Smart Innovation, Systems and Technologies 251

Tomonobu Senjyu
Parakshit Mahalle
Thinagaran Perumal
Amit Joshi *Editors*

IOT with Smart Systems

Proceedings of ICTIS 2021, Volume 2

KES International

Springer

Smart Innovation, Systems and Technologies

Volume 251

Series Editors

Robert J. Howlett, Bournemouth University and KES International, Shoreham-by-Sea, UK

Lakhmi C. Jain, KES International, Shoreham-by-Sea, UK

The Smart Innovation, Systems and Technologies book series encompasses the topics of knowledge, intelligence, innovation and sustainability. The aim of the series is to make available a platform for the publication of books on all aspects of single and multi-disciplinary research on these themes in order to make the latest results available in a readily-accessible form. Volumes on interdisciplinary research combining two or more of these areas is particularly sought.

The series covers systems and paradigms that employ knowledge and intelligence in a broad sense. Its scope is systems having embedded knowledge and intelligence, which may be applied to the solution of world problems in industry, the environment and the community. It also focusses on the knowledge-transfer methodologies and innovation strategies employed to make this happen effectively. The combination of intelligent systems tools and a broad range of applications introduces a need for a synergy of disciplines from science, technology, business and the humanities. The series will include conference proceedings, edited collections, monographs, handbooks, reference books, and other relevant types of book in areas of science and technology where smart systems and technologies can offer innovative solutions.

High quality content is an essential feature for all book proposals accepted for the series. It is expected that editors of all accepted volumes will ensure that contributions are subjected to an appropriate level of reviewing process and adhere to KES quality principles.

Indexed by SCOPUS, EI Compendex, INSPEC, WTI Frankfurt eG, zbMATH, Japanese Science and Technology Agency (JST), SCImago, DBLP.

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

More information about this series at https://link.springer.com/bookseries/8767

Tomonobu Senjyu · Parakshit Mahalle ·
Thinagaran Perumal · Amit Joshi
Editors

IOT with Smart Systems

Proceedings of ICTIS 2021, Volume 2

Springer

Editors
Tomonobu Senjyu
University of the Ryukyus
Nishihara, Okinawa, Japan

Parakshit Mahalle
Sinhgad Technical Education Society
SKNCOE, Pune, India

Thinagaran Perumal
University Putra Malaysia Serdang
Serdang, Malaysia

Amit Joshi
Global Knowledge Research Foundation
Ahmedabad, India

ISSN 2190-3018 ISSN 2190-3026 (electronic)
Smart Innovation, Systems and Technologies
ISBN 978-981-16-3944-9 ISBN 978-981-16-3945-6 (eBook)
https://doi.org/10.1007/978-981-16-3945-6

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

Preface

The Fifth International Conference on Information and Communication Technology for Intelligent Systems (ICTIS 2021) targets state-of-the-art as well as emerging topics pertaining to information and communication technologies (ICTs) and effective strategies for its implementation for engineering and intelligent applications.

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

The conference was held during April 23–24, 2021, digitally on Zoom and organized by Global Knowledge Research Foundation in collaboration with KCCI and IFIP INTERYIT.

Research submissions in various advanced technology areas were received, and after a rigorous peer-review process with the help of program committee members and external reviewer, 160 papers were accepted with an acceptance rate of 16%. All 160 papers of the conference are accommodated in 2 volumes, and also papers in the book comprise authors from 18 countries.

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

We are overwhelmed by our distinguished scholars and appreciate them for accepting our invitation to join us through the virtual platform and deliver keynote speeches and technical session chairs for analyzing the research work presented by the researchers. Most importantly, we are also grateful to our local support team for their hard work for the conference. This series has already been made a continuous series which will be hosted at different location every year.

Nishihara, Japan — Tomonobu Senjyu
Pune, India — Parakshit Mahalle
Serdang, Malaysia — Thinagaran Perumal
Ahmedabad, India — Amit Joshi
April 23–24, 2021

Contents

About the Editors

Dr. Tomonobu Senjyu received his B.S. and M.S. degrees in Electrical Engineering from the University of the Ryukyus in 1986 and 1988, respectively, and his Ph.D. degree in Electrical Engineering from Nagoya University in 1994. Since 1988, he has been with the Department of Electrical and Electronics Engineering, University of the Ryukyus, where he is currently Professor. His research interests include stability of AC machines, power system optimization and operation, advanced control of electrical machines and power electronics. He is Member of the Institute of Electrical Engineers of Japan and IEEE.

Dr. Parakshit Mahalle holds a B.E. degree in CSE and an M.E. degree in Computer Engineering. He completed his Ph.D. at Aalborg University, Denmark. Currently, he is working as Professor and Head of the Department of Computer Engineering at STES Smt. Kashibai Navale College of Engineering, Pune, India. He has over 18 years of teaching and research experience. Dr. Mahalle has published over 140 research articles and eight books and has edited three books. He received the "Best Faculty Award" from STES and Cognizant Technologies Solutions.

Dr. Thinagaran Perumal received his B.Eng., M.Sc. and Ph.D. Smart Technologies and Robotics from Universiti Putra Malaysia in 2003, 2006 and 2011, respectively. Currently, he is Associate Professor at Universiti Putra Malaysia. He is also Chairman of the TC16 IoT and Application WG National Standard Committee and Chair of IEEE Consumer Electronics Society Malaysia Chapter. Dr. Thinagaran Perumal is Recipient of 2014 IEEE Early Career Award from IEEE Consumer Electronics Society. His recent research activities include proactive architecture for IoT systems, development of the cognitive IoT frameworks for smart homes and wearable devices for rehabilitation purposes.

Dr. Amit Joshi is currently Director of the Global Knowledge Research Foundation. He is also Entrepreneur and Researcher, and he holds B.Tech., M.Tech. and Ph.D. degrees. His current research focuses on cloud computing and cryptography. He is Active Member of ACM, IEEE, CSI, AMIE, IACSIT, Singapore, IDES, ACEEE, NPA and several other professional societies. He is also International Chair of InterYIT at the International Federation of Information Processing (IFIP, Austria). He has published more than 50 research papers, edited 40 books and organized over 40 national and international conferences and workshops through ACM, Springer and IEEE across various countries including India, Thailand, Egypt and Europe.

Chapter 1
A Theoretical Approach of Information Measure for Pythagorean Fuzzy Sets

Anjali Munde and H. D. Arora

Abstract Information measures in a fuzzy set compute the amount of vagueness showed in the fuzzy set. The notion of Pythagorean fuzzy sets is pertinent in assessment for its distinctive description of vagueness. Pythagorean fuzzy sets were primarily established through Yager (Yager and Abbasov in Int J Intell Syst 28:436–452, 2013; Yager in Pythagorean fuzzy subsets. Proceedings of the 9th Joint World Congress on Fuzzy Systems and NAFIPS Annual Meeting, IFSA/NAFIPS Edmonton, Canada, pp. 57–61, 2013). The theory was particularly aimed to denote improbability and fuzziness in measurable manner and to deliver a validated instrument for confronting indistinctness to existent obstructions. The foremost characteristic of Pythagorean fuzzy sets is described through three factors specifically, association quantity, non-association quantity, and undetermined quantity, so the total of the square of every factor is one. In the existing paper, an information measure for the Pythagorean fuzzy set has been propositioned including the verification of its rationality. The monotonic performance of the anticipated information measure has been examined and displayed.

1.1 Introduction

The notion of fuzzy arrays was initially propositioned by Zadeh [13] in 1965. Through an extensive application in numerous areas, fuzzy arrays not only offer comprehensive chance to compute ambiguities in effective and rational approach, but also provide with a significant technique to signify imprecise notions in usual language. It is established that maximum techniques created on 'crisp array theory' are in some way hard for managing unclear and ambiguous info. In this view, fuzzy arrays can be consumed to impart enhanced results for additional physical domain setbacks.

A. Munde (✉)
Amity College of Commerce and Finance, Amity University, Uttar Pradesh, Noida, India

H. D. Arora
Amity Institute of Applied Sciences, Amity University, Uttar Pradesh, Noida, India
e-mail: hdaroar@amity.edu

T. Senjyu et al. (eds.), *IOT with Smart Systems*, Smart Innovation, Systems and Technologies 251, https://doi.org/10.1007/978-981-16-3945-6_1

Numerous findings on the use of fuzzy arrays have been approved. Through the enormous inaccurate and ambiguous info in the physical domain, diverse augmentations of fuzzy array have been established by certain investigators.

While fuzzy arrays were created on association quantities or measures among 0 and 1, in actual time positioning it could not be constantly accurate that non-association quantity is equal to (1-association). Consequently, to acquire additional determined consistency and usability, Atanassov [1] oversimplified the notion of 'fuzzy array theory' and suggested intuitionistic fuzzy sets (IFSs) which comprise together association quantity and non-association quantity and quantity of non-determinacy or vagueness where quantity of vagueness = (1-(quantity of association + non-association quantity)). In IFSs, the sets of association classes are indicated by (μ, ω]) fulfilling the prerequisite of $\mu + \omega] \leq 1$.

The foremost gain of the IFS is its property to manage through the uncertainty that can happen owing to vague data. This is accomplished through integrating another function, named a non-association function, besides the association function, of the predictable fuzzy array. The summary of IFSs and the succeeding mechanism on the basics of IFSs, a portion of considerations have been given on emerging vastness and connection degrees among IFSs, as a method to use numerous complications of decisiveness and model identification.

The concept of IFSs is tough, in situations where $\mu + \omega \geq 1$ different from the condition attained in IFSs where $\mu + \omega \leq 1$. This restriction in IFS theory is the rationale for the induction of Pythagorean fuzzy sets (PFSs). Pythagorean fuzzy set (PFS) recommended in Yager [9, 11, 12] is a novel instrument to trade through imprecision studying the association degree, μ, and non-association degree, ω, fulfilling the prerequisite $\mu + \omega \geq 1$. Because of being a simplified array, PFS has similar association with IFS. The notion of PFSs may be consumed to describe indeterminate data further, adequately and precisely than IFSs. From the time when it started, the principle of PFSs has been comprehensively calculated.

Shannon [5] examined the concept of entropy.

Peng [4] proposed Pythagorean fuzzy information measures and their applications along with Hung [2] and Mendel [3] who proposed fuzzy entropy on intuitionistic fuzzy sets. Torra [6] examined hesitant fuzzy sets, and Turksen [7] introduced interval valued fuzzy sets based on normal forms. Wang [8] introduced intuitionistic linguistic fuzzy multicriteria decision-making method based on intuitionistic fuzzy entropy.

1.1.1 Preliminaries

Pythagorean Fuzzy Array: Assume X as a universal array. A Pythagorean fuzzy array P that is defined as an array of systematic arrays above X is outlined through the subsequent:

$$P = \{x, \mu_P(x), \omega_P(x) | x \in X\}$$

where the functions

$$\mu_P(x) : X \to [0, 1] \quad \text{and} \quad \omega_P(x) : X \to [0, 1]$$

describe the amount of association and the amount of non-association, correspondingly, of the part $x \in X$ to A, that is, a subclass of X, and aimed at $x \in X$:

$$0 \le (\mu_P(x))^2 + (\omega_P(x))^2 \le 1 \tag{1.1}$$

Assuming $(\mu_P(x))^2 + (\omega_P(x))^2 \le 1$, next there is an amount of indeterminacy of $x \in X$ to P described through $\pi_P(x) = \sqrt{1 - \left[(\mu_P(x))^2 + (\omega_P(x))^2\right]}$ and $\pi_P(x) \in [0, 1]$. Further, $(\mu_P(x))^2 + (\omega_P(x))^2 + (\pi_P(x))^2 = 1$. Otherwise, $\pi_P(x) = 0$ whenever $(\mu_P(x))^2 + (\omega_P(x))^2 = 1$. The array of all PFS over X is represented through PFS(X).

In this paper, an information measure for the Pythagorean fuzzy set has been propositioned including the proof of its validity. The monotonic performance of the anticipated information measure has been examined and displayed.

The manuscript is structured by describing certain mathematical introductions of fuzzy arrays and IFSs in Sect. 1.1–1.1. In Sects. 1.1–1.2, a parametric information measure for the Pythagorean fuzzy set has been recommended, while Sects. 1.1–1.3 verify the properties and Sects. 1.1–1.4 establish the monotonic performance of the anticipated information measure. Lastly, Sects. 1.1–1.5 recapitulate the developed conclusions of the paper and certain beneficial deductions are extracted.

1.1.2 New Information Measure for the Pythagorean Fuzzy Set

A new information measure for Pythagorean fuzzy set represented by $\boldsymbol{\theta_\alpha^\beta(P)}$ is proposed as follows:

$$\theta_\beta^\alpha(P) = \frac{1}{(1-\alpha)\beta} \sum_{i=1}^{n} \left[\left(\mu_P^2(x_i)^{\alpha\mu_P(x_i)} + \omega_P^2(x_i)^{\alpha\mu_P(x_i)} + \pi_P^2(x_i)^{\alpha\mu_P(x_i)}\right)^{\beta}) - 3^{\beta} \right]$$

where

$$\alpha > 0, \quad \alpha \ne 1, \quad \text{and} \quad \beta > 0 \tag{1.2}$$

1.1.3 Properties of $\theta_\alpha^\beta(P)$

We have supposed that $0^{0\cdot\alpha} = 1$; we study the following properties:

Property 1 $\boldsymbol{\theta_\alpha^\beta}(\mathbf{P}) \geq 0$, i.e., $\boldsymbol{\theta_\alpha^\beta}(\mathbf{P})$ *is nonnegative.*

Property 2 *SHARPNESS*

$\boldsymbol{\theta_\alpha^\beta}(\mathbf{P})$ is minimum iff A is a non-fuzzy set.

For $\mu_P(x_i) = 0,\ _P(x_i) = 0_P(x_i) = 1$, *it implies* $\boldsymbol{\theta_\alpha^\beta}(\mathbf{P}) = 0$.
For $\mu_P(x_i) = 0,\ _P(x_i) = 1_P(x_i) = 0$,*it implies* $\boldsymbol{\theta_\alpha^\beta}(\mathbf{P}) = 0$.
For $\mu_P(x_i) = 1,\ _P(x_i) = 0_P(x_i) = 0$, *it implies* $\boldsymbol{\theta_\alpha^\beta}(\mathbf{P}) = 0$.

Property 3 *SYMMETRY*

Since the compliment of $P = \{\langle x_i, \mu_P(x_i), \nu_P(x_i), \pi_P(x_i)\rangle | x_i \in X\}$ *is* $P^c = \{\langle x_i, \nu_P(x_i), \mu_P(x_i), 1 - \mu_P(x_i) - \nu_P(x_i), \rangle | x_i \in X\}$, *i.e.,* $[\mu_P(x_i)]^c = \nu_P(x_i)$; $[\nu_P(x_i)]^c = \mu_P(x_i)$ *and* $[\pi_P(x_i)]^c = 1 - \pi_P(x_i)$ *for all* $x_i \in X$, *then we have*

$$\boldsymbol{\theta_\alpha^\beta}(\mathbf{P}) = \theta_\alpha^\beta\left(P^c\right)$$

Property 4 $\theta_\alpha^\beta(P^*) \leq \theta_\alpha^\beta(P)$, *where* P^* *is sharpened version of* P.

When $\mu_P(x_i) = \frac{1}{2}$ *and* $\mu_P(x_i)$ *lies between 0 and* $\frac{1}{2}$ *then* $\theta_\alpha^\beta(P)$ *is an increasing function, whereas when* $\mu_P(x_i)$ *lies between* $\frac{1}{2}$ *and* 1 *then* $\theta_\alpha^\beta(P)$ *is a decreasing function of* $\mu_P(x_i)$.

Let P^* *be sharpened version of P which means that*

If $\mu_P(x_i) < 0.5$, *then* $\mu_P * (x_i) \leq \mu_P(x_i)$ *for all* $i = 1, 2, \ldots, n$
If $\mu_P(x_i) > 0.5$, *then* $\mu_P * (x_i) \geq \mu_P(x_i)$ *for all* $i = 1, 2, \ldots, n$

Since $\theta_\alpha^\beta(P)$ *is an increasing function of* $\mu_P(x_i)$ *for* $0 \leq \mu_P(x_i) \leq \frac{1}{2}$ *and decreasing function of* $\mu_P(x_i)$ *for* $\frac{1}{2} \leq \mu_P(x_i) \leq 1$,

$\mu_P * (x_i) \leq \mu_P(x_i)$ *this implies* $\theta_\alpha^\beta(P^*) \leq \theta_\alpha^\beta(P)$ *in* [0,0.5]
$\mu_P * (x_i) \leq \mu_P(x_i)$ *this implies* $\theta_\alpha^\beta(P^*) \leq \theta_\alpha^\beta(P)$ *in* [0.5,1]

Hence, $\theta_\alpha^\beta(P^*) \leq \theta_\alpha^\beta(P)$.

1.1.4 Monotonicity of Proposed Parametric Information Measure for the Pythagorean Fuzzy Set Involving Two Parameters

The monotonicity of the Pythagorean fuzzy measure is verified with the help of the following examples.

Table 1.1 Value of $\theta_\alpha^\beta(P)$ for $\alpha = 1.1$ and $\beta = 2$

$\mu_P(x_i)$	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
$\theta_\alpha^\beta(P)$	43.184	43.787	43.979	44.003	43.911	43.702	43.333	42.720	41.705

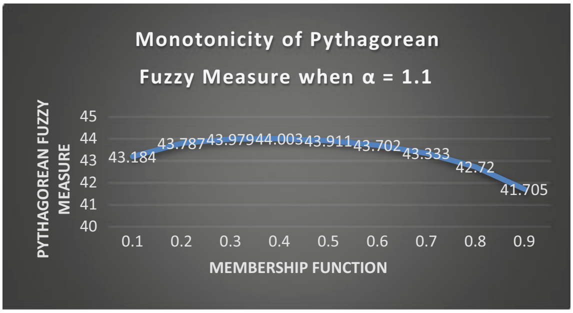
Table 1.2 Value of $\theta_\alpha^\beta(P)$ for $\alpha = 1.2$ and $\beta = 2$

$\mu_P(x_i)$	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
$\theta_\alpha^\beta(P)$	21.672	21.966	22.058	22.069	22.026	21.925	21.745	21.438	20.914

Table 1.3 Value of $\theta_\alpha^\beta(P)$ for $\alpha = 1.3$ and $\beta = 2$

$\mu_P(x_i)$	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
$\theta_\alpha^\beta(P)$	14.496	14.687	14.745	14.752	14.725	14.661	14.545	14.341	13.982

Fig. 1.1 Monotonic behavior of $\theta_\alpha^\beta(P)$ for $\alpha = 1.1$ and $\beta = 2$

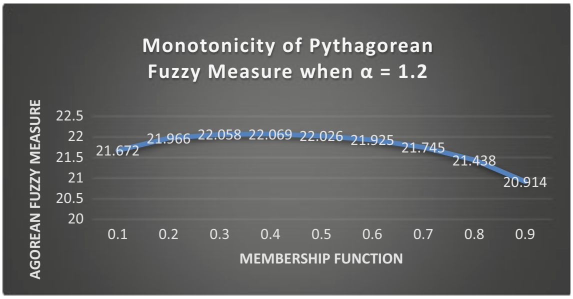


The value of $\theta_\alpha^\beta(P)$ for different values of α and β are given below in the tables namely, Tables 1.1, 1.2, and 1.3 and presented in the form of figures namely, Figs. 1.1, 1.2, and 1.3.

When $\beta = 2$, $\alpha = 1.1$.

When $\beta = 2$, $\alpha = 1.2$.

Fig. 1.2 Monotonic behavior of $\theta_\alpha^\beta(P)$ for $\alpha = 1.2$ and $\beta = 2$



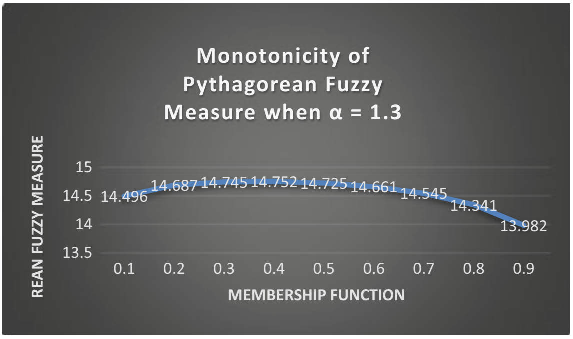


Fig. 1.3 Monotonic behavior of $\theta_\alpha^\beta(P)$ for $\alpha = 1.3$ and $\beta = 2$

When $\beta = 2$, $\alpha = 1.3$.

Thus, now it is distinct from the illustrations that the function is diminishing function because it has presented monotonicity.

In the previous studies, the entropies for Pythagorean fuzzy sets are based on a probability-type involving one real parameter whereas in present study an information measure for the Pythagorean fuzzy set involving two parameters is established.

The proposed information measure for Pythagorean fuzzy sets is suitable and acceptable as they are explained with the properties such as sharpness and symmetry, and further comprehensive interpretations for the results have been explained with the help of monotonic behavior represented by assigning different values of α and β.

1.1.5 Conclusion

In this paper, an information measure for the Pythagorean fuzzy set of order α and type β is proposed. The new proposed measure is simpler and closer to the statistical significance, and it reflects better fuzzy properties. The information measure for the Pythagorean fuzzy set of order α and type β has verified the nonnegative property. Also, it is verified that the sharpness and symmetry hold true for the proposed measure.

Further, the proposed information measure for the Pythagorean fuzzy set of order α and type β has been verified for its rationality and the maximality and monotonic performance with respect to the parameters of the anticipated information measure has been examined and discussed.

By assigning different values of α and β, the monotonic behavior of the proposed information measure for the Pythagorean fuzzy set of order α and type β has been represented with the help of graphs.

References

1. Atanassov, K.T.: Intuitionistic fuzzy sets. Fuzzy Sets Syst. **20**(1), 87–96 (1986)
2. Hung, W.L., Yang, M.S.: Fuzzy entropy on intuitionistic fuzzy sets. Int. J. Intell. Syst. **21**(4), 443–451 (2006)
3. Mendel, J.M., John, R.I.B.: Type-2 fuzzy sets made simple. IEEE Trans. Fuzzy Syst. **10**(2), 117–127 (2002)
4. Peng, X., Yuan, H., Yang, Y.: Pythagorean fuzzy information measures and their applications. Int. J. Intell. Syst. **32**(10), 991–1029 (2017)
5. Shannon, C.E.: A mathematical theory of communication. Bell Labs Tech. J. **27**, 379–423 (1948)
6. Torra, V.: Hesitant fuzzy sets. Int. J. Intell. Syst. **25**(6), 529–539 (2010)
7. Turksen, B.: Interval valued fuzzy sets based on normal forms. Fuzzy Sets Syst. **20**(2), 191–210 (1986)
8. Wang, J.Q., Wang, P.: Intuitionistic liguistic fuzzy multicriteria decision-making method based on intuitionistic fuzzy entropy. Control Decis. **27**(11), 1694–1698 (2012)
9. Yager, R.R.: On the theory of bags. Int. J. General Syst. Methodol. Appl. Educ. **13**(1), 23–37 (1987)
10. Yager, R.R., Abbasov, A.M.: Pythagorean membership grades, complex numbers, and decision making. Int. J. Intell. Syst. **28**(5), 436–452 (2013)
11. Yager, R.R.: Pythagorean fuzzy subsets. In: Proceedings of the 9th Joint World Congress on Fuzzy Systems and NAFIPS Annual Meeting, IFSA/NAFIPS Edmonton, Canada, pp. 57–61 (2013)
12. Yager, R.R.: Pythagorean membership grades in multicriterion decision making. IEEE Trans. Fuzzy Syst. **22**(4), 958–965 (2014)
13. Zadeh, A.: Fuzzy sets. Inf. Comput. **8**, 338–353 (1965)

Chapter 2
Efficient Resource Distribution in Cognitive Radio Network by Fuzzy-Based Cluster Against Attacks

S. Sasipriya, M. Bhuvaneswari, P. Mayil Vel Kumar, and C. Karthikeyan

Abstract In 5G cognitive wireless network, the bandwidth is shared among the primary user (PU) and secondary user (SU) and face security risks from suspicious attacks. Secondary user (SU) networks are enhanced with a security mechanism to support primary users by proposing a mutual protocol for secondary wireless-driven users to collaborate with primary users. During a wireless transmission phase obtained by the secondary user, the cognitive hybrid access point communicates the power signal. Collected energy from primary user is used to communicate with malicious attacks and to acquire transmission capabilities during the wireless information transfer process. In addition, to obtaining optimum performance, a fuzzy-based clustered greedy algorithm is implemented to reduce the potential for interference in PU confidentiality. In the suggested strategy, the effect of injection and the reactive jamming attacks on the wireless transmission phase are analyzed. They can be detected via a convolution neural network to identify and distinguish various attacks. Finally, the simulation results for the proposed protocol and the corresponding resource sharing algorithm not only allow SU to gain transmission opportunities but also boost PU security efficiency in unknown attacks. The results are compared to the existing methods.

