team Blazeclan, AWS and Google AI: An Analytical Comparison - Blazeclan, Blazeclan (2022). https://blazeclan.com/blog/aws-and-google-ai-an-analytical-comparison
22. Karl, T.: Multi-cloud Challenges: Best Practices and Strategies, New Horizons (2024). https://www.newhorizons.com/resources/blog/multi-cloud-challenges
23. Morehouse, J.: The Challenges of Multi-Cloud Management and How Observability Helps Solve Them, Orange Matter (2022). https://orangematter.solarwinds.com/2022/10/06/multi-cloud-management-challenges
24. Piskorz, P.: Top 5 Challenges of Protecting Multi-Cloud Environment | Storware BLOG, Storware (2023). https://storware.eu/blog/top-5-challenges-of-protecting-multi-cloud-environment

VANET Security Through Blockchain Based Location Verification: Innovation and Challenges

S. Sandosh(✉), B. Thriam Bakesvar, J. Subash Saajan, and R. Eniyavasanthan

School of Computer Science, Vellore Institute of Technology, Kelambakkam to Vandalur Road, Chennai, Tamilnadu 600127, India
sandosh.s@vit.ac.in, {thriambakesvar.b2021,subashsaajan.j2021, eniyavasanthan.r2021}@vitstudent.ac.in

Abstract. Vehicular Adhoc Networks (VANET) is an emerging and significant technology which brings in new application fields such as traffic safety and security. Besides VANETs naturally have frequent challenges, like these nodes speeds are variable in nature and the infrastructure is decentralized, which means that those networks need very strong security systems to perform well. Such research study is focused on the analysis of the routing protocols, the use of the applications, the current challenges and security requirements in VANETs, which is terminated with the proposal of innovative solutions designed to ameliorate network performance and minimize the security risks. In particular, it involves location verification which is based on blockchains and making use of geographic routing and reliability tests to assert that the transmission data is secure. Moreover, a secure protocol will be suggested in order to decrease routing overhead and guarantee confidentiality of messages among vehicles that are specified in a group. Conclusively, this study takes the lead in making advancement in the VANET security as it gives the foundation for research to be conducted in this critical area.

Keywords: VANETs · Security · Location Verification · Privacy · Blockchain

1 Introduction

Advanced Vehicle-to-Vehicle communication networks, or Vehicular Ad-Hoc Networks (VANETs), constitute a dynamically changing environment in terms of wireless communication, and transportation systems that can realize ambitious applications to enhance traffic safety and efficiency. Nevertheless, VANETs dynamic nature including fast moving nodes and decentralized infrastructure make them super vulnerable security threats that should be addressed to get the extended benefits. In this scenario, this research paper intends to investigate and address the multidimensional security hurdles of VANET, including location verification and privacy issues, authentication, and malicious attack prevention methods. This introduction introduces the theme of VANETs and security, and presents the increasing value of VANETs and their role in securing their efficient

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
S. Rajagopal et al. (Eds.): ASCIS 2024, CCIS 2428, pp. 375–390, 2025.
https://doi.org/10.1007/978-3-031-86302-8_25

operation. It provides the research?s primary goals, such as studying the current security mechanisms and the problems attached, and the project of innovative solutions that can be used in order to improve the VANET security and performance. Establishing a holistic perception of the security environment within the context of VANETs and introducing novel techniques to handle arising dangers is the purpose of this research and hopes to make a substantial impact on the field leading towards a future where vehicular communication is safer and more secure.

1.1 VANET Design (Issues and Challenges)

VANETs are a dynamic network without the need for any overhead. Thus, designing the VANET is a tedious task with the set of the subsequent issues.

Error (Prone Channel State)
In a wireless network [3], link properties fluctuate, necessitating interaction between routing protocols. It finds an alternative if necessary.

Hidden Problems
Node A and C try to communicate with node B at the same time, but not with each other. A and C may not detect any interference on the wireless network. Thus, the signals clash at node B. Node B is unable to receive signals from either node; this is known as a Hidden issue. The answer to the Hidden terminal problem is for nodes to coordinate transmissions by requesting and giving permission to broadcast and receive data packets. This system is also known as Request to Send/Clear to Send (RTS/CTS). The primary idea is to grab the channel by notifying the nearby nodes of the incoming broadcast. The adjacent nodes can detect the transmission by imitating the receiving node?s output of a brief frame. The adjacent nodes are then required to avoid sending huge data frames.

In the above Fig. 1 describes the solution to solve the hidden problems during the communication between the nodes.

Exposed Terminals
Consider the preceding diagram as a topology with an extra node D, which can only be reached from node C. Suppose node B connects with node A and node C wishes to send a data packet to node D. During the transmission between nodes B and A, node C detects that the channel is congested. Node C incorrectly concludes that it may not send data packets to node D, despite the fact that both transfers are successful. Bad signal reception would only occur between nodes B and C, where neither receiver is situated. This situation is referred to as an exposed terminal problem. When there is a large traffic load, both the hidden and revealed terminal problems reduce network throughput significantly [4].

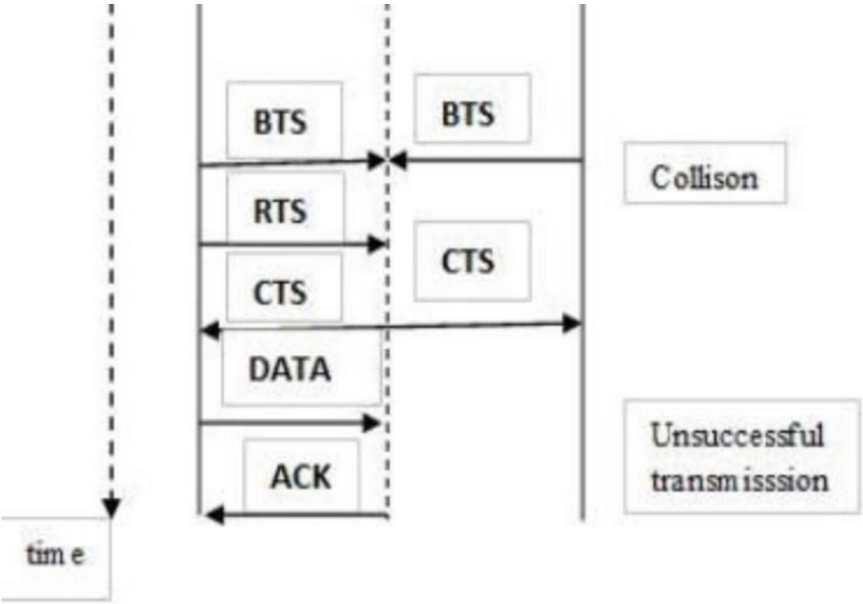


Fig. 1. Hidden problem using three nodes

Bandwidth-Constrained and Variable Capacity Links
Wireless connections have far lesser capacity than conventional networks. Multiple access, fading, noise and interference conditions, and other factors all have an impact on wireless communication performance. The major consequence of wireless networks is congestion, which regularly surpasses network capacity. Mobile ad hoc users will suffer since the mobile network is an infrastructure-free, wireless network.

Energy-Constrained Operation
All nodes in a Software Defined Network are powered by batteries or other finite resources. The system design criterion is the most essential factor in optimizing energy conservation. The node?s energy will be spent during the sending and receiving of data packets. Even in idle mode, the wireless interface consumes a significant amount of energy. Idle energy should be removed or decreased using energy-efficient measures. Experiments have been undertaken to determine power usage during various modes such as send, receive, idle, and sleep.

1.2 Objective of the Study

To investigate the vulnerabilities in VANET - to propose a mechanism to improve privacy and make the protocol more secure.

1.3 Motivation of the Project

The attackers collect and save location data posing new privacy issues, our research eliminates this threat by our proposed idea on VANET communication and location acquisition mechanisms.

2 Results

The study leads to the observation that the proposed protocol provides more security and reduces the routing overhead in the network.

3 Related Works

3.1 Sequence Numbering Techniques

Aiming at the area of blackhole attacks in E-ODMRP K, Aishwarya, Raju, N. K., and Selvan, A. S. drafted their solution [5]. They use a method whereby the basic node does not straight-forwardly accept the first given RREP, but instead puts all of them into a table for some review. Moreover, RREP which has an abnormal high sequence destination number reminds us that the sender is out of order, so we should discard it. The message is delivered while the corrupted node has been neutralized and its subsequent RREPs are ignored anymore. Thus, messaging from the malicious node does not forward and keep the routing table of it. During the building stage when power generation is available at many points, the accumulated RREPs will come to an end once a valid one is received.

Mistry, Jinwala, and Zaveri proposed a solution for AODV network against blackhole attacks [6]. Differing with shortening the retransmission of intermediate and destination node operations, they concentrate on the source node behavior change using an extra function Pre ReceiveReply(Packet P). They make use of a new table Cmg RREP Tab, a timer MOS WAIT TIME, and a variable Mali node, which have been added by them to AODV protocol?s default data structures.

In their procedure, the source node waits for MOS WAIT TIME; then, stores the other RREPs in Cmg RREP Tab. It conducts this analysis on the saved RREPs whose destination sequence number pattern is suspecting the sender as malicious. Once a bad node is found, the solution chooses a reply having the highest destination sequence number from Cmg RREP Tab using Pre ReceiveReply() method. The identity of the malicious node as Mali node, which will make sure that any control message from it will be removed and its routing table is not maintained. Messages containing harmful instructions are not forwarded. The RREP chosen will be eliminated from Cmg RREP Tab in order to maintain freshness. The proposed protocol operation agrees with initial AODV after malicious nodes recognization, the fact is proven by AODV default call.

Shurman et al. [7] suggested two solutions to save VANETs from blackhole attacks. The procedure starts with a pair of routes from the node prescribed as a source to the one designated as a destination. To contact the destination, source node broadcasts a ping packet (RREQ). Subsequently, a predetermined number of RREPs shall be received. Although it sees the sign of a protected channel that it buffered transmits in the packet. The next approach leverages signature of the inconsistent sequence numbers stored in the additional tables in order to prevent rogue nodes from behaving nonsensically. Data backs up our suggestion as reduction of RREQ and RREP is achieved with both methods, with sequence numbers used in the second one as an additional feature to differentiate requesting nodes.

Tamilselvakumar et al. [8] bring to the AODV routing node timers for collecting the requests from the similar nodes. They apply a table concept to store sequence numbers and arrival times of packets, and they make a compliance with a route using a threshold. It is verified by the simulation results where packet delivery ratios (RPD) are higher including less delay and overhead, however, end- to-end (E2E) delay may be variable with the distance between the source node and the destination node.

Sun et al. [9] suggest a self-healing method for VANETs where the neighborhood collaboration detects black hole attacks in the network and the routing of normal traffic is restored by the recovery protocol. By using this technique, they gather neighborhood details and use them to create rules that will force routing entries to take another path rather than a malevolent node. The simulation results demonstrate higher delivery ratios of packet compared to AODV, which also contributes to the latency and routing overhead to an extent.

Mistry et al. [10] classification and upgrade the AODV routing protocol towards black hole attack risks by introducing rule changes to the source node behavior. They exchange a new variable and table as well as a timer with AODV and make use of received RREPs to identify excluding nodes. We obtain increasing levels of packet delivery ratio in the case of black hole attack and a slightly higher delay at the end of end scope.

3.2 Trust Based Techniques

Tsou and co-researchers, for instance, conceived the BDSR (Bait-DSR) algorithm in order to secure VANETs against black hole-based collaboration. In this scenario, the solution has been created by drawing from both up-front and intervening approaches, which have formed a product with features of DSR routing protocol. Through sending bait RREQ messages with fake destination nodes addresses, BBSR attracts malicious nodes which further forges RREPs thus reduces the overall packet delivery times for BBSR than existing DSR and WD methods.

For the purpose of secure routing, Hu and Perrig [12] offer the RoadRunner Protocol as well as another called the AODV-Fi. Arriadne concurs with the use of a relatively efficient authentication mechanism, shared secret hash keys or digital signatures, as it uses both e key and digital signature. The delay hash-key disclosure in Tessla is a mechanism that presents an exceptionally auspicious broadcast authentication system. They eliminate the use of trusted hardware, thus strengthening security without a need for secure processors, powerful enough processors.

Wang and his colleagues [13] design a hash function protocol which allows to give backing of any node and they keep the system secure against collusion attacks. They actually enable nodes to check if packet forwarding it net intact without overstraining computer resources, thus signaling off routing segment attacks.

Djenouri and others [14] provide a plan to use the VANET networks to monitor, find and eliminate black hole attacks. Bayesian method for detecting false nodes as target with their proposed solution considers false accusation vulnerabilities and peaks minimizing by channel conditions and node mobility.

Erasing torture history et al. proposed attack detection algorithm built on the lines of AODV routing protocol [15]. To amend the method, the team included a dynamic list named Data Routing Information (DRI) which was verified against further request (freq) and further reply (FREP). According to the tests, their algorithm outstands the original in terms of both throughput and packet loss.

Chang et al. [16] suggested a distributed cooperative black hole scent detection mechanism that consists of collecting local data, local identification, cooperation detection, and overall response time. These are the primary comparison between POR-++ and AODV in which the former has a significantly higher packet delivery ratio compared to the latter although it comes with a high control overhead.

React Scheme [17] is introduced by Kozma et al., tracking activity of every node with auditing, searching, and identification stages. As opposite, REACT node loses communication load and only in exchange gets bigger delays of identification and may show audit nodes to attackers.

A trust model for dynamic routing of messages in Software Defined Networks (SDNs) is proposed by Liu et al. [18], which aims to ensure security of routing. The model, however, applies the concept of node authentication and peer behavior to calculate trust values which, in turn, are modified to trace out compromised nodes and maintain security protocols.

3.3 Dynamic Learning Techniques

Kurosawa et al. [19] suggest anomaly detection based on the application of dynamic training technology, which includes the periodically updating the training data. They build an elaborate feature vector with many dimensions to show the network condition at each node through elements like the destination sequence number, message statistics, and so on for the aim of the detection attacks. This is done by referencing current states vs. normal and abnormal and then notifying of the errors made in VANET environments.

Raj and Swadas [20] propose DPRAODV (Detective, protective, and reactive AODV), as a solution that will further safeguard AODV against black hole attack. It adopts routine AODV with the RNS mechanism by validating the sequence numbers in received RREP messages as well as adapting the dynamic threshold to the challenging situation. Malicious nodes once located are denied having any participation in ALARM message at all, and their presence in network is notified to controllers with ALARM packets thus spoofing their RREPs and avoiding packet forwarding. Consequently, the implementation of such a middle phase introduces an average increased end-to-end delay and normalized routing overhead.

This concludes the detection and prevention of attacks in routing protocols such as the AODV routing protocol to improve network throughput by adopting several techniques such as sequence numbering, trust management scheme, dynamic learning technique, and so on, but it does not perform efficiently.

4 Motivation

The attackers collect and save location data posing new privacy issues, our research eliminates this threat by our proposed idea on VANET communication and location acquisition mechanisms.

5 Existing System

EACCK - Enhanced Adaptive Acknowledgment [21] technique is used to address three key issues:

Receiver Collision: Occurs when two or more transmitters attempt to effectively relay packets to a single receiver at the same time.

Limited Transmission Power: The node restricts its transmission power in order to conserve its own battery resources. As a result, the packet transfer can be overheard by other nodes but is insufficient to reach the node.

False Misconduct Report: Node A sends a misbehavior report despite the fact that node B delivered the message to node C.

Figure 2 shows that in ACK mode, node S initially transmits an ACK data packet Pad1 to the destination node D. If all intermediary nodes on the route between nodes S and D are cooperative and node D successfully receives Pad1, node D must return an ACK acknowledgment packet Pak1 over the same route but in reverse order. If node S receives Pak1 within a predetermined timeout, the packet transfer from node S to node D is complete. Otherwise, node S will enter S-ACK mode, sending out an S-ACK data packet to detect any misbehaving nodes along the path.

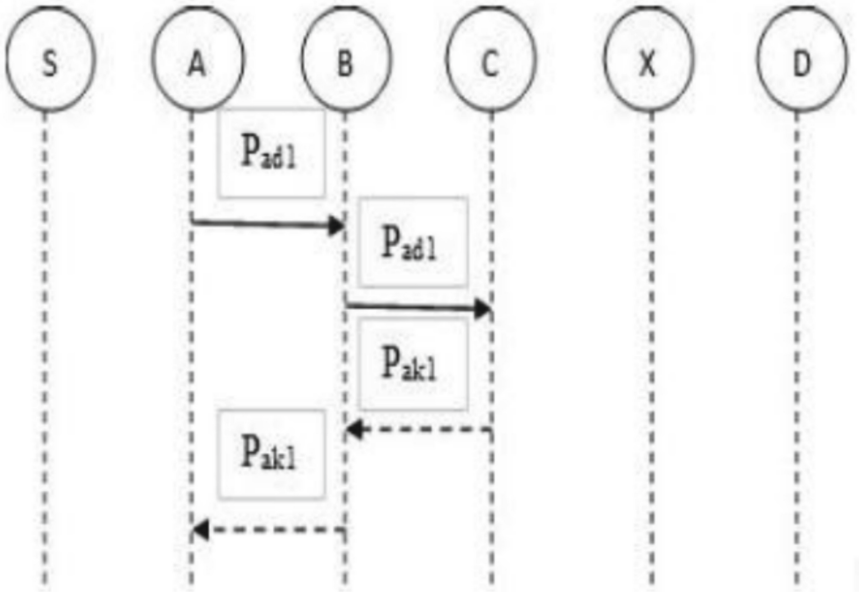


Fig. 2. ACK Scheme

6 Proposed System

6.1 Problem Formulation

Nodes? constant change in VANETs creates a favorable environment for several network attacks. Hence, we present the First Message Authentication Protocol (FMAP), which is used to circumvent unauthorized data usage or collection by attackers in Adhoc On-demand Distance Vector routing algorithm. The flexibility of the approach enables us to model networks where nodes with different energy levels to start with or with various degrees of malicious or self-benefit-oriented behavior are involved.

6.2 Proposed Solution

Hash Key based (symmetric and asymmetric) and certificate-based schemes have high computation and communication overheads because of its Hash Key management issues and certificate revocation lists (CRL). To reduce the overheads, identity-based authentication is used with digital signatures. To further reduce the overheads in the traditional scheme, the frontal encoding scheme separates the signature signing process into offline and online phases. The offline phase is executed by vehicles/RSUs to generate offline signatures whereas online phase is executed by vehicles to generate online signatures while establishing V2V communications.

A secure Frontal Encoding Authentication (FEA) scheme proposed is explained in detail in this chapter. It uses base64 encoding and decoding operations for verification purposes. The use of the FEA scheme helps increase the pairing operations? capability.

Both way communications are first authenticated by FEA mechanism under the proposed First Message Authentication Protocol (FMAP). To ensure privacy-preservation in vehicles, frontal signature generation is used. MAC Address based encoding system with Diffie-Hellman key exchange which generates a unique signature pair from both vehicles is used for authentication (Fig. 3).

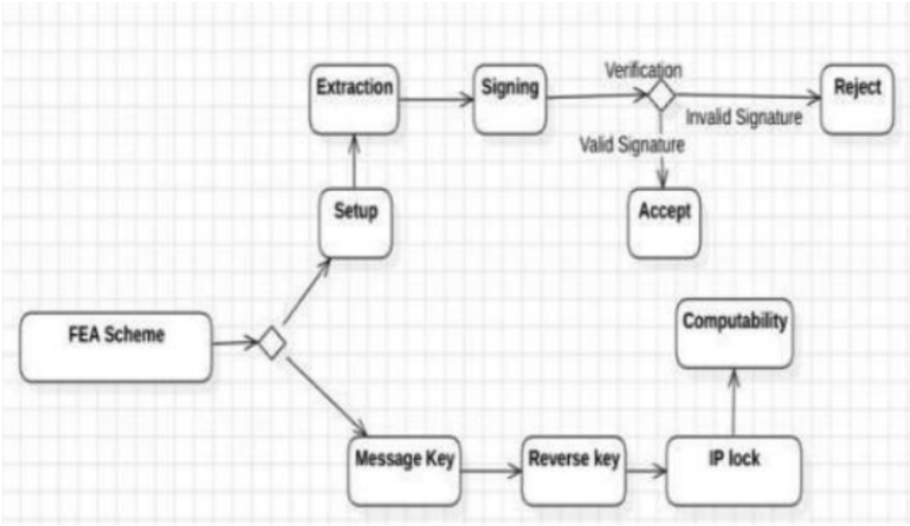


Fig. 3. Frontal Encoding Authentication Scheme

Setup

The systems installed in the vehicles should generate a signature, newSignature?=?encoded(IP_{base64}) + ?:?

Example MAC Addresses considered,
MAC Address 1: 00-B0-D0-63-C2-26
MAC Address 2: 00-B0-D0-63-C2-28

Extraction

The other vehicle should receive the code and its own code should be sent back, the received encoded string should be decoded by base64 decoder in offline mode.

Solution Proposed in Pseudocode Representation

```
    # Device 1 (MAC Address 1: 00-B0-D0-63-C2-26)
    Device1:
        # Private key of Device 1 (randomly generated)
    private_key_1 = 6

    # Compute public key A
    public_key_1 = (g ** private_key_1) % p

    # MAC Address of Device 1
    mac_1 = "00-B0-D0-63-C2-26"

    # Base64 encode public key
    base64_public_1 = base64_encode(public_key_1)

    # Send base64_public_1 and mac_1 to Device 2
# Device 2 (MAC Address 2: 00-B0-D0-63-C2-28)
Device2:
    # Private key of Device 2 (randomly generated)
    private_key_2 = 15

    # Compute public key B
    public_key_2 = (g ** private_key_2) % p

    # MAC Address of Device 2
    mac_2 = "00-B0-D0-63-C2-28"

    # Base64 encode public key
    base64_public_2 = base64_encode(public_key_2)

    # Send base64_public_2 and mac_2 to Device 1

# Device 1 receives base64_public_2 and mac_2
Device1:
    received_base64_public_2 = "MTk="
    received_mac_2 = "00-B0-D0-63-C2-28"

    # Base64 decode public key
    received_public_2 =
    base64_decode(received_base64_public_2)
```

```
    # Verify MAC address
    if received_mac_2 == "00-B0-D0-63-C2-28":
        # Compute shared secret
        shared_secret_1 =
        (received_public_2 ** private_key_1) % p
    else:
        # Drop connection due to MAC address mismatch
        reject_connection()
# Device 2 receives base64_public_1 and mac_1
Device2:
    received_base64_public_1 = "OA=="
    received_mac_1 = "00-B0-D0-63-C2-26"

    # Base64 decode public key
    received_public_1 =
    base64_decode(received_base64_public_1)
    #Verify MAC Address
    if received_mac_1 == "00-B0-D0-63-C2-26":
        # Compute shared secret
        shared_secret_2 =
        (received_public_1 ** private_key_2) % p
    else:
        # Drop connection due to MAC address mismatch
        reject_connection()
# Both devices now have the same shared secret
assert shared_secret_1 == shared_secret_2

# Now use the shared secret for encryption
symmetric_key = shared_secret_1
```

Encoding

```
function base64Encode(input) {
    const chars = <input from external>;
    let output = '';
    let i = 0;
    while (i < input.length) {
        const a = input.charCodeAt(i++);
        const b = input.charCodeAt(i++);
        const c = input.charCodeAt(i++);
        const i1 = a >> 2;
        const i2 = ((a & 3) << 4)|(b >> 4);
        const i3 = isNaN(b)?64:((b & 15) << 2)|(c >> 6);
        const i4 = isNaN(c)?64:c & 63;
        output += chars.charAt(i1) + chars.charAt(i2) +
        chars.charAt(i3) + chars.charAt(i4);
    }
    return output;
}
```

Decoding

```
function base64Decode(input) {
    const chars = <input from external>;
    let output = '';
    let i = 0;
    input = input.replace(/[^A-Za-z0-9\+\/\=]/g, '');
    while (i < input.length) {
        const i1 = chars.indexOf(input.charAt(i++));
        const i2 = chars.indexOf(input.charAt(i++));
        const i3 = chars.indexOf(input.charAt(i++));
        const i4 = chars.indexOf(input.charAt(i++));
        const a = (i1 << 2) | (i2 >> 4);
        const b = ((i2 & 15) << 4) | (i3 >> 2);
        const c = ((i3 & 3) << 6) | i4;
        output += String.fromCharCode(a);
        if (i3 !== 64)
            output += String.fromCharCode(b);
        if (i4 !== 64)
            output +=  String.fromCharCode(c);
    }
    return output;
}
```

Signature Verification

After the signature is set by both vehicles the key will become the unique ID for 2 vehicles to communicate with each other and no other attacker would be able to contact or communicate with the vehicles because the MAC Address would be unique for each device. The encoded MAC address will be helpful in delaying the interceptor to decrypt the data, in a case when interceptor indulges inside, the Diffie-Hellman key exchange is there to ensure safety.

Location Acquisition and Blockchain Storage

The authenticity of device identity within the blockchain network is confirmed by validating the device?s public key, plaintext, and signature data. This verification process is also applied when the service provider requests access to particular device?s location details. This underscores the storage system?s commitment to enabling secure sharing of location information [48]. By the blockchain protocol, we increase the privacy and security, later the authentication systems take over and communication between the vehicles will take place.

Advantages

a. FEA reduces computation and communication overheads compared to traditional authentication schemes.

b. The use of Diffie-Hellman key exchange and MAC Address-based encoding ensures uniqueness and integrity in vehicle authentication.
c. Privacy preservation is upheld, as the MAC Address serves as a vehicle identifier without exposing sensitive information and Diffie-Hellman key exchange playing authentication of real user.

7 Conclusion

VANETS are being developed with improvements in wireless communication technology. To provide a satisfactory performance in the network we have to work more on security protocols and communication systems. Security is the major issue in VANET considering many new types of attacks being generated. Of all, ensuring safety in VANET by securing routing is certainly a challenging responsibility in the network. From a point view research, stable more work is needed in the security area based on location verification, privacy and authentication to improve the performance of the overall system.

This thesis has proposed base64 encoding based First Message Authentication Protocol (FMAP) and Frontal Encoding Authentication (FEA) scheme for communication authentication in VANET, also we recommend to use blockchain protocol for location storage. The proposed protocol uses two stages of Location Verification technique. At the first stage, the packet is confirmed through reliability checks when information is transmitted b/w nodes. Considering the timestamp, transmission range and speed reliability check is performed. Finally, the geographic location is confirmed using the distance bounding technique. The blockchain location storage system ensures the privacy in the side of location acquisition.

A First Message Authentication Protocol (FMAP) with three different stages is proposed for the purpose of minimizing the routing overhead. In the first stage, vehicles in the same transmission range are formed as a group of vehicles. Communication among the vehicles is carried out using an encoded message. The signing and verification of the group are performed by their MAC Address of the vehicle. In the second stage, the packets are sent for reliability check by validating the vehicle?s lifetime and group legitimacy. When a vehicle gets the validation packet, a Co-operative message validation is performed at the third stage. Therefore, the group of vehicles gets verified going these three stages henceforth, it guarantees the safety and security for the correspondence among the grouped vehicles. The study leads to the observation that the proposed protocol provides more security and reduces the routing overhead in the network.

Acknowledgement. I am greatly thankful to Dr. Sandosh S, my guide, from the Department of Computer Science and Engineering, Vellore Institute of Technology, Tamilnadu, India, for believing in me and taking my dream project as his thesis. He was always there to provide necessary guidance, lead me through complex research processes and give me the lasting encouragement that I may need. His valuable suggestions and constant support that have not wavered definitely contributed to have this study done. I also would love to appreciate my friends and companions, Subash Saajan and Eniyavasanthan R, who are from Department of Computer Science Engineering of VIT. I recognize that my advisors and fellow students? contributions in the collaboration groups, brainstorming sessions, and constructive criticism of this paper were pivotal in shaping the ideas that it presents, through their passion and enthusiasm this research project became successful.

References

1. Abboud, K., Zhuang, W.: Stochastic modeling of single-hop cluster stability in vehicular ad hoc networks. IEEE Trans. Veh. Technol. **65**(1), 226?240 (2016)
2. Abumansoor, O., Boukerche, A.: A secure cooperative approach for non line-of-sight location verification in VANET. IEEE Trans. Veh. Technol. **61**(1), 275?285 (2012)
3. Alexey, L.: Heterogeneity in parallel and distributed computing. J. Parallel Distrib. Comput. **73**(12), 1523?1524 (2013)
4. Ameneh, D., Akbar, G.R.: An advanced security scheme based on clustering and key distribution in vehicular ad-hoc networks. Comput. Electr. Eng. **40**(2), 517?529 (2013)
5. Asaar, M.R., Salmasizadeh, M., Susilo, W.: A secure and efficient authentication technique for vehicular ad-hoc networks (2016). https://eprint.iacr.org/2016/822.pdf
6. Ayaida, M., Barhoumi, M., Fouchal, H., Doudane, Y.G., Afilal, L.: Joint routing and location-based service in VANETs. J. Parallel Distrib. Comput. **74**(2), 2077?2087 (2014)
7. Bhoi, S.K., Khilar, P.M.: SIR- a secure and intelligent routing protocol for vehicular ad-hoc networks. IET Netw. **4**(3), 185?194 (2015)
8. Bickel, G.S.: Inter/Intra-Vehicle Wireless Communication (2006) http://www.cse.wustl.edu/?jain/cse574-06/ftp/vehicular~wireless
9. Biswas, S., Misic, J.: Deploying proxy signature in VANETs. In: Global Telecommunication Conference, pp. 1?6 (2010)
10. Bohlooli, A., Jamshidi, K.: A GPS-free method for vehicle future movement directions prediction using SOM for VANET. Appl. Intell. **36**(3), 685?697 (2011)
11. Candido, C.G., Pino, C.G., Jezabel, M.G.: Mutual authentication in self-organized VANETs. Comput. Stand. Interfaces **36**(4), 704?710 (2014)
12. Chang, B.J., Liang, Y.H., Yang, H.J.: Performance analysis with traffic accident for cooperative active safety driving in VANET/ITS. Wirel. Pers. Commun. **74**(2), 731?755 (2013)
13. Chaudhuri, A., DasGupta, S., Saha, S.: Identity based secure algorithm for VANET. Procedia Eng. **38**, 165?171 (2012)
14. Chaurasia, B.K., Verma, S.: Infrastructure based authentication in VANETs. Int. J. Multimedia Ubiquit. Eng. **6**(2), 41?54 (2011)
15. Chaurasia, B.K., Verma, S.: Secure pay while on move toll collection using VANET. Comput. Stand. Interfaces **36**(2), 403?411 (2014)
16. Chen, C.L., Chang, I.C., Chang, C.H., Wang, Y.F.: A Secure ambulance communication protocol for VANET. Wirel. Pers. Commun. **73**(3), 1187?1213 (2013)
17. Chen, T., Mehani, O., Boreli, R.: Trusted routing for VANET. In: 9th International Conference on Intelligent Transport Systems Telecommunications (ITST), pp. 647?652 (2009)
18. Chim, T.W., Yiu, S.M., Hui, L., Li, V.: SPECS: secure and privacy enhancing communications schemes for VANETs. Ad Hoc Netw. **9**(2), 189?203 (2011)
19. Chim, T.W., Yiu, S.M., Hui, L., Li, V.: VANET-based secure taxi service. AdHoc Netw. **11**(8), 2381?2390 (2013)
20. Chim, T.W., Yiu, S.M., Hui, L., Li, V.: VSPN: VANET-based secure and privacy-preserving navigation. IEEE Trans. Comput. **63**(2), 510?524 (2014)
21. Chinnasamy, A., Prakash, S., Selvakumari, P.: Wagon Next Point Routing Protocol (WNPRP) in VANET. Wirel. Pers. Commun. **90**(2), 875?887 (2016)
22. Cunha, F., et al.: Data communication in VANETs: protocols, applications and challenges. AdHoc Netw. **44**, 90?103 (2016)
23. Defuentes, J.M., Manzano, J.M., Tablas, A.I., Blasco, J.: WEVAN ? a mechanism for evidence creation and verification in VANETs. J. Syst. Architect. **59**(10), 985?995 (2013)

24. Dinesh, T., Mayank, B., Ankit, Y., Suman, J.: A novel secure authentication scheme for VANETs. Second Int. J. Comput. Intell. Commun. Technol. **21**(5), 1733?1743 (2016)
25. Dok, H., Fu, H., Echevarria, R., Weerasinghe, H.: Privacy issues of vehicular ad-hoc networks. Int. J. Future Gener. Commun. Network. **3**(1), 197?209 (2010)
26. Domingo, M., Reyes, A.: A clean slate architecture design for VANETs. Wirel. Pers. Commun. **67**(2), 315?333 (2011)
27. Eiza, M.H., Owens, T., Ni, Q.: Secure and robust multi-constrained QoS aware routing algorithm for VANETs. IEEE Trans. Depend. Secure Comput. **13**(1), 32?45 (2016)
28. Fazio, P., De Rango, F., Sottile, C.: A predictive crosslayered interference management in a multichannel MAC with reactive routing in VANET. IEEE Trans. Mob. Comput. **15**(8), 1850?1862 (2016)
29. Feiri, M., Pielage, R., Petit, J., Zannone, N., Kargl, F.: Predistribution of certificates for pseudonymous broadcast authentication in VANET. In: Vehicular Technology Conference (VTC Spring), pp. 1?5 (2015)
30. Feng, X., Li, C.Y., Chen, D.X., Tang, J.: A method for defensing against multi-source Sybil attacks in VANET. Peer-to-Peer Netw. Appl. **10**(2), 305?314 (2016)
31. Feukeu, E.A., Djouani, K., Kurien, A.J.: Doppler shift mitigation in a VANET using an IDDM approach. J. Ambient. Intell. Humaniz. Comput. **7**(3), 321?332 (2016)
32. Fonseca, E., Festag, A.: A survey of existing approaches for secure adhoc routing and their applicability to VANETS. NEC Network Laboratories, pp. 1?28 (2006)
33. Gajbhiye, V.A., Jasutkar, R.W.: Study of efficient routing protocols for VANET. Int. J. Sci. Eng. Res. **4**(3), 1?8 (2013)
34. Ganan, C., Munoz, J.L., Esparza, O., Diaz, J.M., Alins, J.: PPREM: privacy preserving revocation mechanism for vehicular adhoc networks. Comput. Stand. Interfaces **36**(3), 513?523 (2014)
35. Ganan, C., Munoz, J.L., Esparza, O., Diaz, J.M., Alins, J.: EPA: an efficient and privacy aware revocation mechanism for vehicular adhoc networks. Pervasive Mob. Comput. **21**, 75?91 (2015)
36. Garcia, J.S., Campus, J.M.G., Reina, D.C., Toral, S.L., Barrero, F.: On-site driver ID: a secure authentication scheme based on SpanisheId cards for vehicular adhoc networks. Futur. Gener. Comput. Syst. **64**, 50?60 (2016)
37. Ghosh, A., Paranthaman, V.V., Mapp, G., Gemikonakli, O.: Exploring efficient seamless handover in VANET systems using network dwell time. EURASIP J. Wirel. Commun. Network. 1?19 (2014)
38. Hao, Y., Cheng, Y., Zhou, C., Song, W.: A distributed key management framework with cooperative message authentication in VANETs. IEEE J. Sel. Areas Commun. **29**(3), 616?629 (2011)
39. Harsch, C., Festag, A., Papadimitratos, P.: Secure position-based routing for VANETs. In: Vehicular Technology Conference (VTC), pp. 26?30 (2007)
40. Hu, C., Chim, T.W., Yiu, S.M., Hui, L., Li, V.: Efficient HMAC-based secure communication for VANETs. Comput. Netw. **56**(9), 2292?2303 (2012)
41. Huang, D., Misr, S., Verma, M., Xue, G.: PACP: an efficient pseudonymous authentication-based conditional privacy protocol for VANETs. IEEE Trans. Intell. Transp. Syst. **12**(3), 736?746 (2011)
42. Hussain, R., Kim, S., Oh, H.: Towards privacy aware pseudonymless strategy for avoiding profile generation in VANET. In: International Workshop on Information Security Applications, pp. 268?280 (2009)
43. Hussain, R., Oh, H.: On secure and privacy-aware Sybil attack detection in vehicular communications. Wirel. Pers. Commun. **77**(4), 2649?2673 (2014)
44. Hussain, R., Rezaeifar, Z., Lee, Y., Oh, H.: Secure and privacy-aware traffic information as a service in VANET-based Clouds. Pervasive Mob. Comput. **24**, 194?209 (2015)

45. Hussain, R., Rezaeifar, Z., Oh, H.: A paradigm shift from vehicular adhoc networks to VANET-based clouds. Wirel. Pers. Commun. **83**(2), 1131?1158 (2015)
46. Indira, K., Joy, E.C.: Energy efficient IDS for cluster-based VANETS. Asian J. Inf. Technol. **14**(1), 37?41 (2015)
47. Jeon, Y.B., Lee, K.H., Park, D.S., Jeong, C.S.: An efficient cluster authentication scheme based on VANET environment in M2M application. Int. J. Distrib. Sens. Netw. **9**(11), 1?9 (2013)
48. Li, D., Hu, Y., Lan, M.: IoT device location information storage system based on blockchain. Futur. Gener. Comput. Syst. **109**, 95?102 (2020)

Classification and Analysis of Online Business Models: A Framework for Practical Application

Roland Schmuck(✉)

University of Pécs Faculty of Business and Economics, Rákóczi út 80., Pécs 7622, Hungary
schmuck.roland@ktk.pte.hu

Abstract. This research investigates and categorizes online business models through the development and validation of an own online classification for online business models. Adapting Malone et al. (2006) traditional business model framework to the online economy, the study proposes a system categorizing models based on rights (ownership, usage, intermediation) and asset modification in online environments. The framework was tested against datasets comprising the most visited websites and popular mobile applications, demonstrating its robustness in categorizing online business models. The usage rate of online business models was examined in two datasets: websites and mobile applications.

Findings highlight the intellectual property landlord as the most used, dominant online business model, taking over half of the business models in the analyzed two datasets. The own classification system effectively encompassed all identified models from literature review and practical use, showing its suitability for understanding and categorizing online business models.

Keywords: business model · online business model · business model classification

1 Introduction

It has long been known that information systems are of strategic importance in the operation of organizations Porter-Millar 1985; Schmuck 2023. Due to technological advancements and the widespread use of the internet, the issue for organizations today is not the use of information technology but digital transformation, which has become a crucial topic for organizational development Fitzgerald et al. 2013; Kane et al. 2015; Leipzig et al. 2017; Kaufman and Horton 2015; Mursalzade et al. 2023. This digital transformation fundamentally reshapes business processes and affects the strategy, culture, and human resource management of companies Jarjabka 2001, Kane et al. 2015. The higher level of information access, knowledge sharing, and direct market connection brought by digital transformation offers opportunities to improve competitiveness Hornyák et al. 2023. Companies' expectations align with this: they anticipate an improvement in their market position through digital transformation Katona et al. 2023.

According to Atkinson et al. (2010), the significance of the internet can be compared to historical breakthroughs such telephone, electricity of the internal combustion engine.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
S. Rajagopal et al. (Eds.): ASCIS 2024, CCIS 2428, pp. 391–410, 2025.
https://doi.org/10.1007/978-3-031-86302-8_26

These so-called General Purpose Technologies (GPT) have historically appeared roughly every half-century, completely transforming the economy, the production processes and their management. The infrastructure and legislation also adapt to this. GPT industries start with limited spread in raw form and then permeate the entire economy while the costs are decreasing. The are becoming an integral part of industries causing innovation in organizations and business models Atkinson et al. 2010. A newly developing GPT industry could be the AI.

As the functionality of the internet continues to evolve, it not only connects computers but also serves as a communication channel and a platform for legal, economic, and financial transactions Borgulya et al. 2022. Mobile devices, broadband data transmission, WiFi, and high-performance computing capacity are integrated into numerous new devices, all ultimately connecting to a single digital network. Hence, the internet is increasingly omnipresent. The internet has become a "force" bringing about social changes in the lives of people and organizations alike Ruzsa 2018. This process can be compared to the creative destruction described by Schumpeter in the early 20th century, where companies inventing or utilizing new technologies achieve rapid growth by displacing the old ones Nicholas 2003. Digitalization offers good opportunities for disruptive innovation, creating new types of business models Tekic and Koroteev 2019; Gong and Ribiere 2021.

As early as 1987, it was predicted that the use of information technology would change markets Malone et al. 1987. According to the research by Dumagan et al. 2003, between 1989–2001, industries that invested more in information technology saw an average annual productivity increase of 3.03%, compared to the average 0.42% growth in other industries. Rojko et al. (2020) reached similar conclusions in their study examining the periods from 2008–2018. The online economy grows faster than the traditional economy. In Hungary, during the early years of e-commerce from 2004–2010, the volume of e-commerce grew sevenfold, while the industrial average decreased by 10% in the same period Nemeslaki 2012. E-commerce has been continuously growing since then, with the Covid lockdowns providing a boost in recent years. In 2023, 78% of the Hungarian population were online shoppers Lone and Weltevreden 2023. Globally, the volume of e-commerce is estimated to reach USD 3,647 billion by 2024. Growth between 2024–2028 is projected to be 9.83% annually Statista 2024. However, this is only a small slice of digital transformation; Industry 4.0 encompasses much more Ruzsa 2019. Digitalization investments are expected to increase Zaušková et al. 2022.

While the focus of the third industrial revolution was on automation, Industry 4.0 involves the digitalization of all physical devices, connecting them into a common digital ecosystem PwC 2016. In Europe, particularly in Germany, the term Industry 4.0 is closely associated with manufacturing. In the United States and other English-speaking countries, the term Industry 4.0 is generally used in conjunction with concepts such as the "Internet of Things", the "Internet of Everything", and the "Industrial Internet" Deloitte 2014. This research examines the business models of online companies using the latter approach.

Big data offers opportunities for more detailed understanding and analysis of consumer expectations and behavior Acciarini et al. 2023. These opportunities create the chance for companies to focus on their core capabilities. The internet can be used in

many ways. Many new online business models have emerged from traditional business models. Caputo et al. 2021; Obermayer et al. 2023, Schmuck 2019, Tapscott 2011.

Digitalization impacts organizations in many ways, but its innovative impact on the entire organization is noteworthy Deloitte 2017. In rapidly changing environmental conditions, interpreting the impact of value creation processes on the business model shapes the development of companies Łęgowik-Świącik 2023. Companies lagging in digitalization have a lower added value than those participating in digitalization Hartvig et al. 2023. According to a PwC study, which surveyed 2000 companies in 26 countries, companies typically digitalize their internal operational processes and are characterized by the development of new business models through disruptive innovation, leading to increased digital revenues PwC 2016. Disruptive innovation implies the revolutionary change of operations, i.e., the applied business models Christensen et al. 2015; Consumano et al. 2020. Society also needs to adapt to this; acquiring digital skills becomes crucial since disruptive business models impact human resource management Benke 2018; Benke 2020; Bresciani et al. 2021. The Covid pandemic further increased the need for digital skills Dajnoki et al. 2023.

Managing digital transformation requires reviewing the organization's business model, redesigning processes from planning to implementation Jarjabka et al. 2024, Kaufman and Horton 2015; Sahri and Halim 2023. This innovation-based transformation causes changes in organizations, potentially altering the organizational structure or even the more deeply rooted organizational culture Liu et al. 2011. All these can create "unprecedented things" with the use of new technologies Schmuck 2013; Brynjolfsson and McAfee 2014; Tekic and Koroteev 2019.

2 Research Objectives and Methodology

The aim of this research is to examine how the expansion of the online economy affects business models and what types of online business models can emerge in the digital world. The research hypotheses are:

H1. It is possible to develop a classification system for online business models that can categorize the online business models found in practice.
H2. There exists an online business model that constitutes the majority of online business models used in practice.

The research methodology involves first presenting and analyzing the existing classifications of online business models, followed by evaluating them and introducing a comprehensive classification system for online business models. This is followed by testing the developed classification system. Initially, the classification system is compared with previous classifications known from the literature, and then it is tested against online business models found in practice based on two different databases using manual classification. This manual classification was conducted by visually inspecting each website, supported by automated translation tools for non-Latin texts. The classification process followed strict criteria to ensure consistency and was supplemented by statistical analysis to determine the prevalence of certain business models. The research's analytical methodology involves literature analysis followed by manual database analysis. The

sample size is 500 + 200 elements. The databases used are Alexa (2019) with 500 sample elements and Apple (2019) with an additional 200 sample elements. The sampling date is June 6, 2019.

The research replicates a previous study with a new sample. The previous study sampled the Alexa database in 2012 and mobile applications in 2014 Schmuck 2015; Schmuck 2017. Due to the discontinuation of the Alexa database, no newer data is available for this research. The sampling date is also optimal as it is unaffected by the influence of COVID-19.

The topic of this research does not belong to the typical mainstream trends of today. While business informatics researchers focus on corporate governance systems, marketers are more concerned with social media, and business model creators (strategists) focus on traditional business models. This research does not aim to fit into the most commonly researched topics. The scientific examination of online business models has declined over the past decade, making it worthwhile to replicate the earlier research from 2012–2014.

3 Literature Review on the Classifications of Online Business Models

A business model demonstrates how a company creates value for its customers and generates profit using the company's resources Schmuck 2021a; Schmuck 2021b. The business model should describe the operation of an organization concisely and comprehensibly Schmuck 2021c. In this research, the term "online business model" refers to value creation conducted in the online space.

The early Timmers (1998) classification remains one of the most well-known classifications of online business models. Based on case study analysis, Timmers identifies ten online business models, some of which are frequently used while others are less widespread. His models predominantly focus on commercial and corporate support services. Timmers evaluated his models based on the degree of functional integration and innovation. The classification is not closed, allowing for the development of new models beyond the ten described online business models.

Weill and Vitale (2001) classified online business models years before publishing their traditional business model classification. Their goal was to assist traditional "brick and mortar" businesses in entering the online market. Their models help these companies define their activities in the online space. Similar to Timmers' logic, they differentiate between traditional and innovative online business models. However, they acknowledged that their classification does not clearly categorize the models since some businesses may develop new, innovative solutions while using their online business model, leading to hybrid models that are difficult to separate. According to their theory, classification is possible by considering so-called atomic business models, which are fundamental models whose combinations can create new models. The classification cannot be considered comprehensive, presenting only the most important types of online business models as perceived by them. In practice, popular services such as the PayPal payment system or domain traders cannot be classified under Weill and Vitale's models. These shortcomings highlight the model's imperfections, but this cannot be held against the authors since they

designed the model for companies extending their traditional operations into the online business world. The model is suitable for this purpose, but it does not comprehensively classify online business models.

Rappa's (2002) research aims to create a comprehensive taxonomy of online business models. According to him, the list of online business models cannot be considered complete and final, with new models expected to emerge in the future. Rappa classified a total of forty online business model subtypes into nine main types. Since Rappa did not justify the classification with scientific methods, it can be considered more ad-hoc, with its value lying in the large number of described models and its widespread recognition in the literature. Based on Rappa's classification, a Hungarian study was also conducted Nemeslaki et al. 2008.

Laudon and Traver (2004) present twenty-one models of e-commerce business models. Primarily, they categorize the models based on the participating parties (provider and receiver), resulting in B2C, B2B, C2C, P2P, and m-commerce models. Such a level of separation is difficult to achieve in practice.

Eisenmann's (2002) classification of online business models is noted more for its familiarity than its detail. Eisenmann presents the most common online business models through twenty-three case studies. During his research, he distinguishes eight general online business models. Eisenmann's goal was to create case studies, not a comprehensive classification of online business models, which his small sample size would not allow. This is a significant limitation of Eisenmann's classification, as it is not comprehensive.

The classification of Móricz (2009) is based on value proposition of the company. He expanded this classification by including architecture and revenue sources. For the purpose of this research, the classification based on value propositions is relevant, distinguishing four online business models. Móricz's classification follows the principle of simplification.

The above sections have presented the shortcomings of the most well-known online business model classifications found in the literature. Except for Timmers' (1998) classification, all other authors include models that cannot be considered purely online business models. However, Timmers' classification, by his own admission, is not comprehensive. It is evident that each online business model classification has its shortcomings.

4 Own Classification of Online Business Models

Classifications of online business models must meet several important criteria:

- Logical Coherence: The classification must capture the similarities among the business models grouped together.
- Comprehensiveness: It should ensure that all companies can be categorized.
- Clear Definitions: This includes the rules for categorization and their objectivity.
- Conceptual Elegance: It should define as few concepts and conditions as possible, and the terms used must be clear and simple Malone et al. 2006.

Since the existing online business model classifications in the literature do not meet these criteria, this research presents a new, comprehensive online business model classification that fully covers online business activities.

The basis of the new comprehensive classification originates from traditional, non-online business models. This classification is based on the dimensions used by Malone and his colleagues (2006), as their theory systematically classifies traditional (non-online) business models according to predetermined criteria.

Malone et al. (2006) created a classification that meets the criteria outlined above. Their initial sample consisted of the 1,000 largest companies in the United States in the year 2000. According to their theory, the essence of companies can be captured by considering what they sell. This can legally be:

- Ownership Rights: Where the buyer gains ownership of the purchased asset and can do with it as they wish.
- Usage Rights: For a specific period, such as a rented car or hotel room. In this case, the buyer can use the asset under certain conditions, but the ownership remains with the company.
- Intermediary Services: Where the company's goal is to connect sellers with buyers. Examples include real estate agencies.

These fundamental business types are illustrated in Table 1.

Table 1. The four basic business models

What type of rights are involved?	How much is the asset changed by the company?	
	Much	*Not much*
Ownership	Creator	Distributor
Using	Landlord	
Intermediation	Broker	

Malone et al. (2006) p. 30.

Another important dimension of the classification is the nature of the assets involved in the transaction. Four asset categories can be distinguished: (1) physical, (2) financial, (3) intangible, and (4) human. Physical assets include durable and consumer products. Financial assets include money and assets that result in potential future cash flows for their owners, such as stocks, bonds, and insurance. Intangible assets are legally protected intellectual properties, such as patents, copyrights, trademarks, and trade secrets, as well as other intangible assets like knowledge, goodwill, and brand image. Human assets represent people's time and effort. While people are not assets that can be sold or leased, their time and knowledge can be "rented" for money. As shown in Table 2, the combination of rights and assets can create 16 different business models. All of these are logically possible, though some are rare, and two are illegal, as indicated in parentheses in Table 2.

Updating Malone et al.'s framework was necessary to accommodate the unique characteristics of online business models. The changes included adjusting the models to the online environment, checking if they were appropriate for online conditions, and

Table 2. The sixteen ordinary business models

Basic model type	What type of asset is involved?			
	Financial	*Physical*	*Intangible*	*Human*
Creator	Entrepreneur	Manufacturer	Inventor	Human Creator
Distributor	Financial Trader	Wholesaler/Retailer	Intellectual Property (IP) Trader	Human Distributor
Landlord	Financial Landlord	Physical Landlord	Intellectual Landlord	Contractor
Broker	Financial Broker	Physical Broker	Intellectual Property (IP) Broker	Human Resources (HR) Broker

Malone et al. (2006) p. 30.

updating their names, while the main framework remained unchanged. These changes were tested for practical applicability using real-world data to ensure that they aligned with the operational dynamics of modern online businesses.

Before moving on to the own classification system of online business models, it is necessary to clarify what we mean by online business activities. Online business models are used by companies or business units that primarily conduct their activities online, using physical resources only to the extent necessary to perform online activities. These companies or business units:

- Primarily provide or mediate online products, services, or content.
- Are operationally inseparable from the internet, with the internet being their main sales channel.

In practice, a company or business unit can be classified as using an online business model if more than 50% of its cash flow comes from online activities. An online business model can be used not only by a company but also by a business unit Móricz 2009. However, in this analysis, for simplicity, I will refer to them as companies regardless of whether the entire company or the examined business unit uses an online business model. Online business models are not used by:

- Traditional companies with only an informational website.
- Primarily manufacturing companies with an online presence and their own webshop.
- Traditional "brick and mortar" stores with an online presence and supplementary webshops.
- Traditional financial institutions with an online presence, including internet banking.
- Companies leasing physical goods with an online presence, including accommodation providers with online booking systems.
- Public institutions and government websites, even if they offer online services Schmuck 2015.

Table 3. The own classification of online business models

Basic model type	What type of asset is involved?			
	Financial	*Physical*	*Intangible*	*Human*
Creator	(not online operation)	(not online operation)	Intellectual Property Creator	(not allowed by law)
Distributor	Financial Transaction Provider	Online merchant	Intellectual Property Distributor	(not allowed by law)
Landlord	(not online operation)	(not online operation)	Intellectual Property Landlord	Online Service Provider
Broker	Financial Broker	Online Marketplace	Intellectual Property Broker	HR Broker

Own research based on Malone et al. 2006.

Based on the dimensions created by Malone et al. (2006), a unique classification of online business models has been developed, as shown in Table 3. The names of the categories in the classification were chosen to best refer to online activities. Some collective names are quite specific, while others remain general to encompass a broader range of activities.

Some activities are not interpretable within the framework of online dimensions according to Malone's dimensions. These include financial, physical product manufacturers, and "landlords" owning such products. In the case of the latter, rental refers to the physical nature of the asset, which is not primarily considered an online activity. Human creation and distribution activities are not permitted by law, so these cells in Table 3 does not contain online business models. The own developed classification distinguishes ten different online business models. These models are:

- Intellectual Property Creator: Creates and sells or otherwise makes available intellectual properties. Companies using this online business model typically create online content, which they also make available or sell online. Intellectual property creators can also be automated systems, such as statistical traffic measurement systems or mobile applications that create new intangible assets (e.g., edited photos). If intellectual property creators do not sell the property but only provide usage (access) rights, they can be classified as intellectual property landlords (see type 5).
- Financial Transaction Provider: Companies facilitating online financial transactions. This includes providers of online payment solutions and other financial organizations operating online but not traditional banks.
- Online Merchant: Companies selling physical products through the internet. This is a typical online business model, as companies or their business units often extend their sales to online platforms. However, according to the previously defined criteria, only those companies or business units whose online revenue exceeds their traditional

sales revenue belong to this category. Both online wholesalers and retailers fitting this criterion are classified here, with the further condition that they do not conduct sales through common marketplaces, as this would mean they apply an online marketplace business model (see type 8).