2.1 Introduction

Cognitive Wireless Powered Radio Network (CWPRN) which is in recent days widely studied and under research. The increase in spectrum allocation among users is

---

S. Sasipriya (✉) · M. Bhuvaneswari
Sri Krishna College of Engineering and Technology, Coimbatore, Tamil Nadu, India
e-mail: sasipriyas@skcet.ac.in

M. Bhuvaneswari
e-mail: bhuvaneswarim@skcet.ac.in

P. M. V. Kumar · C. Karthikeyan
Karpagam Institute of Technology, Coimbatore, Tamil Nadu, India
e-mail: mayilvel.cse@karpagamtech.ac.in

T. Senjyu et al. (eds.), *IOT with Smart Systems*, Smart Innovation, Systems and Technologies 251, https://doi.org/10.1007/978-981-16-3945-6_2

wider in most areas such as mobile communication, medical sector, unmanned aerial vehicles, and vehicle-to-vehicle networks. In short, wireless network is available in every area. The contact of wireless network has two classes of users in cognitive radio (CR) networks. One is a secondary consumer (SU) that does not have a spectrum of their own. The main user is the primary user (PU), who has a licensed spectrum of his own. Users can essentially be categorized as three key frameworks for the CR network.

These are interweave allocation, underlay allocation, and overlay allocation. The main purpose of this broad classification in cognitive wireless network is to detect licensed band/PU signals reliably by secondary users in order to use idle licensed spectrum. Secondary users may use the band for a less time period and return it to primary users where secondary users may often have to use the band for a longer period of time. There may be a reciprocal allocation between SU and PU to use the licensed spectrum of primary users (PUs). In the secured transmission, there were some confidential primitives known as privacy location of the node sensor, privacy identification of the node sensor, privacy of the packet data, privacy of the path. These privacy primitives enable the sensors to protect the information they hold. However, the attacker can easily catch the packet, obtain this security information, and thus obtain access to the sensor network. In addition, the sensor nodes have limited resources and thus these nodes need to be replaced for a while. Thus, in order to take this sensitive information, the attacker compromises multiple nodes on the network, which is called as malicious node and is managed by the attacker.

The proposed paper is sectioned as: Sect. 2 is the depiction of different traditional methodologies in use. Section 3 provides the detailed narration of the proposed technique. Section 4 is the performance analysis part of the proposed scheme and results. Section 5 is the conclusion and future work.

2.2 Related Works

For multimedia use, traffic sources are shown to be resilient to delay and high bandwidth. It is widely used for the application of cognitive radio networks for high power requirement and restricted bandwidth. There are many proposals to cluster-based, Spectrum-Conscious, Energy-Efficient Multimedia (SCEEM) coordinating show of CRSNs resolve the limitations together. Forming into bunches energizes the quality service and profitability course by obstructive foundation of the partaking nodes [1]. The quantity of clusters is determined in the routing to reduce the distortion of the audiovisual aid output that results owing to packet loss and time delay. Based on energy perception and spectrum, the cluster head is selected. The accessible non-contiguous spectrum bands are clustered together and configured to provide space for modern information and communication.

Effective power and channel management is used by resource-driven primary users who receive energy from secondary systems by transmitting wireless transmission, and through return, the secondary user will have promising option to the available spectrum for a longer period. In view of the temporary channel state information between primary and secondary systems, the problem of efficiency provided by the deterministic constraint [2] has been formulated. In order to fix the issue of resource allocation optimization, a particle swarm optimization technique is put in place to ensure the required data rate for both primary and secondary systems.

In order to improve the safety of cognitive radio networks (CRNs) [3], a new wireless cooperative jammer (CJ) full-duplex (FD)-aided transmission scheme is proposed. The CJ functions in FD phase, allowing it in the meantime to derive energy from the signals emitted by the secondary transmitter (ST) and to send the jamming noise (JN) in order to confuse the potential malicious eavesdropper. Proposed scheme improves the attainable level of confidentiality of the secondary source.

By offering diverse spectrum access, the new technologies adapted by cognitive radio system can possibly resolve accessible radio spectrum. Researchers have been chipping away with the utilization of this radio wave control innovation. Accordingly, the research has gone through enormous development at a fast rate to combat furthermore advancement [4]. Surveys and tutorials are desperately needed to help researchers stay up to date with these innovations. The existing energy detector is the right choice for blind signal detection, although due to noise uncertainty it inflicts from the well-known SNR wall [5]. Deep learning signal detector manipulates the structural details of the message signal and can be seen to perform accurate detection performance, involving no advance knowledge of channel state information or background knowledge. A deep learning cooperative detection structure offers significant performance improvements over traditional cooperative sensing methods.

Nodes in wireless sensor networks normally depend on battery power. Battery replacement is a complex process when there is a larger amount of sensor nodes and a wider area of permeability. This led to propose an energy-efficient medium access control protocol for wireless sensor networks called the Energy Saving Token Ring (ESTR) protocol [6]. The sensor networks are linked together, developing a ring where only the node carrying the token is effective and capable of communicating. This clearly improves the life of the network. ESTR delivers the highest energy performance compared to other MAC protocols.

New cooperative spectrum sensing (CSS)-based machine learning model uses clustering methods to reduce overhead cooperation and effectively improve detection performance. Prior to the collaborative sensing process, radio resource users were correctly clustered to use energy sequence data and the SVM model [7]. All users are split into a series of enhanced groups, each time only one of which is needed for collaborative sensing. The performance of the different methods was assessed in terms of average training, classifier speed, and classification accuracy.

In light of software-defined radio (SDR), cognitive radio system (CRS) offers extra adaptability and improves generally speaking range use productivity [8]. The overview and challenges of the CRS-based Radio Frequency (RF) segment is discussed in this work. Security Implementation for Denial of Service (DoS) attacks on Software-Defined Networks (SDNs) using a mixture of multiple technologies implemented on conventional networks adapted to the SDN [9] architecture. The implementation of the intrusion detection system (IDS) in the application layer environment has resulted in the improvement of network security.

Numerous techniques have appeared in cognitive radio (CR) spectrum sensing [10], and the latest groundbreaking techniques and algorithms have been tested during the development of an intellectual wireless communication device by means of cognitive radio technology. Various MIMO performance enhancement schemes are analyzed for the transmission and reception of the signal [11]. The methods introduced have many advantages in the MIMO. Hybrid access points are deployed at the same time uploading of information and energy, and it is required for harvesting energy by users [12]. Base station is capable of synchronizing the operate in two modes say, Wireless Information Transfer (WIT) and Wireless Energy Transfer (WET) by Radio Frequency (RF) signals. Cognitive radio 5G schemes are discussed in the work [13].

2.3 Proposed Work

Initially, the system model was devised with the required primary user and secondary user with their parameters like frequency, sampling, number of channels, channel state information, etc. A model is proposed for secondary user energy harvesting at wireless power transfer phase and wireless information phase [2]. At the time of the wireless information transfer process, simultaneous transmission is accomplished, thus reducing the incidence of interference. The energy consumed can be minimized, while the transmission and detection of various attacks are carried out using a neural network mechanism. This essence achieves the energy status of the network by having an objective role for various membership functions. The attack detected is then graded using a fuzzy cluster-based greedy algorithm. Detected attacks are classified as DoS, man-in-the-middle, or phishing attacks.

2.3.1 System Model

The situation in which two forms of non-cooperating users are SUs and PUs has been considered. For example, PUs, mobile phones, wireless microphones, or TVs are those to which the sum of wireless spectrum is licensed. Conversely, the SUs are intended to be those without a pre-assigned wireless spectrum. Conversely, cognitive radio-equipped SUs can transmit their individual packets through the opportunities that may occur once the PUs do not use the wireless spectrum that is licensed. In this method, the SUs open wireless spectrum is further divided into multiple channels, all of which have an unchanged number of frequency bandwidths.

The PU energy storage unit should be charged earlier on behalf of the primary timeline. On the other hand, the PU uses energy harvested during the final timeline for subsequent time frames. In order to prevent collisions between the PU and the SU, the SU will not be able to use the authorized spectrum to transmit its own data once the licensed channel is used by the PU because of its priority. In addition, SU will initially absorb energy from ambient radio signals until the licensed channel is being used. In this case, the SU transmitters take a savings step and then transmit a half-duplex energy restriction process.

Rechargeable energy storage system is impossible to discharge and charge simultaneously. For energy harvesting, SUs are uses time spliting techniques in the given slot which is to be consumed for its data communication. After the inactive licensed channel is used by SU, it would broadcast its data using idle licensed band of SU time slot, at some time period, once the PU ends the contact in the time slot of PU. The PU will sustain energy harvesting from environmental sources in the course of this, and the energy collected from PU will be used in subsequent time slots. The initial restriction denotes the causality of the energy limitation that the energy used for the relay transmission must be below the energy harvested by SU. On behalf of PU, SU must take away the extracted energy in the present time slot in an actual time slot. The second limitation applies to the length of the supportive broadcast, and the length of the noncooperative broadcast should be lower than the duration of the time slot.

2.3.2 Fuzzy-Based Cluster Greedy Simultaneous Transmission Optimization Method

The wireless data is transmitted simultaneously using a fuzzy-based cluster greedy optimization algorithm which can reduce the consumption of power or energy harvest at the time of transmission. The energy harvest can be reduced simultaneously in this method. Also, it would minimize the incidence of interference. Following this, the neural network process is initialized to detect the occurrence of attack during the transmission of data to secondary consumers. The initialization of the neural network often classifies the attack type with the available channel state information.

Algorithm: Fuzzy-Based Cluster Greedy Optimization Algorithm

Step 1: Initialize the parameters for power adjustment (no. of subchannel, total power, channel state information, bandwidth, noise variance).
Step 2: Power allocation for primary user and secondary user.
Step 3: Allocation of spectrum for PU and SU user using fuzzy logic-based algorithm.
Step 4: Harvest energy from PU and SU, the BS partitions SU into cluster, and to initialize symbol periods.
Step 5: Greedy algorithm to compute best value.

Step 6: Attack determination.
Step 7: Predicted labels using CNN using the user data features, attack label.

Machine learning technique is used for the process of data analytics, which in turn teaches computers to perform complex functions. The machine learning approach has numerous core systems such as computational image processing, computer vision, computational biology, and many more. The method proposed is aimed at distinguishing phishing emails using key structure features in the phishing site. Thus, machine learning algorithms are used for the method of classification and clustering. In this instance, the network needs to be sensitized. Thus, by attaining the energy state of the network and determining the objective function of the membership function, the classifier effectively classifies the types of attack. The type of attacks can eventually be defined as DoS, man-in-the-middle attack, or phishing attack.

Fuzzy clustering is a method of clustering or grouping of data, where each piece of data may pertain to more than one group. Clustering process requires allocating data points to groupings so that components in the same group are as close as possible. In this proposed work, the data transmitted between the source node to channel hybrid access point and receiver nodes by secondary users is the following fuzzy-based cluster along with greedy algorithm incorporated for maintaining secrecy of primary user while transferring and receiving. Effective spectrum allocation and also energy harvesting have been described as essential and major components for wireless communications.

The real challenge is to work a lot harder for greedy algorithms to know the problems of accuracy. Even with the appropriate method, it is difficult to prove why it is so accurate. Proving that a greedy algorithm is correct seems to be an art than a science.

2.4 Performance Analysis and Results

This session offers the performance analysis of the proposed scheme. Figures 2.1 and 2.2 shows the power adjustment algorithm vs. subchannel indices for man-in-the-middle attack and normal spectrum allocation respectively.

Table 2.1 shows various methods and their outcome when average opportunity discovery ratio. Figure 2.3 displays the average chance discovery ratio vs. the usage of the primary consumer channel, which demonstrates the comparative estimate of the proposed and current techniques. It was clear from the study that the suggested methodology showed greater average discovery of opportunities.

Table 2.2 shows the average opportunity discovery ratio versus $v_{\max}$ which shows the comparative estimate of the proposed and current techniques. It was clear from the study that the suggested methodology showed greater average discovery of opportunities.

It was clear from the study that the suggested methodology showed greater average discovery of opportunities. The comparative estimation of the proposed average

Fig. 2.1 Power adjustment process during attack

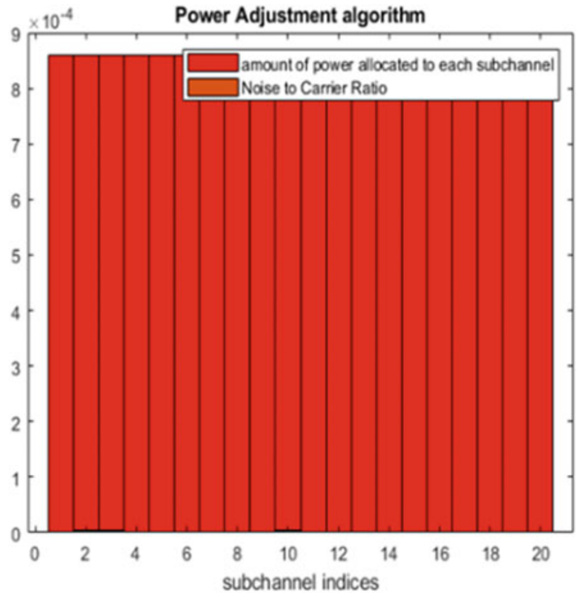
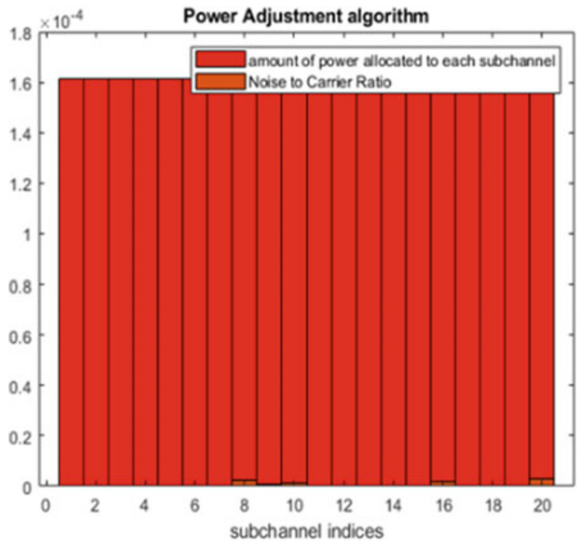


Fig. 2.2 Power adjustment process during normal



opportunity discovery ratio versus average is shown in Fig. 2.4. From the analysis, it was clear that the methodology proposed shows better average discovery of opportunities.

Table 2.1 Average opportunity discovery ratio versus PUs channel utilization when the algorithm is applied

	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
Proposed system	1	1	1	1	1	1	1	1	1	1	0.80
Cluster CMSS policy	0.97	0.96	0.95	0.93	0.90	0.85	0.77	0.72	0.65	0.60	0.57
Greedy noncooperative policy	0.87	0.87	0.86	0.85	0.83	0.77	0.73	0.65	0.58	0.51	0.51
Genie-aided location-aware policy	1	1	1	0.97	0.96	0.94	0.90	0.84	0.79	0.70	0.70

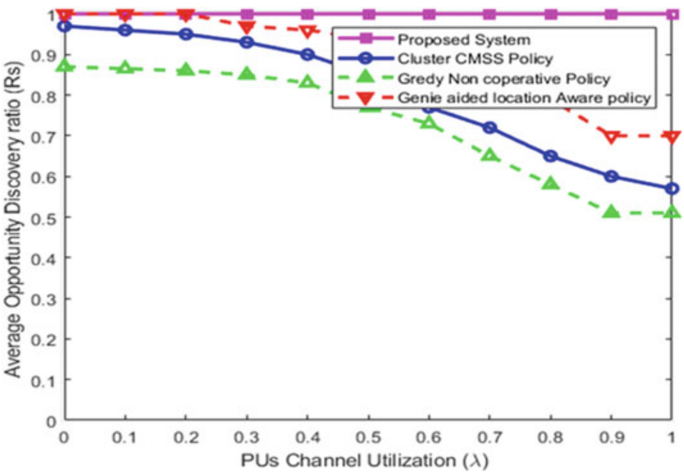


Fig. 2.3 Average opportunity discovery ratio versus PUs channel utilization

Table 2.2 Average opportunity discovery ratio versus V_{max}

	0	5	10	15	20	25
Proposed system	1	1	1	0.91	0.89	0.88
Cluster CMSS policy	0.87	0.86	0.86	0.85	0.84	0.83
Greedy noncooperative policy	0.81	0.78	0.78	0.77	0.77	0.77
Genie-aided location-aware policy	0.96	0.94	0.93	0.93	0.93	0.93

2.5 Conclusion

For wireless powered phase SU with PU collaboration, a new technique fuzzy-based cluster using greedy algorithm was suggested and simulated successfully. The first energy harvest of the SU was carried at the wireless power transfer stage and is broadcast by cognitive hybrid access point. The hybrid access point has the information about the connected devices. In addition, a greedy algorithm based on a fuzzy cluster

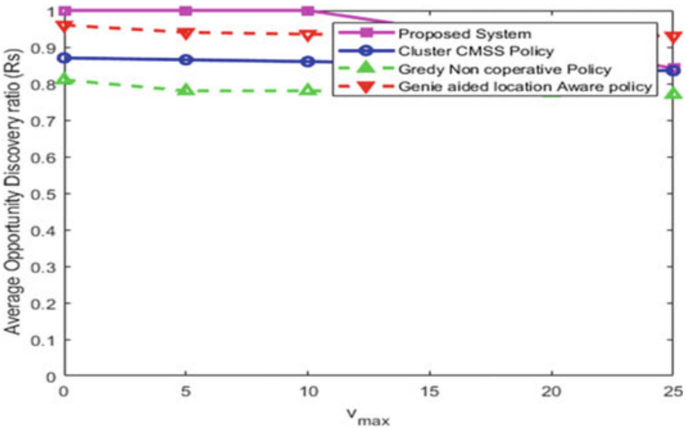


Fig. 2.4 Average opportunity discovery ratio versus $V_{\max}$

is used to reduce the outage of the prospect of PU confidentiality and to give the best optimal values. The power is balanced for SU in this method, and the spectrum allocation is assigned to the primary consumer, followed by SU based on the process of fuzzy logic. To attain the best optimal values, the greedy algorithm is applied.

The efficiency of the proposed system is calculated and compared to current approaches to show efficiency. The simulation result can be verified by creating a network environment and analyzed in 5G MIMO communication and IOT in mm-wave communication.

References

1. Shah, G.A., Alagoz, F., Fadel, E.A., Akan, O.B.: A spectrum-aware clustering for efficient multimedia routing in cognitive radio sensor networks. IEEE Trans. Veh. Technol. **63**(7), 3369–3380 (2014). https://doi.org/10.1109/TVT.2014.2300141
2. Shrestha, A.P., Yoo, S.J.: Optimal resource allocation using a support vector machine for wireless power transfer in cognitive radio networks. IEEE Trans. Veh. Technol. **67**(9), 8525–8535 (2018). https://doi.org/10.1109/TVT.2018.2850799
3. Chen, X., Guo, L., Li, X., Dong, C., Lin, J., Cavalcante, C.C.: Full-duplex wireless-powered jammer aided secure communication for cognitive radio networks. Phys. Commun. **31**, 103–112 (2018). https://doi.org/10.1016/j.phycom.2018.10.003
4. Arjoune, Y., Kaabouch, N.: A comprehensive survey on spectrum sensing in cognitive radio networks: Recent advances, new challenges, and future research directions. Sensors **19**(1) (2019). https://doi.org/10.3390/s19010126
5. Gao, J., Yi, X., Zhong, C., Chen, X., Zhang, Z.: Deep learning for spectrum sensing. IEEE Wirel. Commun. Lett. **8**(6) (2019). https://doi.org/10.1109/LWC.2019.2939314
6. Bagci, F.: Energy-efficient communication protocol for wireless sensor networks. Ad-Hoc Sens. Wirel. Networks **30**(3–4), 301–322 (2016)
7. Li, Z., Wu, W., Liu, X., Qi, P.: Improved cooperative spectrum sensing model based on machine learning for cognitive radio networks. IET Commun. **12**(19), 2485–2492 (2018). https://doi.org/10.1049/iet-com.2018.5245

8. Nguyen, V.T., Villain, F., Le Guillou, Y.: Cognitive radio RF: overview and challenges. VLSI Des. **2012**, 1–13 (2012). https://doi.org/10.1155/2012/716476
9. Fares, A.A.Y.R., De Caldas Filho, F.L., Giozza, W.F., Canedo, E.D., De Mendonca, F.L.L., Nze, G.D.A.: DoS attack prevention on IPS SDN networks (2019). https://doi.org/10.1109/WCNPS.2019.8896233
10. "Network and Complex Systems A Review paper based on spectrum sensing techniques in Cognitive Radio Networks," [Online]. Available: www.iiste.org
11. Bhuvaneswari, M., Sasipriya, S.: Performance analysis of various techniques adopted in a 5G communication network with a massive MIMO system. In: Proceedings of the 5th International Conference on Inventive Computation Technologies (ICICT 2020) (2020)
12. Zhao, Y., Wang, D., Hu, J., Yang, K.: H-AP deployment for joint wireless information and energy transfer in smart cities. IEEE Trans. Vehicular Technol. 7485–7496 (2020). https://doi.org/10.1109/TVT.2018.28219
13. Sasipriya, S., Vigneshram, R. : An overview of cognitive radio in 5G wireless communications. IEEE Xplore, 08 May 2017. https://doi.org/10.1109/ICCIC.2016.7919725

Chapter 3
Promote an Abstract Understanding of the Problem-Solving Knowledge in the Theory of Computation Through Visualization-Based Intelligent Tutor System

Rashmi Dixit, Manisha Nirgude, and Pratibha Yalagi

Abstract Just like heart is to the body, so is theory of computation for computer science and engineering discipline. Course deals with designing of abstract machines like finite automata, push down automata and Turing machines. It not only involves designing but also theoretical concepts like theorems, proofs and examples and to deal with examples accordingly. This course involves visualization, imagination which is difficult for students in the physical classrooms and teacher needs to blend traditional black board teaching method with active learning strategies to make learning easy. A visualization-based intelligent tutor system (ITS) is a computer system that aims to provide immediate and customized instruction to learners, usually without intervention from a human teacher. This instructional strategy is helpful for enabling learning in positive manner as students can "see" what is happening and gets feedback. The active way of learning se using Java Formal Languages and Automata Package (JFLAP) tool. It is a preexisting software tool created by Professor Susan Rodger and her students at Duke University. The control group and experimental group results are presented in this paper partially; supporting the effect of the usage of the visualization-based intelligent tutor system on student learning. The experiment shows that use of VIT enhance learning ability and interest in the course. The usefulness is supported with survey questionnaire.

3.1 Introduction

Though computer science and engineering is a dream discipline nowadays, It requires sound programming knowledge. All courses are linked. Each model in automata theory plays pin roles in several areas. Finite automata are used in text processing, compilers and hardware design. Context-free grammar (CFGs) is used in programming languages and artificial intelligence. Student gets only superficial knowledge by the theoretical concepts.

R. Dixit (✉) · M. Nirgude · P. Yalagi
Walchand Institute of Technology, Solapur, India

T. Senjyu et al. (eds.), *IOT with Smart Systems*, Smart Innovation, Systems and Technologies 251, https://doi.org/10.1007/978-981-16-3945-6_3

The course theory of computation is also focuses on formal languages and automata theory and students cannot easily visualize theoretical constructs by the traditional teaching. In traditional teaching, problem-solving is done by pencil and paper method where the students do not understand actual concept through visualization. They cannot easily evaluate their solved problems to determine if their solutions are correct. Hence, they depend on others for the solutions especially on teachers. This course contains more mathematics and also requires discrete mathematics concepts as prerequisite. So the weaker students need to work on more examples and problem-solving to acquire subject knowledge. The automata theory uses more mathematical notations to represent languages, theorems and proofs.

In traditional "black board method"—teaching method, teacher draws model on the blackboard and students also in their notebook and then teacher analyzes model working by taking different examples and students also try to understand simultaneously with teacher.

Like "All fingers are not same," students of same classroom are also of different essence, different mental ability. When teacher drawing, designing model on board, some students actively participating in analyzing and drawing, some are just coping with gossiping with bench partners and some are even not bothered about what is happening in the classrooms. Negligence may bring losing interest in designing different models which are linked with each other.

Intelligent tutor system a "computer system" based on "artificial intelligence" designed to draw content. The main goal of intelligent tutor system is to interact with students similar to a human tutor. The program personalizes the feedback based on the background and the progress of every individual student. It consists of four basic parts as shown in Fig. 3.1.

1. The domain model with expert knowledge.
2. The student model examines student's level of understanding.
3. The tutoring model gives assignments.

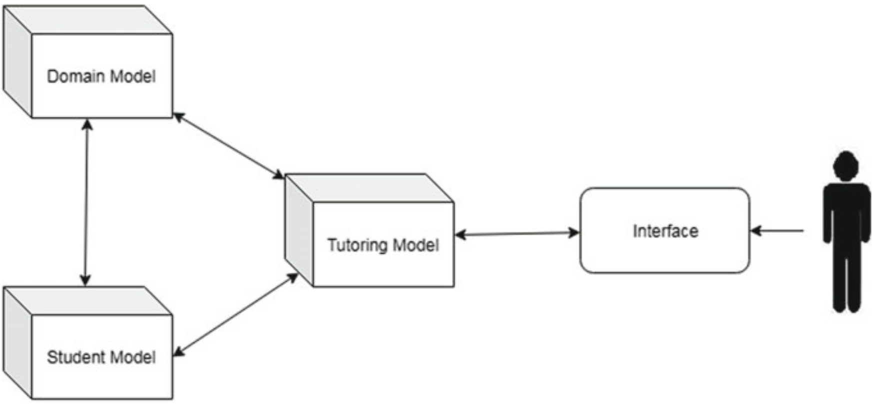


Fig. 3.1 Intelligent tutor model [1]

4. The user interface mode for communication.

Visualization is a way to communicate a message through pictures, graphs, diagrams, and animation, smartly. The student can draw on and actually check the working of the machine for particular input. It shows all the steps of string processing. Visualization provides student a view and flow of machine for particular input. The representation and working of any model in visual form helps users in analyzing. The visualization technique communicates with smarter way. Visualization-based intelligent tutor system not only represents graphically but also provides tutor which gives stepwise working of model and provides feedback. The JFLAP tool provides an interactive visualization for automaton constructs and gives feedback for the construction of automaton like personalized tutor without human.

The experiment is conducted for second year computer science and engineering undergraduate students for the course of theory of computation. The concepts are taught in theory class, and the problems are solved in tutorial sessions with a small batch of students. It is concluded that the student's involvement increases in solving the examples as they can easily visualize the theoretical notations and concepts with proper simulations.

3.2 Literature Review

Doing assignment by own helps students to understands and learn by their own, but if they stuck in between while doing, they will loose interest. Because of intelligent tutor, student gets artificial environment of tutor [2]. Students, nowadays, familiar with many web-based applications, tutors. Human brain having different capacity for audio and video [3]. Mathematica, computer-aided learning tools is mostly used tool for teaching–learning process. Many tools are available for mathematics, discrete mathematics and computation theory [4]. Intelligent tutor system proves stepwise guidance while practicing assignment at any time as compared traditional teaching method [5]. Visualization increases interest as compared to books and provides guidance at any time with the help of intelligence tutor system. Learner learns the design and analysis of different model without additional tutorial sessions. Because of visualization, students can grasp easily the working model of machine which will naturally increase their concentration and interest for subject [6]. The tool also contains Combinatorica [7], with more than 450 algorithms for graph theory and discrete mathematics. Grail+ [8] is also one of the symbolic computation environments for regular expressions, finite-state machines and other formal language theory objects. Grail+ can be used for inputting machines or expressions, convert them from one form to another. It can also be used for minimization, complementing and for making them deterministic. Simpler, non-commercial tool like basic algorithms with animations for various sorting algorithms as well as the source code used [9]. For the subjects like computer networks, animations for data-link layer protocols are available [10]. To make traditional classroom active, teacher requires good presentation skill to keep attention of students throughout the lecture [11]. Tools like Automata Simulator [13] and SoftLab [12] are used in the field of fundamentals of automata theory. Automata Simulator provides interactive generation and simulation of finite-state machines. This simulator supports all types of finite-state machines: automata with binary output (DFA, NFA and NFA-null transitions) and automata with general output.

3.3 Visualization-Based Intelligent Tutor System: Java Formal Languages and Automata Package (JFLAP)

JFLAP, intelligent tutor system, is free open-source tool can be used while teaching formal languages and automata theory. It not only provides interface for designing of various machines like finite automaton, pushdown automaton and Turing machine but also gives stepwise processing of machine. It guides students to analyze the working and processing along with feedback. Using JFLAP complicated examples can be constructed solved stepwise easily. Students can use to check the correctness and accuracy of their model.

It also deals with mealy machine, more machine, grammar, regular expressions and pumping lemma.

JFLAP is act as visualization-based intelligence tutor. It provides an interface to construct automata along step-by-step processing of string in to an FA. It provides menu for defining a grammar and checks whether the given string is a member of that grammar or not. Also for performing various conversions like NFAs in to DFA, context-free grammar to context sensitive grammar. It also has editor to perform minimization of automata. It also has tutorship for combining automata using union and concatenations. It gives visualization of the non-determinism by keeping all its possible configurations. It also has stepwise conversion of a right linear grammar in to NFA menu. Additional facilities are also provided like gaming design for pumping lemma, construction of pushdown automata and checking string membership, as well as acceptance of string by constructing Turing machine, etc.

Blending of visualization-based intelligent tutor with traditional method makes easy for students to study and for teacher to engage them in active way, theory of computation course. It saves time and provides accuracy. As it provides feedback and response immediately to students , they can immediately verify their solved problems and corrects them.

The above figure shows JFLAP screen for String testing membership. The editor with multiple options with stepwise execution table on right-hand side part of editor. As above Fig. 3.2 shows menu available with JFLAP, Figs. 3.3 and 3.4 show use of different editors from JFLAP for different problem-solving.

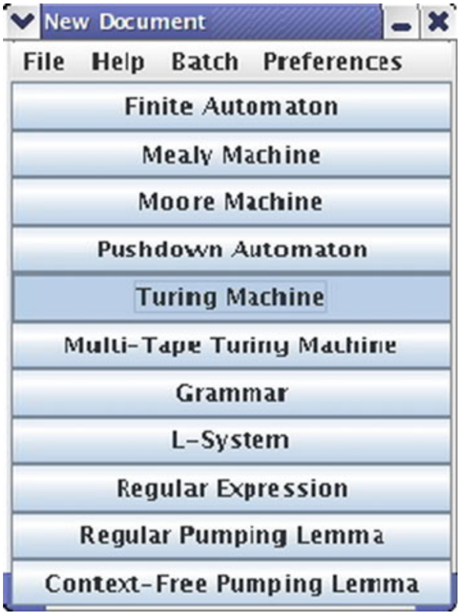


Fig. 3.2 VIT JFLAP screen

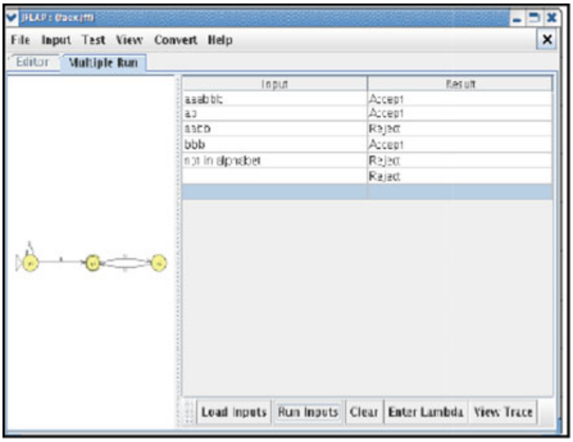


Fig. 3.3 Example screen—string testing for membership

3.4 Implementation

The experiment is conducted for second year B.Tech of computer science and engineering class for the course theory of computation taught in class through traditional blackboard method during lecture session. Usually, the concepts are explained in

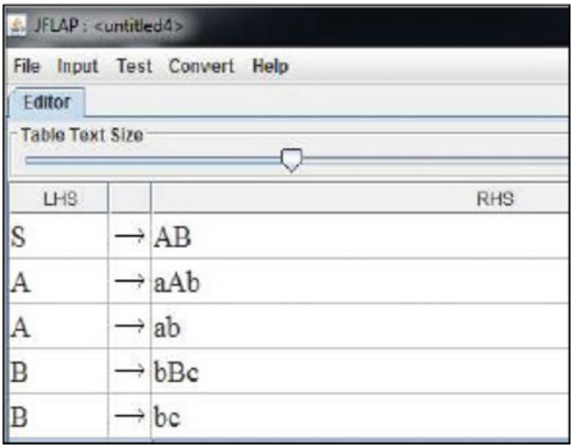


Fig. 3.4 Example screen—grammar representation

a theoretical way with examples. The teacher explains finite automata and elaborates its working for particular string step by step. She uses blackboard to stepwise explanation simultaneously with drawing of machine.

Sample Blackboard

Above figure shows drawing of machine and processing for a particular string stepwise solving by teacher on a blackboard in a traditional approach. Some so-called good or active students draw machine and solve problem concisely stepwise manner with their own comments for better understanding of their own (Fig. 3.5).

Sample Notebook of Good Student

But as we already discussed above in the same paper, class consist of different types of students with different abilities, different perception. Below figure shows

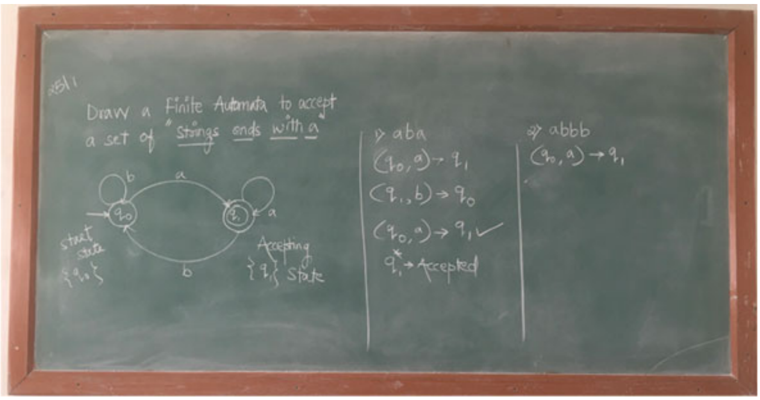


Fig. 3.5 Blackboard representation by teacher

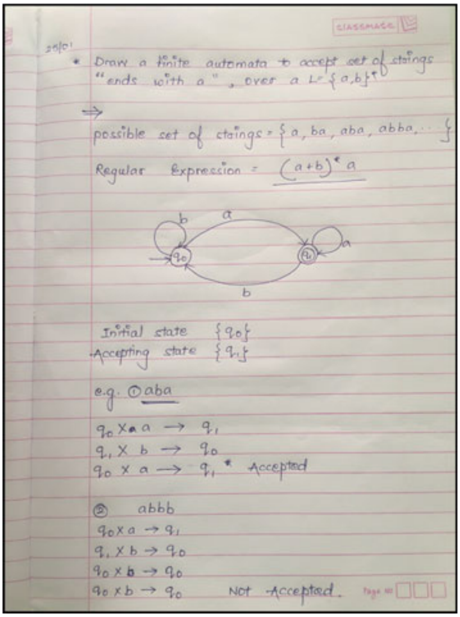


Fig. 3.6 Notebook sample 1

the notebook of candidates who lost interest in the middle and stop at the middle of problem-solving.

Sample Notebook of Another Student [Not Paying Attention While Writing]
During, at the time of exam, students themselves are unable to understand what they wrote. Books give theoretical understanding, and lots of material is available on the net but they explain concept in a theoretical manner. Teacher may not be available. So for better understanding, with their own pace, intelligent tutor is best option which also provides visualization.

During tutorial, assignment is given to student. Two batches used JFLAP during tutorial sessions for assignment. The students use the tool for different exercises based on different concept (Figs. 3.6, 3.7 and 3.8).

The visualization-based intelligent tutor provides visualization as well as tutor for stepwise assistance.

3.5 Experiment and Result

As per university syllabus, course theory of computation is included in the curriculum of second year computer science and engineering first semester. The course is scheduled as 3 h. theory lectures and one hr. tutorial.

Class strength is approximately 71 students, during lecture session. For the tutorial session, four batches are grouped comprising of 18 students scheduled as batch 1,

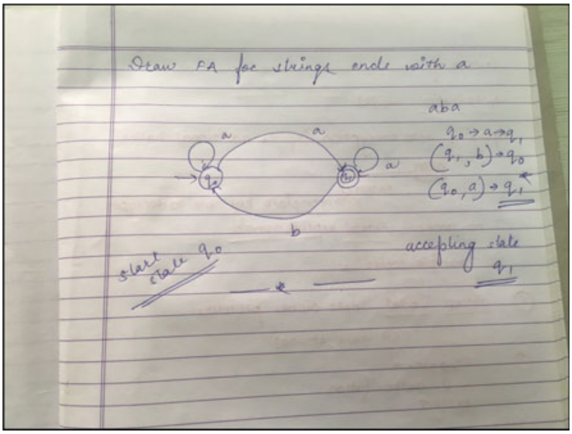


Fig. 3.7 Notebook sample 2

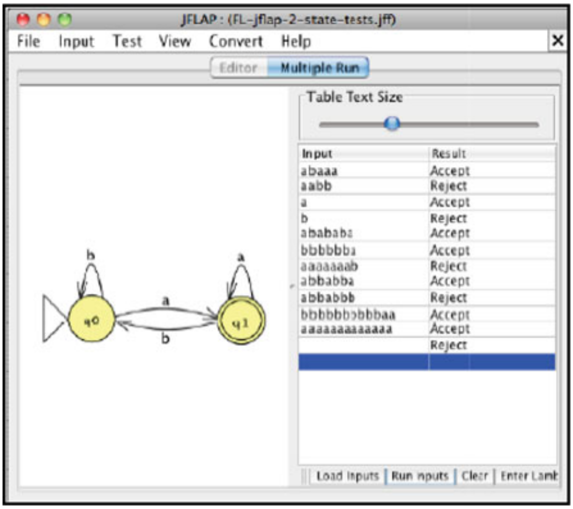


Fig. 3.8 Example string testing for membership for a string ending with a Jflap example

batch 2, batch 3 and batch 4. During the tutorial session, teacher asks batches 1 and 2 to solve examples as exercise assignment with pen and paper. The students do pencil work and complete the assignments in their notebooks. Batch 1 and 2 will act as control group. Students of batches 3 and 4—experimental group—used JFLAP—visualization-based intelligent tutor system to solve the assignment.

After completion of experiment, test is conducted amongcontrol group and experimental group. Same question paper is given to both groups.

By looking at results, it is observed that those were brilliant and topper always perform good, either in control group or in experimental group. But passive student who needs lit bit push, from 3 and 4 batches perform efficiently well as compared to

Table 3.1 Test result analysis

Mark range	Control group	Experiment group
Below or equal to 10	1	0
11–15	23	2
16–20	5	15
21–25	4	13
26–30	1	4

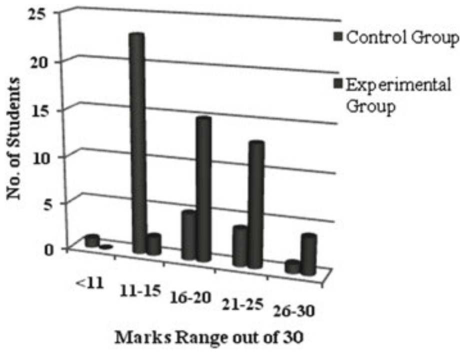


Fig. 3.9 Students marks comparison

1 and 2 batches because they practice through visualization-based intelligent tutor. Their presentation is also good.

Results of control group and experimental group were compared. It is observed that on an average, students of experimental group score more marks than control group.

Table 3.1 shows the analysis of marks among number of students. Numbers of students are calculated for different range of marks. It is observed that students in control group score more marks (Fig. 3.9).

The results are calculated under t-test calculator which gives t-value = -6.15947 and p-value = 0.00001 student test marks for control group and experimental group. It is observed that students of experimental group scored more than 10 marks. Good number experimental group student scored more than 20 marks, approximately 50% of student. Also, range of between 16–20 is near about 45%. And only 2 students get marks between 11 and 15 vice versa in control group 70% student scored marks between 11 and 15. Only 15% student scored marks more than 20.

Students feedback is also collected, and they show a positive response for the use of visualization-based intelligent tutor system as it increases interest in problem-solving.

3.6 Conclusion

Teaching with chalk and board is traditional teaching method which requires good presentation skills of teacher while writing on a board to grab student attention. Naturally, students lost their concentration after 20 min during one hour theory lecture. For proper understanding of course like Theory of Computation, requires elaboration of concepts with proper visualization.

As our brains are tailored to attract toward images by birth only, to improve learning ability of students, visualization is the master key. Additional to visualization, tutor support is like a bonus. Intelligent tutor is available for providing feedback at any time to students. The conclusion after experiment is that the students of experiment group not only developed interest but also improved in designing of different machines like finite automata, push down automata, Turing machine.

References

1. Shanabrook, D.H., Arroyo, I., Woolf, B. P., Burleson, W.: Visualization of student activity patterns within intelligent tutoring systems. In: International Conference on Intelligent Tutoring Systems ITS 2012: Intelligent Tutoring Systems, pp. 46–51 (2012)
2. Merrill, D.C., Reiser, B.J., Ranney, M., Trafton, J.G.: Effective tutoring techniques: a comparison of human tutors and intelligent tutoring systems. J. Learn. Sci. **2**(3), 277–305
3. Brusilovsky, P., Eklund, J., Schwarz, E.: Web-based education for all: a tool for development adaptive
4. Budiselic, I., Srbljic, S., Popovic, M.: RegExpert: a tool for visualization of regular expressions. In: IEEE International Conference on "Computer as a Tool", (EUROCON) (2007)
5. Murray, T.: Authoring knowledge based tutors: tools for content, instructional strategy,student model, and interface design. J. Learn. Sci. **7**(1), 5–64 (1998)
6. Nguyen, T.-H., Khoo, I.-H.: Learning and teaching engineering courses with visualizations. In: Proceedings of the World Congress on Engineering and Computer Science 2009 (WCECS 2009), San Francisco, USA, vol. I, 20–22 Oct 2009
7. Watson, B.W.: The design and implementation of the FIRE engine: a C++ toolkit for finite automata and regular expressions. Computing Science Note 94/22, Eindhoven University of Technology, Netherlands (1994)
8. Grail+ Project homepage. http://www.csd.uwo.ca/Research/grail/index.html
9. Aneva, B., Thiébaut, D.: Sorting algorithms. Smith College. http://maven.smith.edu/~thiebaut/java/sort/demo.html
10. Jacobsen, A.: Data-link network protocol simulation. University of Birmingham, School of Computer Science. http://www.cs.bham.ac.uk/~gkt/Teaching/SEM335/dlsim/Simulation.html
11. Dixit, R., Yalagi, P.: Visualization based intelligent tutor system to improve study of computer algorithms. J. Eng, Educ. Transform. **30**(3) (2017). ISSN 2349-2473, eISSN 2394-1707
12. Rodger, S.H., Finley, T.W.: JFLAP: An Interactive Formal Languages and Automata Package. Jones & Bartlett Publishers, Sudbury (2006)
13. Srbljic, S.S., Milanovic, A.: Extending the textbook: a distributed tool for learning automata theory fundamentals. In: Proceedings of the 9th IEEE International Conference on Electronics, Circuits and Systems (ICECS 2002), Dubrovnik, Croatia, September (2002)

Chapter 4
COVID-19 Outbreak Prediction Using Machine Learning

Jatin Singh, Sachin Yadav, Ketan Chauhan, and Ruchika Malhotra

Abstract The recent global outbreak of the coronavirus has thrown new challenges for the research community. First case was registered in China, and then, it got spread in most of the countries of the world. Initially, the speed of spread was slow but later on, its spread rate was very really high and on analysis it turned out to be exponential. Governments all across the world imposed lockdowns, and people were asked to practice social distancing in order to prevent the spread of the COVID-19 virus. Later on, it was announced as pandemic by World Health Organization (WHO). Machine learning-driven methods can prove to be really vital in predicting risks, effects and parameters of this pandemic. These predictions will help in making strategies to control its spread and understand its nature. More research is beginning to anticipate and a remarkable amount of machine learning models are being talked about to predict COVID-19 cases used by experts or researchers around the globe. In this research project, we have used univariate LSTM model to make predictions. The number of confirmed cases, number of recovered cases and the number of death cases in the coming days are being predicted. Mean absolute error (MAE) is used as the measure of performance metric of the predicted results. The results produced are quite accurate. These results prove that univariate LSTM is a promising model to make predictions for COVID-19 outbreak.

J. Singh (✉) · S. Yadav · K. Chauhan · R. Malhotra
DTU (DCE), New Delhi, India
e-mail: jatinsingh_2k17se50@dtu.ac.in

S. Yadav
e-mail: sachinyadav_2k17se101@dtu.ac.in

K. Chauhan
e-mail: ketanchauhan_2k17se56@dtu.ac.in

R. Malhotra
e-mail: ruchikamalhotra@dtu.ac.in

T. Senjyu et al. (eds.), *IOT with Smart Systems*, Smart Innovation, Systems and Technologies 251, https://doi.org/10.1007/978-981-16-3945-6_4

4.1 Introduction

On the day one, no one is sick, next day some are sick and after some days, a significant number of peoples are sick, that is how the pandemic grows if it is not handled at the right time with a right way. The world got hit by a pandemic in 2019; it is named as COVID-19. The first case was found in December 2019 in the city named Wuhan, which is located in China; after this, many countries were affected by this virus [1]. Some countries like South Korea and New Zealand handle the situation very carefully and were successfully able to show backdoor to the virus, whereas some countries like USA, Brazil, India, etc. are not able to handle situation successfully and the cases are increasing continuously. Till now, the confirmed cases have been more than 60,000,000 (60 million) and deaths cases have been more than 1,000,000 (1 Million) across the globe. The virus transfers from infected person to non-infected person, when infected person sneezes and the droplets of sneeze released in the air are inhaled by a non-infected person [2]. Till now, it is unknown how often virus spreads through contaminated surfaces. Infection mainly happens when people are near each other for a long period of time known as "close contact." People infected with this virus remain infected for up to 10 days in case of moderate case and two weeks in case of severe case. The most common symptoms found in the world's cases include fever, dry cough and tiredness [3]. Less common symptoms include sore throat, diarrhea, headache, a rash on skin, etc., whereas its serious symptoms found in world's cases include difficulty breathing or shortness of breath, chest pain, etc. Person should be given immediate medical attention if he/she have serious symptoms. In general, it takes 5–6 days for the infected person to show virus symptoms; however, it can also take up to fourteen days. Preventive measures that are being used to fight against this virus include social distancing, wearing face mask in public places, washing hand properly with soap and water in home and with alcohol-based hand sanitizer when step outside, covering mouth when sneezing or coughing, disinfecting surfaces, etc. Countries worldwide have implemented restrictions on travel including both domestic and international travels, and some have implemented lockdown either partially or fully [4]. The virus has interrupted the social activities, social gatherings, and it slows the increasing pace of world's economy. According to the estimates of World Bank, in 2020, the world's economy will contract by 5.2% which is the greatest recession seen after II world war [5].

In this paper, we are trying to predict COVID-19 cases of three types that are confirmed cases, recovered cases and death cases across the world by a powerful network of long short-term memory (LSTM). The results will help us to figure out, how the current trend likely to get affected in future days, if it is going to be good, bad or worse, based on the results, government can bring new strategies or rule to affect the expected spread of virus, and the results we achieved are quite accurate when compared with actual count of cases.

4.2 Methods and Materials

4.2.1 Data Resource

The datasets have been collected from GitHub. Three distinct datasets have been used for each confirmed cases, recovered cases and death cases. Datasets are in the form of daily number of confirmed cases, recovered cases and death cases across the world. Each dataset has data of 300 + days starting from January 22, 2020, till November 21, 2020.

4.2.2 Data Analysis

Daily data of worldwide confirmed cases, recovered cases and death cases due to coronavirus has been analyzed. Python libraries like NumPy, Pandas, Matplotlib, Seaborn, etc. have been used to serve the purpose (Fig. 4.1).

In the above plot, *X*-axis is taken to be the total count of cases worldwide, and *Y*-axis represents the total count of days. The plot shows the trend of how number of cases are increasing with passing number of days since January 22, 2020. Blue-colored line represents confirmed cases, red line represents recovered cases and the green line represents death cases. The above plot gives the idea of how fast COVID-19 is spreading across different parts of the world. Within a span of about 300 days, the total number of worldwide confirmed cases have reached to almost 60 million.