- Intellectual Property Distributor: Companies trading intangible assets online. This includes, among others, those dealing with the business sale of domains and websites, as well as domain registrars.
- Intellectual Property Landlord: Provides access to their intellectual property online. This is a very common online business model because internet technology was developed for the publication of data and information. Content providers may charge a subscription fee for their services or finance their operations in other ways, such as displaying advertisements. Typical intellectual property landlords include content-providing websites, which gain people's attention through visits. They sell ad space or charge subscription fees for content. This category includes websites providing text, images, or video content, as well as database operators who own or operate and make online databases accessible. Publishers also fall into this category, primarily including software publishers. Publishers provide limited usage rights to buyers in exchange for a subscription or license fee, while maintaining the right to reissue, copy, and resell their intellectual property.
- Online Service Provider: Organizations providing online services performed by people and charging service fees. This includes, for example, online consulting and search engine optimization. The role of this model decreases with the automation characteristic of Industry 4.0, as services are often performed by software or robots.
- Financial Broker: Brokers financial assets between sellers and buyers. This includes online stock trading software and insurance intermediary websites.
- Online Marketplace: Brings together buyers and sellers of physical assets without owning the assets being traded. Virtual marketplaces and online auction sites operate based on this model.
- Intellectual Property Broker: Mediates intellectual property between creators and users. This includes image and video sharing sites, as well as search engines that mediate information between parties. With the rise of smartphones and social media, the role of these models is becoming more prevalent.
- Human Resource Broker: Companies providing online personnel mediation, job portals, and online dating services.

5 Validating the Own Classification System

A critical test of the own categorization system is whether the types of online business models from previous categorizations can be classified within own system. The comparison between my classification system and others in the previous literature was done based on its ability to categorize modern online business models, with particular attention paid to including purely online activities. For replicability, each theory and framework was analyzed in the same way to ensure consistency in findings. This allows for a comparison of different categorizations and helps determine the emphasis areas of each system as well as the online business models most frequently mentioned in the literature. Based on this logic, Fig. 1 was created, showing the distribution of online

business models according to the own categorization within the most well-known categorizations in the literature. Every online business model in the categorizations can be classified into the own online business model types which is shown in Fig. 1. It must be noted that some less precisely defined online business models may belong to two or even three models into the own categorization. In such cases, calculations were made with half or third counts.

As seen in Fig. 1, most of the previous classifications lacks applicability in online environments, as they miss several online models. Such as, for example, the intellectual property creator or the financial broker is missed by three of them. A key deficiency of these classifications is the lack of completeness. This deficiency was addressed in my own classification system by using a framework that guarantees the completeness and consistency of the classification.

The own classification system was not only validated by comparing it against frameworks from the literature, but additionally to this theoretical validation, the classification was validated using two independent datasets Alexa 2019, Apple 2019, ensuring its robustness across different online environments in practice, which are discussed in the following parts.

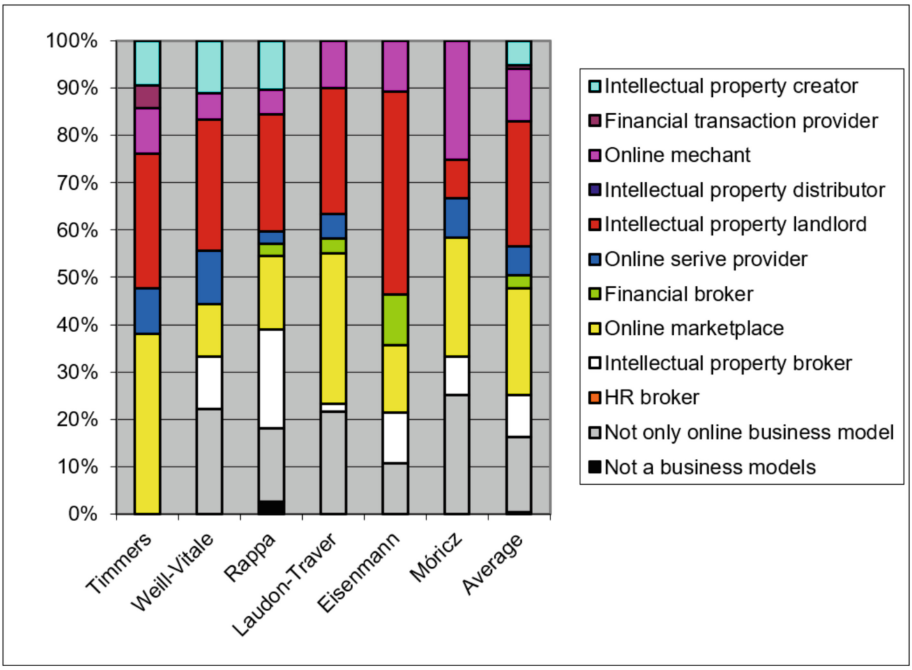


Fig. 1. Comparing the own classification with classifications in the literature [own research]

5.1 Results: Practical Testing of the Own Classification System on a Website Database

For practical examination, a database is needed from which the business model of individual online companies or their organizational units can be determined. Considering that we are dealing with online business models, the examination can be conducted based on websites, as it is reasonable to assume that an online company has a website. Based on the characteristics of business models, it can also be assumed that this can be determined based on the examined website. The purpose of a company's website is to present its products, services, or make them available online. In the case of an online business model, the latter is applicable, so analyzing the website through which the online activity is conducted clearly shows the underlying basic business model. Similar research logic has already been applied in Hungary. Nemeslaki and colleagues conducted a manual analysis of 125 websites and analyzed 6,800 websites with a robot by examining the expressions on the website Nemeslaki et al. 2008. Out of the 6,800 websites, business models were identified for 1,409 websites, but 79.27% of the sites did not have an online business model by nature (informational websites of "traditional" companies, individuals, or non-profit organizations). The robot survey operated with 80% reliability based on manual control tests, and the research did not provide information on encountering problems during individual analysis Nemeslaki et al. 2008. The current research only uses manual examination, which is considered highly reliable.

The practical analysis of websites is based on the Alexa Top 500 database Alexa 2019, which contains the world's most visited websites. Alexa's measurement is conducted using a browser script, updating the database daily. The sampling date was June 6, 2019.

In some cases, a company may have multiple websites listed in the Alexa Top 500 database, but this does not affect the research results, as each website's business model is classified separately, and the database is cleaned of actual duplicates before the analysis. Duplications include, for example, Google.com and Google.de, which are the same service of the same company in different languages. Different non-duplicate websites of a company are interpreted as if a company has multiple business units. Often, a company applies the same business model to different websites, but it is conceivable that it uses several different business models. In the manual analysis, the company or organizational unit is classified into the most characteristic model.

The sample also includes online appearances of traditional companies. Based on previous assumptions, these companies are not considered online companies. As previously explained, companies are considered online companies if more than 50% of their cash flow comes from online activities. The same logic can be applied to business units. In the absence of financial data, this can unfortunately only be estimated in some cases by weighing the available information, but this solution is adequate for the research purposes. The classification is mostly clear in such cases, while additional information may be required in other cases. The situation is similar for companies using mixed business models, where the more significant business model is the one generating the larger portion of the incoming cash flow. In some cases, where the activity belongs to different websites (e.g., Google–YouTube), several online business models may belong to one company. These appear in the sample in small numbers based on practical analysis, as

only the largest companies can get multiple websites onto the top lists. This is often achieved through different business units.

The classification of websites' business models is based on manual examination of each website. Every website is checked visually, and foreign language websites are translated using Google Chrome's built-in translation algorithm to facilitate easier classification, especially since the sample contains websites in various languages. This is particularly important for websites written in non-Latin scripts, such as the significant proportion of Chinese websites in the sample. The classification of websites is done with the due diligence expected in research, but this does not exclude possible misclassification due to human factors, although its rate is minimized and, given the large sample size of 500, it does not significantly affect the research results. The sample size was reduced to 388 after excluding duplicates (58), which were the same services of the same companies under different domains, traditional company websites not using an online business model (52), and websites where the business model could not be identified due to website unavailability or lack of other information (2). For this reduced sample, the methodology first determines what asset the transaction on the website involves, which can be financial, physical, intangible, or human. Then it examines what the company does with this asset: creates it, distributes it, owns it (landlord), or intermediates it (broker). This methodology has shown that every website can be classified into ten online business model types, which are shown in Fig. 2 including the results of the classification of the sample of 388 websites.

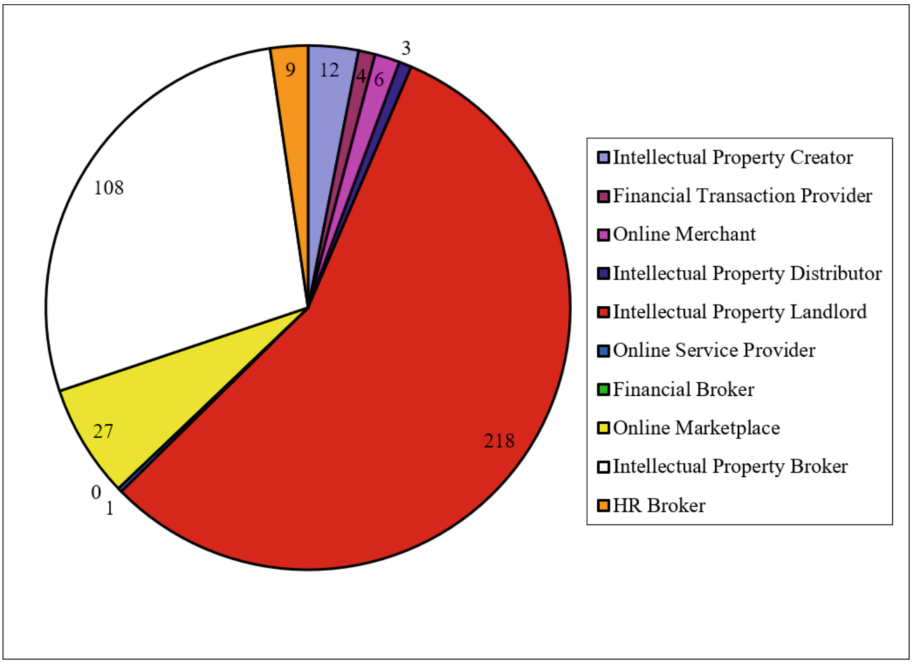


Fig. 2. Online business models of websites in the Alexa database [own research based on Alexa 2019]

According to the research results, some online business models are significantly more frequently used among the websites in the database. Since the database is practically a visitor ranking list, the online business models found here clearly show which models can be considered successful. These are primarily the intellectual property landlord (56.2%) and the intellectual property broker (27.8%) models. Online commerce is also important but less significant among online business models (although it is presumably more important in terms of revenue). The third most common model is the online marketplace (7%), followed by the intellectual property creator (3.1%). All other models are below 1.5% in the sample. Online business models related to financial activities account for only 1% of the sample, although several traditional bank websites are included, but these banks do not use online business models.

5.2 Results: Practical Testing of the Own Classification System on a Mobile Apps Database

The previously used methodology assumes that online activities are conducted through websites. In reality, this is not necessarily the case, as websites are the most commonly used interfaces on the internet, but it is not the only one. Mobile applications is another technology worth to consider in this research. Therefore, the testing of the self-developed classification is extended to include the currently popular mobile applications.

Several platforms exist for accessing mobile applications, the most well-known being Apple iOS, Google's Android, and Microsoft Windows operating systems. Considering that this is merely a difference in operating systems, analyzing a single operating system (platform) will provide sufficient results for the research, assuming that the online business models used are platform-independent. Since smartphones were first successfully popularized by Apple with the iPhone, the content of the Apple AppStore is examined, analyzing the top 100 free and top 100 paid applications, with the database provided by the iTunes Charts Apple 2019.

The iTunes Charts Apple 2019 database includes 100 + 100 applications. These applications can be run on iPhone, iPad, or Apple Watch devices. Mobile applications can be downloaded through the AppStore application on these devices, henceforth referred to as AppStore applications – as is commonly used. The AppStore opened on June 10, 2008, and currently contains approximately 1.8 million applications Costello 2020.

The database access date was June 6, 2019. It should be noted that the list dynamically changes even within a day, but this was the state at the time of download. The significance of the list change is minimal since this is not a comprehensive but a sample study – the sample is the top 100 free and top 100 paid applications available at the time of sampling on the iTunes Charts Apple 2019.

Classification into online business models assumes that the application found in the database reflects the business model of the company or business unit that created it. While for websites, ownership information of the domain name and legal statements on the sites provided clear information, for mobile applications in the AppStore, only the developers' names are available. Unlike websites, where various services often run from subdomains within a domain (e.g., maps.google.com, etc.), each application in the AppStore appears under a separate name. Therefore, overlaps were not considered here – unlike websites. Often, individual applications are related only to specific services

of the company that created them (e.g., Gmail, Google Maps, Yahoo Mail, etc.), which were treated as business units. This detailed analysis is not conducted here; the aim is merely to determine whether all applications in the sample can be classified into the self-developed online business model categories, thus testing the appropriateness of the self-developed classification. Excluding elements that do not use an online business model from the 200-element sample, a 197-element sample remains. Each of these could be clearly classified into the self-developed online business model framework. The results are shown in Fig. 3.

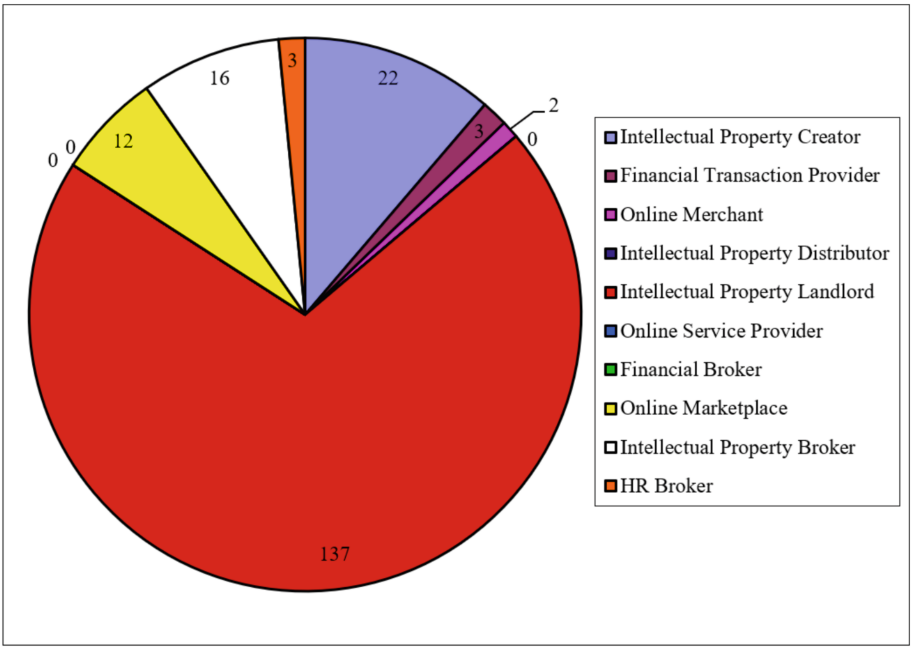


Fig. 3. Online business models in the iTunes Charts [own research based on Apple 2019]

The results indicate that the most commonly used online business model among AppStore applications is the intellectual property landlord (70.3%). The second most common is the intellectual property creator (11.3%), followed by the intellectual property agent (8.2%), and then the online marketplace (6.2%). The usage rate of other online business models does not exceed 1.5%. If we examine the free and paid applications separately, shifts in these proportions can be observed. In the sample, all online marketplace applications are free, while the vast majority of intellectual property creator applications are paid. This is shown in Table 4.

Table 4. Distribution of online business models in the iTunes Charts (2019) top 100 free and top 100 paid apps.

	Ratio among free apps	Ratio among paid apps	Ratio among all apps
Intellectual Property Creator	3.1%	19.4%	11.3%
Financial Transaction Provider	3.1%	0.0%	1.5%
Online Merchant	2.1%	0.0%	1.0%
Intellectual Property Distributor	0.0%	0.0%	0.0%
Intellectual Property Landlord	61.9%	78.6%	70.3%
Online Service Provider	0.0%	0.0%	0.0%
Financial Broker	0.0%	0.0%	0.0%
Online Marketplace	12.4%	0.0%	6.2%
Intellectual Property Broker	14.4%	2.0%	8.2%
HR Broker	3.1%	0.0%	1.5%
Total	100.0%	100.0%	100.0%

Own research based on Apple 2019

6 Conclusions

The research hypotheses were the following:

- H1. It is possible to develop an online business model classification into which the practically existing online business models can be categorized.
- H2. There exists an online business model that constitutes the majority of the practically applied online business models.

In examining the H1 hypothesis, the most well-known scholarly online business model classifications were analyzed: Timmers (1998), Weill-Vitale (2001), Rappa (2002), Laudon-Traver (2004), Eisenmann (2002) and Móricz (2009). None of these classifications fully met the criteria set by Malone and his colleagues (2006) regarding logicality, comprehensiveness, clarity, and conceptual "elegance." A new classification was introduced, which categorizes online business models based on Malone and his colleagues' (2006) framework for traditional business model classifications. The basis of this categorization is the rights (ownership, usage, mediation) and the extent to which the involved asset is altered. Based on this, theoretically 16, and practically, considering legal conditions, 10 online business model types can be distinguished in the online space. The classification was tested in two ways. First, the new classification was compared with the aforementioned scholarly sources, resulting in all online business models found

in the literature being categorizable into the new classification. In the second step, a practical analysis was conducted using samples from Alexa (2019) with 500 elements and iTunes Charts Apple 2019 with 200 elements. All online business models found in these databases could be categorized into the new online business model classification. Based on these findings, the H1 hypothesis is accepted, meaning it is possible to develop an online business model classification into which the practically existing online business models can be categorized.

The H2 hypothesis was examined using data from the Alexa (2019) and iTunes Charts Apple 2019 databases, applying the methodology of a newly introduced online business model classification. In both samples, the intellectual property landlord model was the most prevalent, appearing in 56.2% of cases in the Alexa database and 70.3% in the iTunes Charts database. This model's frequency exceeded 50% in both databases, indicating it is the majority among practically applied online business models. The second most common model in each database had significantly lower frequencies, with the intellectual property agent model at 27.8% in the Alexa database and the intellectual property creator model at 11% in the iTunes Charts database. Based on the analysis of these two databases, the H2 hypothesis is accepted, confirming the existence of an online business model that predominates among practically applied online business models: the intellectual property landlord model.

The research focused on analyzing the most visited websites and the most popular mobile applications. This approach has its limitations, as it does not account for "long tail" companies Anderson 2007 that cater to specialized market segments and do not appear in top lists. Despite this, the new online business model classification should be applicable to these companies as well. Another limitation is that the research only examined websites and mobile applications, without considering other technologies. However, since these two technologies are the most popular and widely used, the research results still meet the set objectives despite this limitation.

As Industry 4.0 and digitalization continue to advance, the online economy is expected to grow as well. In this context, the research results are particularly valuable, offering guidance for companies aiming to enter the online market by identifying prevalent models and their usage. Additionally, this research can serve as a reference for other studies. The newly introduced online business model classification can be applied in further research, as it logically and comprehensively categorizes the online business models in practical use.

7 Summary

The primary goal of this research was twofold: first, to develop a comprehensive framework for categorizing and classifying online business models based on practical applications; and second, to determine whether there exists a mostly used dominant online business model among the online business models.

The study began by reviewing existing literature on online business model classifications, including seminal works by Timmers (1998), Weill-Vitale (2001), Rappa (2002), Laudon-Traver (2004), Eisenmann (2002) and Móricz (2009). These frameworks were compared to a new, own classification system developed on the basics of Malone et al.

(2006) framework for ordinary non-online business models adapted for online contexts. This own classification system categorizes online business models based on rights (ownership, usage, intermediation) and the extent of change to the involved assets.

To validate the own classification, the study conducted practical analyses using two datasets: the Alexa (2019) database, comprising 500 elements, and the iTunes Charts Apple 2019, comprising 200 elements. Both datasets were manually analyzed to categorize each identified online business model according to the proposed framework.

The findings confirmed that the own classification system is complete and right. Across both datasets, all identified online business models from literature reviews and practical analyses were successfully categorized within this framework. Specifically, the intellectual property landlord emerged as the most prevalent online business model, constituting over 50% of the models observed in the analyzed samples (Alexa: 56.2%, iTunes Charts: 70.3%).

The strength of this study lies in using the comprehensive own framework, which not only categorizes online business models but also offers practical insights for businesses aiming to adapt to the digital economy. By using empirical data, the own classification system demonstrates its applicability across diverse online platforms, making it a useable framework for both scientific and practical applications.

Despite these findings, the study has the following limitations allowing to conduct further research on this topic. It primarily relied on data from the most visited websites and popular mobile applications, potentially overlooking niche market companies (also called long tail businesses) not represented in these datasets. While this study relies on the Alexa (2019) and iTunes Charts Apple 2019 datasets, the discontinuation of Alexa and the fixed sampling date of June 2019 limit the generalizability of the findings. Future research could benefit from updated datasets and a broader scope to understand trends in online business models. Therefore it can be expanded by other datasets and long-tail businesses in its sample, exploring the applicability of the own classification system across different technologies than websites and mobile apps.

As the study confirmed, this research provides a robust framework for categorizing online business models and identifies the intellectual property landlord as the predominant online business model. The findings offer valuable insights for businesses entering online markets, emphasizing the strategic importance of certain online business models over others.

References

Acciarini, C., Cappa, F., Boccardelli, P., Oriani, R.: How can organizations leverage big data to innovate their business models? Syst. Liter. Rev. Tech. **123**(102713), 1–18 (2023)

Alexa: Alexa Top 500. http://www.alexa.com/topsites, adatbázis dátuma: 2019. június 6. (2019)

Anderson, C.: Hosszú farok. A végtelen választék átírja az üzlet szabályait. Budapest: HVG Books (2007)

Apple, 2019: iTunes Charts. https://www.apple.com/itunes/charts/free-apps/és, https://www.apple.com/itunes/charts/paid-apps/, date of database: 6 June 2019 (2019)

Atkinson, R.D., Ezell, S.J., Andes, S.M., Castro, D.D., Bennett, R.: The internet economy 25 years after .com. transforming commerce & life. The Inf. Technol. Innov. Found. (2010)

Benke, M.: A motiváció és tanulás kapcsolatának vizsgálata: fókuszban az önmeghatározás elmélet. [PhD dissertation] Pécs: University of Pécs Faculty of Business and Economics (2018)
Benke, M.: Tanuláselméletek és összehasonlításaik. Közép-Európai Közlemények **13**(3), 181–215 (2020)
Borgulya, I., Balogh, G., Jarjabka, Á.: Communication management in industrial clusters: an attempt to capture its contribution to the cluster's success. J. East Eur. Manage. Stud. **27**(3), 179–209 (2022). https://doi.org/10.5771/0949-6181-2022-2-179
Bresciani, S., Huarng, K.H., Malhotra, A., Ferraris, A.: Digital transformation as a springboard for product, process and business model innovation. J. Bus. Res. **128**(február), 204–210 (2021)
Brynjolfsson, E., Mcafee, A.: The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies. W.W. Norton & Company, Inc., New York, N.Y. (2014)
Caputo, A., Pizzi, S., Pellegrini, M.M., Dabić, M.: Digitalization and business models: where are we going? A science map of the field. J. Bus. Res. **123**, 489–501 (2021). https://doi.org/10.1016/j.jbusres.2020.09.053
Christensen, C.M., Raynor, M.E., McDonald, R.: What is disruptive innovation? Harv. Bus. Rev. **93**(12), 44–53 (2015)
Consumano, M.A., Yoffie, D.B., Gawer, A.: Future of platforms. MIT Sloan Manage. Rev. Spec. Issue Disrupt. **2020**, 61304 (2020)
Costello: How Many Apps Are in the App Store? (2020). https://www.lifewire.com/how-many-apps-in-app-store-2000252, Accessed 28 Feb 2024
Dajnoki, K., et al.: Characteristics of crisis management measures in the HR area during the pandemic in hungary – results of a countrywide survey of organizations. Acta Polytechnica Hungarica **20**(7), 193–210 (2023). https://doi.org/10.12700/APH.20.7.2023.7.11
Deloitte: Industry 4.0: Challenges and Solutions for the Digital Transformation and Use of Exponential Technologies (2014)
Deloitte.: Understanding talent, technology and transformation: Digital Disruption Index. November 2017 (2017)
Dumagan, J., Gill, G., Ingram, C.: Industry-level effects of information technology use on overall productivity. chapter 4. Digital Economy 2003. Economic & Statistics Administration, U.S. Dept. of Commerce (December 2003), pp. 45–60 (2003)
Eisenmann, T.R.: Internet Business Models. McGraw-Hill, Text and Cases (2002)
Fitzgerald, M., Kruschwitz, N.; Bonnet, D., Welch, M.: Embracing digital technology: a new strategic imperative. MIT Sloan Management Review. Research Report 2013. (2013)
Gong, C. Ribiere, V.: Developing a unified definition of digital transformation. Technovation **102**, 102217 (2021). https://doi.org/10.1016/j.technovation.2020.102217
Hartvig, Á.D., Madari, Z., Pap, Á., Wimmer, Á., Oroszné Csesznák, A.: A digitalizáció hatása a vállalati hozzáadott értékre Magyarországon. Közgazdasági Szemle **70**, 672–689 (2023)
Hornyák, M., Kruzslicz, F., Lányi, B.: A kis- és középvállalatok digitális transzformációja - az online jelenlét és a versenyképesség összefüggései. Közgazdasági Szemle **70**(május), 517–543 (2023)
Jarjabka, Á.: A stratégia fogalmának modern jelentéstartalma. Vezetéstudomány **32**(7–8), 25–32 (2001)
Jarjabka, Á., Sipos, N., Kuráth, G.: Quo vadis higher education? Post-pandemic success digital competencies of the higher educators – a Hungarian university case and actions. Hum. Soc. Sci. Commun. **11**(310), 1–11 (2024).https://doi.org/10.1057/s41599-024-02809-9
Kane, G.C., Doug, P., Phillips, A.N.: Strategy, not technology, driver digital transformation. MIT Sloan Management Review and Deloitte University Press, July 2015 (2015)
Katona, A., Birkner, Z., Németh, K., Péter, E.: Ipari digitalizációra való felkészülés eltérő méretű hazai cégeknél. Vezetéstudomány/Budapest Manage. Rev. **54**(6), 47–59 (2023)