Figure 4.2 shows the sum count of total confirmed cases for the last 90 days. It is clearly that the total count of confirmed cases on 220th day (starting from January 22, 2020) were about 28 million which increased to about 55 million on the 300th day. This shows that rate of spread of COVID-19 is exponential. In a similar fashion, total count of recovered cases and death cases have been plotted for the last 90 days

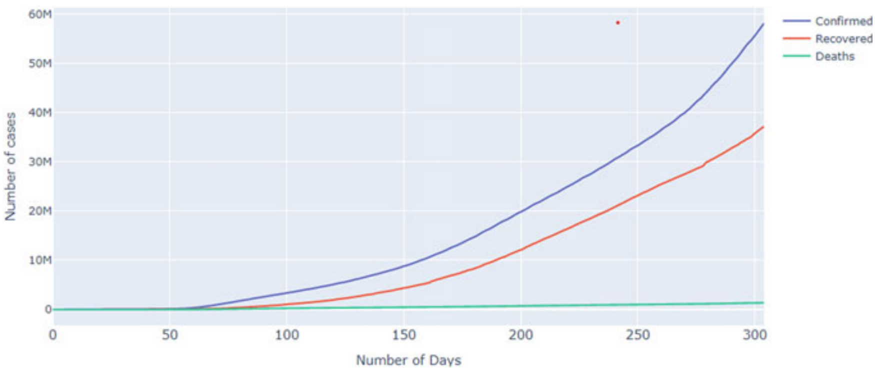


Fig. 4.1 Confirmed, recovered and death cases

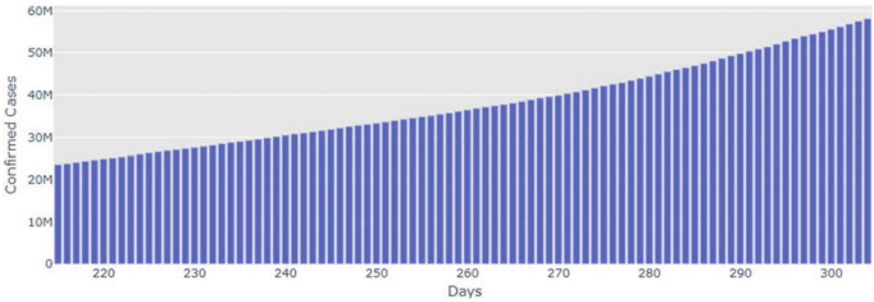


Fig. 4.2 Confirmed cases of last 90 days

in Figs. 4.3 and 4.4. And we can see that the death cases are rising at a slower pace than confirmed and recovered cases.

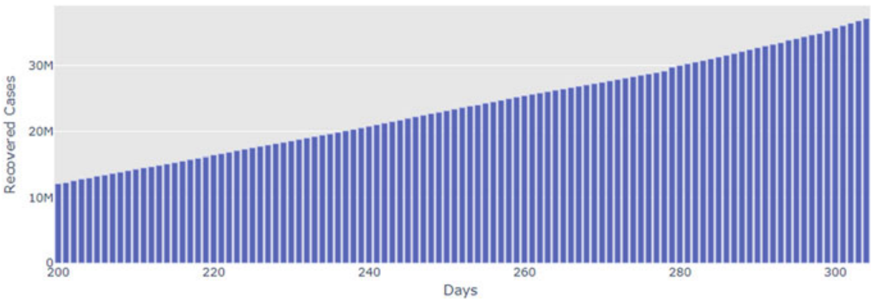


Fig. 4.3 Recovered cases of last 90 days

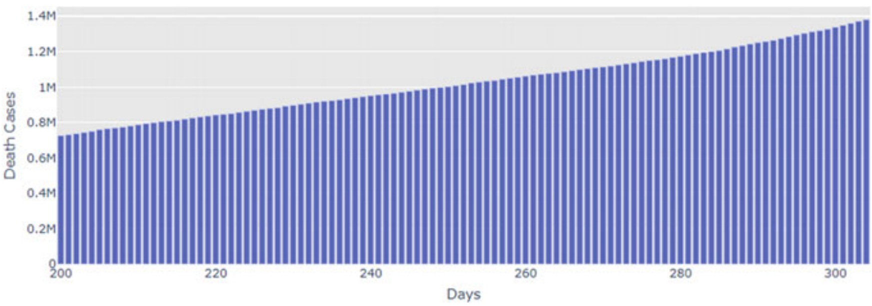


Fig. 4.4 Death cases of last 90 days

4.2.3 Model

Univariate long short-term memory (LSTM) model has been used to make predictions for worldwide confirmed, recovered and death cases due to coronavirus in the forthcoming days. LSTM networks are a form of recurrent neural network capable of learning about order dependency in problems of sequence prediction [6]. For processing, classifying and predicting values based on time series data, the LSTM model is useful [7]. Mean absolute error (MAE) has been used as a metric to measure the performance of the model. The formula to calculate MAE is given below.

$$\mathrm{MAE} = \frac{\sum_{i=1}^{n} |y_i - x_i|}{n}$$

MAE mean absolute error.
y_i prediction.
x_i true value.
n total number of data points.

4.3 Results and Discussion

In both the above figures, the representation on X-axis is the number of days and on Y-axis is the total count of confirmed cases. Figure 4.6 is the magnified view of Fig. 4.5 for the last ten days. Blue-colored line shows the actual count of confirmed cases, and the red line from 295 to 305th day shows the predicted confirmed cases by the model. As stated earlier, we have used MAE as a performance metric to evaluate the model. The MAE value is around 210,356. Key observation to be notice here that predictions are quite accurate because the original cases have touch 60 million bar; therefore, predictions are quite in alignment with the actual cases (Fig. 4.6).

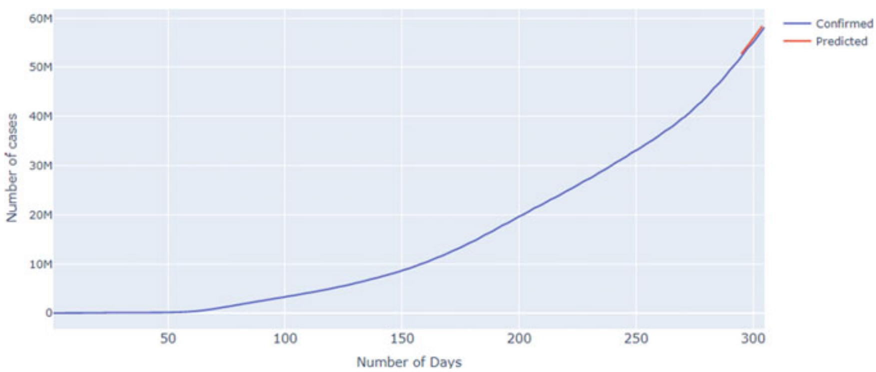


Fig. 4.5 Predicted confirmed versus actual confirmed cases

Fig. 4.6 Predicted confirmed versus actual confirmed cases of last 10 days

In both the above figures, the representation on X-axis is the total count of days and on Y-axis is the total count of recovered cases. Figure 4.8 is the magnified view of Fig. 4.7 for the last ten days. The MAE value is around 82,465. The recovered cases are almost 38 million, and the predictions are quite accurate with respect to actual recovered cases.

In both the above figures, the representation on X-axis is total count of days and on Y-axis is the total count of death cases reported. Figure 4.10 is the magnified view of Fig. 4.9 for the last ten days. The MAE value is around 2836. The recovered cases are almost 1.4 million, and our predictions are predicting somewhat more cases, which is good and helps us to get prepare for the worst scenario.

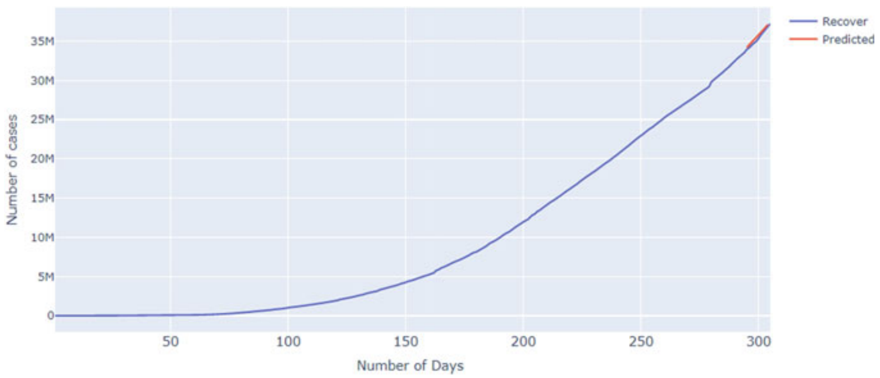


Fig. 4.7 Recovered cases versus predicted recovered cases

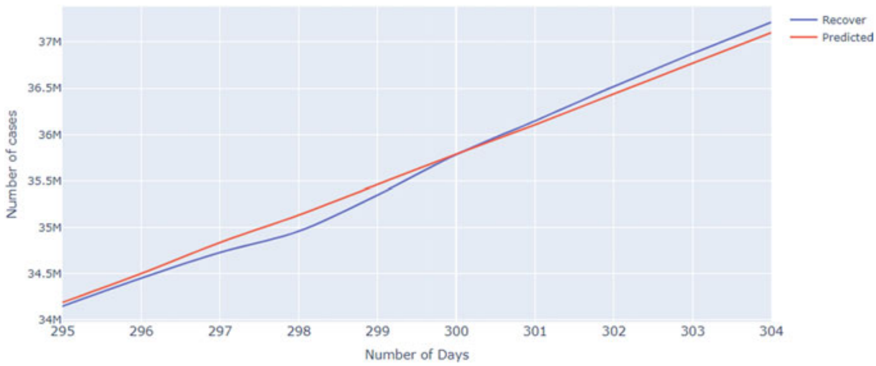


Fig. 4.8 Recovered cases versus predicted recovered cases of last ten days

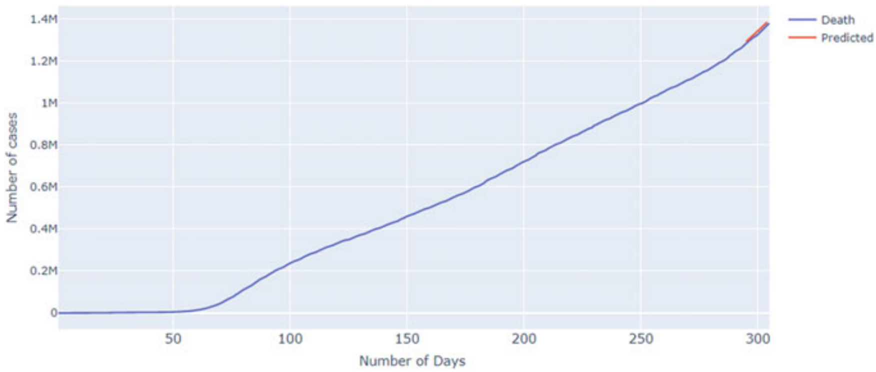


Fig. 4.9 Death cases versus predicted death cases

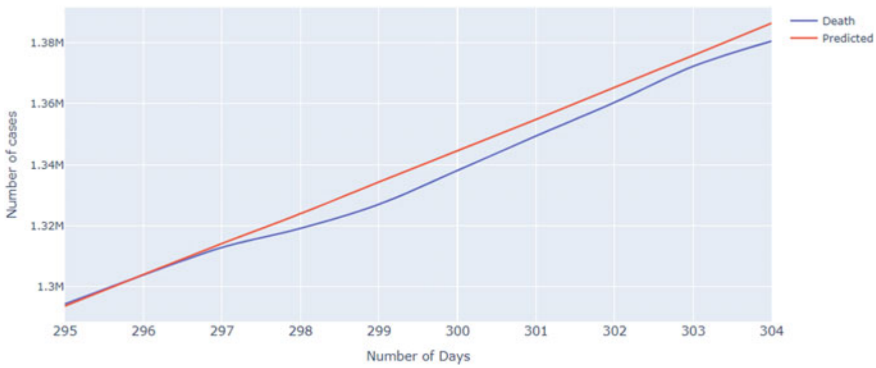


Fig. 4.10 Death cases versus predicted death cases of last ten days

4.4 Conclusion

In this paper, prediction of confirmed, recovered and death cases of COVID-19 across the world in the forthcoming days has been successfully done using univariate LSTM models. The results are then compared with actual values, and they turned out to be quite accurate. The accuracy of the obtained results proves that LSTM models are beneficial for carrying out the time series analysis on COVID-19 cases.

References

1. Monica, G., Bharathi Devi, M.: Using machine learning approach to predict Covid-19 progress. Int. J. Modern Trends Sci. Technol. **6**(8S), 58–62 (2020)
2. Pasayat, A.K., et al.: Predicting the COVID-19 positive cases in India with concern to lockdown by using mathematical and machine learning based models
3. Sujatha, R., Chatterjee, J.M., Hassanien, A.E.: Correction to: a machine learning forecasting model for COVID-19 pandemic in India. Stoch. Environ. Res. Risk Assess. **34**, 1681 (2020). https://doi.org/10.1007/s00477-020-01843-8
4. Fanelli, D., Piazza, F.: Analysis and forecast of covid-19 spreading in China, Italy and France. Chaos Solitons Fractals **134**, 109761 (2020)
5. Novel coronavirus diagnosis protocol. https://en.wikipedia.org/wiki/2019%E2%80%9320_coronavirus_pandemic
6. Ardabili, S., Mosavi, A., Ghamisi, P., Ferdinand, F., Varkonyi-Koczy, A., Reuter, U., Rabczuk, T., Atkinson, P.: COVID-19 outbreak prediction with machine learning (2020). https://doi.org/10.20944/preprints202004.0311.v1.
7. https://www.researchgate.net/publication/340782507

Chapter 5
SLAM Using Neural Network-Based Depth Estimation for Auto Vehicle Parking

S. Karamchandani, S. Bhattacharjee, D. Issrani, and R. Dhar

Abstract Simultaneous localization and mapping (SLAM) has been implemented as the standard technique for introducing autonomy in mobile robots or vehicles irrespective of the geographical environment. The unique characteristics of SLAM towards simultaneously and recursively solving the problem of localization and mapping make it an ideal approach without any human interference. Nonetheless, there are multiple cases where the camera pose estimation error increases significantly in dynamic scenes which results in the vehicle losing its location. This paper presents a novel approach towards SLAM using neural network-based depth estimation technique. We integrate the deep stereo ConvNet architecture which replaces the convolution layers with deconvolution layers and pooling with upscaling. We have evaluated our approach along with a comparison of performances from existing implementations on the TUM RGB-D dataset. The proposed implementation would be ideal algorithm for angle, perpendicular or parallel parking, and would be a helpful aid in autonomous street parking, parking lots, garage parking or any other parking spaces.

5.1 Introduction

The challenge of simultaneous localization and mapping (SLAM) asks if it is feasible for cars to be positioned in an unfamiliar area at an unknown position and for the car to create a coherent map of this environment incrementally while evaluating its position inside this map at the same time. For the mobile robotics culture, a response to the SLAM issue has been seen as a "holy grail" as it will have the capacity to make a robot fully autonomous. One of the significant achievements of the robotics society over the past decade has been the "solution" of the SLAM dilemma. In a variety of different ways, SLAM has been conceived and resolved as a theoretical concern. From indoor robotics to outdoor, aquatic and airborne devices, SLAM has also been introduced in a variety of different realms. SLAM should now be considered

S. Karamchandani (✉) · S. Bhattacharjee · D. Issrani · R. Dhar
Dwarkadas J. Sanghvi College of Engineering, Mumbai, India

T. Senjyu et al. (eds.), *IOT with Smart Systems*, Smart Innovation,
Systems and Technologies 251, https://doi.org/10.1007/978-981-16-3945-6_5

a fixed issue at a theoretical and logical basis. Major problems exist, however in the functional realization of more general SLAM solutions and in the development and use of visually rich maps as part of the SLAM algorithm.

The parallel calculation of the condition of a vehicle fitted with sensors and the creation of a model (map) of the area sensed by the sensors is part of SLAM. The state of the vehicle is defined by its location (position and orientation) in specific cases, though some variables can be included in the condition, such as velocity, sensor errors and metrics of calibration. The map is a depiction of facets of interest that define the area in which the vehicle resides (e.g. location of landmarks and obstacles).

There are two explanations for the need to use an ecological map. First to facilitate other functions, the map is also necessary; for example, a map may advise route planning or provide a human operator with an intuitive visualization. Second, the map makes the mistake incurred in calculating the vehicle's state to be reduced. Dead-reckoning can easily wander over time in the absence of a map; on the other hand, the vehicle would "reset" its position mistake by revisiting known places by using a map, e.g. a collection of distinguishable landmarks (so-called loop closure). Therefore, in all cases in which a prior map is not usable and has to be constructed, SLAM seeks applications. The position of a group of references is known a priori in certain industrial application. For example, with a manually designed map of synthetic detectors in the area, a robot working on a factory floor may be created. The situation where the vehicle has connection to GPS is another illustration. SLAM may not be needed in such situations if localization can be performed effectively with respect to recognized landmarks. The prominence of the SLAM problem is attributed to the advent of indoor mobile robotics applications. Indoor activity renders out the use of GPS to attach the position error; in addition, SLAM offers an enticing solution to user-built maps, illustrating that in the absence of an ad hoc localization framework, robot operation is feasible. In helping vehicular systems find their way through an unknown world, simultaneous localization and mapping (SLAM) has a crucial role to play, and this has many repercussions in different industries. For example, without previous knowledge of space, drones may be designed to navigate their way into a logistics warehouse in order to collect information from a specific shipment. With a low-cost solution that is increasingly gaining popularity, the fact that this can be achieved using a single camera extends the realm of possibilities SLAM can provide. We extend this algorithm for autonomous vehicle.

Depth maps would need to be produced based on images captured on the camera in order to incorporate SLAM so that precise monitoring can be accomplished. Due to the high precision that can be obtained in the image recognition space, convolutional neural networks (CNNs) are currently commonly used in image perception for monocular cameras. Relatively low neurons transmit information to higher-level neurons about the objects in an image, where they conduct additional convolutions using characteristics found in the picture to classify these items. CNNs, however, have a serious restriction that requires them to recognize only if the objects remain in the image; they do not encode the objects' position and location. Geoffrey Hinton, one of the founding fathers of deep learning, discussed this problem and suggested an alternative deep neural network that could resolve this downside.

In this paper, we present a novel approach towards optimization of simultaneous localization and mapping for vehicles. The methodology makes use of depth maps trained on the stereo ConvNet which provides improvement in terms of accuracy. The research is divided into five sections. Section two gives an overview of the advancements in the SLAM problem. Section three describes the techniques implemented along with equations. Section four showcases a comparison of the results which validates the performance optimized in our method.

5.2 Related Work

SLAM has started to grow rapidly over the previous decades since the advent of the SLAM methodology in robotics computer technology. In terms of accuracy, prediction and error elimination, the advancement of SLAM remains even today to enhance the efficiencies and features of the SLAM algorithm.

At the IEEE Robotics and Automation Conference in San Francisco, the study conducted in 1986 proposed the idea of applying the calculation of spatial ambiguity [1]. It became the building block for the pre-development of the system of SLAM. Then in 1991 [2], developed the SLAM method based on historical work done in that used a probabilistic model to identifying the issue of SLAM. The implementation of the expanded Kalman filter approach is implemented in this work, which later implements the first EKF-SLAM algorithm. In 2001 [3], the use of millimeter waves (MMW) was suggested to construct relative maps related to the environmental mapping phase of the robot system.

5.2.1 FastSLAM as an Alternate

The FastSLAM algorithm was introduced in 2002 [4]. Among the most known SLAM approaches after EKF-SLAM is this SLAM algorithm. FastSLAM's suggestion is to use a hybrid approach that incorporates the particle filter and the expanded Kalman filter approach. This implementation brings greater data precision to FastSLAM, which is famous for it. FastSLAM evolved a year back, which culminated in a second version of FastSLAM dubbed FastSLAM 2.0 [5]. The adjustment introduced in FastSLAM 2.0 is that the transmission of the proposal would be based on both the previous estimate of the pose and the real calculation of the robot, whereas the first iteration of FastSLAM depended only on the robot's previous estimation of the pose.

The smoothing method called square root smoothing and mapping (SAM) was proposed in 2006 [6] for the development of the mobile robot mapping operation. The solution uses the square root knowledge smoothing approach to solve the SLAM problem to increase the performance of the mapping process. In 2008 [7], introduced a new methodology called UFastSLAM that offers a robust algorithm based on scale-unscented transformation. By using the scale-unscented transformation algorithm strategy, the methodology was proposed to strengthen the FastSLAM process. In 2009 [8], a new methodology called differential evolution technique was developed to solve the SLAM problem. Chatterjee [9] worked to use an EKF-based neuro-fuzzy method for SLAM disorders and Ventura [10] for carrying out cell phone localization. Most of these projects are mainly carried out using a key frame-based approach, where if a good range of points are found in the image stream to be monitored in the next frame, those points are marked as key points. The Harris corner detector [11] is considered a strong key frame algorithm used in previous works. The basic concept is to create a picture area that, if there are adjustments in all directions, may be used for monitoring. In addition, a few more algorithms are also good at detecting key points, such as the FAST corner detector proposed by Roston and Drummond in 2006 and the Gaussian Laplacian (LoG) blob detector, which identifies blobs on the smoothed picture instead of corners found using Laplacian. In addition, CNN-based SLAM [12] can be used to define the depth projections and feed data to the key frame initializer and to fuse the semantic mark, but much of the analysis is performed using the depth or stereo camera effectively. In particular, the expected depth map is used as a source for the equation for Kellar's point-based fusion (RGB-D SLAM), but due to blurring artefacts, it lacks sharp data.

The majority of the SLAM algorithm implemented deals with single-robot projection and fixed environment. Modern SLAM innovation, however, reveals that various frameworks have been introduced in SLAM to handle increasingly complex environments and numerous mobility of robots. From the very first SLAM algorithm adopted to the new SLAM creation, the SLAM algorithm continues to boost and strengthen its accuracy. Similar SLAM algorithms aim to fix and improvise numerous facets of SLAM issues. The ultimate aim of the SLAM algorithm is to realize the difficulty of efficient navigation, localization and mapping capable of self-exploration by vehicle without human intervention over a period of time in complex environments.

5.3 Methodology

In this section, by reconstructing the disparity map from the single image, we clarify the outline of the proposed system for map creation and it is combined with the observed key frame and optical flow measurement to create the global map of the atmosphere. Figure 5.1 details the flow diagram of the implemented methodology.

Interestingly, we explain how the depth of the image may be found when using a monocular image. It is slow to use the CNN-based solution, and a lot of image pre-processing is needed because of a lack of CNN information transfer. In order to produce a good depth accuracy, we recommend a depth map on the whole frame using deep stereo ConvNet architecture. Moreover, by calculating the pixel-wise trust of each disparity generated, an uncertainty map is created.

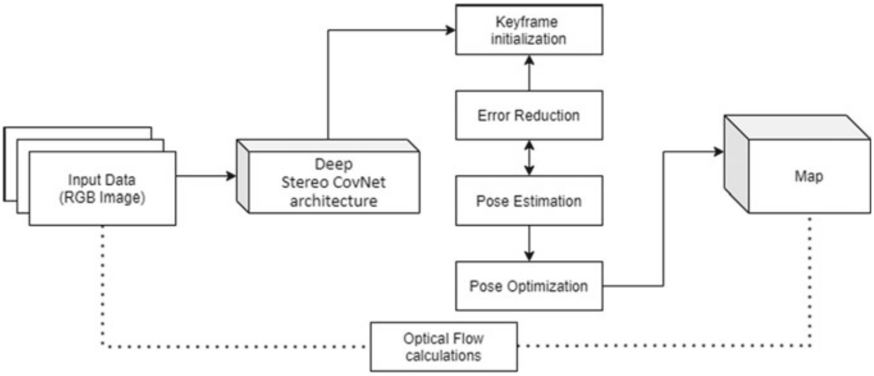


Fig. 5.1 Flow diagram of the algorithm using the deep ConvNet architecture

5.3.1 Pose Estimation

The calculation of camera poses is based on the key frame method. In particular, the picture and predicted camera location C_0 are projected onto all potentially recognizable landmarks, and a warped prototype is created for each landmark from the depth map D_k for each key frame $k_1 \ldots k_n \in K$. The camera location C_t^{ki}, i.e. the transition between the closest point k_i and frame t, generated by the square rotation matrix and a 3D translation vector, is determined at each frame t. It was then possible to define the intended camera pose as in (5.1).

$$E\left(C_t^{ki}\right) = \sum \rho\left(\frac{r\left(u, C_t^{ki}\right)}{\sigma\left(r\left(u, C_t^{ki}\right)\right)}\right) \tag{5.1}$$

where ρ is the Huber norm and σ is a function measuring the residual uncertainty. r is the optical residual defined by unmapped pixels from 2D points to 3D coordinates.

5.3.2 CNN Architecture

Stereo fully convolutional neural network using stereo images for depth map prediction. The framework is introduced using the Lasagne neural network library, and with the aid of the newly developed batch normalization method, training time/over-fitting is greatly decreased. The network is completely convolutional, which takes a few stereoscopic grayscale images concatenated along the axis of the channel and outputs a single image representing the map of depth. A series of convolution and maximum pooling layers followed by a series of upscaling and deconvolution layers allow the network to extract smaller-scale picture discrepancy characteristics (object edges) and produce a smooth approximation of the larger-scale depth map (full object).

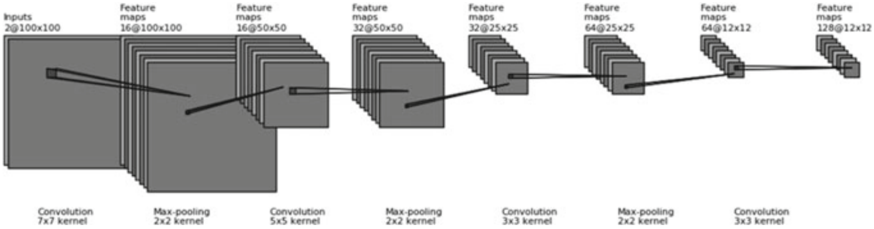


Fig. 5.2 CNN architecture block diagram

The key advantage of this approach over other computer vision analysis approaches (based on an explicit image difference calculation) is its robustness, in particular the fact that it is capable of generating smooth depth map calculations even on less field texture. Various orientations of the 3D model generated by the Blender software tool create the images and ground truth depth maps used for preparation, validation and research. We are using stereo ConvNet as our first step, and the first half of the network is seen below. The mirror image of the last convolution layer is the second half of the network, combining convolution with deconvolution and pooling with upscaling. Since the input image consists of concatenated left and right image pairs, it is taken as two independent images by the network. In deeper stereo ConvNet, input stays stable, but an additional convolution and deconvolution layer modifies the design. To collect further data, the depth of the filters is also increased.

5.3.3 Extended Kalman Filter

The Kalman filter is a popular method for fusing noise measurement of a dynamic system to get a good estimate of current state. It assumes the system to be linear and any dynamic model could be predicted efficiently if said system is linear. But since most real-world problems are nonlinear systems, the extended Kalman filter (EKF) seems to be more appropriate for handling such problems (Fig. 5.2).

5.4 Results

In this paper, we address different experiments conducted using different algorithms to obtain state-of-the-art outcomes. We compared our findings to ORB-publicly SLAM's accessible outcome [13] and CNN-SLAM, both of which are used for feature-based approaches in which one depends on key frames based on image outlines and the other on features generated using the CNN network and classification of images. We have contrasted our outcome with Direct Sparse Odometry SLAM, which is a heavier computational variant of dense SLAM. As a benchmark

Table 5.1 Depth estimation (correctly estimated) results on multiple architectures

Input data	Our method	ORB-SLAM	CNN-SLAM
TUM sequence 1	14.832	0.539	12.364
TUM sequence 2	23.441	1.007	22.896
TUM sequence 3	31.361	3.012	34.763

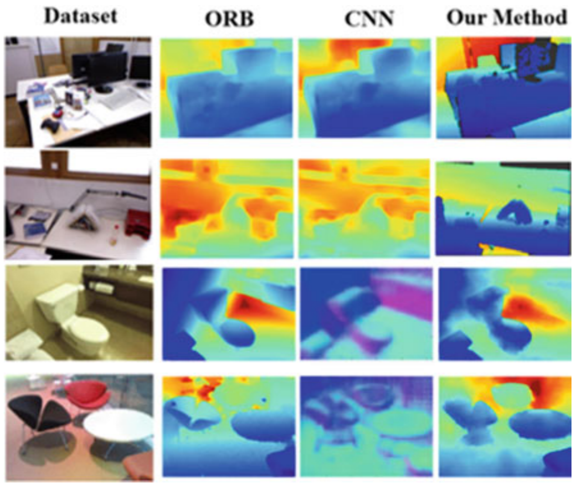


Fig. 5.3 Depth maps

for our SLAM process, data from the Technical University of Munich (TUM) was used. In the Faculty of Informatics, the data collection prepared by the Computer Vision Community comprises 50 video sequences with a total length of more than 100 min, captured in different real-world settings—both small indoor corridors and large outdoor areas. The findings in.

Table 5.1 demonstrates that with the exception of TUM sequence 3, our system achieves a higher degree of accuracy in most situations. In general, we can conclude that the stereo ConvNet does a better job of forecasting depth than the other strategies, and we can, for example, relate this outcome to the above-mentioned benefits of using the network relative to CNN (Fig. 5.3).

5.5 Conclusion

From our research, we have seen how the use of stereo ConvNet over CNNs has a lot of potential in optimizing vehicle depth estimation for SLAM issues. Also, when dealing with single-camera device constraints, depth maps acquired from the system can be fairly accurate for use. In order to increase overall performance, potential enhancements could require the fusion of additional sensors. The accelerometer, for instance, will identify the vehicle's position and path, and the gyro sensor can be used to preserve orientation. In order to create another series of depth map from a single angle, additional cameras can also be used. It would be important to process the information obtained from the other sensors differently from that of the images.

References

1. Smith, R.C., Cheeseman, P.: On the representation and estimation of spatial uncertainty. Int. J. Robot. Res. **5**(4), 56–68 (1986)
2. Beom, H., Cho, H.: Mobile robot localization using a single rotating sonar and two passive cylindrical beacons. Robotica **13**(3), 243–252 (1995). https://doi.org/10.1017/S026357470001777X
3. Dissanayake, G., et al.: A solution to the simultaneous localization and map building (SLAM) problem. IEEE Trans. Robotics Autom. **17**, 229–241 (2001)
4. Montemerlo, M., Thrun, S., Koller, D., Wegbreit, B.: FastSLAM: a factored solution to the simultaneous localization and mapping problem. In: Eighteenth National Conference on Artificial Intelligence, pp. 593–598. American Association for Artificial Intelligence, Menlo Park, CA, USA (2002)
5. Montemerlo, M., Thrun, S., Roller, D., Wegbreit, B.: FastSLAM 2.0: An improved particle filtering algorithm for simultaneous localization and mapping that provably converges. In: Proceedings of the 18th International Joint Conference on Artificial Intelligence, pp. 1151–1156. Morgan Kaufmann Publishers Inc., San Francisco (2003)
6. Dellaert, F., Kaess, M.: Square Root SAM: simultaneous localization and mapping via square root information smoothing. Int. J. Robot. Res. **25**(12), 1181–1203 (2006). https://doi.org/10.1177/0278364906072768
7. Kim C., Sakthivel R., Chung W.K.: Unscented FastSLAM: a robust algorithm for the simultaneous localization and mapping problem. In: Proceedings of the 2007 IEEE International Conference on Robotics and Automation, Roma, Italy, 10–14 April 2007, pp. 2439–2445
8. Moreno, L., Garrido, S., Blanco, D., Muñoz, M.L.: Differential evolution solution to the SLAM problem. Robot. Auton. Syst. **57**, 441–450 (2009)
9. Chatterjee, A., Matsuno, F.: A neuro-fuzzy assisted extended Kalman filter-based approach for simultaneous localization and mapping (SLAM) problems. IEEE Trans. Fuzzy Syst. **15**(5), 984–997 (2007)
10. Ventura, J., Arth, C., Reitmayr, G., Schmalstieg, D.: Global localization from monocular SLAM on a mobile phone. IEEE Trans. Vis. Comput. Graph. **20**(4), 531–539 (2014)
11. Shuvalova, L., Petrov, A., Khithov, V., Tishchenko, I.: Interactive markerless augmented reality system based on visual SLAM algorithm. In: Robot Intelligence Technology and Applications 4. Advances in Intelligent Systems and Computing, p. 447. Springer, Cham (2017)
12. Tateno, K., Tombari, F., Laina, I., Navab, N.: CNN-SLAM: real-time dense monocular SLAM with learned depth prediction. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, pp. 6565–6574 (2017)
13. Godard, C., Mac Aodha, O., Brostow, G.J.: Unsupervised monocular depth estimationwith left-right consistency. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017

Chapter 6
Cross-View Gait Recognition Using Deep Learning Approach

Jyoti Bharti and Lalit Lohiya

Abstract Walking speed is used principally to assess the status of human health. Recent gait recognition systems often experience difficulties, including variations in the viewing angle and enormous variations in the intraclass. For decades, computer vision-based approaches are in great demand and more effective in clinical gait analysis. Used referenced dataset is under well-controlled conditions. Gait movement rate is the predominant human biomechanical determinant. This paper proposes a scientific way of determining the pattern of gait movement at a particular speed of healthy people. Based on deep learning technology, our proposed model comprises of a fully convolutional neural network accompanied by batch normalization and max pooling. Proposed models on large-scale gait datasets were being extensively tested and have been compared with the benchmark approaches. The proposed model's overall average accuracy is about 93.0%. Where various clothing and transport conditions have been encountered, the method is also robust for such conditions.

6.1 Introduction

Gait is one of the most popular biometric aspects of humans because it can be bona fide excluding subject cooperation at a distance from a lens. In the gait recognition model, the speed difference between matching pairs plays an important role and the gait mode of walking or running makes it more challenging. The distinctive features of the human body, such as handwriting, voice, and gait, have been highly studied for many years to make excellent progress in biometrics [1]. Every individual has a particular way of walking that happens because of facilitated and worked together activities of the skeletal muscles and sensory system. This makes the biometric step a ground-breaking marker to decide obsessive behavior caused by actual injury, maturation, or related issues. These inside and outside variables legitimately influence the movement and activity of the body and result in walk impedance [1]. Biomechanical examples of human movement are by and large speed-dependent, the adequacy of

J. Bharti (✉) · L. Lohiya
Maulana Azad National Institute of Technology, Bhopal, India
e-mail: jyotibharti@manit.ac.in

T. Senjyu et al. (eds.), *IOT with Smart Systems*, Smart Innovation, Systems and Technologies 251, https://doi.org/10.1007/978-981-16-3945-6_6

explicit development normally levels with the development speed (stride velocity is a determining factor in the nature of phases) [1, 2].

In an average walk examination, patients perform stride preliminaries at their agreeable speed and their step designs are regularly contrasted, and a reference design is from a regulating information base. Nonetheless, the impact of stride speed is commonly not represented when the step example of obsessive people is contrasted and sound ones who do not walk at an equal speed [2].

A repeated pattern consisting of steps and strides is the gait loop. A step phase began with one foot's initial contact and finishes with the other foot's initial contact. The stride phase begins with one foot's initial contact and finishes with the same foot's next initial contact, which comprises of two steps. There are two main phases in the gait cycle: stance phase comprises of 60% of the entire gait cycle and swing phase comprises of 40 percent of the full gait cycle. The step is the time when the foot is in touch with the ground and the weight is borne by the limb. The swing process is the time during which the reference foot is not in touch with the surface and swings in the air, indicating the weight of the body is borne by the contralateral limb [1, 3–5].

6.2 Related Work

6.2.1 Gait Recognition System

Human authentication based on gait has significant importance and a wide scope of research. The science of recognizing or identifying people by their physical characteristics is known as biometrics. One form of biometric technology is gait recognition [5]. The authors are working on gait recognition biometric technology to analyze the movements of body parts such as the foot, the knee, the shoulder, and so on. There are various gait acquisition technologies such as cameras, wearable sensors, and non-wearable sensors. Figure 6.1 demonstrates the gait recognition system.

Any recognition system is developed based on the training phase and the testing phase. There are two approaches, such as authentication and recognition. Automatic video processing was the basis for the first gait recognition [3, 6, 7]. It generates a mathematical model of motion. This method is the most common and requires the study of video samples of walking subjects, joint trajectories, and angles. In order to perform detection, the second approach uses a radar system to monitor the gait period that is then compared with the other samples [7].

Fig. 6.1 Gait recognition system

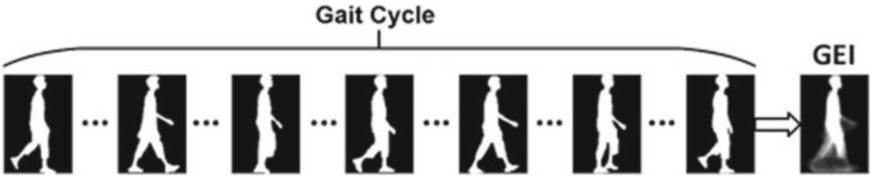


Fig. 6.2 Gait energy image extraction

A potential solution for this issue would be to gather a few walking preliminaries at different walking rates to construct a reference information base for essentially any conceivable step speed [8–11]. The tedious idea of such an assortment of knowledge, however, would be cost-restrictive and unviable. Scientists have suggested regression techniques as a reachable alternative for anticipating phase limits based on exploratory facts to resolve this barrier. The expectation data depends purely on the normal, slow, and fast walking speeds for stable subjects but only 10% time frame phase period when the complete step cycle was considered. [1, 2, 7, 12, 13]

The gait energy image (GEI) is an enhanced spatiotemporal feature to analyze a subject's gait and posture characteristics. Binary silhouette sequence normalization was used instead of direct silhouette sequence to provide robustness. GEI can be stated as,

$$\mathrm{GEI} = \frac{1}{S_f} \sum_{S=1}^{S_f} I_o(u, v, s) \tag{6.1}$$

where S_f states the total number of silhouette images in the gait cycle and s is an image number at an instant. I_o (u, v) denotes the silhouette frame at an instant. Figure 6.2 shows the extracted GEI image from full gait cycle.

6.3 Proposed Work

6.3.1 Implementation Details of Proposed Model and Model Summary

CASIA B Gait Dataset contains silhouette sequence for each of the view angle. This silhouette sequence of images has been converted to gait energy image. So, there are in total 13,640 images. 11 angle view * 10 categories * 124 subjects = 13,640 GEIs. To evaluate the proposed method, all the sequences have been converted to a gait energy image (GEI), then set of GEIs are fed to gait recognition proposed model. Figure 6.3 shows the various GEIs under three conditions and 11 view angles.

One layer consists of a 2D convolutional layer followed by batch normalization and max pooling in the proposed deep learning model for cross-view gait recognition.

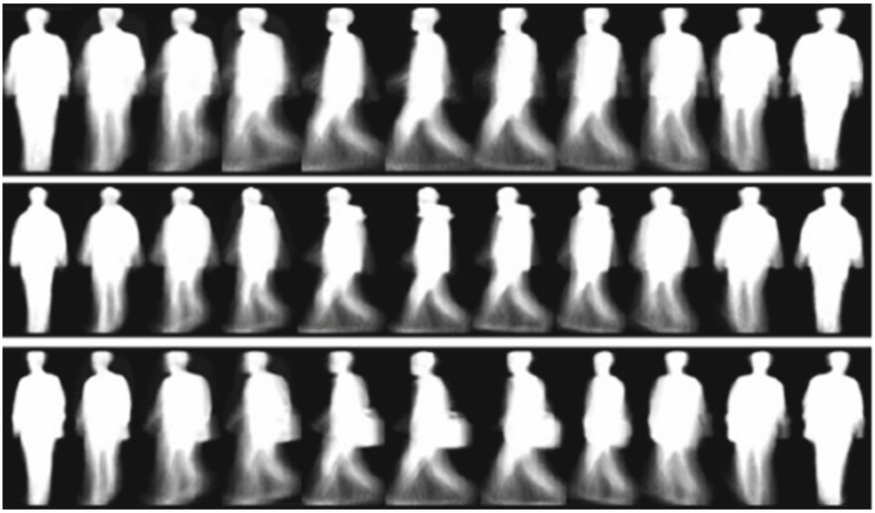


Fig. 6.3 GEI examples from an individual's CASIA B dataset under various conditions from view 0° to 180° with an 18° interval. Top row: regular walking, middle row: coat-wearing, and bottom row: bag-wearing

Fig. 6.4 Proposed model

As such, three stacks of layers are available in total. Figure 6.4 shows the design of the model proposed. Table 6.1 shows description of the model proposed.

6.3.1.1 Convolutional 2D Layer

A convolutional layer is a part of the deep neural network that can intake an input image, attribute importance to the different aspects/objects in the frame by learning weights and biases, and distinguish one from the other. The preprocessing needed for the convolutional neural network is much lower than for other classification algorithms. The size of input image is (240 × 240), and size of convolutional kernel is (7 × 7). The input image is downsampled, and output of convolutional 2D layer is about (234 × 234). Output of convolution layer is nothing but the feature map.

Table 6.1 Summary of proposed model (BN—batch normalization, MP—max pooling, and Conv—convolution layer)

Layers	Output size	No. of filters	Filter size	Padding	Params
Conv 1	234 × 234 × 16	16	7 × 7	1	2368
BN	234 × 234 × 16	–	–	–	64
MP 1	117 × 117 × 16	16	2 × 2	2	0
Conv 2	111 × 111 × 16	16	7 × 7	1	12,560
BN	111 × 111 × 16	–	–	–	64
MP 2	55 × 55 × 16	16	2 × 2	2	0
Conv 3	49 × 49 × 16	16	7 × 7	1	12,560
BN	49 × 49 × 16	–	–	–	64
MP 3	24 × 24 × 16	16	2 × 2	2	0
Conv 4	18 × 18 × 64	64	7 × 7	1	50,240

6.3.1.2 Batch Normalization

Batch normalization helps each network layer to learn much more independently of other layers. Higher learning values can be used because batch normalization assures that no activation has actually went high or really low. It reduces overfitting due to a slight regulation effect. To stabilize a neural network, it normalizes the performance of a previous activation unit by deducting batch average and dividing by batch confidence interval.

6.3.1.3 Pooling 2D

An issue with a feature map output is that they are sensitive to input feature position. To fix this sensitivity, sample the function maps is one approach. Pooling layers offer an approach to sample feature maps by summarizing features in the feature map kernel. Pooling is needed to detect feature in feature maps. We used max pooling. Results are pooled function maps highlighting the kernel's most present feature. The kernel size is (2 × 2).

6.3.1.4 Dropout

To simulate having a large number of distinct network architectures, a single model can also be used by arbitrarily lowering out nodes during training. This is known as dropout and provides a very inexpensive and effective regularization impact for reducing overfitting and enhancing generalization error in deep neural networks.

There are two sets of GEIs known as gallery set and probe set. While training the model gallery set is used and for evaluation probe set is used. As there are 11 views, there will be 11 models trained and tested.

6.4 Experiments and Result Analysis

6.4.1 CASIA-B Gait Dataset

This is a large multiview gait database; the information was gathered from 11 views and consists of 124 subjects. There are three varieties of variations clothing and carrying conditions including wide range of view angles and also consist of extracted silhouettes.

With 124 subjects, including both genders, 93 males and 31 females in an indoor environment, the data was gathered at 25 fps and a size of frame is 320×240. Therefore, there are 13,640 sequences in total. Figure 6.5 shows how data was collected from different angles [1–6, 14]. CASIA-B dataset contains 124 subjects; data is captured from 11 angle views from 0° to 180° with 18° interval. There are in total

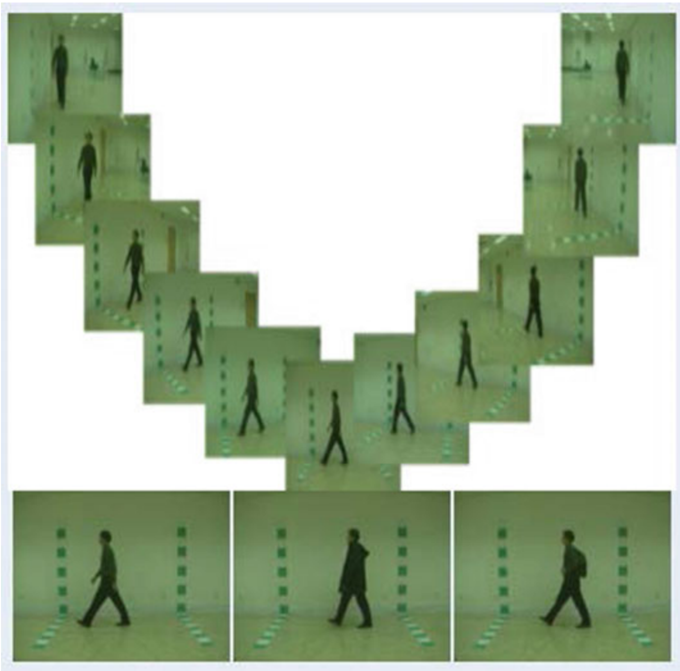


Fig. 6.5 11 angle view of CASIA-B dataset

10 categories such as normal walking with 6 subsets, walking while carrying bags with 2 subsets, and walking while wearing coat with 2 subsets.

6.4.2 Experimental Environment

To perform our experiment, we have used Python 3.7, Jupyter Notebook, and Anaconda Environment. TensorFlow 2.0 is used. The dataset used is CASIA-B dataset.

6.4.3 Result Analysis on CASIA-B Dataset

Proposed model is trained at 11 angle views, and each trained model is evaluated against different viewing angles probe set. Table 6.2 shows the experimental results on probe set.

It should be noted that the proposed method achieves the best recognition accuracies at almost all angles. This is the result of gait energy image over the silhouette sequence. For evaluation of the proposed model, we compared our model with GEI template methods, as input to model is GEIs. This are the models—Gait-Net [5], L-CRF [15], LB [16], RLTDA [17], CPM [12], and DV-GEIs [14]. Table 6.3 shows the comparison of the benchmark approaches and state-of-the-art models with the proposed approach with respect to some of the probe evaluation CASIA-B gait dataset.

Table 6.2 Shows cross-view gait recognition accuracy on CASIA-B dataset

Gallery	Probe										
	0°	18°	36°	54°	72°	90°	108°	126°	144°	162°	180°
0°	–	95.71	94.86	94.25	95.97	95.98	96.03	95.98	95.98	95.99	96.03
18°	95.27	–	94.38	95.50	95.98	95.94	95.98	95.94	95.89	95.76	95.72
36°	95.51	95.69	–	91.92	91.51	90.08	92.39	94.73	94.79	95.97	95.63
54°	93.50	92.24	89.29	–	77.47	73.39	75.78	85.42	90.84	94.45	93.63
72°	95.62	95.57	95.70	95.63	–	94.56	93.65	94.86	95.57	95.86	95.60
90°	96.00	95.90	96.01	95.76	96.50	–	95.99	95.78	95.99	96.02	95.95
108°	95.93	95.87	96.09	96.25	96.64	96.68	–	96.17	96.11	96.02	95.97
126°	96.04	95.73	95.86	96.07	96.30	96.44	96.51	–	96.15	95.99	96.06
144°	93.29	89.35	89.71	91.23	90.13	88.09	85.84	87.40	–	90.67	90.29
162°	93.80	92.73	89.15	84.83	84.35	83.15	81.94	86.97	89.26	–	94.17
180°	94.35	90.64	88.35	86.80	85.74	85.37	84.27	86.56	90.72	95.37	–

Table 6.3 Comparison with under different walking view angles on CASIA-B by accuracies

Probe	Gallery	Proposed model	GaitNet [5]	L-CRF [38]	LB [37]	RLTDA [17]
54°	36°	91.92	91.6	93.8	92.7	80.8
54°	72°	95.63	90.0	91.2	90.4	71.5
90°	72°	94.56	95.6	94.4	93.3	75.3
90°	108°	95.99	87.4	89.2	88.9	76.5
126°	108°	96.17	90.1	92.5	93.3	66.5
126°	144°	87.40	93.8	88.1	86.0	72.3
	Mean	93.61	91.4	91.5	90.8	73.8

Table 6.4 Comparison of recognition cross-view accuracy on CASIA-B dataset

Methods	Average accuracy (%)
CPM [12]	24.1
GEI-SVR [13]	42.2
CMCC [18]	43.9
ViDP [19]	45.4
LB [16]	56.9
L-CRF [15]	67.8
Gait-Net [5]	81.8
DV-GEIs [14]	83.4
Proposed model	93.0

It should be noted from Table 6.2. The average accuracy of the proposed model is around 93.0%, which outperforms the GaitNet with 12.2% and DV-GEI with 9.6%. Compared the output with state-of-the-art methods like L-CRF, use GEIs instead of.

Silhouette sequence as an input to the model, this is another proof that our gait characteristics strongly illustrate all major gait variations. Table 6.4 shows the comparison of cross-view recognition accuracy on CASIA-B dataset.

6.5 Conclusion

The deep learning approach is proposed in this paper. The proposed method can efficiently extract spatial and temporal parameters. In this model, we fed gait energy image (GEIs) as input to the model rather than gait silhouette series. This method is efficient than conventional methods. The gait energy image keeps the dynamic and static facts of a gait sequence, although time is not properly considered, this is the one limitation of the proposed method. The proposed model is robust in variations including obstructive objects, clothes, and viewing angles. Eventually, gait awareness develops stronger. Our model was tested on a broad CASIA-B dataset. Max pooling layer is used to adapt the spatial information, improving the mapping efficiency. Recognition results obtained demonstrated our model's dominance over well-known approaches.