Kaufman, I., Horton, C.:. Digital transformation: leveraging digital technology with core values to achieve sustainable business goals. European Financial Review, December 2014-January 2015, pp. 63–67 (2015)
Laudon, K.C., Traver, C.G.: E-commerce. Business. Technology. Society. 2nd Edition, Pearson, USA (2004)
Łęgowik-Świącik, S.: Creating the value network and converting the business model of technological companies. Sci. Papers Silesian Univ. Technol. Organ. Manage. Ser. **183**, 285–301 (2023)
Leipzig, T., et al.:. Initialising customer-orientated digital transformation in enterprises. Procedia Manuf. **2017**(8), 517–524 (2017)
Liu, D.-Y., Chen, S.-W., Chou, T.-C.: Resource fit in digital transformation: lessons learned from the CBC bank global e-banking project. Manag. Decis. **49**(10), 1728–1742 (2011)
Lone, S., Weldevreden, J.W.J.: European E-Commerce Report 2023. Amsterdam University of Applied Sciences & Ecommerce Europe, Amsterdam/Bruxelles (2023)
Malone, T.W., Yates, J., Benjamin, R.I.: Electronic markets and electronic hierarchies. Commun. ACM **30**(6), 484–497 (1987)
Malone T.W., et al.: Do Some Business Models Perform Better than Others? MIT Sloan School of Management, MIT Sloan Working Paper 4615–06, May 2006 (2006)
Móricz, P.: Élenjáró magyarországi internetes vállalkozások fejlődése az üzleti modell nézőpontjából. [PhD disseration] Budapest: Budapesti Corvinus University (2009)
Mursalzade, H., Molnár, L., Saraswatei, H.S.: Digitalization and value co-creation in the context of social entrepreneurship. Vezetéstudomány/Budapest Manage. Rev. **54**(11), 2–14 (2023)
Nemeslaki, A.; Urbán, Z., Trestyén, A.: Alapvető e-business modellek működése és magyarországi elterjedtségük. Vezetéstudomány/Budapest Manage. Rev. **39**(12), 4–15 (2008)
Nemeslaki, A.: Vállalati internetstratégia. Akadémiai Kiadó, Budapest (2012)
Nicholas, T.: Why Schumpeter was right: innovation, market power, and creative destruction in 1920s America. J. Econ. History **63**(4), 1023–1058 (2003)
Obermayer, N., Máhr, T., Banász, Z.: A digitalizáció társadalmi hatásainak vizsgálata hazai szakértők véleményének tükrében. Vezetéstudomány/Budapest Manage. Rev. **54**(6), 32–46 (2023)
Porter, M.E., Millar, V.E.: How information gives you competitive advantage. Harv. Bus. Rev. **63**(4), 2–13 (1985)
PwC:. Industry 4.0: Building the Digital Enterprise. 2016 Global Industry 4.0 Survey (2016)
Rappa, M.: Managing the Digital Enterprise - Business Models on the Web (2002)
Rojko, K.; Erman, N., Jelovac, D.: Impacts of the transformation to industry 4.0 in the manufacturing sector: the case of the U.S. Organizacija, **53**(4), 287–305 (2020). https://doi.org/10.2478/orga-2020-0019
Ruzsa, Cs.: "Z" generáció fő jellemzői és a várható munkerő-piaci kihívások. Közép-Európai Közlemények **11**(3), 149–157 (2018)
Ruzsa, Cs.: Industry 4.0 - expected technological impacts on companies. In: Simic, M.L., Crnkovic, B. (ed.) 8th International Scientific Symposium "Economy of Eastern Croatia - Vision and Growth, pp. 1344–1355 (2019)
Sahri, Z., Halim, Z.: Bibliometric analysis of online digital business for indigenous and ethnic community using VOSViewer. J. Islamic, Soc. Econ. Dev. **8**(57), 120–136 (2023)
Schmuck, R.: Factors affecting the online economy on a regional basis. Reg. Formation Dev. Stud. **3**(11), 217–224 (2013)
Schmuck, R.: Online üzleti modellek. [PhD dissertation]. Pécs: University of Pécs Faculty of Business and Economics (2015)
Schmuck, R.: Online üzleti modellek vizsgálata Magyarországon. Taylor: Gazdálkodás- és Szervezéstudományi Folyóirat: A Virtuális Intézet Közép-Európa Kutatására közleményei **9**(2), 197–202 (2017)

Schmuck, R.:. Vállalati stratégiák a virtuális térben. Taylor: Gazdálkodás- és Szervezéstudományi Folyóirat: A Virtuális Intézet Közép-Európa Kutatására közleményei **11**(1), 73–83 (2019)
Schmuck, R.: The definition of the business model concept. Közgazdász Fórum **23**(144), 39–58 (2021)
Schmuck, R.: The use of online business models. Procedia Manuf. **54**, 45–51 (2021)
Schmuck, R.: Formulating your business model: does it stand the elevator test? Academia Lett. **442**, 1–5 (2021)
Schmuck, R.: Comparison of the use of ICT resources of ISO 9001 certified SMEs and non-certified SMEs in Hungary. Int. J. Prod. Qual. Manage. **38**(4), 545–562 (2023). https://doi.org/10.1504/IJPQM.2021.10041082
Statista:. eCommerce- Worldwide. https://www.statista.com/outlook/emo/ecommerce/worldwide Accessed 28 February 2024 (2024)
Tapscott, D.: Rethinking Strategy in a Networked World (or Why Michael Porter is Wrong about the Internet). Strategy+Business, Issue 24, 2011 Third Quarter (2011)
Tekic, Z., Koroteev:. From disruptively digital to proudly analog: a holistic typology of digital transformation strategies. Bus. Horizons **62**(6), 683–693 (2019). https://doi.org/10.1016/j.bushor.2019.07.002
Timmers, P.: Business models for electronic markets. Electron. Mark. **8**(2), 3–8 (1998)
Weill, P., Vitale, M.: Place to Space - Migrating to eBusiness Models. Harvard Business School Press, Boston (2001)
Zaušková, A., Kusá, A., Kubovics, M., Ščepková, S., Urmínová, M.: Current state and prediction of the future of digitization as a part of industry 4.0. Serbian J. Manage. **17**(1), 111–123 (2022)

A Comprehensive Analysis of Low-Energy Adaptive Clustering in Underwater Acoustic Networks

N. Karthika(✉) and D. Kavitha

KG College of Arts and Science, Coimbatore, Tamil Nadu, India
karthika.n@kgcas.com, d.kavitha@kgisliim.ac.in

Abstract. This survey study investigates the junction of Machine Learning (ML), networking, and the Internet of Things (IoT) within the framework of Low-Latency, Low-Energy Adaptive Clustering Hierarchy (LL-LEACH) protocols for Underwater Acoustic Networks (UANs). The deployment of UANs has special difficulties because of the hostile underwater environment, limited energy resources, and need of real-time data transfer. LL-LEACH systems optimize energy economy, reduce latency, and improve network performance by means of adaptive clustering techniques enhanced by machine learning algorithms to address these challenges. Emphasizing their fundamental features, advantages, and limitations, this work explores and assesses many recently proposed LL-LEACH methods. This study addresses decreasing latency and possible computer-assisted solutions for LEACH protocols in underwater acoustic networks. It also addresses how machine learning techniques may enhance network management, data processing, and decision-making capabilities as well as how IoT devices might be added into UANs. By means of this extensive analysis, we want to provide insights on the state-of- the-art LL-LEACH systems in UANs and suggest future research routes at the intersection of machine learning, networking, and IoT for underwater communication systems.

Keywords: Adaptive Clustering · Low-Latency · Low-Energy Internet of Things · ML · Networking

1 Introduction

Demand for wireless sensor networks (WSNs) has grown as they have lately merged with innovative technologies. Apart from that, WSNs are now linked with new IoT applications that may join many devices online and exchange massive amounts of data [1–6]. The limited lifespan of the battery determines the number of sensor nodes in a WSN, therefore generating several problems and restrictions [7–9]. Therefore, the main objective constantly is to solve these problems by lowering energy use and raising lifespan. One of the most important and useful communication technologies for data transmission and reception, WSNs are rapidly acquiring momentum in the scientific and technological landscape of today [10–13]. A WSN gathers data using several small sensor nodes

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
S. Rajagopal et al. (Eds.): ASCIS 2024, CCIS 2428, pp. 411–421, 2025.
https://doi.org/10.1007/978-3-031-86302-8_27

spread throughout the surroundings. These sensors can detect and record environmental variables such as sound, light, humidity, temperature, and pressure because they comprise sensing, processing, and trans-receiver electrical components Evaluate long-term energy savings in underwater networks by establishing metrics for energy efficiency and network longevity, monitoring energy use and performance, and comparing with baseline data. Implement and adjust energy-saving measures, conduct regular reviews, and assess scalability and technological impacts to ensure sustained effectiveness [14–17].

Wireless Sensor Networks (WSNs) with minimal latency are required when real-time data transmission is crucial. This is especially true for intelligent infrastructure, healthcare monitoring, and industrial automation [18, 19]. Low latency means reducing the time between data production at sensor nodes and reception at its intended destination, such as a data sink or a control center. To ensure efficient and timely data transmission, attaining low latency necessitates enhancing numerous network components, including communication protocols, routing algorithms, and hardware capabilities [20, 21]. LL-LEACH protocols face security vulnerabilities such as unauthorized access, node impersonation, denial of service, data integrity issues, Sybil attacks, compromised cluster heads, and replay attacks. Mitigating these risks requires encryption, authentication, anomaly detection, and secure clustering techniques. In networking, "hierarchy protocols" refer to structuring and communicating scattered systems to improve efficiency while minimizing complexity [22, 23].

2 Background Study

Agarwal, K. et al. (2018) these writers simulated and investigated the working approach of the LEACH algorithm using MATLAB. Several important WSN characteristics, such as network lifespan, throughput, packet flow to BS, etc., were used to study this protocol. According to the results, choosing each node at random to become CH reduces energy usage.

Al-Shalabi, M. et al. (2018) state that when designing WSNs, it is critical to consider the routing protocol architecture. Numerous methods existed to categorize these protocols, but clustering and hierarchical routing stood out. Across the many benefits of clustering routing are improved network scalability, more equitable allocation of energy across nodes, and less strain on individual nodes (Table 1).

Alvarado, G. et al. (2017) Using the energy distribution of the mesh, the author presented the LEACH protocol and the ONCH algorithm as a practical solution to sensor network clustering wherein she found and selected a node with a high energy density as the cluster leader. Among the factors taken into account were size, package length, energy properties, and data aggregation.

Asad, M. et al. (2019) The author presents in this work a novel approach to WSNs: multi-adaptive clustering (MAC) with Internet of Things (IoT) support for freeing sensor node resources. Based on a novel hybrid network architecture combining homogeneous and heterogeneous data with distributed and centralized nodes mixed with each other with the base station (BS) at its center, the recommended routing technique was proposed.

Dong, W. et al. (2021) these authors research investigated the clustering problem for direct-to--- device (D2D) multicast communication in settings of short video applications. The author proposed a CHs control approach based on deep reinforcement learning

Table 1. Analysis of Low-Energy Adaptive Clustering

Author	Year	Methodology	Advantage	Limitation
Agarwal, K. et al	*2018*	*Simulation and analysis using MATLAB*	*Comprehensive study on the (LEACH) protocol*	*Limited discussion on real-world implementation challenges*
Al-Shalabi, M. et al	*2018*	*LEACH protocol*	*Provides insights into various LEACH variants, identifying their strengths and weaknesses*	*Limited experimental validation of proposed variants*
Ghazy, A. S. et al	*2023*	*Development of LL-LEACH protocols for UANs*	*Addresses low-latency and low-energy requirements for underwater acoustic networks*	*Limited discussion on scalability and deployment challenges in real-world underwater environments*
Jamal, T. & Butt, S. A	*2017*	*Enhancement of Leach Protocol for military security operations*	*Provides improvements for security in military operations using Leach Protocol*	*Limited discussion on energy consumption optimization and scalability for large-scale deployments*

to increase the distribution efficiency and user enjoyment of the system. To introduce cluster formation and the CHs selecting approach, the author closely investigated the physical interest distance and the distance-power weight.

Hassan, H. M., & Priyadarshini, R. (2021) the limitations of Wireless Sensor Networks, particularly their memory, speed, and energy consumption, have been the subject of much research and development. Many efforts have been made to address the inadequacies of the LEACH protocol, which is considered one of the most prominent strategies.

Jamal, T., & Butt, S. A. (2017) Although LEACH was a routing protocol that aims to improve the energy economy, it has several limitations that make it unsuitable for military security operations. Issues with CH selection, time distribution, CM distance from CH, data aggregation, and base station transmission were the root of these problems. The LEACH protocol must be modified to be used for military security purposes.

Kang, J. et al. (2020) Using ML and meta-heuristics, the author developed approaches to reduce energy consumption inside the WSN framework while considering environmental issues. Furthermore, the author demonstrated the superiority of the suggested framework via many experiments.

Liu, F., & Chang, Y. (2019) these authors research presents the Energy-Aware Adaptive Kernel Density Estimation Algorithm (EAKDE), a novel uneven clustering method for WSNs, to ensure that none of the sensor nodes are overworked. EAKDE consisted of four main elements: cluster creation, cluster radius computation, cluster head election, and routing method.

Mittal, M. et al. (2020) The author of this paper addresses using neural networks and machine learning to investigate security and energy concerns in WSNs. The author started this effort concentrating on energy assessment. This was accomplished by including into the LEACH process for CH head selection the Levenberg-Marquardt Neural Network (LMNN) technique, a fast learning mechanism. One after the other shortest path finding techniques including Dijkstra, Bellman-Ford, and BFS were used to enhance the energy parameter. The method chooses the quickest route for every CH sensor node to go to the central station.

Sefati, S. et al. (2022) Combining LEACH and BA algorithms, the suggested method lowers power consumption and increases the lifetime of the network. Developing Flying ad-hoc network (FANET) load balancing and clutter selection techniques first focused mostly on reducing UAV power consumption. Therefore, it was crucial to provide a system for choosing an appropriate and energy-efficient CH to guarantee an equal energy consumption throughout the network's nodes.

Wibowo, F. et al. (2020) these authors research suggests a novel approach to find the ideal cluster combining cluster identification with adaptive ML clustering. The elbow technique was used for detecting as this would drive energy usage more. ML may also be used to choose cluster heads as it is a strong technique of increasing network lifecycle performance in wireless sensor networks.

Mei, H. et al. (2024) Even if the network's complexity is reduced, performance suffers greatly due to congestion that happens when the intra-cluster load exceeds the upper limit of the intra-cluster information transmission capacity. To prevent congestion, we suggest LLCP. These authors research goal is to minimize network complexity and prevent congestion as much as possible by maximizing the number of cluster members according to node load and location. In order to find the sweet spot between having too many nodes in a cluster and too few, we first provide an optimization method that takes node load and location into account.

Lakshmi, K. et al. (2024) a key component of UAWSN is energy optimization algorithms, which help lower power consumption, boost network lifespan and scalability. Specifically, this study suggests using the DEEC clustering method for UAWSN topology management. Considerations such as network design, Cluster Head (CH) node selection, data packet transfer from CH to sink, number of live nodes, and network lifespan are all part of the DEEC clustering method. The simulation results show that compared to the current methods, the DEEC clustering algorithm is better.

Lowes, G. J., & Neasham, J. (2024) In order to detect and localize vessel activity, this research set out to design a low-energy device that is readily deployable at sea. The plan all along was to build a single node and use it in a US Navy system that uses triangulation based on auditory communication and information exchange between nodes (Fig. 1 and Table 2).

3 Existing Methodology

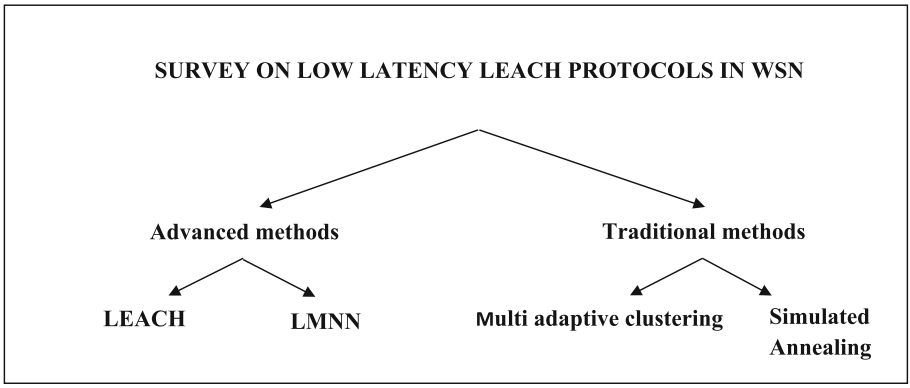


Fig. 1. Advanced and Traditional Methods

3.1 Simulated Annealing

SA's iterative method analyzes adjacent options based on an objective function considering characteristics including data aggregation efficiency, inter-cluster communication latency, and cluster head energy usage. Using probabilistic acceptance criteria and a cooling schedule, SA finds a balance between exploring and exploiting the complex and dynamic underwater acoustic environment, all while improving network performance and energy efficiency (Fig. 2).

3.2 Leach

The Low-Energy Adaptive Clustering Hierarchy (LEACH) systems provide low-energy, low-latency solutions for underwater acoustic networks in major part. LEACH is designed for the rigorous underwater environment; hence it optimizes energy consumption and enables real-time communication. It applies strategies of adaptive clustering. Dynamic cluster head rotation and node clustering allow the constant energy consumption distribution around the network. By allowing data to be collected locally and forwarded to cluster leaders, this hierarchical technique lets sink nodes convey the aggregated data, hence lowering latency (Fig. 3).

Table 2. Comparative Analysis of Clustering Algorithms Merits and Demerits

Algorithm	Merits	Demerits
LEACH *(Agarwal, K. et al. (2018))*	*LEACH optimizes network lifetime by spreading the job of cluster head across sensor nodes in a cycle, therefore lowering power consumption per node*	*The probabilistic approach of LEACH for head selection might cause unequal energy depletion and possible network hotspots*
LMNN *(Mittal, M. et al. (2020))*	*LMNN aims to learn a distance measure to maximize the margin separating several class data points. By motivating more cluster separation, this enhanced discrimination helps clustering algorithms handle complex data distributions or situations where clusters cross-over*	*Training LMNN might be computationally costly in feature spaces spanning many dimensions or huge datasets*
Multi-adaptive clustering *(Asad, M. et al. (2019))*	*Multi-adaptive clustering's make data distribution, cluster dynamics, and network features easier to change. Given its flexibility, clustering in dynamic environments might be more effective and strong*	*Multi-adaptive clustering systems might be more complex than ordinary clustering systems as they create dynamic changes by means of adaptive techniques. Application and optimization of such complexity might call for more understanding and computational tools*
Simulated Annealing *(Kang, J. et al. (2020))*	*The capacity of simulated annealing to avoid local optima and traverse a large search space distinguishes it. This feature makes it appropriate for addressing challenging non-convex optimization issues such cluster development in low-energy adaptive clustering and provide high-quality answers*	*Simulated annealing may be computationally costly for large search regions or multidimensional optimization problems. The approach restricts computation by using probabilistic updates and analyzing several viable answers*

3.3 Multi-adaptive Clustering

Using multi-adaptive clustering, we derive adaptive clustering hierarchy protocols for low-energy, low-latency UWAN networks. These systems ensure real-time communication and increase energy economy by constantly changing clustering methods based

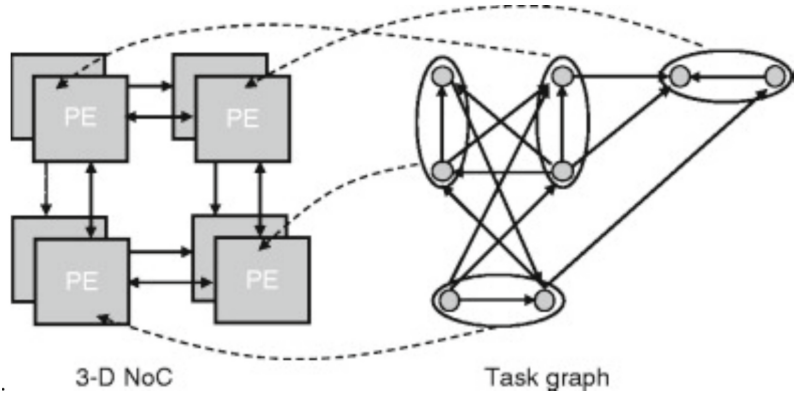


Fig. 2. Simulated Annealing architecture

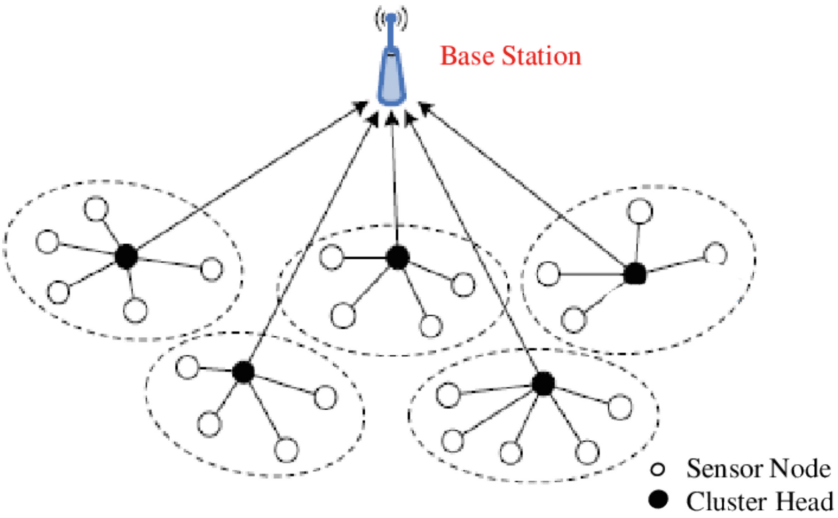


Fig. 3. LEACH protocol architecture

on several factors like network design, data traffic patterns, and node energy levels. By allowing network nodes automatically build and rearrange clusters, multi-adaptive clustering techniques reduce latency, hence boosting data aggregation and transmission. Since underwater acoustic networks may vary their operating needs and environmental surroundings, they rely on multi-adaptive clustering to provide energy-efficient and responsive communication (Fig. 4).

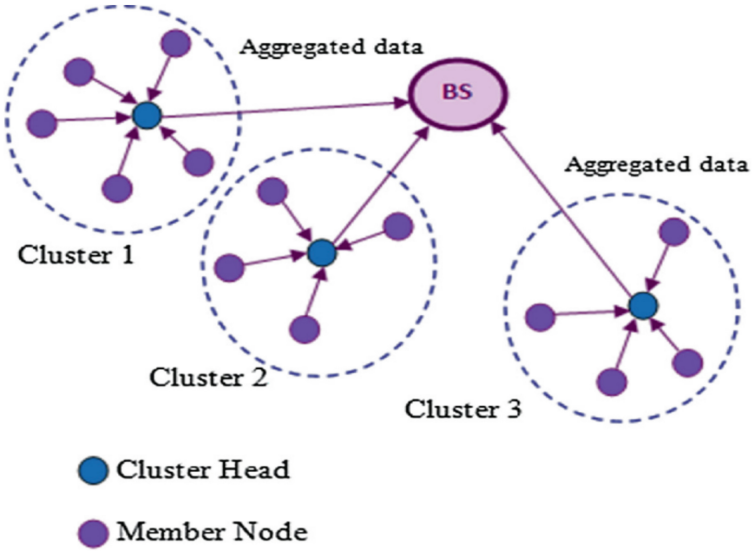


Fig. 4. Multi Adaptive clustering architecture

4 Discussion

Within the context of UWANs, this review research investigated potential methods for combining ML, networking, and the Internet of Things (IoT) to construct LL-LEACH protocols. LL-LEACH systems use adaptive clustering approaches reinforced by machine learning to optimize energy economy, reduce latency, and improve network performance in real-time data transmission situations. These methods are used in response to specific challenges like limited energy and the harsh underwater surroundings. This paper clarifies the state of the art in LL-LEACH systems by means of thirty research papers, therefore underlining their characteristics, advantages, and disadvantages.

The performance of adaptive clustering techniques in underwater acoustic networks is significantly influenced by varying environmental conditions such as water salinity, temperature, depth, and turbidity. Water salinity affects signal propagation and absorption, necessitating dynamic adjustments in clustering parameters to maintain effective communication. Temperature fluctuations impact sensor accuracy and signal speed, requiring compensation mechanisms to ensure precision. Depth affects signal attenuation and travel time, demanding depth-aware clustering algorithms. Turbidity reduces signal clarity, requiring noise filtering and enhanced signal processing. To address these challenges, adaptive clustering techniques can integrate real-time environmental sensing, dynamically adjust parameters, and employ robust communication protocols to maintain network reliability and efficiency.

Integrating machine learning into LL-LEACH protocols significantly increases computational overhead. Supervised learning adds high training costs but improves adaptability, unsupervised learning involves moderate training and lower inference costs, reinforcement learning has very high training overhead with moderate inference costs,

and hybrid approaches combine multiple methods, leading to high training and variable inference costs. Balancing these costs with performance benefits is crucial for effective deployment in underwater acoustic networks. Integrating Internet of Things (IoT) technology with Low-Latency, Low-Energy Adaptive Clustering Hierarchy (LL-LEACH) protocols in underwater acoustic networks enhances network efficiency by incorporating IoT devices for real-time data collection and analysis. IoT devices, such as sensors and actuators, provide critical environmental and operational data that enable adaptive clustering decisions, optimizing cluster head selection and routing to improve data accuracy, reduce latency, and extend network lifetime. This integration allows LL-LEACH protocols to dynamically adjust to changing conditions and operational needs, leveraging machine learning for better data analysis and network management, resulting in more intelligent and responsive underwater networks.

5 Conclusion

This survey paper explores and evaluates the intersection of ML, networking, and the Internet of Things (IoT) within the LL-LEACH systems for UANs. Supported by machine learning, adaptive clustering's techniques seek to address problems like improving network performance, reducing latency, and optimizing energy economy. These solutions aim to solve two issues: limited energy and dangerous underwater settings. We also discussed how Internet of Things devices may improve decision-making and network management. This detailed examination defines the present state of the art in LL-LEACH systems and leads towards additional research to develop underwater communication systems. Future research on LL-LEACH systems for underwater acoustic networks (UANs) should focus on several key areas. First, the integration of advanced machine learning techniques, such as deep learning and reinforcement learning, could further enhance adaptive clustering and energy optimization. This would allow UANs to autonomously adjust to more complex and dynamic underwater environments. Second, there is potential for improving real-time data transmission by exploring low-latency communication protocols that maintain energy efficiency without sacrificing network performance.