References

1. Sepas-Moghaddam, A., Etemad, A.: View-invariant gait recognition with attentive recurrent learning of partial representations. IEEE Trans. Biometrics Behavior Identity Sci.
2. Sepas-Moghaddam, A., Ghorbani, S., Troje, N.F., Etemad, A.: Gait recognition using multi-scale partial representation transformation with capsules (2020). arXiv preprint arXiv:2010.09084
3. Lin, B., Zhang, S., Yu, X., Chu, Z., Zhang, H.: Learning effective representations from global and local features for cross-view gait recognition (2020). arXiv preprint arXiv:2011.01461
4. Fan, C., Peng, Y., Cao, C., Liu, X., Hou, S., Chi, J., Huang, Y., Li, Q., He, Z.: GaitPart: temporal part-based model for gait recognition. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 14225–14233 (2020)
5. Zhang, Z., Tran, L., Yin, X., Atoum, Y., Liu, X., Wan, J., Wang, N.: Gait recognition via disentangled representation learning. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 4710–4719 (2019)
6. Chao, H., He, Y., Zhang, J., Feng, J.: Gaitset: regarding gait as a set for cross-view gait recognition. In: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 33, pp. 8126–8133 (2019)
7. Moissenet, F., Leboeuf, F., Armand, S.: Lower limb sagittal gait kinematics can be predicted based on walking speed, gender, age and BMI. Sci. Rep. **9**(1), 1–12 (2019)
8. Nandy, A., Chakraborty, R., Chakraborty, P.: Cloth invariant gait recognition using pooled segmented statistical features. Neurocomputing **191**, 117–140 (2016)
9. Anusha, R., Jaidhar, C.D.: Human gait recognition based on histogram of oriented gradients and Haralick texture descriptor. Multimedia Tools Appl. 1–22 (2020)
10. Hasan, M.M., Mustafa, H.F.: Multi-level feature fusion for robust pose-based gait recognition using RNN. Int. J. Comput. Sci. Inform. Secur. (IJCSIS) **18**(1) (2020)
11. Janković, M., Savić, A., Novičić, M., Popović, M.: Deep learning approaches for human activity recognition using wearable technology. Medicinski podmladak **69**(3), 14–24 (2018)
12. Chen, X., Weng, J., Wei, Lu., Jiaming, Xu.: Multi-gait recognition based on attribute discovery. IEEE Trans. Pattern Anal. Mach. Intell. **40**(7), 1697–1710 (2017)
13. Kusakunniran, W., Wu, Q., Zhang, J., Li, H.: Support vector regression for multi-view gait recognition based on local motion feature selection. In: 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 974–981. IEEE (2010)
14. Liao, R., An, W., Yu, S., Li, Z., Huang, Y.: Dense-view GEIs set: view space covering for gait recognition based on dense-view GAN (2020). arXiv preprint arXiv:2009.12516
15. Recognition based on attribute discovery. IEEE Trans. Pattern Anal. Mach. Intell. (TPAMI) **40**(7), 1697–1710 (2018)
16. Wu, Z., Huang, Y., Wang, L., Wang, X., Tan, T.: A comprehensive study on cross-view gait based human identification with deep CNNs. IEEE Trans. Pattern Anal. Mach. Intell. **39**(2), 209–226 (2017)
17. Haifeng, Hu.: Enhanced gabor feature based classification using a regularized locally tensor discriminant model for multiview gait recognition. IEEE Trans. Circ. Syst. Video Technol. **23**(7), 1274–1286 (2013)
18. Kusakunniran, W.: Recognizing gaits on spatio-temporal feature domain. IEEE Trans. Inf. Forensics Secur. **9**(9), 1416–1423 (2014)

19. Hu, M., Wang, Y., Zhang, Z., Little, J.J., Huang, Di.: View-invariant discriminative projection for multi-view gait-based human identification. IEEE Trans. Inf. Forensics Secur. **8**(12), 2034–2045 (2013)

Chapter 7
An Overview of Electrical Load Classification and Prediction Methods

Mncedisi S. Figlan and Elisha D. Markus

Abstract For decades, the topic of load prediction has existed, and various approaches have been followed to achieve adequate results. Most of these techniques are based on a traditional or contemporary method. This article describes a compacted background of scientific work by engineers and compares traditional and new methods with the short-term electrical load prediction. Here, six various ways of predicting loads and their short outline are discussed. Every technique gets its own methodology; all the positive and negative aspects are covered, and the benefits and drawbacks listed. Suggestions for future research and observations are made.

7.1 Introduction

As electrical system networks grow steadily and their complexity increases, many elements have played a vital role in the generation, demand, and administration of electrical energy. For simple power plant operations, forecasting is aimed to assist planners make informed decisions on unit involvement, hydrothermal coordination, interchange reviews, and safety assessments, etc. Statistical approaches or intelligent systems, such as system of experts, fuzzy logic, neural networks, and regression, are used in a large number of prediction models [1].

In generalisation and mastery of nonlinear relationships between variables, artificial neural networks (ANNs) have proven to be successful, and therefore, ANN-based strategies are frequently preferred for short-terminal load forecast (STLF) problems.

Typical models of load prediction can be divided into two main groups: dynamic and daytime models. A technique that is non-dynamic is the time-of-day model, expressing the load as a specific time series which is based on the forecasting method, predicts that each of the forecast day value will fluctuate from hour to hour. The

M. S. Figlan (✉) · E. D. Markus
Center for Sustainable SMART Cities, Department of Electrical, Electronic and Computer Engineering, Central University of Technology, Free State, Bloemfontein, South Africa

E. D. Markus
e-mail: emarkus@cut.ac.za

T. Senjyu et al. (eds.), *IOT with Smart Systems*, Smart Innovation, Systems and Technologies 251, https://doi.org/10.1007/978-981-16-3945-6_7

dynamic model, on the other hand, recognises that the load is affected not only by the daytime, but also by the load's past behaviour. Based on various variables such as time, weather, and customer classes, the load can be mathematically represented. In a standard optimization method, the predicted electricity use is provided by

$$\hat{L} = L_n + L_w + L_s + L_r \tag{7.1}$$

where $\hat{L}$ is the estimated total load, L_n is the normal part of the load, Lw is the weather-sensitive part of the load, L_s is a special event component that causes a major deviation from the standard load pattern, and L_r is a complete random phrase, which is the noise component. Prior findings indicate that good forecasting results can be achieved by taking energy prices into account in mathematical equations.

The expression of a multiplicative method can be expressed in the form

$$\hat{L} = L_n \cdot F_w \cdot F_s \cdot F_r \tag{7.2}$$

where L_n is the normal (base) load, and the correction factors F_w, and F_r positive numbers that may increase the overall load or decrease it. Such modifications are established on current weather (F_w), special events (F_S), and random fluctuations (F_r), including factors such as the price of energy or the growth factor of load.

The approaches used for load prediction are the same day approach, models of regression, fuzzy logic, series of time, artificial intelligence, learning from statistic approach, and expert of systems. Such approaches can be categorised according to their statistical analytical levels in the forecast models.

- Methods of regression
- Series of time
- Daytime techniques
- Similar day method
- Stochastic time series structures
- System focused on intelligent networks (using GA and ANN).

7.2 Load Prediction Techniques and Comparison

7.2.1 Methods of Regression

This is among the most widely used methods for predicting electrical load. By defining a mathematical equation, regression methods attempt to model the connection between the influencing factors such as changes in climate and day form and electricity load frequency. Regression is a commonly used statistical approach that is relatively simple to implement. Regression is utilised to represent the relationship between electricity use and other variables such as temperature, day types, and customer groups. Such a method is based on the assumption that a standard pattern based on loads and a design can be divided linearly based on such load-influencing factors. The regression-based algorithms, with a high computational load and lengthy computational time, are of high complexity.

The expression is

$$\hat{L}(t) = L_n(t) + \sum_{i=1}^{n} a_i x_i(t) + \varepsilon(t) \tag{7.3}$$

where $L_n(t)$ is the normal or standard load at time t, a_i is the estimated slowly varying coefficients,$x_i(t)$ are the independent influencing factors such as weather effect, $\varepsilon(t)$ is a white noise component, and N is the number of observations, usually 24 or 168 h.

7.2.2 Series of Time

The time series strategy can be described as a successive collection of data, such as hourly, or weekly loads, calculated over time. The fundamental principle of forecasting is to first establish as accurately as possible a format recognition of accessible information and then to calculate the forecast values using the approved model with respect to time.

Methods of time series assume that the data have an internal structure like autocorrelation, pattern or seasonal change. The first impetus is the precise assembly of the available data to fit trends, and the time value predicted using the given model.

The model's specifications can be defined as

$$\hat{L}(t) = L_n(t) + S(t) + R(t) \quad t = \ldots - 1, 0, 1, 2 \tag{7.4}$$

where the normal load value is denoted by $L_n(t)$, the seasonal term is $S(t)$, and the abnormal or random part is $R(t)$.

7.2.3 Daytime Techniques

The simplest type of load prediction is this method. The model forecasts the current week's load based on the previous week's actual load trend. In addition, a series of electricity usage are stored for typical week with different weather conditions. To establish the prediction, they are then integrated computationally. The model's formula could be expressed as

$$\hat{L}(t) = L_n(t) + \sum_{i=1}^{n} a_i f_i(t) + \varepsilon(t) \tag{7.5}$$

where load at time *t*, $\hat{L}(t)$ is known as the sum of explicit time functions, $f_i(t)$ normally sinusoids for 24 h or 168 h based on the lead time,a_i are coefficients that differ slowly based on time, and $\varepsilon(t)$ is an error term.

7.2.4 Similar Day Method

The "similar day" method takes into account a "similar" day in the historical data to the one predicted. It is routinely implemented in industrial applications due to its simplicity. The similarities are typically based on the patterns of calendars and weather. A linear combination or regression method that involves many similarities on the days can be the forecast. A similar days' approach can be formulated using the Euclidean norm. The load is taken by this model as a row vector, wherein adjusted parameters are used to test the similarity between the predicted days to search previous days. Assume the vector $\mathbf{x} = [x_1, x_2, \ldots, x_n] \in R^n$, then the equivalent Euclidean norm can be defined in R^n space by the formula:

$$\|x\| = \sqrt{x_1^2 + \cdots + x_n^2} \tag{7.6}$$

7.2.5 Stochastic Time Series Structure

The methods of stochastic time series are based on the assumption that data have an internal structure such as trends or autocorrelation. The methodology for time series techniques is created on the basis of the past load data. The future load is then forecast based on the developed model. Reference [2] developed a method to face problems with stochastic composition optimisation with two expected value functions. An inner goal function was incorporated into an outer one by the approach.

In relation to its previous value, the method of the stochastic time series presents the existing load linearly, and the zero mean and variance white noise sequence, contrasting with the classic forecasting techniques.

The backshift operator is introduced by this representation and enables the technique to partially control the difficulty of complex load prediction. The prototype is shown as

$$x_t = \emptyset_1 x_{t-1} + \emptyset_1 x_{t-2} + \emptyset_p x_{t-p} + \varepsilon_t \tag{7.7}$$

where the actual load value is x_t, x_{t-1} and x_{t-2} are the past values of the load, and ε_t is a zero mean white noise. A backshift operator is the basic concept of this method.

7.2.6 *Systems Focused on Intelligent Networks*

The thread that unifies so many different concepts are woven from the interpretation of the intelligent system. In practice, an intelligent system is used to comply with some or all of its computer requirements by means of artificial intelligence (AI). Many studies have shown that intelligent system methods are superior to load forecast models. Some of the widely used artificial intelligence strategies are now discussed briefly in the following section.

Neural Network based on Artificial Systems

An artificial neural network (ANN) is a computational model built on the biological neural system. The network comprises interconnected memories and arranges data using a generative computing technique. Neuron-distributed simulation results in intelligent outcomes. The ANN structure learns to complete the required function using specialised training principles directly from examples. Figure 7.1 illustrates the basic structure of an ANN process. This technique is mostly seen as an integrated approach that adjusts its design on the basis of external or internal network data during training.

A Multilayer Perceptron

Figure 7.2 shows a network establishment of a three-layer feedforward system. The inputs are fed and multiplied by interconnection weights into the input layer, before moving to the next layer, and then passed through an activation function.

Expert (Methods) Systems

This approach is using high-level machine learning. It is built by computer programmers and specialists by means of close interactions and experience. The scientific objective of AI is to grasp intelligence by designing software applications that show inventive behaviour. The ideas and techniques for computer reasoning or symbolic

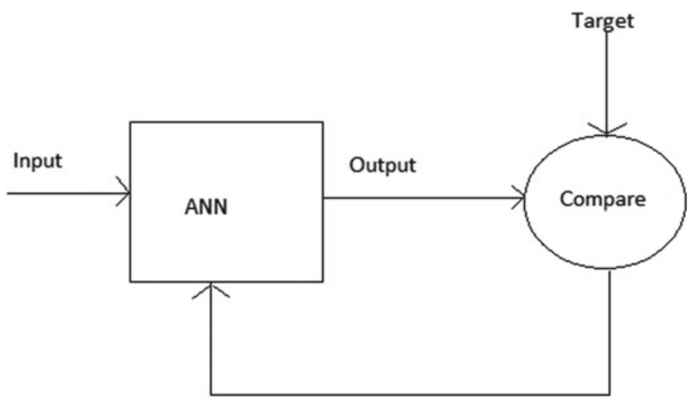
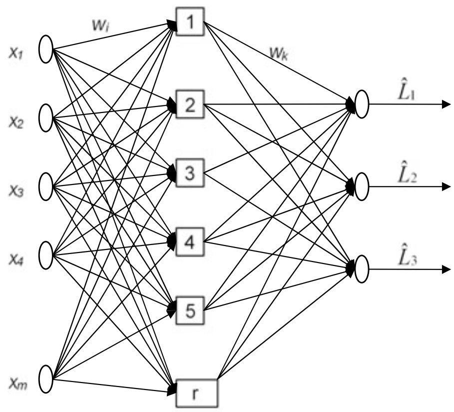


Fig. 7.1 Simple structure of ANN

Fig. 7.2 Single hidden layer and multi-output feedforward neural network



inference are addressed and how the information used to generate these results is depicted inside the machine.

The basic principle here is the manipulation and encapsulation of high-level information to emulate an expert's behaviour.

Expert systems or knowledge-based expert systems (KBES) are recent heuristic techniques resulting from progress in the artificial intelligence (AI) field. No specific technique structure or historic pattern is needed for the ES.

Once the electricity use and the variables that impact it are identified and extracted, expert systems can apply a parameter-based rule. This rule is of the form "if–then", plus some mathematical expressions. This rule can be used on a daily basis to generate the forecasts.

Fuzzy Approach

In the growth of fuzzy set theory, the use of the term "Fuzzy Logic" became widespread during the 1960s. A fuzzy logic model is a logical–mathematical mechanism that imitates the human viewpoint in the simple machine form based on an "if–then" rule structure. There are typically four modules in a fuzzy rule structure:

(a) Input fuzzification—which changes the "zingy" input method to a fuzzy one.
(b) Fuzzy rules—if–then logical rule that connects the input and output variables.
(c) Fuzzy inference—a system that illustrates and incorporates the effects of rules.
(d) Output defuzzification—the tool that converts the fuzzification into a fuzzy output number.

Developmental Computing

Recent STLF literature indicates that one of the appropriate methods is the genetic algorithm (GA), particularly for optimising the network model of load prediction. The dependency on earliest conditions, lengthy preparation, routing protocol configuration, etc., are most of the disadvantages associated with conventional experts computing and can be easily solved using this method.

It was deemed appropriate to provide a brief introduction to the theory behind the method due to the expected dominance of GAs in STLF.

Optimization Based on Genetics

A number of straightforward optimisation problems can be solved using the basic regulations based on back propagation rules. Nonetheless, their efficiency falls quickly as problem complexity increases. Other disadvantages include problems such as lengthy instruction time, one-point quest, weighing, and reliance on earliest conditions. The genetic algorithm (GA) is, however, seen as a backup solution.

The first discovery of this method was by John Holland at Michigan University in the mid-1970s. The key concept was to develop artificial systems that maintain natural systems' robustness and adoption properties. Since the beginning, other researchers have enhanced these methodologies, and in various fields (business, research, engineering, etc.), they are now commonly used to resolve a spectrum of problem maximisation outside the reach of conventional toolbox multiplication.

In a given N-dimensional potential number of solutions, GA emulates physiological mechanisms to run a selection of mechanisms. Through a given search space for an optimisation problem, one must try to determine the right answer. Darwin's evolutionary theory (survival of the fittest) inspires the idea behind the GA concept.

7.2.7 *Comparison of Approaches*

The papers that are viewed show a number of solutions that are based on load prediction problems related to different approaches, specifically for short-term load predictions. Therefore, a global boundary has been made between the limitations of different techniques, and the approach proposed may be an ideal attempt.

Different academics show that the literature uses various techniques to respond to a load forecast. To compare the accuracy of ANN prediction using a time series model, Ref. [3] implemented a feedforward, multilayer neural network model. The model based on ANN gave fair results. The validity of the forecasting models mainly depends on the training and the prediction time.

In [4], in a load forecasting model, the authors evaluated the influence of electricity prices. For areas with sudden adjustments in the energy tariff, this evaluation would typically be ideal because it greatly affects the predictive accuracy. The relation between load and price is highly nonlinear and difficult to model. And more so, a supervised model based on the neurotic network was used to predict the load in the Nigerian national grid in [5]. However, due to environmental conditions, the analysis did not consider the impact of climate, so the accuracy could be enhanced. One must be mindful of the number of the neurons to be included in the hidden layer as excessive neurotics can result to overspecialisation and consequently losing generalising ability. However, Ref. [6] built a model with regard to the weight-space probability distribution (pdf) function. This formation solves a few of the modelling instabilities, so further progress could thus improve forecast model efficiency. It is important to establish inconsistency in model inputs.

In addition, Ref. [7] proposed a multi-context artificial neural network feedforward and feedback (FFFB-MCANN) as a realistic load forecasting approach. In order to achieve improved accuracy, they suggested using the rate values rather than the absolute. Recurrent ANN models are difficult and often need excellent training.

7.3 Conclusion

A comprehensive literature review was conducted on current load prediction methods. In addition to this, a comparative analysis of some specific models was conducted. The study reports on comparisons of intelligent system-based structures with traditional methods of load forecasting were compared. Ideally, such evaluations are meant to show the strengths of classical ANN-based models and pay less interest to the limitations of modern devices. A majority of papers analysed have not specifically highlighted the emphasis on network model optimisation, and therefore, it is immense that similar model network groups are evaluated and network optimisation for load prediction is explored. Research also stated that intelligent system-based models outperform classical methods, so an ANN method is used for this project.

In comparison, a number of artificial neural network models have some disadvantages including long training times, reliance on input values and the design of an optimised system architecture for better load prediction are the three main problems encountered. ANN was used to forecast the load at Transnet Port Terminal (TPT) in East London, South Africa.

7.4 Future Work

In particular, the focus of the project was limited to off-line training. However, certain problems need to be dealt with during the model creation process for real-time applications. The fact that the historical load curve and the weather data have bad data (outliers) can adversely influence the accuracy of the prediction.

References

1. Markus, E.D., Okereke, O.U., Agee, J.T.: Predicting telephone traffic congestion using multi layer feedforward neural networks. Adv. Mater. Res. **367** (2012)
2. Nijhuis, M., et al.: Stochastic household load modelling from a smart grid planning perspective. In: IEEE International Energy Conference(ENERGYCON) (2016). https://doi.org/10.1109/energycon.2016.7513954

3. Sharif, S.S., Taylor, J.H.: Short-term load forecasting by feed forward neural networks. In: Proceedings of IEEE ASME First Internal Energy Conference (IEC), Al Ain, United Arab Emirate (2005). http://www.ee.unb.ca
4. He, F., et al.: A hybrid short-term load forecasting model based on variational mode decomposition and long short-term memory networks considering relevant factors with Bayesian optimization algorithm. Appl. Energy **237**, 103–116 (2019). https://doi.org/10.1016/j.apenergy.2019.01.055
5. Adepoju, G.A., Ogunjuyigbe, S.O.A, Alawode, K.O.: Application of neural network to load forecasting in Nigerian electrical power system. Pacific J. Sci. Technol. **8**(1) (2007)
6. Lauret, P., Fock, E., Randrianarivony, R.N., Manicom-Ramsamy, J.-F.: Bayesian neural network approach to short time load forecasting. Energy Convers. Manage. **49**(5), 1156–1166 (2007)
7. Rashid, T., Kechadi, T.: A practical approach for electricity load forecasting. In: Proceedings of World Academy of Science, Engineering and Technology, vol 5 (2005). http://www.waset.org

Chapter 8
A Survey of Energy-Efficient Electro-hydraulic Control System for Collaborative Humanoid Robots

Teboho Ntsiyin, Elisha Didam Markus, and Lebohang Masheane

Abstract The demand for everyday humanoid robots is growing, as well as the ease of their use. It is therefore imperative that these robots use their energy as efficiently as possible in order to ensure maximum operation. This paper lays out the different methods of actuation used in driving humanoid robots as well as the necessity for human–robot collaboration (HRC). The study presents various methods that have been used by different researchers for robot actuation as well as advantages and disadvantages of some of the commonly used methods. Furthermore, the research discusses results found by different studies based on their preferred method of actuation whether its hydraulic, pneumatic, or electric. Findings show that hydraulic actuation for robots seems to be the more viable solution for the actuation of legged robots as the system has the highest load capacity compared to other systems with one of the studies showing up to 60% reduction in operator load by using a robot that can operate in either active or passive mode depending on the operator's needs.

8.1 Introduction

Human–robot collaboration (HRC) is a topic that is spreading widely in the research community as well as in the industrial world. This trend has pushed researchers to find robots that can handle high payloads as well as do repetitive tasks. There are still some hurdles that are faced in human–robot collaboration. One of the problems is the interaction and communication between human and robot resources, and the other is increasing the energy density of these robots. There are various methods used for robot joint actuation in robots to tackle these problems, and three of the most common ones are electric motor actuation, pneumatic actuation, and hydraulic

T. Ntsiyin (✉) · E. D. Markus · L. Masheane
Central University of Technology, Bloemfontein, Free state, South Africa

E. D. Markus
e-mail: emarkus@cut.ac.za

L. Masheane
e-mail: lmasheane@cut.ac.za

T. Senjyu et al. (eds.), *IOT with Smart Systems*, Smart Innovation, Systems and Technologies 251, https://doi.org/10.1007/978-981-16-3945-6_8

actuation while others that are not very popular include soft actuators and memory alloys though these are used in micro- and macroapplications [1, 2]. For many years, majority of robots that have been developed were driven by electric motors though recent in years research has shown that these robots have a low power density as well as small output torque and that a combination of electronics and hydraulics in robots comes with advantages such as lightweight, small volume, high torque, low cost, high speed, and bearing capacity [1, 3]. This is a research field better known as electro-hydraulic manipulation. Electro-hydraulic systems have been around for many years, though they were widely used only in heavy-duty machinery such as forklifts, excavators, tractor loader backhoes (TLB), and many other industrial machineries [4]. This trend has proved that the need for machines and systems that can assist humans in their daily lives has and will always be there. As humans, we have always tried finding ways in which we can make daily tasks easier and much faster, and this can be traced back starting with the invention of mechanical power transmission that employs various kinematic mechanisms such as belts, chains, pulleys, bar linkages, and gear systems which are driven by a manual motion from human, animal as well as environmental sources such as wind and water [2]. This trend continues as we see more research going into robots that can assist and even replace humans in their day-to-day tasks. These robots include assistive mobility robots, hazardous environment robots, prosthetic robots, medical robots, manufacturing robots, construction robots, and even cooking robots. Today, assistive mobility is one of the fields of great interest that is discussed when it comes to HRC. This is a field that emphasizes the importance of developing devices that can support and aid the elderly in their daily lives as the human race is an ever aging society [5].

8.2 Analysis of Different Actuation Method for Different Robot

In this section, three of the most common types of actuation, namely hydraulic, pneumatic, and electric actuation will be discussed. Table 8.1 shows the advantages and disadvantages of using these three methods.

Though hydraulics are only trending now, they have been around even longer than electrical systems. They can be traced as far back as the 287 BC though they only became more popular during the eighteenth century when the London Hydraulic Power Company was set up to provide water hydraulic systems that would power theatrical scenery, lifts, and other systems. Figure 8.1 shows the trend in hydraulics and the two types of hydraulics that have been used throughout time.