References

1. Abu-Baker, A., Alshamali, A., Shawaheen, Y.: Energy-efficient cluster-based wireless sensor networks using adaptive modulation: performance analysis. IEEE Access **9**, 141766–141777 (2021)
2. Agarwal, K., Agarwal, K., Muruganandam, K.: Low energy adaptive clustering hierarchy (LEACH) protocol: simulation and analysis using MATLAB. In: 2018 International Conference on Computing, Power and Communication Technologies (GUCON) (2018). https://doi.org/10.1109/gucon.2018.8674988
3. Al-Shalabi, M., Anbar, M., Wan, T.C., Khasawneh, A.: Variants of the low-energy adaptive clustering hierarchy protocol: survey, issues and challenges. Electronics **7**(8), 136 (2018)
4. Alvarado, G., Bosquez, C., Palacios, F., Córdoba, L.: Low-energy adaptive clustering hierarchy protocol and optimal number of cluster head algorithm in a randomized wireless sensor

network deployment. In: 2017 International Conference on Electrical, Electronics, Communication, Computer, and Optimization Techniques (ICEECCOT), pp. 1–4. IEEE, December 2017
5. Asad, M., Aslam, M., Nianmin, Y., Ayoub, N., Qureshi, K.I., Munir, E.U.: IoT enabled adaptive clustering based energy efficient routing protocol for wireless sensor networks. Int. J. Ad Hoc Ubiquitous Comput. **32**(2), 133–145 (2019)
6. Behera, T.M., Mohapatra, S.K., Samal, U.C., Khan, M.S., Daneshmand, M., Gandomi, A.H.: Residual energy-based cluster-head selection in WSNs for IoT application. IEEE Internet Things J. **6**(3), 5132–5139 (2019)
7. Cui, Z., Cao, Y., Cai, X., Cai, J., Chen, J.: Optimal LEACH protocol with modified bat algorithm for big data sensing systems in Internet of Things. J. Parall. Distrib. Comput. **132**, 217–229 (2019)
8. Dong, W., Li, Z., Chen, X.: Deep reinforcement learning-based adaptive clustering approach in short video sharing through D2D communication. In: 2021 IEEE International Conference on Parallel & Distributed Processing with Applications, Big Data & Cloud Computing, Sustainable Computing & Communications, Social Computing & Networking (ISPA/BDCloud/SocialCom/SustainCom), pp. 745–752. IEEE, September 2021
9. El Khediri, S., Khan, R.U., Nasri, N., Kachouri, A.: MW-LEACH: low energy adaptive clustering hierarchy approach for WSN. IET Wirel. Sens. Syst. **10**(3), 126–129 (2020)
10. El Khediri, S., Khan, R.U., Nasri, N., Kachouri, A.: Energy efficient adaptive clustering hierarchy approach for wireless sensor networks. Int. J. Electron. (2020). https://doi.org/10.1080/00207217.2020.1756454
11. Ghazy, A.S., Kaddoum, G., Singh, S.: Low-latency low-energy adaptive clustering hierarchy protocols for underwater acoustic networks. IEEE Access (2023)
12. Hassan, H.M., Priyadarshini, R.: Optimization techniques for IoT using adaptive clustering. In: 2021 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS) (2021). https://doi.org/10.1109/icccis51004.2021.9397128
13. Ilyas, M., et al.: Trust-based energy-efficient routing protocol for internet of things–based sensor networks. Int. J. Distrib. Sens. Netw. **16**(10), 1550147720964358 (2020)
14. Jamal, T., Butt, S.A.: Low-energy adaptive clustering hierarchy (LEACH) enhancement for military security operations. Proc. J. Basic Appl. Sci. Res. ISSN, 2090–4304 (2017)
15. Jurado-Lasso, F.F., Jurado, J.F., Fafoutis, X.: A centralized reinforcement learning framework for adaptive clustering with low control overhead in IoT networks (2024). arXiv preprint arXiv:2401.15767
16. Kang, J., Kim, J., Kim, M., Sohn, M.: Machine learning-based energy-saving framework for environmental states-adaptive wireless sensor network. IEEE Access **8**, 69359–69367 (2020)
17. Li, W., Wang, Y., Sun, Y., Mao, J.: Research on Low-energy adaptive clustering hierarchy protocol based on multi-objective coupling algorithm. KSII Trans. Internet Inf. Syst. (TIIS) **14**(4), 1437–1459 (2020)
18. Liu, D., Liang, C., Mo, H., Chen, X., Kong, D., Chen, P.: LEACH-D: a low-energy, low-delay data transmission method for industrial internet of things wireless sensors. Internet Things Cyber-Phys. Syst. **4**, 129–137 (2024)
19. Liu, F., Chang, Y.: An energy aware adaptive kernel density estimation approach to unequal clustering in wireless sensor networks. IEEE Access **7**, 40569–40580 (2019)
20. Mittal, M., Iwendi, C., Khan, S., Rehman Javed, A.: Analysis of security and energy efficiency for shortest route discovery in low-energy adaptive clustering hierarchy protocol using Levenberg-Marquardt neural network and gated recurrent unit for intrusion detection system. Trans. Emerg. Telecommun. Tech. (2020). https://doi.org/10.1002/ett.3997

21. Neji, W., Othman, S.B., Sakli, H.: T-leach: threshold sensitive low energy adaptive clustering hierarchy for wireless sensor networks. In: 2020 20th International Conference on Sciences and Techniques of Automatic Control and Computer Engineering (STA), pp. 336–342. IEEE, December 2020
22. Ngangbam, R., Hossain, A., Shukla, A.: Improved low energy adaptive clustering hierarchy and its optimum cluster head selection. Int. J. Electron. 1–13 (2019). https://doi.org/10.1080/00207217.2019.1661023
23. Perera, P., Patel, V.M.: Efficient and low latency detection of intruders in mobile active authentication. IEEE Trans. Inf. Forensics Secur. **13**(6), 1392–1405 (2017)
24. Saha, A.K., Khan, M.S.U., Rahman, M.M.: Region based low energy adaptive clustering hierarchy (R-LEACH) protocol: prolonging the lifetime of WSN. In: Proceedings of the 6th International Conference on Networking, Systems and Security, pp. 127–132, December 2019
25. Sahib, H.A., Kurnaz, S., Mohammed, A.H., Sahib, Z.A.: Network of low energy adaptive clustering protocols. In: 2020 4th International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT), pp. 1–5. IEEE, October 2020
26. Sefati, S.S., Halunga, S., Farkhady, R.Z.: Cluster selection for load balancing in flying ad hoc networks using an optimal low-energy adaptive clustering hierarchy based on optimization approach. Aircr. Eng. Aerosp. Technol. **94**(8), 1344–1356 (2022)
27. Wang, H., Huang, X., Wu, Y.: GD3N: adaptive clustering-based detection of selective forwarding attacks in WSNs under variable harsh environments. Inf. Sci. 120375 (2024)
28. Wibowo, F.W., Sediyono, E., Purnomo, H.D.: An adaptive framework using machine learning in wireless sensor network. In: 2020 IEEE International Conference on Communication, Networks and Satellite (Comnetsat), pp. 234–239. IEEE, December 2020
29. Zamry, N.M., Zainal, A., Rassam, M.A.: LEACH-CR: Energy saving hierarchical network protocol based on low-energy adaptive clustering hierarchy for wireless sensor networks. In: 2021 3rd International Cyber Resilience Conference (CRC), pp. 1–6. IEEE, January 2021
30. Mei, H., et al.: Node load and location-based clustering protocol for underwater acoustic sensor networks. J. Marine Sci. Eng. **12**(6), 982 (2024)
31. Tian, K., Zhou, C., Zhang, J.: Improved LEACH protocol based on underwater energy propagation model, parallel transmission, and replication computing for underwater acoustic sensor networks. Sensors **24**(2), 556 (2024)
32. Lakshmi, K., Hossain, A., Mummaneni, K.: A novel 3d distributed energy efficient clustering algorithm for underwater acoustic wireless sensor network (2024)
33. Lowes, G.J., Neasham, J.: PADAL-passive acoustic detection and localisation: low energy underwater wireless vessel tracking network. Comput. Netw. Netw. **241**, 110216 (2024)

Research Perspectives on Load Balancing Strategies in Serverless Computing

Yashwanth Balan[1(✉)] and R. Arokia Paul Rajan[2]

[1] Christ University, Bangalore 560029, India
yashwanth.balan@res.christuniversity.in
[2] Christ University, Bangalore, India

Abstract. Serverless computing, a groundbreaking trend in cloud computing, has transformed how applications are deployed and managed by abstracting the infrastructure layer. Serverless computing enables developers to concentrate exclusively on their code while cloud providers care for server provisioning, maintenance, and scaling. Services like AWS Lambda, Google Cloud Functions, and Azure Functions exemplify this model, offering s ubstantial advantages in terms of reduced operational complexity and cost. However, one persistent challenge in this domain is load balancing. Effective load balancing in serverless computing ensures efficient resource utilization, optimal performance, and cost-effectiveness. Unlike traditional load balancing, which typically relies on long-lived server instances, load balancing in serverless environments must accommodate the stateless and ephemeral nature of serverless functions. Traditional techniques are not directly applicable because serverless architectures functions that are instantiated on-demand in response to incoming requests. This paper surveys various strategies and approaches developed to address the unique load balancing challenges in serverless computing, providing a comprehensive overview of the current state of research and practice. The paper extends further research on serverless computing by analyzing the survey papers. The paper highly focuses the research areas in the field of edge computing, hybrid cloud models and distributed load balancing for the future usage.

Keywords: Cloud computing · Serverless computing · FaaS · Load balancing

1 Introduction

In the modern era of digital technology, cloud computing has emerged as an innovation [1], offering organizations and individuals unprecedented flexibility, scalability, and efficiency in managing their computing resources. At its essence, cloud computing refers to delivering computing services – including servers, storage, databases, networking, software, and more over the Internet on a pay-as-you-go basis [2].Unlike traditional computing models that rely on physical hardware and on-premises data centers, cloud computing leverages the power of remote servers hosted in data centers operated by cloud providers.The function as service approach is adapted in the serverless computing

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
S. Rajagopal et al. (Eds.): ASCIS 2024, CCIS 2428, pp. 422–432, 2025.
https://doi.org/10.1007/978-3-031-86302-8_28

to manage the code and use the resources. While this paradigm offers scalability, cost savings, and ease of deployment, it introduces challenges in load balancing. Efficient load balancing is crucial to handle varying. Workloads, ensure high availability, and optimize resource usage. This paper reviews existing load balancing techniques in serverless computing, examining their strengths and limitations. Several primary load balancing techniques are explored as follows: Round Robin, Least Connections, Resource-Based Load Balancing, Latency-Based Load Balancing, and Cost-Aware Load Balancing. Round Robin is a straightforward method that distributes requests sequentially across available function instances. The application is easy to implement and recognizable. When there is a deviation in the workload round robin will not be accountable for the characteristics of the situation. At this situation the least connection will be directing the request to the active connections with the help of these two methods the load will be balanced equally.to improve the performance and resource utilization in the long-lived connections these two methods will be the pioneer to make it successful. To monitor the CPU and its memory and related metrics, resource utilization is based on the load balancing which focuses on resource based.

Resource based techniques will be the proper method to rectify the error and increase reliability and performance [1]. In the geographically distributed situations, the latency-based load balancing algorithms provides an impact to reduce the latency for the crucial times. To overcome the huge resource cost and optimize the distribution based on the request is available the coast aware load balancing makes different approach. To maintain the performance of the resources the cost aware load balancing algorithms plays a vital role [2]. The paper completely focuses on the load balancing techniques which applies son the serverless computing.The scaling and other resource management with resource utilization demand with the load balancing techniques also focuses in the paper [3]. The modern technologies which uses AWS lamda and google cloud and Azure functions gives the control in the load balancing mechanism for the serverless computing for a reliable environment to adapt the crucial situations.

The load balancing can be utilized using the event trigger mechanism with the cold start, this methos will be abide by the Azure functions to get a proper result [4]. This paper completely focuses on the different mechanism to enhance the knowledge of future reference in the serverless computing which applies the load balancing efficiently. The artificial intelligence and machine learning mechanism plays a vital role in the area now a days. Using these two technologies in the load balancing for serverless computing will be able to monitor the analytics development and reduce the latency and cold start. The technology helpful to take a proper decision on multi cloud load balancing and in the multi cloud deployments [5]. It concludes the correct scenario of load balancing technology using in serverless computing lead to the impact of performance and utilization of the resources and its cost can also be in higher. By using the artificial intelligence and machine learning can be an added advantage in the future to enhance the ground work for achieving new goals in the load balancing area to apply on serverless computing [6]. The major difference between the load balancing techniques which uses in serverless architecture and the traditional load balancing is mainly coined by the infrastructure in the serverless and traditional. The second difference focuses on the autoscaling approach in serverless architecture. Latency and cost also make a difference in the serverless

architecture as compared to the traditional load balancing technique. Resource allocation and monitoring of the performance also have a huge impact in the serverless architecture as compared to the traditional load balancing (Fig. 1).

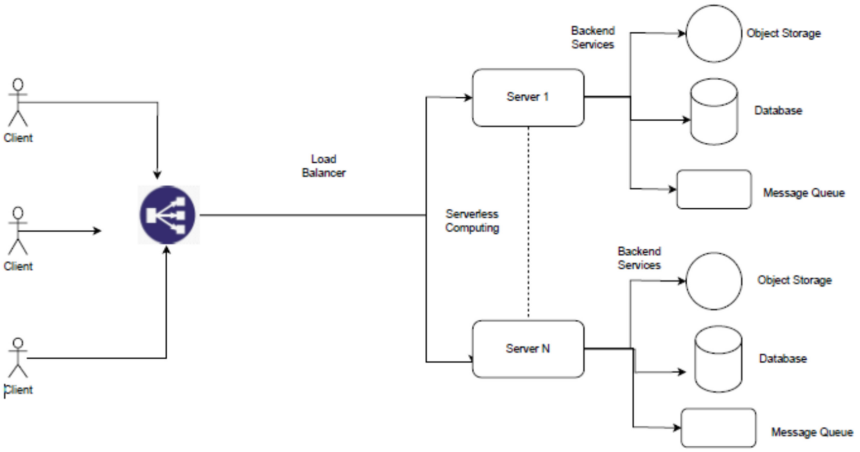


Fig. 1. Architectural diagram of serverless computing using load balancing

1.1 Serverless Computing

The technology of serverless computing relies on the deployed and managed applications which related to the infrastructure of servers. The serverless computing allows the users of completely concentrate on the programming part rather than the difficulties of relying on the scaling and provisioning and the back-end resource management. There are different applications which made a remarkable improvement in the area of serverless computing which includes Google cloud, Azure and AWS lamda [7]. The main advantage of using serverless computing it has the ability to scale automatically according to the situations. The serverless computing allows to automatically scale in the situations when there is a huge traffic and it will adjust the resources and increase the load and ensure the responses based on the demand without the human intervention.

There are situations the elasticity become very crucial when the work load becomes very high in nature. These situation demands the perfect load balancer to accommodate the large number of requests which rises and distribute among the servers to overcome the crucial elasticity. In the serverless computing the load balancing which applies effectively makes a huge difference to manage and resource identification of the different servers. Through effective load balancer serverless can be working properly and the business can be improved in large manner. The load balancing in serverless computing ensure the robustness and scalability of the applications. The modern application demands the serverless computing with the effective load balancer to manage the resource in an effective manner [8].

1.2 Load Balancing

The number of requests on an online application is increasing day by day. The greater number of requests which receives on the server it should have the capabilities of ensuring all the request are responded in a proper manner the servers are not overwhelmed. To make e situation in a stabled manner and the servers executes the request there should be a load balancer to control it. The load balancer plays vital role in managing the request dynamically which arrives and distributed among the servers. The efficiency of the load balancer describes how well each request is responded.

The traditional load balancers work with servers which has a fixed servers for each of the request to take care. But the serverless load balancing approach the request in a different with the help of latency throughout and managing the cost. The request arrives in a particular manner and the load balancer should identify and evaluates give the proper information of which function should be processing the particular request. The load balancer makes an impact in the resource utilization with the help of managing the performances to the peak. The number of requests arrives on the load balancer and there can be a big traffic but adaptability of the load balancer will ensure all the request are considered in balanced response time and the servers are available for the next request [9]. Serverless platforms like AWS Lambda, Azure Functions, and Google Cloud Functions inherently support this dynamic load balancing. They automatically manage the creation and termination of function instances based on demand, abstracting this complexity from the developer. By leveraging these capabilities, organizations can build highly scalable and resilient applications without the overhead of managing the underlying infrastructure. Thus, load balancing in serverless computing not only enhances efficiency but simplifies operational management, allowing developers to focus more on innovation and less on infrastructure [9].

2 Related Works

This section reviews previous studies that have conducted literature reviews on the application of serverless computing and cloud computing to machine learning. Serverless computing is experiencing a surge in popularity. Wen et al. [11] conducted a mining and analysis of serverless-related questions on Stack Overflow, revealing a growing interest in this topic. Their study also identified several challenges that developers face when using serverless computing. These challenges include issues related to programming language support, database connectivity, resource configuration, and security concerns Several studies have investigated the challenges of serverless computing. For instance, Khatri et al. [10] reviewed potential bottlenecks and measured the performance of serverless computing, The scenarios which controls the peak and spike with a possibility of limitations. They high-lighted the difficulties in testing and performance measurement of serverless applications and explored how machine learning can monitor and predict performance. The research paper gives the vision of how machine learning can give a good impact to overcome the difficulties of test the applications on serverless computing to ensure the performance. Hassan et al. [13] conducted a survey of 275 papers to find the challenges in the serverless computing to give a future direction of using the serverless for other applications. Wu et al. [11] recommends the data scientist focus on

the practical side of the serverless computing for the scalability and to improve making a cost effective serverless platform. Scheuner and Leitner [12] has analyzed 112 studies to state that serverless computing will easily adapt as function as a service for all the applications. The studies states that the reproducibility plays huge impact on the serverless computing. Sadaqat et al. [13] also conducted a multivocal literature review to outline the core components, benefits, and challenges of serverless computing. They identified common challenges such as vendor lock-in, the need for skilled workers, testing complexity, and monitoring. Their work also discussed the anticipated evolution of serverless computing, including its adoption by market growth project in the companies.

32 papers researched by Tiabi et al. [14] to conclude a multi vocal survey paper and selected 24 papers among them which focuses on the serverless computing. He has organized the paper pattern into different categories with event management and management availabilities for the serverless computing communication. It categorized these patterns into orchestration, aggregation, event management, availability, communication, and authorization. They noted that patterns could vary depending on the serverless provider; for example, AWS Lambda adapted its queue service (SQS) to enable FIFO messages, whereas FIFO messages still require manual management in Azure. Their work serves as a pattern catalog, offering valuable insights for practitioners and researchers in serverless computing.

3 Dynamic Load Balancing Principles

3.1 Round Robin

Round Robin is a straightforward algorithm that distributes requests sequentially across available instances. While simple to implement, it does not consider the current load or performance characteristics of instances, which can lead to suboptimal performance under varying workload scenarios with long-lived connections [15]. AI and ML are revolutionary for enhancing load balancing, conventional methods such as Round Robin and Least Connections remain essential as foundational strategies. In serverless computing, these traditional techniques can be incorporated or adapted into more sophisticated AI/ML-based models, often serving as benchmarks or fallback options. Here's how they could be applied to the research focus on Intelligent Load Balancing Using Reinforcement Learning for Serverless Computing.

3.2 Least Connection

The Least Connections algorithm routes incoming requests to the instance with the fewest active connections. This method helps to balance the load more evenly than Round Robin, particularly in scenarios with long-lived connections [16]. Round Robin and Least Connections are well-established and effective load balancing techniques. In the context of serverless computing, they can function as backup strategies or be incorporated into more advanced AI/ML-driven systems. These methods help distribute traffic efficiently, avoid overloading, and offer a reference point for comparing performance in reinforcement learning-based optimization models.

3.3 Resource Based Load Balancing

Resource-based load balancing monitors the CPU, memory, and other resource metrics of instances to make informed distribution decisions. This approach aims to prevent overloading any single instance, enhancing performance and reliability [17].

3.4 Latency Based Load Balancing

Latency-based load balancing directs traffic to instances that respond the fastest. This technique is particularly useful in geographically distributed environments where minimizing latency is critical for user experience [18].

3.5 Cost Aware Load Balancing

Cost-aware load balancing optimizes the distribution of requests based on the cost of resource usage. Serverless platforms typically charge based on the number of requests and execution time, so balancing load to minimize cost while maintaining performance is crucial [22].

3.6 Application of Artificial Intelligence and Machine Learning in Serverless Computing

AI and ML play a pivotal role in crafting flexible, efficient, and scalable load-balancing strategies tailored for serverless computing. By utilizing techniques such as reinforcement learning, predictive modeling, anomaly detection, and deep learning, we can greatly enhance resource distribution, reduce operational costs, and optimize system performance. These approaches empower serverless environments to adapt dynamically to fluctuating workloads and traffic patterns, making them ideal for modern cloud infrastructures. In essence, AI and ML are key to unlocking smarter, more responsive resource management, ensuring both cost-effectiveness and high performance in serverless architectures. The areas covered in the AI-ML of serverless computing for intelligence are reinforcement learning for dynamic decision making, predictive analysis for traffic forecasting, Anomaly detection and auto healing optimization with deep learning, Auto tuning and load balancing parameters, intelligent function placement. These areas provide the more value to the intelligent load balancing in serverless computing

4 Challenges in Load Balancing for Serverless Computing

Serverless computing has emerged as a revolutionary paradigm in cloud computing, allowing developers to focus on code execution without managing the underlying infrastructure. Despite its advantages, serverless computing introduces unique challenges in load balancing due to its dynamic and ephemeral nature. This paper explores these challenges, including cold starts, state management, latency variability, and cost optimization, and discusses potential solutions and future research directions [19]. Serverless computing platforms, such as AWS Lambda, Azure Functions, and Google Cloud Functions, offer scalable and cost-efficient solutions for running applications. The abstraction

of server management and automatic scaling are key benefits. However, the stateless and ephemeral nature of serverless functions presents distinct challenges for load balancing. Effective load balancing is crucial for ensuring high performance and reliability in serverless environments.

4.1 Cold Start

One significant challenge is the cold start problem. A cold start occurs when a new function instance is initialized in response to an incoming request, resulting in higher latency compared to warm instances. Cold starts can significantly impact user experience, especially for latency-sensitive applications. Load balancers must manage cold starts by predicting demand patterns and pre-warming function in-stances, though this approach can increase operational costs [21].

4.2 State Management

Serverless functions are inherently stateless, which complicates state management. Maintaining state across multiple function invocations requires external storage systems, adding latency and complexity. Load balancers must efficiently route requests while considering the state storage mechanism to ensure consistency and performance. Techniques such as session affinity and distributed caching can help mitigate these challenges but require careful implementation [22].

4.3 Scalability

Serverless environments must handle rapid scaling up and down in response to fluctuating demand. Load balancing solutions need to be highly scalable to maintain performance and efficiency during peak loads [27].

4.4 Heterogenous Workloads

Serverless applications often run diverse workloads with varying resource requirements. Effective load balancing must adapt to these heterogeneous workloads to optimize resource utilization [23].

4.5 Latency Variability

Latency variability is another challenge in serverless environments. Factors such as network latency, instance initialization times, and resource contention can cause significant fluctuations in response times. Load balancers must dynamically adapt to these variations, employing real-time metrics to distribute requests effectively. Advanced algorithms that consider historical performance data and current load conditions are essential for minimizing latency variability [24].

4.6 Cost Optimization

Cost optimization is a critical concern in serverless computing, as billing is typically based on the number of function invocations and execution duration. Load balancing strategies must optimize the use of function instances to reduce costs while maintaining performance. This involves balancing the trade-off between resource utilization and execution efficiency. Techniques such as auto-scaling policies and efficient function orchestration can aid in cost optimization [25] (Table 1).

4.7 Comparison of Various Load Balancing in Serverless Computing

Table 1. Comparison between different load balancing techniques in serverless computing.

Criteria	Round Robin	Leas Connection	Resource based	Latency based	Cost aware
Description	Distributes requests evenly in a circular order	Routes requests to the server with the fewest active connections	Allocates based on available resources such as CPU and memory	Directs requests based on the geographic location of the server	Adjusts distribution dynamically based on real-time metrics
Complexity	Low	Medium	High	Medium	High
Scalability	Good	Good	Excellent	Excellent	Excellent
Latency	Variable	Low	Low	Low	Low
Resource Utilization	Fair	Good	Excellent	Good	Excellent
Implementation Difficulty	Easy	Moderate	Difficult	Moderate	Difficult
Fault Tolerance	Low	Medium	High	High	High
Best Use Case	Homogeneous workloads	Variable connection load	Resource-intensive tasks	Global user base	Highly dynamic environments

The table provides a clear comparison of various load balancing strategies like Round Robin, Least Connections, Resource-based, Latency-based, and Cost-aware methods. Here's a breakdown of their performance based on specific criteria. Round Robin is simple to implement and offers good scalability, but it tends to have higher latency and lacks strong fault tolerance. Least Connections helps balance traffic more efficiently, though it may struggle in environments with rapidly changing workloads. Resource-based and Cost-aware techniques perform exceptionally well in terms of scalability, resource utilization, and fault tolerance, but their complexity makes them harder to implement.

Latency-based approaches are ideal for minimizing response times for globally distributed users but are less focused on resource efficiency compared to other dynamic methods. In serverless computing, advanced strategies like Resource-based and Cost-aware bring significant advantages, particularly in large-scale, dynamic environments, though at the cost of increased complexity. Each method has its strengths depending on workload characteristics. Simpler methods like Round Robin are easier to implement but become less efficient as system demands grow more complex. More sophisticated methods, while harder to deploy, provide better fault tolerance and resource optimization, making them ideal for dynamic, distributed systems.

5 Conclusions

Serverless computing represents a transformative shift in the deployment and management of applications, enabling developers to focus solely on their code while cloud providers manage the underlying infrastructure. Services such as AWS Lambda, Google Cloud Functions, and Azure Functions illustrate the substantial benefits of this model, including reduced operational complexity and cost. Despite these advantages, load balancing remains a significant challenge in serverless environments due to serverless functions' stateless and ephemeral nature. When the load balancing strategy is need to be applied the traditional load balancing is difficult in the areas. The new approach has to be implanted to show case the identifiable challenges.

The research paper has gone through different approaches to the serverless computing which applies loa balancing. The ideology works on highlighting the new methodology to compare and check the proper usage of resource utilization and the performance of optimal and managing the cost effectively. The AI methodology concatenated with round robin and least connection is to improve the latency in the techniques. Through this serverless will be in dynamic in nature. Load balancing is an effective method of controlling and taking care of the request which arises in the serverless computing and to maximize the usage of it. The future enhancement of this technology will lead to the best result based on the load balancing which is applied in the serverless computing. The era is changing into artificial intelligence and machine learning. Adapting this two wide are of technologies into load balancing strategy can be a bench mark in the future direction. Through this serverless computing can be a major role in the area as well. The advantages are vast and peculiar through the AI and ML based load balancing strategies on serverless computing.

5.1 Future Directions

Integrating AI and machine learning into load balancing algorithms can enhance their predictive capabilities, allowing for more efficient resource management and reduced latency. Integrating AI and machine learning into load balancing algorithms can enhance their predictive capabilities, allowing for more efficient resource management and reduced latency. Developing techniques to reduce cold start latency, such as pre-warming function in-stances based on predictive analytics, can significantly improve performance in serverless environments. Advanced monitoring and analytics tools can provide deeper

insights into function performance and resource utilization, enabling more informed load balancing decisions. The increasing adoption of multi-cloud strategies, developing load balancing solutions that can operate seamlessly across different cloud providers will be essential. The paper extends the possibility of further research to focus on load balancing strategies on serverless computing.

References

1. Adhikari, V., et al.: Survey on load balancing in serverless computing. J. Cloud Comput. **9**(1), 15–30 (2020)
2. Ailamaki, A., et al.: Efficient resource management for serverless computing. IEEE Trans. Cloud Comput. **10**(2), 456–470 (2021)
3. Chandrasekaran, S., et al.: Load balancing techniques for serverless applications. In: Proceedings of the 2019 IEEE International Conference on Cloud Computing, pp. 123–130 (2019)
4. Kumar, P., et al.: State-of-the-art load balancing approaches in serverless architectures. ACM Comput. Surv. **54**(4), 65–78 (2022)
5. Zhang, Y., et al.: Dynamic load balancing in serverless computing: challenges and solutions. Future Gener. Comput. Syst. **118**, 160–175 (2021)
6. Varghese, B., Buyya, R.: Next generation cloud computing: new trends and research directions. Future Gener. Comput. Syst. **79**, 849–861 (2018)
7. Lee, J.B., Yoo, T.H., Lee, E.H., Hwang, B.H., Ahn, S.W., Cho, C.H.: High-performance software load balancer for cloud-native architecture. IEEE Access **9**, 123704–123716 (2021). https://doi.org/10.1109/ACCESS.2021.3108801
8. Dasoriya, R., Kotadiya, P., Arya, G., Nayak, P., Mistry, K.: Dynamicload balancing in cloud a data-centric approach. In: 2017 International Conference on Networks Advances in Computational Technologies (NetACT), pp. 162–166 (2017). https://doi.org/10.1109/NETACT.2017.8076760
9. Chen, M., Huang, H.: Cross-task dynamic load balancing strategy. In: 2018 IEEE 3rd Optoelectronics Global Conference (OGC), pp. 39– 42 (2018). https://doi.org/10.1109/OGC.2018.8529907
10. Jamal, F., Siddiqui, T.: Comparative analysis of load balancing techniques in cloud computing, based on LB Metrices. In: 2021 5th International Conference on Information Systems and Computer Networks (ISCON), pp. 1–5 (2021). https://doi.org/10.1109/ISCON52037.2021.9702508
11. Khatri, D., Khatri, S.K., Mishra, D.: Potential bottleneck and 1244 measuring performance of serverless computing: a literature study. In: Proceeding of 8th International Conference on Reliability, Infocom Technologies and Optimization (ICRITO), p. 161164, June 2020
12. Hassan, H.B., Barakat, S.A., Sarhan, Q.I.: Survey on serverless computing. J. Cloud Comput. **10**(1), 129 (2021)
13. Wu, Y, et al.: Serverless data science are we there yet? Acase study of model serving, (2021). arXiv:2103.02958
14. Scheuner, J., Leitner, P.: Function-as-a-service performance evaluation: a multivocal literature review. J. Syst. Softw. **170**, 10708 (2020)
15. Sadaqat, M., Colomo-Palacios, R., Ricardo, K., Knud sen, L.E.S.: Serverless computing: a multivocal literature review (2018). http://hdl.handle.net/11250/2577600
16. Taibi, D., El Ioini, N., Pahl, C., Niederkoer, J.R.S.: Serverless cloud computing (function-as-a-service) patterns: a multivocal litera ture review. In: Proceeding of 10th International Conference on Cloud Computing and Services Science (CLOSER), p. 112 (2020)

17. Filho, C.P., et al.: A systematic literature review on distributed machine learning in edge computing **22**(7) (2022). https://doi.org/10.3390/s22072665
18. Zhu, L., Cui, J., Xiong, G.: Improved dynamic load balancing algorithm based on Least-Connection Scheduling. In: 2018 IEEE 4th Information Technology and Mechatronics Engineering Conference (ITOEC), pp. 1858–1862 (2018). https://doi.org/10.1109/ITOEC.2018.8740642
19. Cui, Y., Long, Z.: Load balancing scheduling algorithm based on improved particle swarm optimization. In: 2021 IEEE International Conference on Artificial Intelligence and Computer Applications (ICAICA), pp. 1044–1047 (2021).
20. https://doi.org/10.1109/ICAICA52286.2021.9498058
21. Benefits of cloud computing for your business Globaldots. Accessed 16 April 2024. https://www.globaldots.com/resources/blog/cloud-computing-benets-key-advantages-for-your-business/
22. Castro, P., Ishakian, V., Muthusamy, V., Slominski, A.: The rise 955 of serverless computing. Commun. ACM **62**(12), 4454,956 (2019). https://doi.org/10.1145/3368454
23. Kanagaraj, K., Geetha, S.: A hybrid framework for effective prediction of online streaming data. J. Phys., Conf. **1767**(1), 959 (2021). Art. no. 012016
24. Zhang, C., Yu, M., Wang, W., Yan, F.: MArk: exploiting cloud 1032 services for cost-effective, SLO-aware machine learning inference serving. In: Proc. USENIX Annual Technical Conference Renton, WA, USA: 1034 USENIX Association, pp. 10–49, July 2019. https://www.usenix.org/conference/atc19/presentation/zhang-chengliang
25. Ali, A., Pinciroli, R., Yan, F., Smirni, E.: BATCH: machine learning inference serving on serverless platforms with adaptive batching. In: 1072 Proceeding of International Conference for High Performance Computing, Networking, Storage and Analysis, pp. 115–1034, November 2020
26. Xu, F., Qin, Y., Chen, L., Zhou, Z., Liu, F.: DNN: Achieving predictable distributed DNN training with serverless architectures. IEEE Trans. Comput. **71**(2), 45–52 (2022)
27. Patros, P., Spillner, J., Papadopoulos, A.V., Varghese, B., Rana, O., Dustdar, S.: Toward sustainable serverless computing. IEEE Internet Comput. **25**(6), 42–50 (2021)
28. Cassel, G.A.S., Rodrigues, V.F., Righi, R.D.R., Bez, M.R., Nepomuceno, A.C., Costa, C.A.D.: Serverless computing for internet of things: a systematic literature review. Future Gener. Comput. Syst. **128**, 299316 (2022)
29. Scheuner, J., Leitner, P.: Function-as-a-service performance evaluation: a multivocal literature review. J. Syst. Softw. **170**, 110708 (2020)
30. Wu, Y., Dinh, T.T.A., Hu, G., Zhang, M., Chee, Y.M., Ooi, B.C.: Serverless data science are we there yet? A case study of model serving (2021). arXiv:210302958

Real-Time Node's Power-Aware Kubernetes Scheduler in a Cloud Environment

Kiran Kumari[1,2(✉)] and Dhananjay Dakhane[3]

[1] Department of Computer Engineering, Ramrao Adik Institute of Technology, D.Y. Patil Deemed to Be University, Nerul, Navi Mumbai, India
kir.kum.rt21@dypatil.edu, kirankumarisinha@somaiya.edu
[2] Faculty at K. J. Somaiya School of Engineering, Somaiya Vidyavihar University, Mumbai, India
[3] Department of Computer Science and Engineering, Ramrao Adik Institute of Technology, D.Y. Patil Deemed to Be University, Nerul, Navi Mumbai, India
dhananjay.dakhane@rait.ac.in

Abstract. Kubernetes is a popular container orchestration system for deploying containerized applications in the cloud environment due to the popularity of Container-as-a-Service (CaaS). The Kubernetes default scheduler does not consider node's power consumption for scheduling the jobs that result in cluster power consumption. To minimize the power consumption of a Kubernetes cluster, a real-time Node's Power-Aware Kubernetes Scheduler (NPAKS) is proposed. Further, NPAKS is implemented using the Kubernetes scheduling framework and deployed on a bare-metal EKS cluster. The experimental results show that NPAKS scheduler minimizes cluster power consumption with respect to the Kubernetes default scheduler. Finally, future research opportunities are discussed.

Keywords: EKS · Kubernetes · Power · Scheduling

1 Introduction

Container management and scaling can be facilitated by Docker Swarm, Mesos, Kubernetes, etc. Kubernetes is the most preferred open-source option among developers and businesses. Production-level Kubernetes clusters are increasing in public cloud environments such as AWS, Azure, Google, etc. In Kubernetes, the default scheduler, known as "Kube-scheduler," schedules the incoming jobs over the nodes for execution. The researchers contributed to scheduling strategies that manage resource utilization (CPU, memory, disk), image transmission time, resource interference, power predictions, etc. to improve cluster performance metrics such as response time, makespan time, power consumption, etc. Only 6% of research contributions are observed from a power perspective [1, 2].

The power consumption of every node in the Kubernetes cluster may have a unique manufacturing configuration and may use power differently depending on the load [3, 4]. Measuring the power of Kubernetes nodes in a cloud environment is challenging because

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
S. Rajagopal et al. (Eds.): ASCIS 2024, CCIS 2428, pp. 433–443, 2025.
https://doi.org/10.1007/978-3-031-86302-8_29

only some instances support power measurement, and only software tools (hardware power meters can't be attached) can be used to measure the power. Red Hat has developed "Kepler" in 2022, a Kubernetes power exporter. It collects kernel data using the interface RAPL (Running Average Power Limit) [5, 6]. Adaltas has developed "Scaphandre" in 2020, a bare metal energy consumption monitoring tool for virtual machines with a Qemu or KVM hypervisor [6, 7]. These tools create extra overhead on the cluster. The proposed Kubernetes NPAKS has low overhead because it uses the Linux utility "Turbostat," which provides accurate and real-time power consumption for each Kubernetes node.

The proposed real-time Kubernetes Node's Power-Aware scheduler is implemented using the Kubernetes scheduling framework. It is a pluggable architecture. The scheduling features are customized by implementing custom APIs while keeping the core default scheduler lightweight and maintainable. It consists of extensible and internal APIs. Extensible APIs have customizable filters and score functions, as shown in Fig. 1. Filters are used for hard constraints, aka pre-emption of nodes, and scores are used for soft constraints to pick the best node for scheduling. The proposed NPAKS extends custom score and normalized score functions.

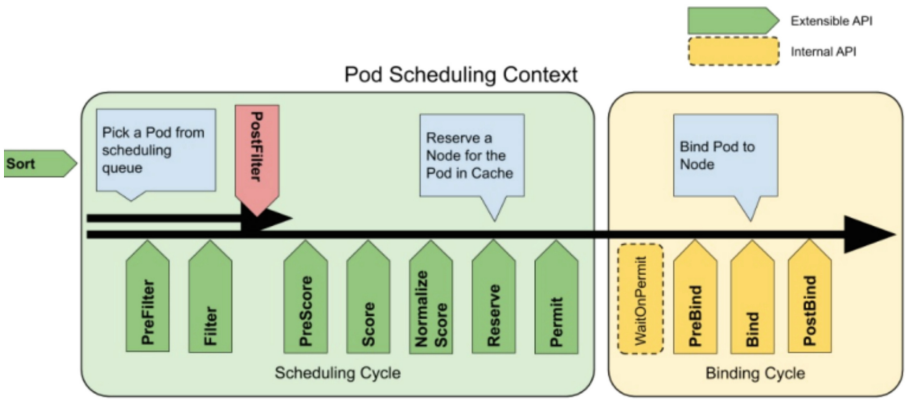


Fig. 1. Scheduling Framework

An EKS cluster of two bare metal instances is created on the AWS cloud for NPAKS deployment, and its performance is compared with the default scheduler considering cluster power consumption.

The structure of the paper is organized as follows: Sect. 2 explains the related survey; Sect. 3 describes the proposed system; Sect. 4 discusses result analysis and discussion; and Sect. 5 concludes with a conclusion and future contributions.

2 Related Survey

Research contributions have been observed on container scheduling from a power perspective, with the main aim being to minimize cluster power consumption.

A scheduler is proposed based on the concept that "energy consumed is proportional to resource utilization and energy efficiency starts degrading at high levels of resource

utilization" for Docker Swarm [4]. A ML-based prediction is proposed on Kubernetes in a server-less computing environment. An overall 8% power reduction is achieved [8]. AI-based multi-criteria (cloud condition and user requirements) scheduling strategy is formulated for node selection [9]. TOPSIS (Technique for the Order of Prioritization by Similarity to the Ideal Solution), a multi-criteria optimization method is discussed, which gives equal weight to six parameters, like power consumption, memory, CPU and storage utilization rate, number of containers on each node and image transmission time in an emulation environment. 5–10% improvement in power is achieved [10]. Resource interference is discussed by co-allocating many jobs on a single node and assuming renewable energy sources (solar, wind, etc.). The overall energy improvement observed is 14.42%, against First-Cum-First-Serve (FCFS) [11]. A neural network-based power estimation model is used for container scheduling using resource utilization information about the containers [12]. A plugin is developed to replace the default Kubernetes scheduler by considering both software and hardware models. For the software model, a predictive analysis is implemented based on the power consumption of the incoming containers. A 10–20% reduction in power consumption is observed over default Kubernetes [13]. The authors proposed a machine-learning scheduling approach to estimate the node's power consumption. This approach selects node that has lowest power consumption for Docker SwarmKit [14]. An improved genetic algorithm is proposed to save energy compared with the First Fit conventional GA and the Particle Swarm Optimization (PSO) algorithm [15]. The brownout concept is discussed to inactivate the containers so that the nodes can be protected from overload to diminish energy consumption, response time, and energy while ensuring QoS and SLA. A 6% to 12% reduction in power consumption is suggested [16]. A renewable energy-based container scheduling scheme is presented. 15% energy savings is achieved [17]. Container image pulling costs are considered and 30% energy reduction is discussed [18]. GENPACK, a framework for container scheduling, is proposed for the Docker setting. It articulates garbage collection on cluster nodes to deploy containers by applying the best-fit scheduling technique. 23% energy efficiency is observed compared to Docker Swarm [19].

All the above researches contributed to improve power to some extent, but real-time Kubernetes node's power is not considered by scheduler. This is an important parameter that needs to be addressed to minimize power consumption.

3 Proposed System

The system architecture is presented in the Fig. 2. A production-level EKS Kubernetes cluster is created with two bare metal, c5d.metal and c5n.metal Ubuntu instances on the AWS cloud. Model-Specific Register (MSR), used for CPU performance monitoring, is present in these bare metal processors. Bare Metal supports the "Turbostat" tool, a Linux utility to get the CPU measurement from MSR. In Kubernetes, the power consumption metric of a node is not available with the node manifest; even a pod can't access the node's hardware information directly because the Kubernetes architecture is such that all pods execute in isolation. To access the MSR value from the nodes, a privileged daemon set is proposed. The privileged daemon pod works as a power metric collector to pull real-time power consumption from each node's MSR using the Turbostat tool and stores it in the node's environmental variable in real-time, recursively.

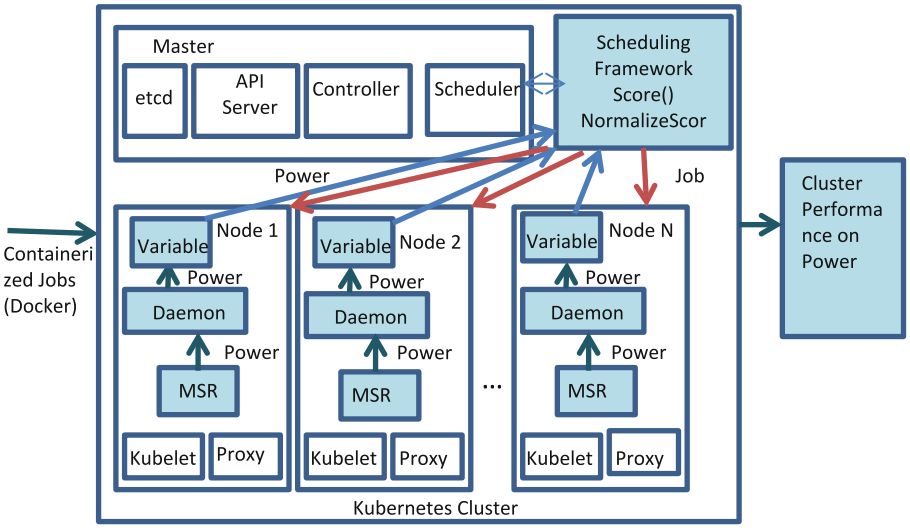


Fig. 2. Proposed System

A power-daemonset.yaml is created to execute the image of power-daemonset.go in privileged mode with node environmental variable and volume mount to host volume. Pseudo code for power-daemonset.go is given below:

- Create Kubernetes client config and clientset
- Fetch the node name from environment variable
- Fetch node object
- While
  - o Execute Turbostat and update node variable with PkgWatt
  - o Update node object
- End

A proposed Kubernetes scheduler is implemented using Kubernetes Scheduling Framework. A new "PowerScore" score plugin interface (API) is implemented with Score() and NormalizeScore() functions. Score() function pulls the power from each node's environment variable, the real-time power collected by Daemon Pods. NormalizeScore() function normalizes the values in the range 0 to 100 in reverse order because a low-power node can have the highest normalized value. Finally, the scheduler schedules the jab on the node with the highest value, which is the one with the lowest power consumption. To implement proposed "PowerScore" Score plugin interface, the Kubernetes Scheduling Framework downloaded from GitHub repository, and the following steps are followed: [20].

Step 1: A powerscore.go code is implemented in the scheduler-plugins /pkg/powerscore folder as shown below:

|- scheduler-plugins
|- pkg
|- powerscore
|- powerscore.go
- PowerScore struct for handle of framework
- Initializes a new plugin
- Score() function for ScorePlugin interface
- NormalizeScore() function for range of 0 to 100
- New function for PluginFactory interface

Step 2: ScorePowerArgs is added to apis/config and following codes are modified:
|- scheduler-plugins
|- apis/config
|- types.go
|- v1beta3
|- types.go
|- defaults.go
|- register.go
|- register.go
Step 3: ScorePower plugin is registered in the cmd/scheduler/main.go
Step 4: [./ hack/update-codegen.sh] command is executed to prepare DeepCopy
Step 5: Configuration file prepared to enable only ScorePower plugin
The profile of configuration file is given below:

```
profiles:
  - schedulerName: power-scheduler
    plugins:
      score:
        enabled:
          - name: ScorePower
        disabled:
          - name: "*" // disable all default plugins
```

Step 6: Scheduler Image is built with configuration file
Step 7: Service Account, ClusterRole and ClusterRoleBinding are created
Step 8: Scheduler Image is deployed with service Account on the Kubernetes Cluster

4 Result Analysis and Discussion

The EKS cluster setup consists of two worker nodes, c5d.metal and c5n.metal instances. PkgWatt is watts consumed by CPU package of a machine at a specified time. The Nodes and cluster power consumption are show in the Fig. 3. Power consumption of nodes in idle state (no load) is between 40 and 50 watts for c5d.metal and between 30 and 40 watts for c5n.metal. Power consumption of nodes at the base line (after deployment of the proposed scheduler and daemon set) is between 40 and 55 watts for c5d.metal and between 35 and 50 watts for c5n.metal. Power consumption of cluster is between 70 to 90 for Idle state and 80 to 100 for Base Line.

Cluster power consumption considered is total power consumption (PC) of both worker nodes at a given time:

$$PC_{Cluster} = PC_{c5d} + PC_{c5n} \quad (1)$$

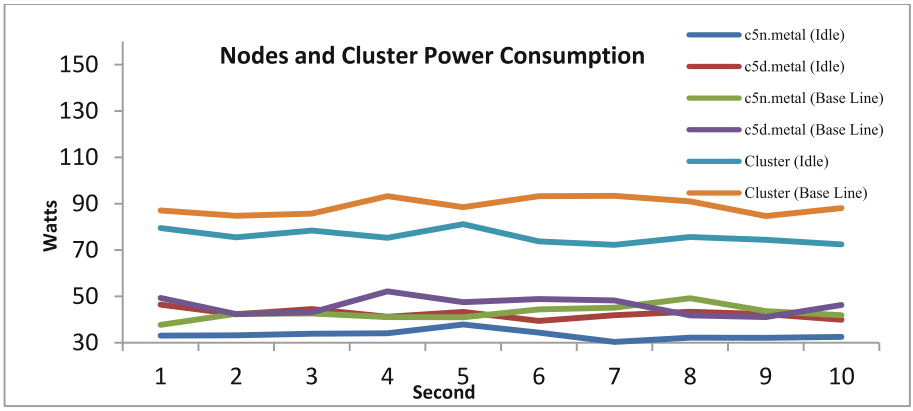


Fig. 3. Power consumption of nodes and Cluster at Idle and Base line

Deamon-set collects power of each node and stores it into the node's variable. The power scheduler schedules the jobs to the node with the lowest power consumption at a specific time. Most of the jobs are scheduled on low power consumption nodes, resulting in minimization of cluster power consumption.

To evaluate the performance, an experiment is performed with three sets of jobs of different time complexity. In the experiments, jobs are deployed in the interval of one second (sequential order). All three sets of jobs are deployed in the cluster with power scheduler and default scheduler independently. The sets of jobs are given below:

Input-set-01: 10 jobs, factorial of 10
Input-set-02: 10 jobs, bubble sort of 10 random numbers
Input-set-03: 10 jobs, file write and read

4.1 Case 1: Input-Set-01: 10 Jobs, Factorial of 10

Figure 4 shows the trend of cluster power consumption when input-set-01 is deployed on cluster with power scheduler and default scheduler individually. A peak is an increase in cluster power consumption when jobs are assigned to the nodes. Lower peaks for the power scheduler show that most of the jobs are scheduled on the c5n.metal instance (having low power consumption), resulting in low cluster power consumption.

Figure 5 compares the proposed scheduler with the default scheduler. The default scheduler consumes greater power than the power scheduler. Power saving of the proposed scheduler is 180 watts over the default scheduler for 10 jobs in 10 s.

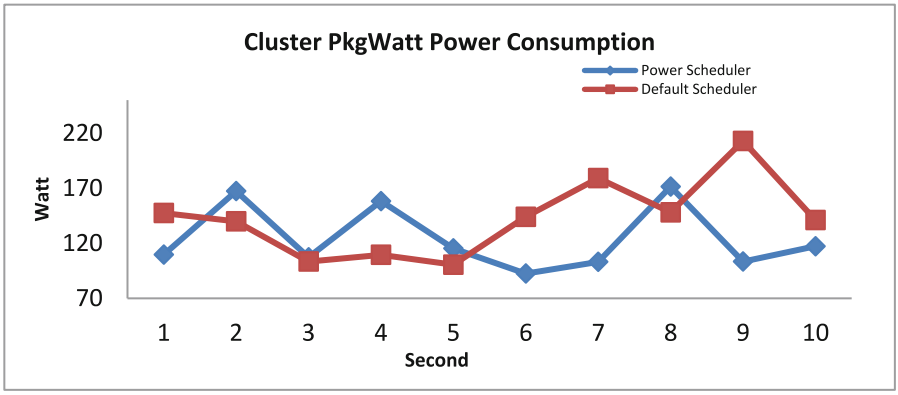


Fig. 4. Factorial of 10 sequential jobs

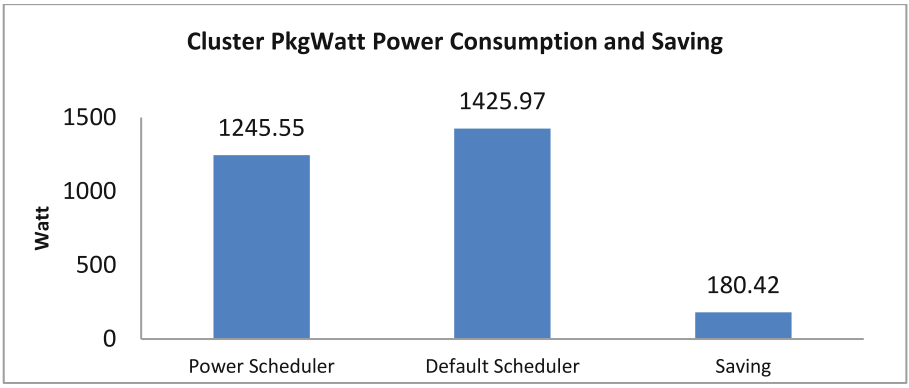


Fig. 5. Power consumption vs saving for Factorial of 10

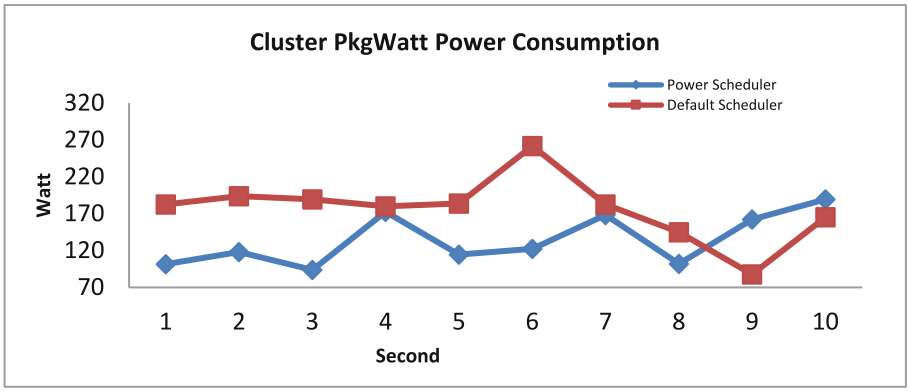


Fig. 6. Bubble sort for sequential jobs

4.2 Case 2: Input-Set-02: 10 Jobs, Bubble Sort of 10 Random Numbers

When input-set-02 is deployed on a cluster with a power scheduler and a default scheduler separately, Fig. 6 displays the trend of cluster power usage. When jobs are issued, the cluster power usage peaks. The power scheduler's lower peaks indicate less power consumption of cluster than default scheduler.

The default scheduling and the power scheduler are compared in Fig. 7. Comparatively, the power scheduler uses less power than the default scheduler. When 10 bubble sort jobs are completed in 10 s, the suggested scheduler saves 426 watts power than the default one.

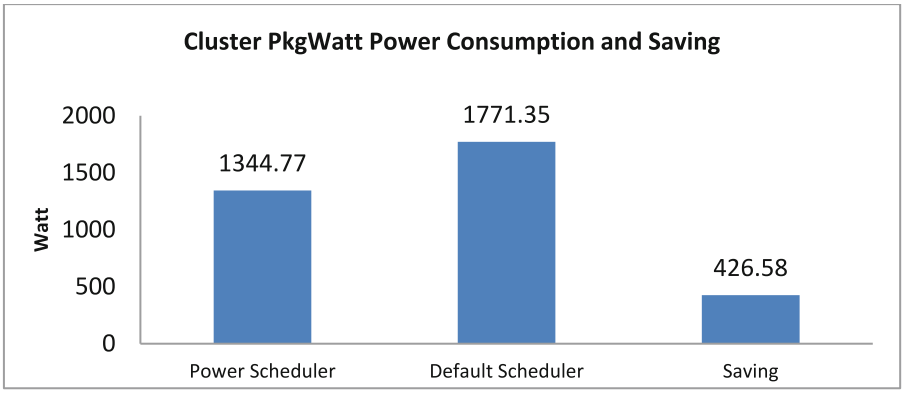


Fig. 7. Bubble sort power consumption vs saving

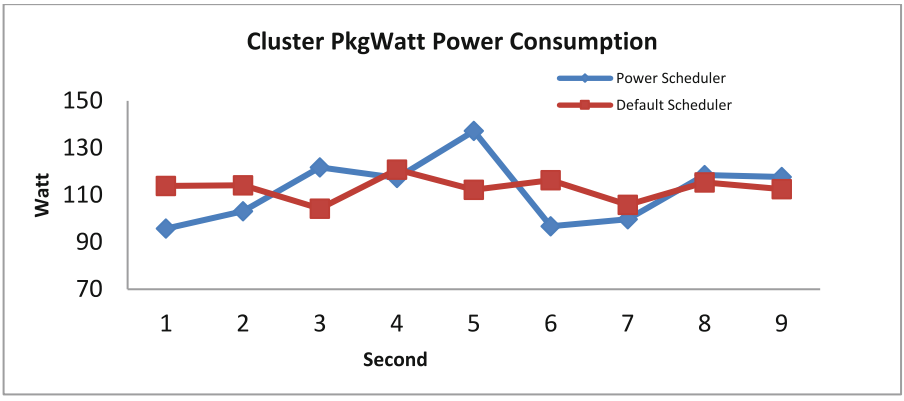


Fig. 8. File R/W for sequential jobs

4.3 Case 3: Input-Set-03: 10 Jobs, File Write and Read

Upon deployment of input-set-03 on the cluster with separate power scheduler and default scheduler, Fig. 8 displays the tendency of cluster power usage. When tasks are deployed on the cluster, the power scheduler usage is less at most of the time.