Since then the use of hydraulics has been on the rise and a lot of advancements have been made in the hydraulic technology that has made hydraulics the go-to option when it comes to joint actuation. For this, an electro-hydraulic actuation system (EHAS) is designed for the specific robot. In an EHAS, a single electric motor is used to drive a fixed displacement hydraulic pump which in turn drives

Table 8.1 Advantages and disadvantages of using these three methods

Actuation method	Advantages	Disadvantages
Hydraulic	• Small components with predictable power output to pass massive power • Actuators can be controlled with high precision • Has the ability to move even under high initial loads • Under varying loads, it creates even and smooth motions since the fluids are not compressible and valves can control flow rates accurately • It provides consistent power compared to pneumatic systems at moderate speeds • Heat can be dissipated easily and quickly	• High maintenance cost due to regular fluid change and filters • More prone to failure due to oil leaks • Aeration which occurs from air entering the system can cause foaming and erratic actuator movements • Can be expensive to purchase • Rapturing of high pressure lines can cause injuries • Performance can be affected by temperature changes
Pneumatic	• More efficient and cheap to use as they use air, which is abundant and can easily be stored and transported • Can operate at very high speeds • No risk of leaks that can be messy • Cheaper than hydraulics • Safe to operate • Can function in harsh environments	• Less powerful than hydraulics • Generate a lot of noise during operation • Due to air being compressible pneumatics are less accurate • Before use, air needs to be cleaned by removing water as well as any dust particles
Electrical motor	• Easy to use and integrate than pneumatics or hydraulics • High level of precision and control • Safest method of actuation • Costs less to run as well as to maintain • Easy to connect and assemble with only a few wires • They can be modified for almost any purpose or force requirements	• More likely to overheat during operation • Average failure rate can be high • Increased wear of reduction gears • Some electric actuators can have difficulty holding a locked position and some issues with backlash

the hydraulic actuators controlled by the hydraulic servo valve [6]. Sakurai (2000) discusses the use of a load sensing hydraulic system as its one of the most effective systems of energy saving in hydraulic systems [7] while Zhang [3] proposed a method of optimizing and improving the stability of electro-hydraulic robots by analyzing the driving system of these robots. In the latter, the proposed system was able to satisfy the use of multiple actuators at the same time without interference while having load-sensitive function in pressure, being energy efficient and has the ability to respond quickly. The migration into hydraulics will however require among many things the development of new components that do not exist yet such as new cylinders, drive

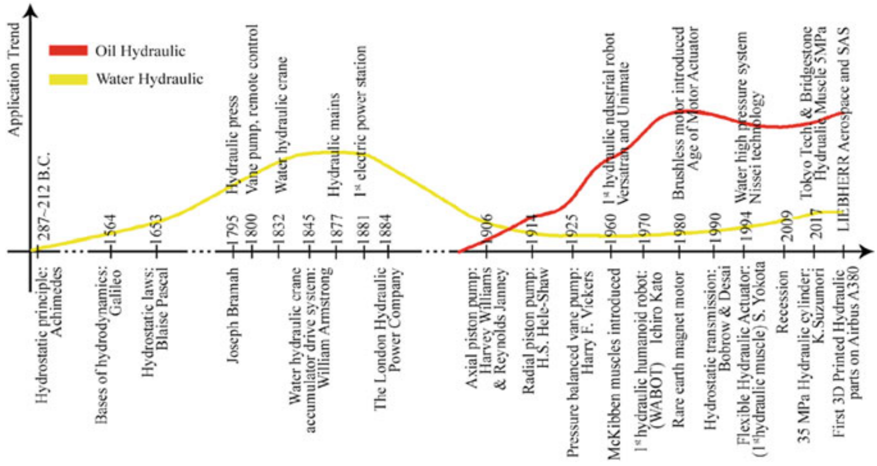


Fig. 8.1 Trend in hydraulic technology [2]

systems, and other different components that will be used specifically for robotic. Such works have been observed in many hydraulic robot designs such as BigDog, Hy-Mo, IHEA, ISAv4, and ISAv4 to name a few. Authors in [8] discuss a novel, highly integrated hydraulic servo actuator with an additive manufactured titanium body. In this paper, two versions of the actuator are discussed; these are integrated servo actuator (ISA) developed by Moog in collaboration with IIT. The first version of the robot was designed to fit into the HyQ quadruped robot, weighs 0.92 kg, and has a maximum force of 4 kN while the second version was for the HyQ2Max, weighs 1.15 kg, and can produce a maximum force 6.2 kN. Its body is made of titanium allow through the use of additive manufacturing (AM). These smart actuators incorporate a lot of components all built in, which are a hydraulic cylinder, servo valve, sensors such as pressure sensor, position sensor, temperature sensor as well as all the electronics, communication module, and overload protection. Figure 8.2 shows the design of the cylinder discussed above [8–10] and discusses a design solution of efficient and compact electro-hydraulic actuators. The actuator as shown in Fig. 8.3 has an overall size of 300 mm × 90 mm × 70 mm with a maximum force of 2000 N and a maximum travel speed of up to 20 mm/s.

Some researchers are now are trying to replace traditional hydraulic actuation systems to electro-mechanical actuation systems (EMAS) due to the ease of use of the system. An EMAS system is one that uses gearbox to lower the actuation torque to allow the use of small electric motors [10] and further discusses the efficiency as well as the limitations of EMAS. One of the downsides of using this system is that the gearbox increases the overall weight and size of the system as well as having some disadvantageous effects on the overall performance of the system since the inertia will be recognized from the motor. The performance of a hydraulic system is influenced by the controlling servo valve in the system. Due to this, it is important

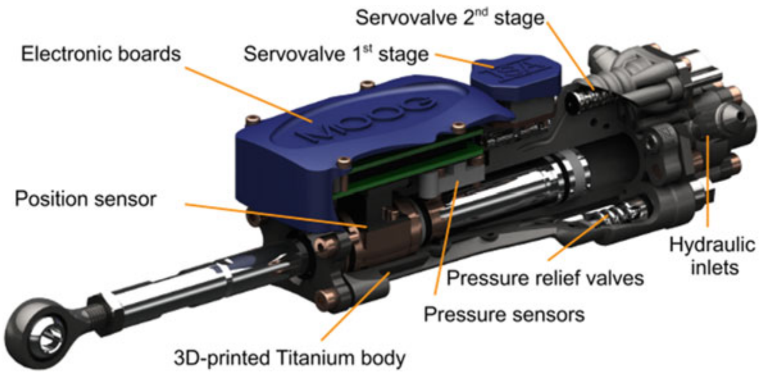


Fig. 8.2 CAD rendering of the ISA v2 with a cut-out section to illustrate the main features of the actuator and integrated components [8]

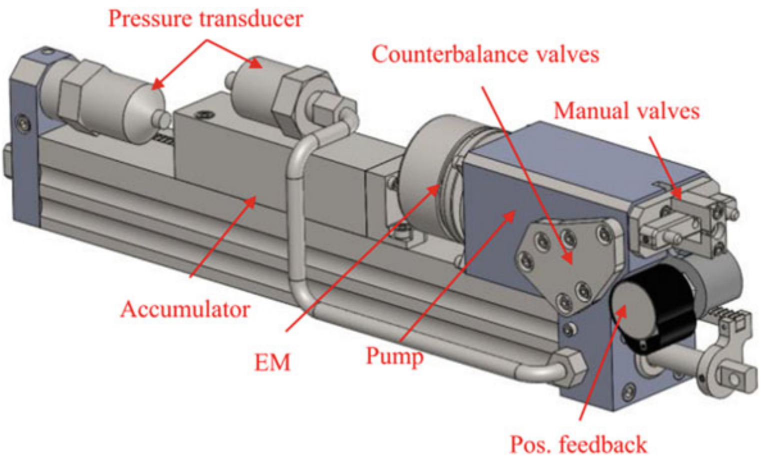


Fig. 8.3 Design of a compact EHA system proposed in [10]

that the servo valve is of high quality and has accurate control with precision feedback. Moreover, the control system needs to be properly designed to ensure that the maximum overshoot of the system is less than 5% with. [11] A nozzle flapper type 2 stage valve with spring mechanism for spool feedback is used. The valve has a flow rate of 63L/m @ 210 bar and has a current consumption of 40 mA @ 10 VDC. Position sensor has a maximum travel of 1 mm with a resolution of 1 μm. The setup can be seen in Fig. 8.4. The experimental results showed that a PID control system can drastically improve the control performance as well as the frequency response of hydraulic systems. In addition to this, oscillations within the system can be solved by including a low pass filter.

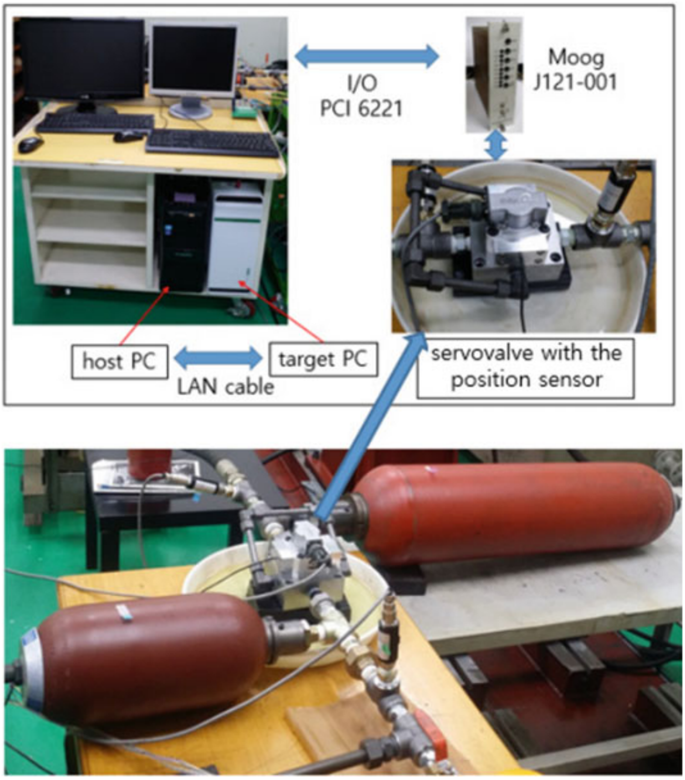


Fig. 8.4 Nozzle flapper type servo valve with electric position sensor [11]

Dongyoung Lee discusses the design of an electro-hydraulic exoskeleton, and the layout of the system as shown in Fig. 8.5 shows the components used in the design of the exoskeleton robot. The system uses an EHA system for actuation, with hydraulic cylinders used for the joints [12]. The authors, however, address issues faced in traditional EHA systems. According to the authors, these systems have a lot of internal leakage during operation which decreases the overall efficiencies of the overall system. The main components responsible for this are the solenoid valve and the gear pump. By designing a new type of pump that can change fluid direction as well as having higher efficiencies over traditional EHA systems, the pump designed is a variable speed piston pump that was further integrated into a combined system with the pump, actuator, and motor as a single system.

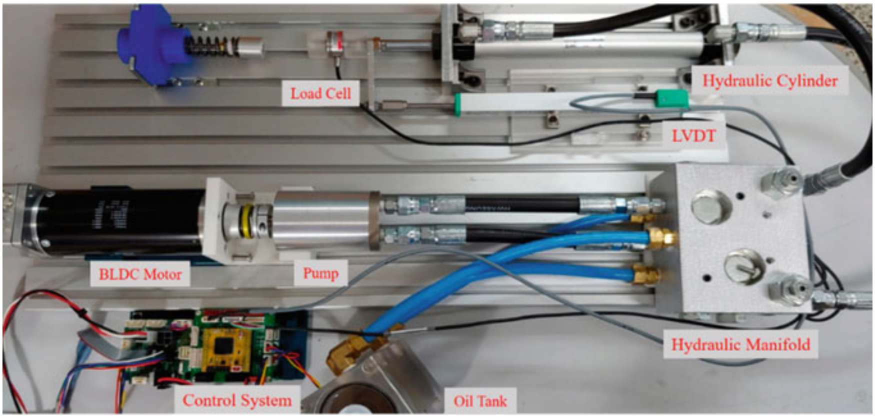


Fig. 8.5 Control of an electro-hydraulic actuator system for an exoskeleton robot [12]

8.2.1 Electric Systems

Electrical motors serve as the main drive system of most robots that exist today. Several techniques are used as actuation when using electric motors, and they can be achieved through the use of servo motors, stepper motors, brushless DC (BLDC) motor, and EMAS [6]. Different manufacturers use one of the above or a combination of them to achieve actuation in their robots. KUKA has nearly 40 years of designing and manufacturing industrial robot arms. The first robot arm to use the six axes was the FAMULUS robot which is now set as their standard. The six axes of the robot are as electric motor driven even up to date with their latest robot designs [13]. Since the development of their first robot, industrial robots have been dominant in large industries and have further become much faster, more efficient, more accurate, and most important of all, more accessible and affordable. The main focus of these robots is more accuracy and repeatability for manufacturing. In 1996, Honda announced the development of its first ever humanoid robot with two arms and two legs and the robot named P2 [14]. Its research had begun 10 years prior, and the desired goal was to develop a robot that can coexist and collaborate with humans and do tasks that humans cannot do. P1 has 30 degrees of freedom, and each joint is controlled with a servo motor. The latest design from Honda is ASIMO. After the development of the P1 robot, Honda developed several other robots P2, P3 before releasing its latest, ASIMO. The actuation systems that these robots use have been investigated, and results have shown that they have a very low power density which more than often results in poor load capacity, low speed, and limited strength. Due to these factors, developing practical legged robots is still an issue. Moreover, these factors play a huge role in the overall performance of any type of robot and thus need to be resolved if a perfect robot is to be built [1]. Another major key in developing a good humanoid robot is to develop a system that can drive the robot with a minimum cost of transport (COT) as possible. COT can be defined as the power consumption divided by the weight, multiplied by velocity [15]. The main advantage of using electric motors over any other system is the ease of use, as the wiring can be simplified a lot depending on what motor is used; moreover, higher accuracies and precision can be achieved.

8.2.2 HRC

As the need for human assisting devices increases, several studies have been conducted to determine suitable interaction strategies to assist humans. These studies showed that the one thing needs to be determined in order to build a meaningful robot that can assist humans to find out how people perform under load while being assisted or working in collaboration with a robot. This study focuses on a design system for a collaborative humanoid robot that can work remotely from the human by mimicking all the human motions.

A study conducted by Tanim discusses the development of ARMAR-6 which is a collaborative humanoid robot for the industrial environment. The main purpose of the robot is to perform maintenance tasks in warehouse environments. The robot works by aiding and supporting technicians that perform maintenance tasks; thereafter, the robot then recognizes that the human worker needs help and will then provide the appropriate assistance. ARMAR-6 is mainly constructed using electric motors of different sizes to achieve 27 degrees of freedom. The robot stands 192 cm high and has a width of 310 cm. It can further stretch 40 cm more in height using its prismatic joint actuator situated in its torso. It has three different actuators that are custom designed and contain a brushless DC motor each. The motors can only achieve a payload capacity of 10 kg per arm with torque capacities of 176 Nm, 123 Nm, and 63 Nm, respectively. The actuators are then powered by a 48 V power unit. The overall weight of the robot is 162 kg without a battery pack with a nominal power consumption of 460 and 1050 W peak (Fig. 8.6).

The robot is artificially intelligent and thus does not require any human input for control. Robots such as this one need to comply with Asimov's three laws of robotics which state that:

1. A robot may not injure a human being or, through inaction, allow a human being to come to harm.
2. A robot must obey the orders given to it by human beings except where such orders would conflict with the first law.
3. A robot must protect its own existence as long as such protection does not conflict with the first or second law.

As the robotic industry is growing more rapidly than ever, more research is being done to find better and more effective ways in which humans can not only work alongside robots but can quickly and easily program industrial robots. Research has shown that the programming phase of robots can be long and time consuming especially in industries that need to constantly change their manufacturing environment.

Fig. 8.6 Overall dimensions of ARMAR-6

Ferraguti discusses a software control architecture to obtain human–robot cooperation during the programming phase of a robotic task. The method discussed is a walk-through programming which is a method that enables humans and robots to coexist in the same workspace and cooperate to achieve a common task. This method of programming can easily be achieved and is done by manually guiding a robot to teach it a specific task [16]. As mentioned previously, the approach taken in this research focuses mainly on creating a software platform that can be used on existing industrial robots to convert them to collaborative ones. Newer robots however are built with this in mind and thus require no additional software.

There are still a few factors that need to be taken into consideration for effective human–robot collaboration, safety being on the top of the list. In addition to this, a properly designed intuitive user interface that will take full advantage of both robot and human abilities being high levels of accuracy, speed, reliability from robots together with flexibility as well as cognitive skills of humans. In [17], three different types of human–robot interactions are discussed and can be seen in Fig. 8.7.

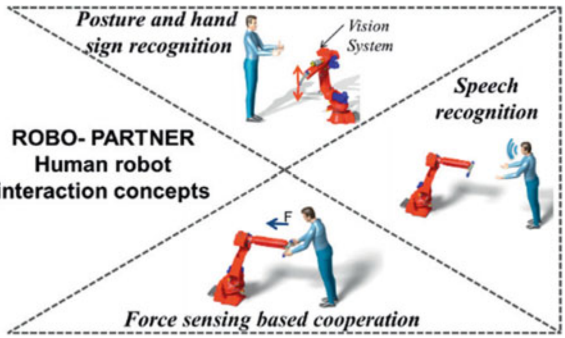


Fig. 8.7 Human–robot interaction concepts [17]

A 2016 paper by the international federation of robotics (IFR) studied at the trend of industrial robots for 4 years and found out that industrial robots increased annually by an average of 13% with a final estimate of 2.6 million industrial robots at the end of 2019. Furthermore, other researches showed that the automotive industry showed the highest demand of industrial robots with the electronic industry showing the second highest demand. This is because of the affordability and ease of use of collaborative robots [18].

Statistics of the Occupational Health and Safety published by the department of health in the US department of labor indicate that more than 30% of European manufacturing workers are affected by lower back pain, and this results in enormous social and economic costs. It is for this reason that some researchers have embarked on a quest to designing assembly line robots for workers that will reduce ergonomic concerns that arise due to on-the-job physical and cognitive loading, while at the same time improving safety, quality, and productivity [17]. Cherubini [19] conducted a study on a Peugeot Citroën assembly line, where Rzeppa homo-kinetic joints are manufactured and the process still utilizes manually labor. In the study, they found that the workers were experiencing muscular pains and further showed that this part of the assembly generated more musculoskeletal disorder (MSDs) than any other part of the assembly. Furthermore, due to the complexity of the manufacturing process, a human worker could not be replaced and thus the researchers found this would be the ideal scenario for collaborative humanoid research and so they proposed a novel, collaborative human–robot design that would outline the following [19]:

- A framework that would successfully manage direct physical contact between robot and human as well as between robot and environment
- A robot that alternates between passive and active behaviors during assembly to lighten the burden on the operator when in passive mode and to comply with the operator when in active mode.
- The approach taken should apply to standard robots, as well as non-torque-controlled robots.

The results of this research indicate that the proposed system reduced the operator load by about 60%. However, the system's cycle time was much slower than that of manual assembly and the solution to this problem would be to increase the number of robots a single operator interacts with. This would result in high installation costs but will reduce MSDs injuries and associated costs [19] (Fig. 8.8).

8.3 Discussion

Extensive research has been done through the literature survey shown above, and based on this, it is evident that research in robot actuation is still on going as the need for robots increases steadily. From pneumatics, hydraulics, EHAS, and EMAS, various authors have shown that robot actuation is of high importance and that replacing humans in with robots in a workplace is not always the best automation

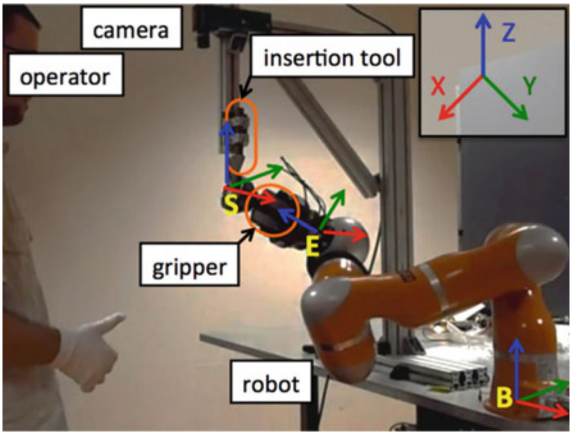


Fig. 8.8 Top: experimental setup for collaborative Rzeppa joint assembly with: insertion tool support (S), end effector (E), and robot base (B) reference frames

solution, but rather a combination of both in some instances is a necessity. Furthermore, one of the key features to designing a good robot that will be human-friendly in the same workspace is designing a robot with the ability to have back drivability of the actuators. This enables users of the system to be able to move the robot joints with ease while not causing any harm to the robot joints. Moreover, these studies have shown that hydraulic actuation for robots seems to be the more viable solution for the actuation of legged robots as the system has the highest load capacity of all the above systems, although it still carries the problem of difficulty in designing and manufacturing the necessary parts as the manufacturing processes can be tedious. Hydraulic systems do promise much stronger robots with a small form factor, and for this, new manufacturing technologies such as additive manufacturing are now being used in the manufacturing process of most robotics as more complex and efficient designs can be achieved at a scale of manufacturing that has never been done before. This technology is becoming better and more accessible than in the past. From heavy-duty machinery, exoskeleton robots, to legged robots, hydraulics have come a long way and will continue to improve actuation in everyday machinery and robotics. On the other hand, implantation of HRC and EMAS introduces robots that are lightweight and are more powerful while being compact. The issue that is currently faced with EMAS is that back drivability is difficult to achieve due to the design principle of these systems, and as mentioned, back drivability enables ease of use for the robot.

8.4 Conclusion and Future Work

This paper set out to analyze and compare the different actuation methods used to drive different humanoid robots, as well as realizing the importance of HRC. Though the debate continues as to which actuation methods are the best for humanoid type robots, more and more researchers are migrating toward EHAS. These systems bring about more capabilities in robots that could not be achieved with electrical systems. To summarize, more powerful robots with a small form factor could be built as well. Moreover, with the availability of technologies such as additive manufacturing, new design method has been achieved to design actuators that could not be designed with traditional design methods like those designed in [8] and [10]. On the other hand, the researchers still recommend in-depth research to still be conducted before designing any form of robot, as requirements may differ from robot to robot. When it comes to robot control, different researchers are still working to finding more robust and energy-efficient ways, as well as a user-friendly control system for robots. The future direction of this study aims at designing more viable and compact hydraulic systems that can be used for assistive mobility as well as finding better and cheaper solutions for designing robots that can lift heavy, while having a small foot print and can be controlled with ease.

References

1. Yang, K., Zhou, L., Rong, X., Li, Y.: Onboard hydraulic system controller design for quadruped robot driven by gasoline engine. Mechatronics **52**, 36–48 (July 2017) (2018). https://doi.org/10.1016/j.mechatronics.2018.03.010
2. Suzumori, K., Faudzi, A.A.: Trends in hydraulic actuators and components in legged and tough robots: a review. Adv. Robot. **32**(9), 458–476 (2018). https://doi.org/10.1080/01691864.2018.1455606
3. Zhang, C.Y., Chen, G.Z., Miao, C.W., Zhao, Y.F., Zhong, S.: Design and stability factors analysis of electro-hydraulic driving system for load-sensing electro-hydraulic robot. IOP Conf. Ser. Mater. Sci. Eng. **392** (6) (2018). https://doi.org/10.1088/1757-899X/392/6/062040.
4. Hanafusa, H., Asada, H., Mikoshi, T.: Design of electrohydraulic servo systems for articulated robot arm control. In: IFAC Proceeding, vol. 13. Pergamon Press, Oxford, U.K. (1981). Tech. Session V, Paper 1, pp. 223–228. ISBN 0-08-027317-3 (1981). https://doi.org/10.1016/S1474-6670(17)65145-0.
5. Martins, M.M., Santos, C.P., Frizera-Neto, A., Ceres, R.: Assistive Mobility Devices Focusing on Smart Walkers: Classification and Review (2012). https://doi.org/10.1016/j.robot.2011.11.015
6. Ntsinyi, T., Markus, E.D., Masheane, L.: Prospective Electro-hydraulic Control System for a Collaborative Humanoid Robot. 2019 Open Innovations (OI). IEEE (2019)
7. Sakurai, Y.: Calculation of Dynamic Overall Efficiency of a Load Sensing Hydraulic System by Bondgraphs. https://doi.org/10.1109/IECON.2000.972508
8. Semini, C. et al.: A Brief Overview of a Novel, Highly-Integrated Hydraulic Servo Actuator with Additive-Manufactured Titanium Body, pp. 4–7
9. Barasuol, V.: Highly-Integrated Hydraulic Smart Actuators and Smart Manifolds for High-Bandwidth Force Control, vol. 5 (June, 2018). https://doi.org/10.3389/frobt.2018.00051
10. Altare, G., Vacca, A.: A design solution for efficient and compact electro-hydraulic actuators. Procedia Eng. **106** (2015). https://doi.org/10.1016/j.proeng.2015.06.003
11. Istanto, I., Lee, I.Y., Huh, J.Y., Lee, H.C.: Controller design for a nozzle-flapper type servo valve with electric position sensor. J. Phys. Conf. Ser. **1303**(1), 29–35 (2019). https://doi.org/10.1088/1742-6596/1303/1/012094
12. Lee, D., Song, B., Park, S.Y., Baek, Y.S.: Applied Sciences Development and Control of an Electro-Hydraulic Actuator System for an Exoskeleton Robot (2019)
13. Shepherd, S., Buchstab, A.: KUKA Robots On-Site, pp. 373–380 (2014). https://doi.org/10.1007/978-3-319-04663-1

14. Hirai, K., Hirose, M., Haikawa, Y., Takenaka, T.: Hirai 1998 Development of a Honda humanoid robot.pdf, pp. 1321–1326 (1998)
15. Xue, Y., Yang, J., Shang, J., Wang, Z.: Energy efficient fluid power in autonomous legged robotics based on bionic multi-stage energy supply. Adv. Robot. **28**(21), 1445–1457 (2014). https://doi.org/10.1080/01691864.2014.946447
16. Nolli, M., Ferraguti, F., Pesamosca, M., Fantuzzi, C., Secchi, C., Landi, C.T.: Walk-through programming for industrial applications. Procedia Manuf. **11**(June), 31–38 (2017). https://doi.org/10.1016/j.promfg.2017.07.126
17. Michalos, G., Makris, S., Spiliotopoulos, J., Misios, I., Tsarouchi, P.: ROBO-PARTNER: seamless human-robot cooperation for intelligent, flexible and safe operations in the assembly factories of the future. Procedia CIRP **23**, 71–76 (2014). https://doi.org/10.1016/j.procir.2014.10.079
18. Villani, V., Pini, F., Leali, F., Secchi, C., Fantuzzi, C.: Survey on human-robot interaction for robot programming in industrial applications. IFAC-PapersOnLine **51**(11), 66–71 (2018). https://doi.org/10.1016/j.ifacol.2018.08.236
19. Cherubini, A., Passama, R., Crosnier, A., Lasnier, A., Fraisse, P.: Collaborative manufacturing with physical human-robot interaction. Robot. Comput. Integr. Manuf. **40**, 1–13 (2016). https://doi.org/10.1016/j.rcim.2015.12.007

Chapter 9
Gesture Recognition-Based Interaction with Smartwatch and Electric Wheelchair for Assistive Mobility and Navigation

Elisha Didam Markus, Teboho Ntsinyi, and Eric Monacelli

Abstract Smartwatches are becoming increasingly accessible to end users, thereby placing the power of mobile technology at their fingertips. On the other hand, electric powered wheelchairs are designed to assist people with mobility impairments to move with ease. However, people with cognitive difficulties may not be able to navigate the wheelchair as required and may require additional support from wearable devices. This paper presents a gesture recognition-based navigation for the wheelchair using a smartwatch. Users are able to navigate the wheelchair using gestures that would be easily interpreted by the smartwatch. This capability opens up new mobility possibilities never before envisaged.