Figure 9 shows a comparison between the power scheduler and the default scheduler. In contrast to the default scheduler, the power scheduler consumes less energy. When 10 R/W jobs are performed in ten seconds, the proposed scheduler saves 99 watts less power than the default one.

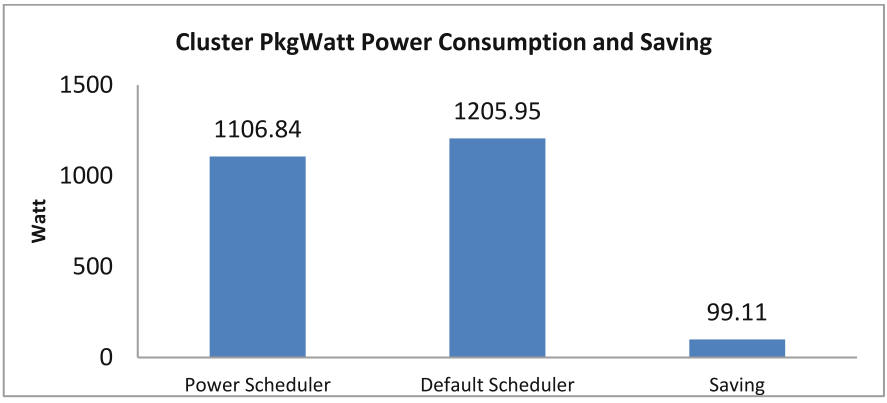


Fig. 9. File R/W power consumption vs saving

4.4 Percentage Power Savings

Figure 10 shows comparison on percentage of power savings. All three sets of jobs are deployed several times on EKS cluster. The top three results on percentage savings for Factorial jobs are between 7 to12%. Bubble sort jobs save between 11 to 24% and File operation jobs save between 2 to 8% for three iterations. The power scheduler saves average power between 5% and 16% over the default scheduler for three best results of sequential job input sets. Percentage of saving depends on type of nodes, type of jobs, number of iterations, and duration of execution. Variation in result is because of power instabilities of nodes.

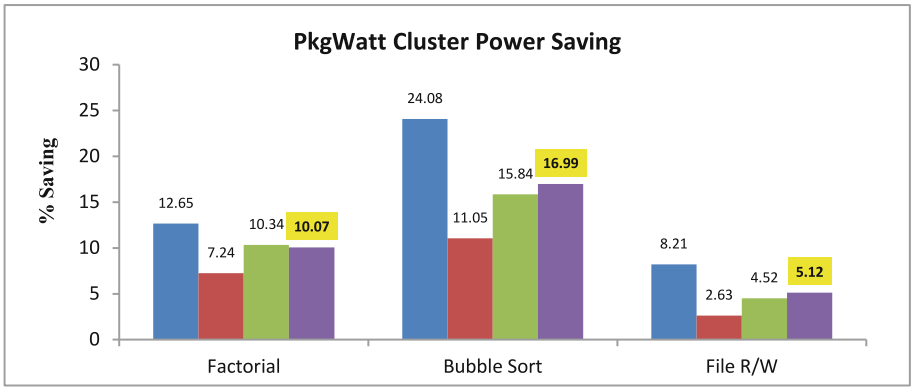


Fig. 10. Percentage of three iterations and average power savings for three input sets of NPAKS compared to default scheduler

5 Conclusion and Future Contribution

The result demonstrates that the proposed NPAKS saves average cluster power consumption between 5% and 16% with respect to the default scheduler for three job-sets. The scheduling trends and power savings show that NPAKS schedules jobs on low power consuming nodes with respect to the default scheduler, which spreads the jobs on the nodes. The cluster power savings may change based on the types of nodes, types of jobs, number of iteration and duration of the study. The limitation of NPAKS is that nodes should support "Turbostat". NPAKS's main focus is to decrease the cluster power consumption. Other cluster parameters like quality of service, nodes resource utilization, response time and carbon emissions etc. will be considered in future to balance the cluster overall performance.

References

1. Senjab, K., Abbas, S., Ahmed, N., et al.: A survey of Kubernetes scheduling algorithms. J. Cloud Comp. **12**, 87 (2023). https://doi.org/10.1186/s13677-023-00471-1
2. Rejiba, Z. and Chamanara, J.: Custom scheduling in Kubernetes: a survey on common problems and solution approaches. ACM Comput. Surv. **55**(7), 1–37 (2022). Article No.: 151, https://doi.org/10.1145/3544788
3. Peaks power efficiency aware Kubernetes scheduler. https://github.com/sustainable-computing-io/peaks, Accessed June 2024
4. Mendes, S., Simao, J., Veiga, L.: Oversubscribing micro-clouds with energy-aware containers scheduling. ACM ISBN 978-1-4503-5933-7/19/04 (2019) https://doi.org/10.1145/3297280.3297295
5. https://github.com/sustainable-computing-io/kepler, Accessed June 2024
6. Carlo, C., et al.: Impact of power consumption in containerized clouds: a comprehensive analysis of open-source power measurement tools. Comput. Networks **245** (2024). https://doi.org/10.1016/j.comnet.2024.110371
7. https://hubblo-org.github.io/scaphandre-documentation/tutorials/getting_started.html, Accessed June 2024

8. Chiorescu, R., Djemame, K.: Scheduling energy-aware multi-function serverless workloads in OpenFaaS. In: Naldi, M., Djemame, K., Altmann, J., Bañares, J.Á. (eds.) Economics of Grids, Clouds, Systems, and Services. GECON 2024. LNCS, vol. 15358, pp. 137–149. Springer, Cham (2025). https://doi.org/10.1007/978-3-031-81226-2_13
9. Jorge-Martinez, D., Butt, S.A., Onyema, E.M., et al.: Artificial intelligence-based Kubernetes container for scheduling nodes of energy composition. Int. J. Syst. Assurance Eng. Manage. (2021). https://doi.org/10.1007/s13198-021-01195-8
10. Menouer, T.: KCSS: kubernetes container scheduling strategy. J Supercomput. **77**, 4267–4293 (2021). https://doi.org/10.1007/s11227-020-03427-3
11. Kaur, K., et al.: KEIDS: kubernetes-based energy and interference driven scheduler for industrial IoT in edge-cloud ecosystem. IEEE Internet Things J. **7**(5) (2020). https://doi.org/10.1109/JIOT.2019.2939534
12. Mehta, H.K., Harvey, P., Rana, O., Buyya, R., Varghese, B.: WattsApp: power-aware container scheduling. In: IEEE/ACM 13th International Conference on Utility and Cloud Computing (UCC), Leicester, UK, vol. 2020, pp. 79–90 (2020). https://doi.org/10.1109/UCC48980.2020.00027
13. Townend, P., et al.: Improving data center efficiency through holistic scheduling in kubernetes. In: IEEE International Conference on Service-Oriented System Engineering (SOSE), IEEE Computer Society, pp. 156–166 (2019). ISBN 978-1-7281-1443-9, https://doi.org/10.1109/sose.2019.00030
14. Menouer, T., Manad, O., Cérin, C., Darmon, P.: Power efficiency containers scheduling approach based on machine learning technique for cloud computing environment. In: Esposito, C., Hong, J., Choo, KK. (eds.) Pervasive Systems, Algorithms and Networks. I-SPAN 2019. Communications in Computer and Information Science, vol 1080. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-30143-9_16
15. Zhang, R., Chen, Y., Dong, B., Tian, F., Zheng, Q.: A genetic algorithm-based energy-efficient container placement strategy in CaaS. **7**, 121360–121373 (2019). https://doi.org/10.1109/ACCESS.2019.2937553
16. Xu, M., Buyya, R.: BrownoutCon: a software system based on brownout and containers for energy-efficient cloud computing. J. Syst. Softw. (2019). https://doi.org/10.1016/j.jss.2019.05.031
17. Kumar, N., Aujla, G.S., Garg, S., Kaur, K., Ranjan, R., Garg, S.K.: Renewable energy-based multi-indexed job classification and container management scheme for sustainability of cloud data centers. IEEE Trans. Industr. Inf. **15**(5), 2947–2957 (2019). https://doi.org/10.1109/TII.2018.2800693
18. Zhang, D., Yan, B.H., Feng, Z., Zhang, C., Wang, Y.X.: Container oriented job scheduling using linear programming model. In: 3rd International Conference on Information Management (2017). https://doi.org/10.1109/INFOMAN.2017.7950370
19. Havet, A., Schiavoni, V., Felber, P., Colmant, M., Rouvoy, R., Fetzer, C.: Genpack: a generational scheduler for cloud data centers. In: IEEE International Conference on Cloud Engineering (IC2E), pp. 95–104. IEEE (2017). https://doi.org/10.1109/IC2E.2017.15
20. Scheduling framework, Scheduling Framework | Kubernetes, scheduling-framework. https://github.com/kubernetes/enhancements/tree/master/keps/sig-scheduling/624-scheduling-framework. Accessed June 2024

EAPTEEN: Energy-Aware Adaptable Routing for Continuous Monitoring in Event-Driven Wireless Sensor Networks the EAPTEEN Approach

D. Jeya Rani[1,2(✉)] and Nagarajan Munusamy[1]

[1] KSG College of Arts and Science, Coimbatore, India
jeyarani1@gmail.com
[2] KG College of Arts and Science, Coimbatore, India

Abstract. Due to WSN's many uses in agriculture, mechanical assembly, military surveillance, ecological checking, clinical and medical care, disaster recovery duties, and so on, advances in the WSN sphere have gained speed in recent years. To facilitate communication between themselves, the wireless sensor nodes in a WSN are structured in a completely random fashion. One of the main limitations of WSN is energy since the nodes are battery-powered.This paper proposes the EAPTEEN Method for Energy-Aware Adaptable Routing for continuous monitoring in WSN. The "Enhanced Adaptable Periodic Threshold Sensitive Energy Efficient Sensor Network Protocol (EAPTEEN) Approach" introduces a new routing protocol that maximizes energy efficiency and maintains continuous monitoring in event-driven WSNs. Current protocols, like TEEN, prioritize data transmission when thresholds are exceeded, which can result in gaps in monitoring during inactive periods and the possibility of missing important events. Additionally, EAPTEEN enhances TEEN by enabling sensor nodes to consistently send data to the base station at set intervals, regardless of any threshold conditions. The suggested method may respond to changes in the network including topological dynamic changes, node failures, and energy level variations. Focus of the discussion of experimental findings is the present method and assessment of network performance criteria including latency, energy, and throughput.

Keywords: Continuous monitoring · EAPTEEN · Energy efficiency · Routing protocols · Wireless sensor networks

1 Introduction

In newly developing sectors such smart agriculture, environmental monitoring, and healthcare, WSNs are very vital. Given their fundamental qualities, it is essential to carefully balance in these networks security, reliability, energy economy, and quality of service (QoS [1, 2]). Achieving this equilibrium, however, may be very difficult given the limited resources of sensor nodes and the often shifting and erratic character of their

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
S. Rajagopal et al. (Eds.): ASCIS 2024, CCIS 2428, pp. 444–460, 2025.
https://doi.org/10.1007/978-3-031-86302-8_30

working conditions [3]. By enabling continuous monitoring in many contexts, wireless sensor networks (WSNs) have revolutionized data collecting [4, 5]. These networks of autonomous sensors coordinate monitoring of environmental or physical parameters like motion, sound, humidity, and temperature. Geologically, they are scattered [6]. WSNs provide applications requiring remote and real-time data collecting scalability and adaptability [7]. In event-driven WSNs, sensors react to certain events or environmental changes, hence minimizing resource use by only producing data when very necessary [8, 9]. This method best uses energy, thereby supporting sustainability and prolonging the lifetime of sensor networks.Constant monitoring made feasible by event-driven WSNs is essential in many diverse fields, including environmental conservation, industrial process automation, and healthcare monitoring [10, 11].

Installing and integrating WSNs into pre-existing systems might be especially difficult when operating in a temporary site [12, 13]. For instance, a monitoring network may only be utilized for a short period in one location before being transferred to another [14, 15]. KPIs for most applications include network longevity and energy usage as most of the nodes—sensors—use batteries [16, 17]. Significantly, these routing techniques decide the lifetime and energy consumption of the network. But in pragmatic settings where data loss might have disastrous effects [18, 19], network dependability is very crucial. And constantly changing transmission rates and routing patterns, protocols like TEEN help to increase data delivery efficiency and network longevity and thus optimizing energy use [20, 21]. This flexible approach helps the network to keep constant monitoring in event-driven conditions by lowering energy consumption during idle periods [22–24].

The main contribution of the paper is.

- Enhanced Adaptable Periodic Threshold Sensitive Energy Efficient Sensor Network Protocol

Many authors cover a wide range of continuous monitoring techniques in Sect. 2. Section 3 discussed the EAPTEEN model. The results and discussion were discussed in Sect. 4. Finally, Sect. 5 displays the outcome and potential future endeavors.

1.1 Study Motivation

Improving energy efficiency and resolving the problems with continuous monitoring in event-driven WSNs drives the research. Conventional systems such as TEEN, which rely on thresholds to send data, may miss important events during idleness. EAPTEEN is working on a protocol to get past these constraints and guarantee that data can still be sent continuously even in cases with threshold criteria satisfied. EAPTEEN increases reliability and longevity by allowing WSNs to respond to network changes and properly control energy resources. This qualifies them better for long-term use in environments with limited resources and ongoing change.

2 Background Study

Balakiruthiga, B. et al. [2], the author provides a strategy for energy-efficient routing in a physically dispersed and conceptually centralized Open Daylight(ODL) controller system based on the Multipath TCP (MPTCP) technique. The performance of the proposed solution has been examined in this study using the Mininet virtual network framework.

Giordano M. et al. [4] discussed that the energy-aware method was designed and implemented. The technique has low complexity. The approach was tested on a real-world application use case and compared to the top adaptable sampling techniques.

Marchese, M., &Patrone, F.[10] due mainly to their decreased cost, small Low Earth Orbit (LEO) satellites are becoming increasingly attractive for various applications, including the construction of SatCom networks capable of integrating with the forthcoming 5G terrestrial infrastructure. Considering the hardware limits and available resource constraints resulting from lower prices and size is essential.

Sundarasekar, R. et al. [17] these authors research uses Adaptable Energy-Aware Quality of Service (AEA-QoS) for consistent data transmission in Underwater Acoustic Sensor Networks (UWASN). AEA-QoS selects data from sensor nodes using deep learning algorithms and discrete-time stochastic control processes in order to lower energy consumption in UWASN and increase throughput and network longevity.

Shanmugavalli, R., & Subashini, P. [18] the author provide the MAODV routing protocol for Wireless Sensor Networks, which can find the best route to the target node and increase the network's lifetime. This protocol finds the residual energy and threshold of nodes using the max-min energy idea, which decreases energy usage. In this research, two routing protocols—AODV and MAODV—are compared.

Muthulakshmi, K. et al. [19] to achieve improved EE and lifespan in WSN, a new AWDO-EACS model was created. Previous clustering methods, such as GA-LEACH and MW-LEACH, were outperformed by the AWDO-EACS model, which successfully optimizes both energy efficiency and end-to-end latency. Longer network lifetime and lower data transmission latencies are achieved by the effective formation and maintenance of clusters made possible by the adaptive wind-driven optimization algorithm (Table 1).

2.1 Problem Definition

In event-driven WSNs, a major difficulty is effectively managing power consumption and guaranteeing consistent data transfer. Designed especially for reactive networks in WSNs, the TEEN protocol emphasizes applications needing constant data sensing but only transmission when a certain threshold is reached. While TEEN is made with energy conservation in mind by limiting transmissions, when the detected characteristic often exceeds the threshold it may not yield in significant energy savings compared to alternative protocols. Managing and configuring a network may be somewhat complex as it requires continuously modifying thresholds to fit evolving environmental conditions and application requirements. For this project, therefore, I am showcasing the EAPTEEN.

Table 1. EAR Techniques in Wireless Sensor Networks

Author	Year	Methodology	Advantage	Limitation
Al-Turjman, F., & Kilic, K. I	2019	Energy-aware routing (EAR) for wireless nano-sensor networks	Simplified routing for nano-sensors, energy-efficient	Limited scalability for more extensive networks means they cannot handle dynamic network conditions effectively
Ekler, P. et al	2022	E.A. IoT routing algorithms in an intelligent city environment	Adapted for smart city IoT environments, it considers energy constraints	Due to complex routing algorithms, high computational overhead cannot be suitable for real-time applications
Giordano, M., et al	2023	E.A. Adaptable Sampling for IoT Self-sustainability	Enhances device longevity and supports energy-neutral operations	Requires sophisticated energy modeling, sensitive to parameter tuning
Han, Y. et al	2022	Trust-based secure routing in WSNs using genetic algorithms	Integrates security with energy-aware routing, adaptable to varying trust levels	Genetic algorithm convergence issues can require extensive tuning for optimal performance
Musaddiq, A., et al	2018	E.A. trickle timer for RPL-based routing in IoT	Efficient energy utilization integrates with the RPL protocol stack	Dependency on protocol-specific implementation cannot generalize well across different IoT platforms

3 Materials and Methods

This chapter shows the EAPTEEN method to optimize energy efficiency and ensure continuous monitoring in event-driven wireless sensor networks (WSNs).

3.1 Network Model

Operating within a WSN, the EAPTEEN protocol uses sensor nodes to track environmental events. Important elements include.

One of the basic formulas explaining the periodic data transmission in EAPTEEN may be shown as follows.

$$T_{periodic} = T_{base} + \in \tag{1}$$

where:

$T_{periodic}$: Interval of periodic transmission from sensor nodes to the base station.

T_{base}: Base transmission interval computed from network circumstances.

$\in$: Changeable values for optimal periodic transmission depending on energy level and data significance (Fig. 1).

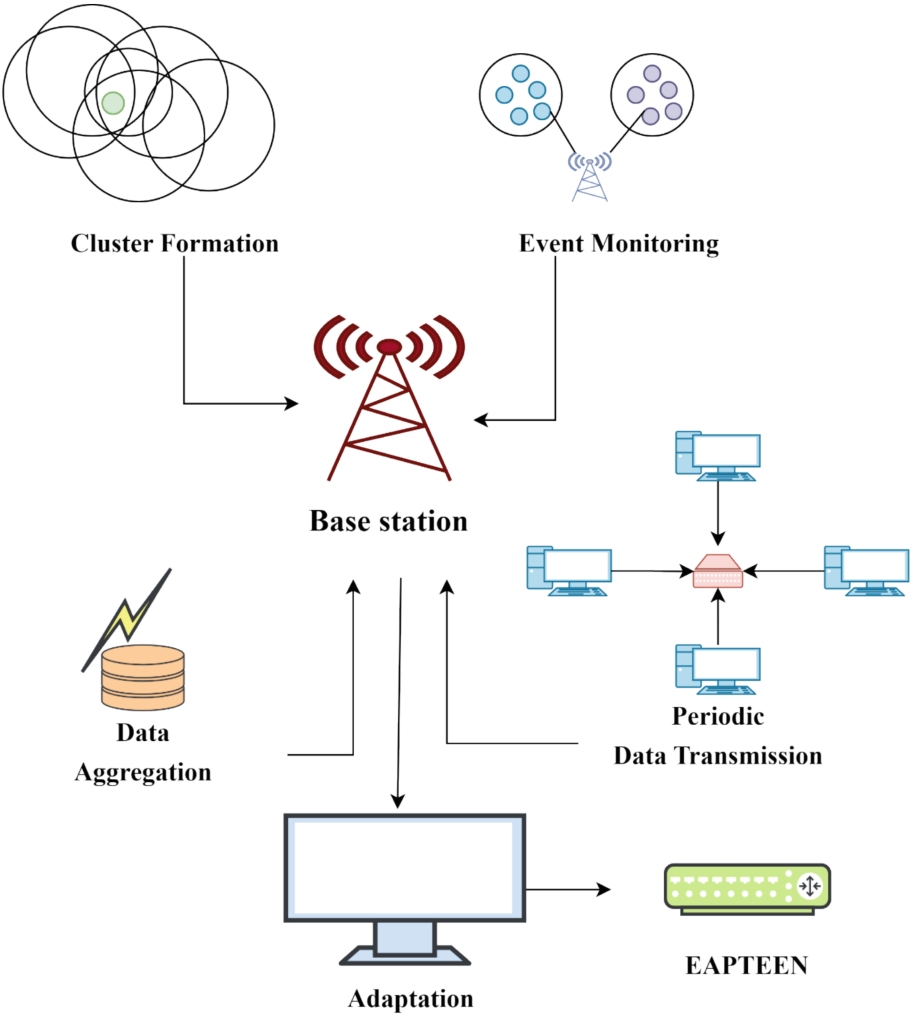


Fig. 1. EAPTEEN workflow architecture

3.2 System Model

Cluster Heads (CH_i): selected from the sensor node network according to reliability of their connections, distance from the hub, and available power.

Energy Management Module**:**applies methods to optimize energy consumption across sensor nodes, hence increasing the lifetime of the network.

One of the critical equations describing the transmission rate adjustment in EAPTEEN can be formulated as:

$$T_{transmission} = T_{base} + f(E_i) \tag{2}$$

3.3 Enhanced Adaptable Periodic TEEN

Guaranteeing reliable data transmission and lowering power consumption are two main challenges in event-driven WSNs. Two aging schemes with notable flaws are EAPTEEN and Secure. Although EAPTEEN increases reliability and load balancing across numerous channels, its significant energy consumption to maintain several active pathways limits network lifetime.Secure And Energy-Aware Heuristic-Based Routing typically generates bad routing choices despite its heuristic character, which results in higher latency and excessive energy usage under dynamic network environments.

3.4 TEEN

A CH node's power consumption is mainly affected by its distance from the B.S. In cases when the distance exceeds dc, the transmission power loss is represented by the multipath fading model $b\varepsilon_{mp}d^4$, whereas in other cases, the free space model $b_{\varepsilon fs}d^2$ is used. Assigning the amplifier low power processed in the empty-space model is denoted as fs, whereas it is abbreviated as *mp* in the multipath model. The following Equation represents the energy lost during the LEACH process.

$$E_T(b,d) = \{bE_{elec} + b_{\varepsilon fs}d^2, > dcbE_{elec} + b\varepsilon_{mp}d^4, d > dc \tag{3}$$

$$E_{T-elec}(b) = bE_{elec} \tag{4}$$

Algorithm 1: TEEN

Input:

b: Quantity of data bits to be sent

d: The separation between the C.H. and B.S. nodes

d_c: Distance that spans

E_{elec}: Power loss in electronic parts per bit

$b_{\varepsilon fs}$: Amplification efficiency in the free-space model

$b_{\varepsilon mp}$: Max power consumption of the amplifier in the multipath fading model per bit processed

Algorithm Overview:

In the case when d > dc: a. Use the multipath fading model to determine the energy dissipation

: $E_T(b,d) = \{bE_{elec} + b_{\varepsilon fs}d^2, > dcbE_{elec} + b\varepsilon_{mp}d^4, d > dc$

If the value of d is equal to or less than dc: 1. Determine the dissipation of energy by using the free space model.

: $E_{T-elec}(b) = bE_{elec}$

Return the total energy consumption.E_T

Outcome:

E_T: The total amount of energy lost

3.5 EAPTEEN

LEACH improves one reactive clusterering routing method: EAPTEEN. The data of every cluster is collected under its C.H. Following data fusion and processing, the BS or higher level C.Hs acquire the processed data. Data aggregation falls to the C.H.s using clustering routing systems; all nodes are only needed to broadcast data to their C.H. Node mobility impacts the EAPTEEN protocol by causing frequent changes in network topology, which can disrupt established routing paths and increase routing overheads as the protocol continuously updates routing tables. This mobility can challenge data delivery reliability, leading to difficulties in ensuring packets reach the base station.EAPTEEN optimizes energy usage by enabling sensor nodes to consistently send data at set intervals, regardless of event detection, and is adaptable to dynamic network changes such as node failures, topological adjustments, and energy variations.

In the EAPTEEN protocol, energy thresholds are dynamically adjusted in response to network changes through several mechanisms. When a node detects significant changes in its energy levels, such as depletion due to high communication or processing loads, it can trigger updates to its energy threshold settings. These thresholds are recalibrated based on current energy reserves, network traffic patterns, and node performance.

It keeps energy costs down to ensure energy consumption is distributed equitably, with each node taking turns being the C.H. A random selection process is conducted for the C.H. election. After forming a cluster, C.H.s schedules data transmission times for each cluster member.While the protocol is designed to enhance energy efficiency and ensure continuous data transmission, the specifics of how it adapts energy thresholds to maintain optimal network performance under varying conditions remain underexplored.

For the time being, for nodes to transmit sensed data to C.H., the H.T. value, which is derived from the range of user interest values, must exceed the present value of the sensed attribute. The Sensed Value (S.V.) is an internal variable that stores the detected characteristic. Data transmission begins only when the next value differs from S.V. by a total equal to the Sensed Threshold (S.T.); nodes constantly monitor the attribute.The protocol demonstrates improvements in energy efficiency and continuous monitoring; there is a lack of analysis on how EAPTEEN performs as the network size increases or in environments with a higher number of sensor nodes and more complex topologies.

The S.T. lowers the data transmission frequency by forgoing relatively small changes in perceived attributes. Users can adjust the value of S.T. to suit their requirements. Although it increases energy consumption, the network becomes more precise when a smaller S.T. is set. Therefore, to manage the energy economy vs accuracy trade-off, users need to change the size of S.T.A more comprehensive simulation setup, reflecting these diverse conditions, would provide a more accurate assessment of EAPTEEN's adaptability and effectiveness in practical applications.

After computing a threshold formula T(n), each node is given a random integer between 0 and 1. Once the random number falls below T(n), the node chooses to become the C.H. To get T(n), use Eq. (5).

$$T(n) = \{\frac{p}{1 - p^*\left[rmod\left(\frac{1}{p}\right)\right]}; n \in G0 \qquad otherwise \tag{5}$$

This Equation determines the threshold value $T(n)$ Used by each node to decide whether to become a C.H. It depends on a random number r, the probability p of a node becoming a C.H., and a correction factor. p^*. If a node's random number r falls below $T(n)$, it becomes a C.H.

This is where the near round number (r) and the proportion of nodes becoming C.H.s in each round (P) fit. In the final 1/p rounds, nodes not chosen as C.H. are grouped as G. To ensure that every node has an equal chance of being selected as C.H., the C.H. election procedure is repeated after each round. It makes sure that nodes use energy uniformly. The simulation environment details need to be more specific, including the network topology, node parameters such as density, initial energy, communication range, and data generation rates. The simulation settings should specify the duration, time steps, and any event-driven conditions (Fig. 2).

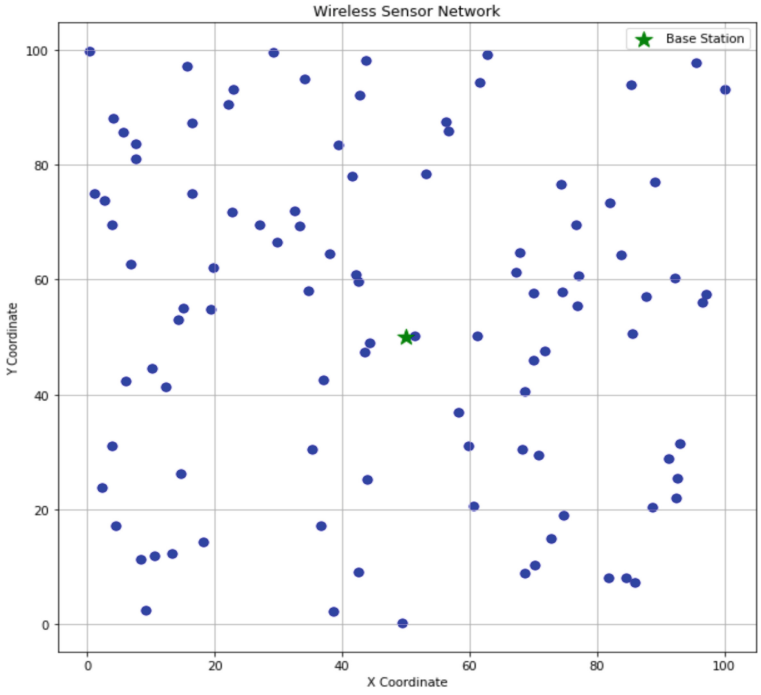


Fig. 2. EAPTEEN architecture

One of the drawbacks of the EAPTEEN routing protocol is C.H. random election. The C.H. election is marked in considerable part by unpredictability and lack of control. This paper presents the EAPTEEN protocol, a method for improving routing. Measuring the electromagnetic environment noise after the application needs helps one decide the bandwidth and available channels. When calculating the likelihood of a node becoming the cluster head, EAPTEEN considers the number of available channels. The possibility of a node becoming C.H. increases as its awareness of idle channels increases.

The EAPTEEN Method enhances energy efficiency and continuous monitoring in WSNs by improving the TEEN protocol. Unlike TEEN, which triggers data transmission based on thresholds, EAPTEEN allows nodes to send data at regular intervals, avoiding gaps in monitoring. It adapts to network changes, node failures, and energy variations. Experimental results show improvements in latency, energy consumption, and throughput.

$$P_i(t) = \{\frac{k}{N - k^*\left(r \, mod \, \frac{N}{k}\right)}; C_i(t) = 10 \quad ; C_i(t) = 0 \tag{6}$$

This Equation calculates the probability that node i becomes a C.H. at time t. It depends on the number of available channels. $C_i(t)$, the total number of nodes N, and the parameter k^*. If $C_i(t) = 1$ (indicating available channels), $P_i(t)$ It is computed based on some of the nodes k and r.

We can predict the quantity of C.H.s in the network (N) and the number of nodes (k) in the network (k) using the current round and these variables in Eq. (6). The probability that node i will be selected as C.H. is only connected to the projected quantity of C.H.s and the total digit of nodes.

$$P_i(t) = min\left(ka\frac{C_i}{m}, 1\right) \tag{7}$$

To determine if a node is a cluster member or cluster leader in CWSN, the EAPTEEN method employs Eq. (7). The CH that detects a more significant number of open channels can then set up communication pathways with other cluster members.

Algorithm 2: EAPTEEN

START

- N: Set of sensor nodes
- T: Event threshold
- E_{th}: Energy threshold for cluster head selection
- E_{min}: Minimum energy required for a node to operate
- P: Period for periodic data transmission
- d_{th}: Distance threshold for cluster formation
- C: Set of clusters
- $E_{initial}$: Initial energy level of each sensor node
    - Assign initial energy levels to each sensor node N_i from$E_{initial}$.
    - Set $C = \{\}$ (empty set).
    - For each sensor node N_i with energy E_i:
        - If$E_i > E_{th}$, designate N_i as a CH.
        - Form a cluster C_j with N_i as the cluster head and neighboring nodes within d_{th} distance as members.
        - Update$C = C \cup C_j$.
    - Each sensor node $N_i$monitors its environment for events.
    - If an event occurs or a predefined threshold T is exceeded, the sensor node N_i Collects data.
    - For each sensor node N_i in a cluster C_j:
        - If the time for periodic transmission P has elapsed:
            - N_i Transmits data to its CH.
            - Cluster head aggregates received data from component nodes.
    - Each CH C_j:
        - Aggregates data received from component nodes.
        - Compresses aggregated data to reduce transmission overhead.
        - Transmits aggregated data to the base station BS.
    - Monitor energy levels of sensor nodes and CH.
    - If a cluster head's energy drops below E_{th}:
        - Re-elect a new cluster head within the cluster.
    - If a sensor node's energy drops below E_{min}:
        - Exclude it from cluster formation and update cluster membership accordingly.

END

- C: Set of clusters with elected cluster heads
- Data transmission from SNs to CHs and from CHs to the base station

The EAPTEEN protocol faces several security implications, including risks to data integrity, as it can lack robust mechanisms to prevent data tampering or injection attacks. Its reliance on periodic data transmission can make it vulnerable to replay attacks, where intercepted data is retransmitted maliciously. Compromised nodes could manipulate energy thresholds and routing decisions, leading to network disruption or unauthorized data access. Additionally, inadequate authentication and authorization could allow unauthorized nodes to infiltrate the network, while denial of service (DoS) attacks could

exploit the protocol's energy efficiency focus to deplete node resources. Privacy concerns also arise if the protocol does not incorporate encryption or privacy measures, potentially exposing sensitive data to eavesdroppers.

4 Results and Discussion

This section presents the simulation results of the EAPTEEN protocol and evaluates it against the existing MPTCP and SEHR routing protocols using Network Simulator (NS-2.34). We analyze the outcomes obtained from experiments to assess its performance in optimizing energy consumption while ensuring uninterrupted monitoring in event-driven Wireless Sensor Networks.

$$\text{Throughput} = \frac{Number\ of\ Packet\ Size}{Time\ duration * Successful\ average\ Packet\ size} \tag{8}$$

Table 2. Throughput comparison table

Packet Size	Throughput levels		
	MPTCP	SEHR	EAPTEEN
50	0.238	0.256	0.384
100	0.476	0.512	0.769
150	0.714	0.769	1.153
200	0.952	1.025	1.538
250	1.190	1.282	1.923

Table 2 and Fig. 3 illustrate that across packet sizes, MPTCP, SEHR, and EAPTEEN protocols have different performance trends. Starting at 0.238 for a packet size of 50 and consistently rising to 1.190 for a packet size of 250, MPTCP shows initially poorer throughput. Beginning with somewhat greater throughput than MPTCP at 0.256 for a packet size of 50, SEHR peaks at 1.282 for a packet size of 250. Starting at 0.384 for a packet size of 50 and to 1.923 for a packet size of 250 EAPTEEN constantly shows the best throughput values. These results show that when packet size rises, all systems get better throughput.

$$\text{Energy} = \frac{Number\ of\ Sensor\ nodes}{Energy\ consumption\ for\ sending\ packets\ at\ a\ times} x100 \tag{9}$$

Table 3 and Fig. 4 show the energy levels in joules for MPTCP, SEHR, and EAPTEEN protocols across varying numbers of nodes, hence illuminating their energy consumption efficiency. MPTCP uses 62.5 J from 10 nodes, and increases to 625 J with 100 nodes. Starting with a much lower energy usage of 58.8 J when there are 10 nodes, SEHR rises similarly to 589 J when there are 100 nodes. Beginning at 54 J with 10 nodes and scaling

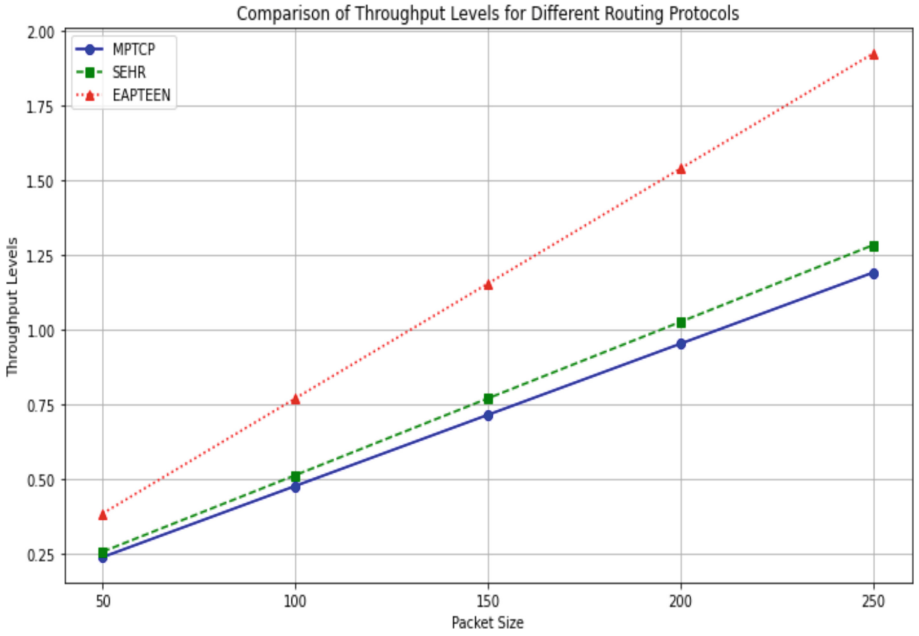


Fig. 3. Throughput value comparison chart

Table 3. Energy

Number of Nodes	The energy level in joules		
	MPTCP	SEHR	EAPTEEN
10	62.5	58.8	54
20	125	117	108
40	250	235	216
60	375	352	324
80	500	470	432
100	625	588	540

up to 540 J with 100 nodes, EAPTEEN constantly displays rather effective behavior. The findings show that EAPTEEN improves energy management over MPTCP and SEHR. This makes it a great option for settings needing effective energy consumption, including wireless sensor networks for continuous monitoring.

$$\text{Time Delay} = \frac{\textit{Number of Sensor nodes}}{\textit{energy consumption for sending packets at a times x forwarding time in ms}} \quad (10)$$

Table 4 and Fig. 5 show, over varying numbers of nodes, the end-to--end time delay in milliseconds for MPTCP, SEHR, and EAPTEEN protocols, hence illuminating their

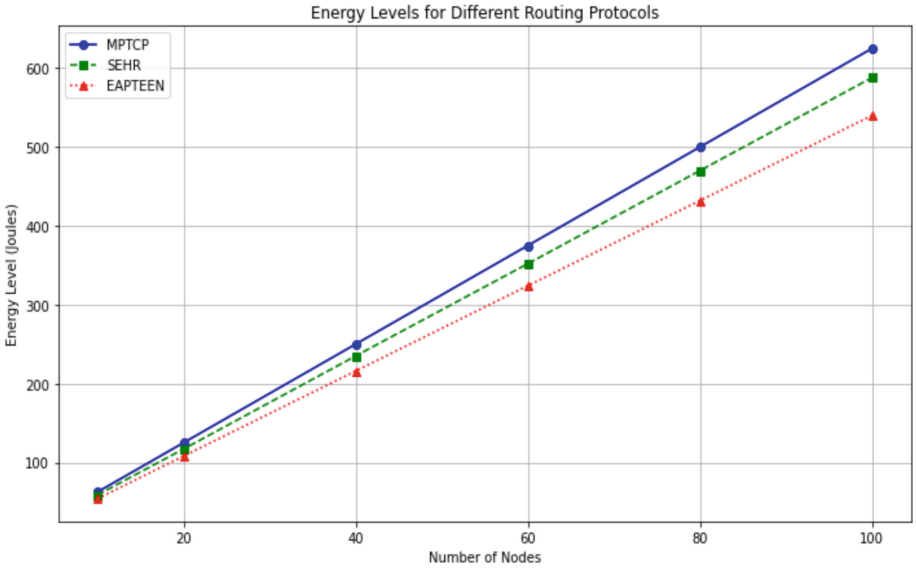


Fig. 4. Energy value comparison chart

Table 4. End to End delay comparison

Number of Nodes	End to End Time Delay (ms)		
	MPTCP	SEHR	EAPTEEN
10	0.058	0.057	0.056
20	0.116	0.114	0.112
40	0.232	0.228	0.225
60	0.348	0.342	0.337
80	0.465	0.457	0.450
100	0.581	0.571	0.563

latency performance. MPTCP's latency climbs from 0.058 ms to 0.581 ms progressively as the number of nodes increases. Starting with 10 nodes, SEHR exhibits a little decrease in latency at 0.057 ms; thereafter, scaling up to 100 nodes generates a similar increase in delay to 0.571 ms. Starting at 0.056 ms with 10 nodes, EAPTEEN routinely demonstrates the lowest time delay—only slightly increasing to 0.563 ms with 100 nodes. The results show, independent of node count, EAPTEEN routinely offers shorter latency than MPTCP and SEHR. This suggests EAPTEEN is suitable for applications in wireless sensor networks requiring low-latency communication.

$$\text{PDR} = \frac{Total\ Number\ of\ Packets\ Receives}{Total\ Packet} * 100 \tag{11}$$

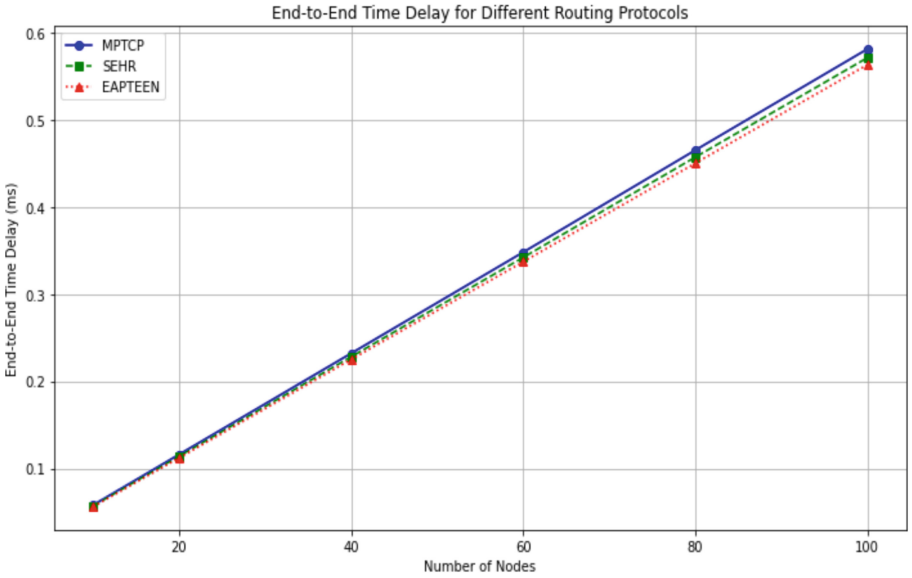


Fig. 5. End-to-end delay value comparison chart

Table 5. Packet delivery ratio

Number of packets	Packet Delivery Ratio		
	MPTCP	SEHR	EAPTEEN
50	97.2	98.2	98.6
100	98.6	99.1	99.3
150	99	99.4	99.53
200	99.3	99.5	99.65
250	99.4	99.64	99.72

Table 5 and Fig. 6 show the PDR for MPTCP, SEHR, and EAPTEEN protocols at many packet volumes, thereby highlighting their remarkable capacity to transport data without loss. Beginning with a PDR of 97.2% for 50 packets, the MPTCP rises steadily to 99.4% for 250 packets. For 50 packets, SEHR demonstrates improved starting performance with a PDR of 98.2%; it rises to 99.64% for 250 packets. EAPTEEN consistently obtains exceptional packet delivery rates from 98.6% for 50 packets to 99.72% for 250. These results show how constantly all approaches retain high delivery rates as packet count increases. EAPTEEN guarantees the highest reliability in data transmission, so it is ideal for uses needing high packet delivery in WSNs.

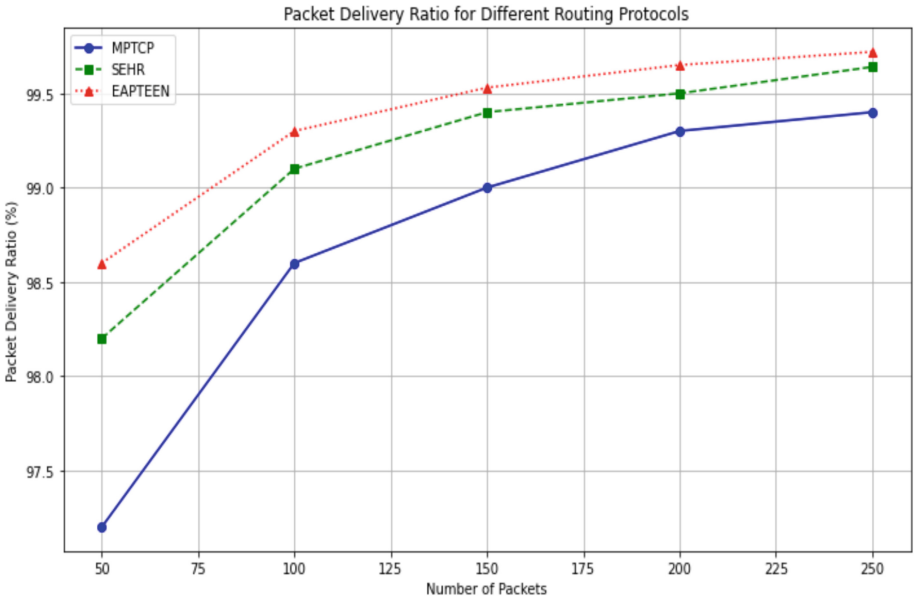


Fig. 6. PDR chart

5 Conclusion

The EAPTEEN protocol is a significant development in energy-aware adaptive routing overall for event-driven WSNs. EAPTEEN saves energy by way of constant monitoring guaranteed by periodic data transmission and dynamic adaptation to network conditions. It solves the flaws of TEEN's and other traditional methods. By allowing the network to identify effective routing methods and maximize energy use among nodes, its lifetime and dependability are much enhanced. This is very necessary for guaranteeing ongoing performance in challenging surroundings. Our approach guarantees dependable performance and long-term durability in real-world implementations, therefore greatly improving the efficiency and efficacy of WSNs in applications including continuous monitoring. Furthermore, investigating the scalability of EAPTEEN in large-scale WSN installations and its adaptability for different sensor settings might increase its relevance in many IoT uses.

References

1. Al-Turjman, F., Kilic, K.I.: LaGOON: a simple energy-aware routing protocol for wireless nano-sensor networks. IET Wirel. Sens. Syst. **9**(3), 110–118 (2019)
2. Balakiruthiga, B., Deepalakshmi, P., Mohanty, S.N., Gupta, D., Kumar, P.P., Shankar, K.: **Segment** routing based energy aware routing for software defined data center. Cogn. Syst. Res. **64**, 146–163 (2020)
3. Ekler, P., Levendovszky, J., Pasztor, D.: Energy aware IoT routing algorithms in smart city environment. IEEE Access **10**, 87733–87744 (2022)

4. Giordano, M., Cortesi, S., Mekikis, P.V., Crabolu, M., Bellusci, G., Magno, M.: Energy-aware adaptable sampling for self-sustainability in resource-constrained IoT devices. In: Proceedings of the 11th International Workshop on Energy Harvesting & Energy-Neutral Sensing Systems, pp. 65–71, November 2023
5. Han, Y., Hu, H., Guo, Y.: Energy-aware and trust-based secure routing protocol for wireless sensor networks using adaptable genetic algorithm. IEEE Access **10**, 11538–11550 (2022)
6. Haseeb, K., Almustafa, K.M., Jan, Z., Saba, T., Tariq, U.: Secure and energy-aware heuristic routing protocol for wireless sensor network. IEEE Access **8**, 163962–163974 (2020)
7. Jecan, E., Pop, C., Ratiu, O., Puschita, E.: Predictive energy-aware routing solution for industrial IoT evaluated on a WSN hardware platform. Sensors **22**(6), 2107 (2022)
8. Kanoun, O., et al.: Energy-aware system design for autonomous wireless sensor nodes: a comprehensive review. Sensors **21**(2), 548 (2021)
9. Li, N., Seah, W. K., Hou, Z., Jia, B., Huang, B., Li, W.: An energy aware adaptable clustering protocol for energy harvesting wireless sensor networks. In: Proceedings of the 18th International Symposium on Spatial and Temporal Data, pp. 161–170, August 2023
10. Marchese, M., Patrone, F.: E-CGR: Energy-aware contact graph routing over nanosatellite networks. IEEE Trans. Green Commun. Network. **4**(3), 890–902 (2020)
11. Mothku, S.K., Rout, R.R.: Adaptable fuzzy-based energy and delay-aware routing protocol for a heterogeneous sensor network. J. Comput. Netw. Commun. **2019**(1), 3237623 (2019)
12. Musaddiq, A., Zikria, Y.B., Kim, S.W.: Energy-aware adaptable trickle timer algorithm for RPL-based routing in the Internet of Things. In: 2018 28th international telecommunication networks and applications conference (ITNAC), pp. 1–6. IEEE, November 2018
13. Qureshi, K.N., Bashir, M.U., Lloret, J., Leon, A.: Optimized cluster-based dynamic energy-aware routing protocol for wireless sensor networks in agriculture precision. J. Sens. **2020**(1), 9040395 (2020)
14. Sangaiah, A.K., et al.: Energy-aware geographic routing for real-time workforce monitoring in industrial informatics. IEEE Internet Things J. **8**(12), 9753–9762 (2021)
15. Selem, E., Fatehy, M., Abd El-Kader, S.M.: MobTHE (mobile temperature heterogeneity energy) aware routing protocol for WBAN IoT health application. IEEE Access **9**, 18692–18705 (2021)
16. Subramani, N., Perumal, S.K., Kallimani, J.S., Ulaganathan, S., Bhargava, S., Meckanizi, S.: Controlling energy aware clustering and multi-hop routing protocol for IoT assisted wireless sensor networks. Concurrency Comput. Pract. Experience **34**(21), e7106 (2022)
17. Sundarasekar, R., et al.: Adaptable energy aware quality of service for reliable data transfer in underwater acoustic sensor networks. IEEE Access **7**, 80093–80103 (2019)
18. Shanmugavalli, R., Subashini, P.: Energy aware routing using AODV protocol for low energy consumption in WSN. ADBU J. Eng. Technol. **13**(1) (2024)
19. Muthulakshmi, K., PrakashBalaji, S., Stephe, S., Vijayalakshmi, J.: Adaptable wind driven optimization based energy aware clustering scheme for wireless sensor networks. Tehničkivjesnik **31**(2), 466–473 (2024)
20. Savitha, M., Jayanthi, B.: Novel enhanced power-efficient gathering in sensor information systems (Nepegis) for energy aware data routing in IotWsn. Educ. Admin. Theor. Pract. **30**(5), 13040–13047 (2024)
21. Lei, C.: An energy-aware cluster-based routing in the Internet of things using particle swarm optimization algorithm and fuzzy clustering. J. Eng. Appl. Sci. **71**(1), 135 (2024)
22. Janarthanan, A., Srinivasan, V.: Multi-objective cluster head-based energy aware routing using optimized auto-metric graph neural network for secured data aggregation in wireless sensor network. Int. J. Commun. Syst. **37**(3), e5664 (2024)
23. Mythili, D., Duraisamy, S.: Coswarm-LEACH-optimal throughput and energy efficiency routing protocol for wireless sensor networks. Naturalista Campano **28**(1), 1833–1842 (2024)

24. Bria, R., Wahab, A., Alaydrus, M.: Energy efficiency analysis of TEEN routing protocol with isolated nodes. In: 2019 Fourth International Conference on Informatics and Computing (ICIC), pp. 1–5. IEEE, October 2019

Author Index

A
Acharya, Anuradha 161
Agrawal, Shikha 183
Aldarwish, Abdulla J. Y. 235
Ananth, Viswanath 309

B
Bakesvar, B. Thriam 375
Balan, Yashwanth 422
Bhatt, Meghna 69
Bhogayata, Ankur 212

C
Charuhasini, T. M. 289
Chitra, R. 276

D
Dakhane, Dhananjay 433
Deekshitha, V. 309
Dubey, Shirish Mohan 183
Durga, N. 145

E
Eniyavasanthan, R. 375

G
Gupta, Kamlesh 183

H
Hanirex, D. Kerana 132
Hlomani, Hlomani 95

I
Ilyas, F. Mohamed 26

J
Jariwala, Nikisha 247

K
Kannan, Rathimala 261
Karthika, N. 411
Kaur, Pankaj Deep 335
Kaur, Tarandeep 335
Kavitha, D. 411
KeranaHanirex, D. 145
Kumari, Kiran 433

M
Mahajan, Lochana 3
Meenakshi, C. 57
Meva, Divyakant 212
Mokgetse, Tshepiso L. 95
Munusamy, Nagarajan 444
Muthukumaravel, A. 145, 322

N
Nagarajan, V. 57

P
Patel, Kalyani 235
Perumal, Arun Pandiyan 365
Popat, Kalpesh 40, 247
Prema, R. 26
Priscila, S. Silvia 26
Priya, A. Shanmuga 351

R
Radhika, C. 132
Radhika, R. 322
Rajagopal, Sridaran 3, 69
Rajan, Arul 289
Rajan, R. Arokia Paul 422
Rajgopal, Sridaran 121
Ramakrishnan, Kannan 261
Rangwala, Husain 195
Rani, D. Jeya 444
Raval, Kinjal 69

© The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
S. Rajagopal et al. (Eds.): ASCIS 2024, CCIS 2428, pp. 461–462, 2025.
https://doi.org/10.1007/978-3-031-86302-8

S
Saajan, J. Subash 375
Sandosh, S. 375
Sathish, M. 276
Schmuck, Roland 391
Shah, Rujuta 3
Sheela, K. 26
Sini, V. 121
Solanki, Jitesh 40
Sona, K. V. 26
Sridaran, Rajagopal 95
Srividya, V. 107
Stervin, Jone 276
Sujatha, R. 309
Susana, D. 107

T
Thakkar, Kenil 161
Thoriya, Amit 212

U
Uma Maheshwari, B. 309

V
Vaitheki, K. 351
Vijaykumar, N. 289
Vimala Roselin, J. 26
Vora, Tarak 195, 212

Y
Yaseen, Aqeel A. 235
Yassin, Ali A. 235