9.1 Introduction

Wearable devices like smartwatches are becoming increasingly popular and more accessible with improvement in technology and a reduction in their cost. These watches are equipped with various sensors such as gyroscopes, heart rate monitors, blood sugar monitors, ECG monitors, etc. As they are interactive in nature, they enable their users to make calls, send messages, take photos, and just to mention a few. Recently, researchers are taking advantage of the ubiquitous nature of these devices to improve on existing technologies such as smart wheelchairs.

E. D. Markus · T. Ntsinyi (✉)
Department of Electrical, Electronic and Computer Engineering, Central University of Technology, Bloemfontein, Free state, South Africa

E. D. Markus
e-mail: emarkus@cut.ac.za

E. Monacelli
Laboratoire d'Ingénierie des Systèmes de Versailles (LISV), Université de Versailles Saint-Quentin France, Versailles, France

T. Senjyu et al. (eds.), *IOT with Smart Systems*, Smart Innovation, Systems and Technologies 251, https://doi.org/10.1007/978-981-16-3945-6_9

It has been reported that the wearable technology market is growing fast. This is reflected by a similar rapid growth rate of smartphones and their respective connectivities. According to the report Universally, the wearable electronics market is expected to reach US$19 billion in 2020 [1].

This paper presents a study on the use of a smartwatch for assistive mobility. The next section discusses related work in the area of the use of wearables for smart devices and also in monitoring systems. Section 9.3 discusses the system architecture of the proposed design. Section 9.4 presents the results of the study and is followed by the conclusion of the paper.

9.2 Related Work

In recent times, smartphones have evolved with many capabilities. This is so as they are now equipped with various sensors and abundant connectivity. These capabilities have endeared many researchers to unravel their potential. In some cases, smartphones have been used to gather environmental and location data for various purposes. Smartphones have also been widely used in sports and fitness activities to record activities in real time. Other studies have reported their use in health care such as in gathering data for diagnostic and monitoring purposes. Other applications are in assistive mobility such as to help wheelchair users gather data for health applications.

Early attempts to use smart devices to improve wheelchair use includes [2] where the authors designed a smartphone application to evaluate and advance accessibility for wheelchair users. Their proposed system employed a dedicated smartphone application that records many sensor measurements and also allowed for obstacle avoidance.

Then, the information gathered from accessibility is conveyed on maps which could compute the most reachable route for the users while of the wheelchair taking into account their respective personal information and abilities.

Figure 9.1 gives a summary of various wearable devices and their applications.

In this paper, gesture implementation and control are classified into smartphone and smartwatch use cases.

9.2.1 Gesture Control with Smartphone

Smartphones have been popular in gesture-related control of wheelchairs. For instance, in [4], a new application of leap motion sensor in wheelchair navigation was explored. The wheelchair navigation system was designed to assist amputee users, stroke, and rehabilitation patients. The system was able to eliminate the reliance on finger movements for patients who are paraplegic. More so, the system did not depend on the use of gaze and, hence, did not also require voice or gesture control. This created more independence for the users with specific medical conditions or

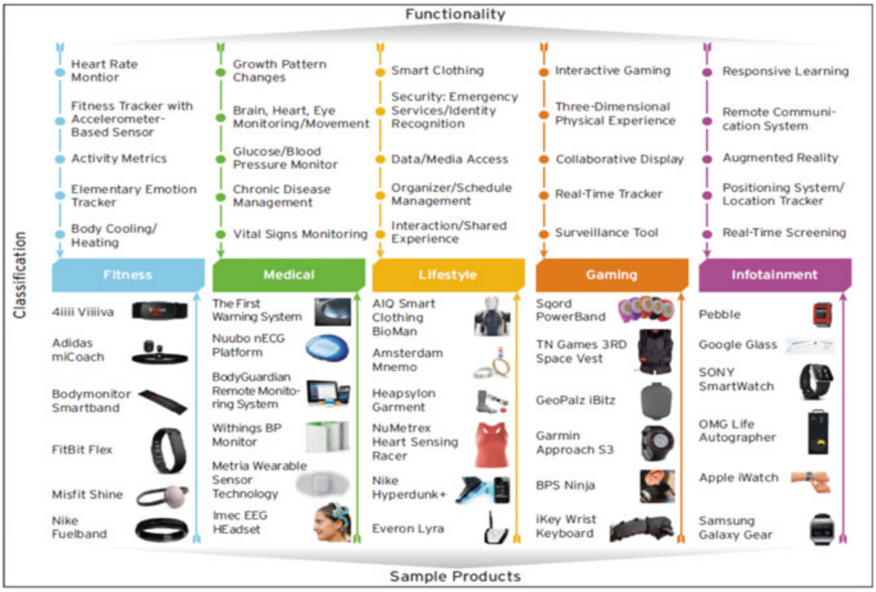


Fig. 9.1 Overview of wearable devices and applications [3]

for those undergoing rehabilitation or treatment. Their results demonstrated a low latency in wheelchair command to action, which allowed for a better control for its users, thereby reducing the incidence of control-related accidents. More so, the wheelchair navigation system did not require Internet connection, which allowed for more autonomy in range for wheelchair users as compared to other available cloud-based models.

In [5], a smart wheelchair (SMW) ecosystem for autonomous navigation in urban environments was designed. The design comprised of two components: a mapping service used for navigation and the wheelchairs itself used as a client. The SWS integrated a 3D LIDAR for outdoor mapping of its environments and autonomous motion without the use of GPS. The drawback here is that the system has high computational power requirements.

In a similar vein, the authors in [6] presented an android controlled SMW for disabilities. This system uses a control architecture for the motorized wheelchair as well as an embedded system for monitoring critical patients. The embedded system also uses biometric features of the user to monitor critical and hostile scenarios that may occur. The wheelchair would produce an alert by raising the alarm with the measurement of the heartbeat at a particular interval.

Furthermore, the authors in [7] addressed the problem of automatic detection of municipal topographies based on movements of wheelchair users. This is important in cities as providing navigation directives to people who are

incapacitated involves knowing the topologies and landscapes like curb ramps, steps or other obstacles that could act as impediments to users. These urban features may not necessarily be obtainable from maps and could often change with time. Initial studies on wheelchair users showed that an automatic crowdsourcing mechanism was needed, thereby avoiding users' participation. In their contribution, they presented a solution to crowdsource urban features using inertial sensors that have been connected to the wheelchair. The technique recognized activities based on information coming from a set of sensors. Experiments were carried out based on data assembled from ten wheelchair users moving in an outdoor setting.

Furthermore, [8, 9] designed smartwheels for detecting urban features for use by wheelchair users. The smartwheels were created by inserting three standalone inertial sensors to the wheels of the wheelchair. These sensors provided information on the required urban features. Since these features are not obtainable from maps and can also vary with time, the authors used crowdsourcing to obtain this information from the users. This information gathering is required to be automatic in nature. Hence, the smartwheels were used to detect these urban features by examining data emanating from wheelchair movements of inertial sensors. These activity recognition methods were also used to extract data from the sensors.

In [10], a quick response (QR) code localization system was used to provide localization and navigation for smart wheelchairs. The application was mainly for indoor use as it was designed to enhance the independence and improve the quality of life for its users. It provides some form of autonomy for the user where they were able to navigate to a place of their choice within the room. Meanwhile, the GPS signal is not available indoors; hence, alternative sensors were essential. The system was designed based on QR codes. These codes provide the absolute position of the landmark. Also available is a vision system that recognizes the codes, estimates the relative position of the robot with respect to the codes, and then calculates the robot position based on data provided in the codes.

In [11], an android phone was used to connect to a wheelchair using a speech recognition algorithm. The set up enabled a motorized, speech recognized protocol to be implemented on the wheelchair. This is required by physically disabled wheelchair users who are unable to control the joystick of the wheelchair due certain ailments affecting their limbs such as paralysis, amputations, etc. Such users are not able to exert external manual force. The robotic system uses raspberry pi for the speech recognition and Bluetooth interface between the android phone and Raspberry Pi.

9.2.2 Gesture Control with Smartwatch

Other studies have described various control technologies to make wheelchairs more flexible and easier to use than traditional joystick-controlled ones. Their objectives are to make the wheelchair smarter for disabled users especially in developing countries like South Africa. Apart from control using android operated mobile devices, smartwatches have been tested. Recent efforts on this technique have looked at taking advantage of the rolling out of these devices for various use cases. In [12], a smartwatch application was designed to help students who are disabled in an IoT-enabled

Fig. 9.2 Electric wheelchair used for the project

campus. It is common in these environments that students with disabilities often have limited access to various buildings and facilities such as classrooms, libraries, and this impinges of their learning experience and full participation in university life. Based on the foregoing, there has been an increasing interest in technologies that would help to integrate the Internet of Things (IoT) into buildings. The authors discussed how IoT and wearable technology could benefit these class of students and improve the quality of their university experiences. They used the smartwatch to gain access and controlling in-building devices as well as monitoring environmental conditions.

In another development [13], A smartwatch was used to enhance the navigation for the visually impaired. The system provided an improved system for the chosen routes, which were based on distance, luminosity, and noise along the respective paths. While keeping the user conversant with the path, the study gathered a lot of environmental data through a smartwatch. These interactions and notifications were based on the smartphone vibration patterns. These assisted the user to navigate his path without the need of looking at the device. However, this was just conceptual and was not implemented on any mobility device.

Similarly, [14] presented an autonomous wheelchair navigating under a predefined environment. The study utilized an inner feedback linearizing loop to formulate the control problem allowing for real time computationally efficient implementation (Fig. 9.2).

9.3 System Description and Implementation

Figure 9.3 shows the architecture of the smartwatch gesture-based control. The system comprises of a smartwatch, a Bluetooth module, Arduino board, and an electric wheelchair. The smartwatch is worn on the wrist and is equipped with a gyroscope that is used to measure the *X*, *Y*, and *Z* orientation. The values of each

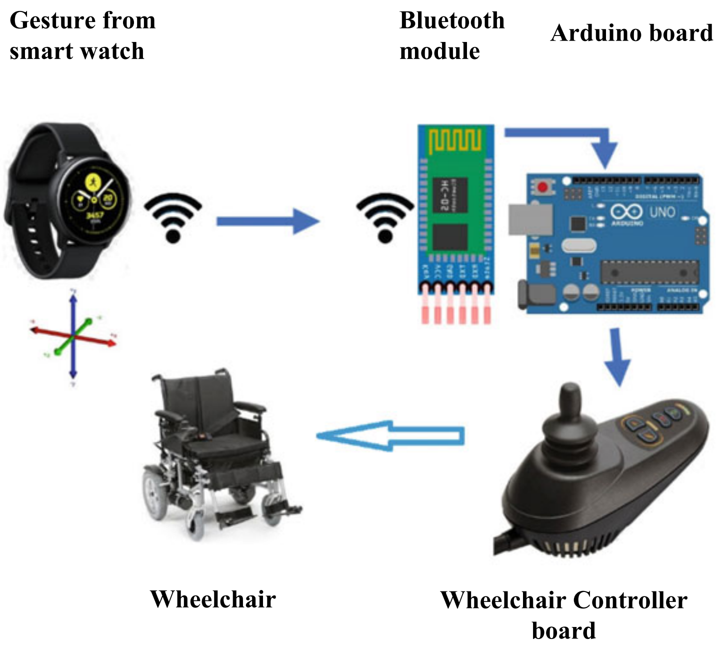


Fig. 9.3 Architecture of smartwatch gesture-based control

axis range from -1.00 when tilted in one direction to + 1.00 when tiled in the opposite direction. However, for driving the wheelchair, only two of the axes are needed, being the *X* and *Y* axes. The measured values are sent via Bluetooth to the Bluetooth transceiver connected on the wheelchair for it to know which direction to drive. The Bluetooth module on the wheelchair has been configured to be a slave to connect strictly to the single remote Bluetooth device that is controlling it. The controller on the wheelchair then converts the values to a voltage level ranging from 0 to 5 V using pulse width modulation (PWM). The wheel joystick of the wheelchair has a voltage of 2.5 V on both the *X* axis and *Y* axis to keep it in its halted position. If the voltage is set to be below the halt voltage on the *X* axis, the wheelchair will drive backwards, and if it is set above the halt voltage, it will drive forward. On the *Y* axis, a higher voltage will turn it to the right and a lower voltage will turn it to the left. The higher the voltage difference is to the halt voltage, the faster the wheelchair.

9.4 Results and Discussion

Preliminary results show that the gesture communication between the smartwatch and the wheelchair via Bluetooth is effective. Different modalities of human robot interaction were studied. In the first instance, the user of the wheelchair would point

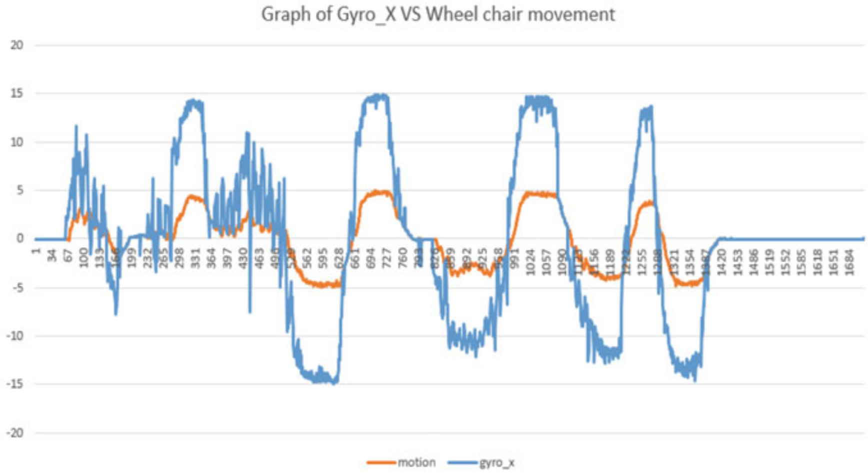


Fig. 9.4 Acceleration of the wheelchair based on commands received from gyroscope with *Y* axis kept constant

to a target like a place or an object instead of driving from joystick. The wheelchair will move from this movement to the target. Another application is at hospital beds or at home where the user calls the wheelchair by hand to the bed side. These movements could be personalized to various needs as may be determined by the user.

A gyroscope was mounted on the wheelchair's right arm rest to record the movements and acceleration of the wheelchair. The transmitter/controller transmits gesture movements every 10 ms which is then received by the wheelchair at the same rate. The gyroscope that is mounted on the wheelchair then records the wheelchair's movement every time the information is received. The graph in Fig. 9.4 shows the acceleration of the wheelchair as it receives the gyroscope commands and executes them. The Y axis in this instant is kept as a constant on the wheelchair, and the controlling gyroscope is moved to control the wheelchair in all directions.

Figure 9.5 shows the wheelchair movements with the *X* axis kept constant. It can be noticed that some of the movements are lost during transmission as the wheelchair tends to be oriented in the *Y*-direction.

The graphs show that the wheelchair responds almost instantly as the commands are received. The only parameter that is not accounted for is latency in transmission between the two Bluetooth devices which may range with the transmission distance.

9.5 Conclusion

This paper presented a study on the use of gesture from a smartwatch to steer a wheelchair for assistive mobility. Results show that the wheelchair is able to move

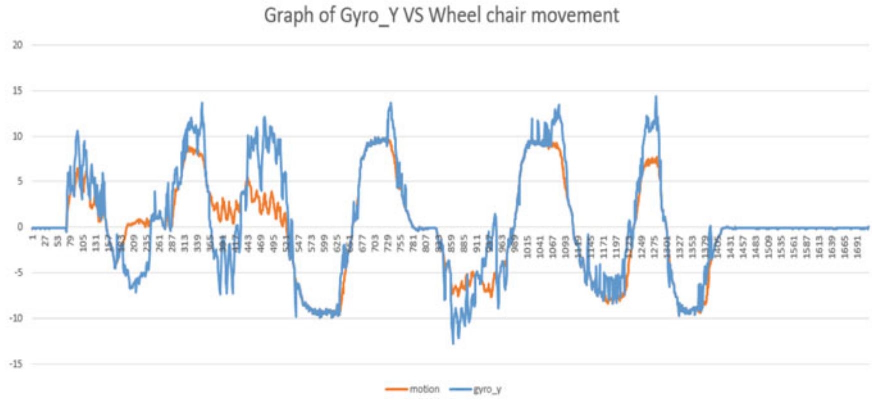


Fig. 9.5 Acceleration of the wheelchair based on commands received from gyroscope with *X* axis kept constant

from the gesture commands provided by the smartwatch on the wrist of the user. This application is useful for users who are not able to use their fingers but can move them in a gesticular way. These users could include paraplegics or patients with autism. Another application is at hospital beds or at home where the user calls the wheelchair by hand to the bed side. These movements could be personalized to various needs as may be determined by the user.

The future work involves including navigation capabilities for the wheelchair to make it autonomous. Direction of this study aims at designing more viable and compact hydraulic systems that can be used for assistive mobility as well as finding better and cheaper solutions for designing robots that can lift heavy, while having a small foot print and can be controlled with ease.

References

1. Wearables, statistics. https://www.statista.com/outlook/319/100/wearables/worldwide (Accessed Nov 2020)
2. Mourcou, Q., Fleury, A., Dupuy, P., Diot, B., Franco, C., & Vuillerme, N.: Wegoto: A Smartphone-based approach to assess and improve accessibility for wheelchair users. In: 2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp. 1194–1197. IEEE (2013, July)
3. Bhat, A., Badri, P., Reddi, U.S.: Wearable Devices: The Next Big Thing in CRM. Cognizant 20–20 Insights (2014)
4. Fereidouni, S., Sheikh Hassani, M., Talebi, A., Rezaie, A.H.: A novel design and implementation of wheelchair navigation system using Leap Motion sensor. Disabil. Rehabil. Assistive Technol. 1–7 (2020)
5. Schwesinger, D., Shariati, A., Montella, C., Spletzer, J.: A smart wheelchair ecosystem for autonomous navigation in urban environments. Auton. Robot. **41**(3), 519–538 (2017)
6. Debnath, T., Abadin, A.Z., Hossain, M.A.: Android controlled smart wheelchair for disabilities. Glob. J. Comput. Sci. Technol. (2018)

7. Civitarese, G., Mascetti, S., Butifar, A., Bettini, C.: Automatic detection of urban features from wheelchair users' movements. In: 2019 IEEE International Conference on Pervasive Computing and Communications (PerCom), pp. 1–10. IEEE (2019, March)
8. Mascetti, S., Civitarese, G., El Malak, O., Bettini, C.: SmartWheels: Detecting urban features for wheelchair users' navigation. Pervasive Mobile Comput. 62, 101115 (2020)
9. Zang, K., Shen, J., Huang, H., Wan, M., Shi, J.: Assessing and mapping of road surface roughness based on GPS and accelerometer sensors on bicycle-mounted smartphones. Sensors **18**(3), 914 (2018)
10. Cavanini, L., Cimini, G., Ferracuti, F., Freddi, A., Ippoliti, G., Monteriù, A., Verdini, F.: A QR-code localization system for mobile robots: application to smart wheelchairs. In: 2017 European Conference on Mobile Robots (ECMR), pp. 1–6. IEEE (2017, September)
11. UPase, S.U.: Speech recognition based robotic system of wheelchair for disable people. In: 2016 International Conference on Communication and Electronics Systems (ICCES), pp. 1–5. IEEE (2016, October)
12. Kim, J.E., Bessho, M., Sakamura, K.: Towards a smartwatch application to assist students with disabilities in an IoT-enabled campus. In: 2019 IEEE 1st Global Conference on Life Sciences and Technologies (LifeTech), pp. 243–246. IEEE (2019, March)
13. Gaggi, O., Palazzi, C.E., Ciman, M., Bujari, A.: Stepbywatch: A smartwatch-based enhanced navigation system for visually impaired users. In: 2018 15th IEEE Annual Consumer Communications and Networking Conference (CCNC), pp. 1–5. IEEE (2018, January)
14. Bardaro, G., Bascetta, L., Ceravolo, E., Farina, M., Gabellone, M., Matteucci, M.: MPC-based control architecture of an autonomous wheelchair for indoor environments. Control. Eng. Pract. **78**, 160–174 (2018)

Chapter 10
OSTBC-MIMO Based Radio Propagation System Over Irregular Terrain in the Northern Cape Province of South Africa

Elisha Didam Markus and Xolani B. Maxama

Abstract The Northern Cape Province in South Africa, along the Orange River valley, has radio signal reception challenges due to high mountain ranges. The South African Electricity Authority- Eskom has assets to monitor in this region. However, due to radio signal reception challenges in this area, it is currently impossible to monitor their high-voltage assets. This paper addresses this setback by making use of the orthogonal space-time block codes, multiple in, multiple out (OSTBC MIMO) technology. The study further presents simulated results of this technology using two different OSTBC MIMO transceiver systems over this mountainous area of the Northern Cape Province of South Africa. The simulation results are generated using MATLAB, Pathloss 4 and Atoll wireless network software. The results generated show that employing a low-frequency OSTBC MIMO transceiver system, in rough terrain environments, can greatly improve radio signal reception and link reliability.

10.1 Introduction

Owing to the high mountain ranges along the Orange River valley in the Northern Cape Province of South Africa, all forms of radio communication, including the (SCADA), are not available [1]. As a result, it impossible for the power utility, Eskom to control and monitor Substations remotely via the wireless communication network.

This has rendered Eskom incapable of monitoring its substations in this region. It is therefore a huge challenge and needs to be addressed timeously. This study seeks to find a way to propagate these radio signals successfully in irregular terrains. If not addressed, the high voltage equipment, the environment and personnel could be at risk. The terrain, from Vioolsdrift Radio Site (the transmit site) to Henkriesmond Substation (the receive site), is an uneven mountainous surface as seen in Fig. 10.1.

E. D. Markus (✉) · X. B. Maxama
Central University of Technology, Blemfontein, Free State, South Africa
e-mail: emarkus@cut.ac.za

T. Senjyu et al. (eds.), *IOT with Smart Systems*, Smart Innovation, Systems and Technologies 251, https://doi.org/10.1007/978-981-16-3945-6_10

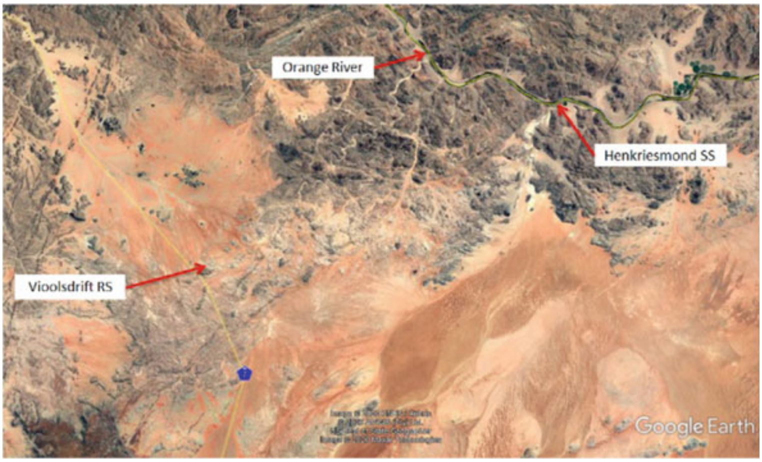


Fig. 10.1 Google earth image of the mountainous terrain in the Northern Cape, South Africa

One of the softwares chosen for this study is the Atoll wireless network because of its rural environment propagation features to help predict radio propagation over rough terrains for our area of concern. The model used for the study is the Rayleigh fading model [2].

The terrain of the area of interest (between Vioolsdrift Radio Site and Henkriesmond Substation) has been extracted from Google Earth as shown in Fig. 10.1. The image from Google Earth is imported to Atoll wireless network software to obtain a coverage plot diagram of the area, as shown in Fig. 10.2.

As seen in Fig. 10.2, there is no radio signal coverage in the area of Henkriesmond Substation, from Vioosdrift Radio Site. This is so because of the high mountain peaks, reflecting high frequency radio signals in that area.

10.2 Related Work

This study is a part of an ongoing project on radio transmission in challenging environments [3]. A lot of study has been done on transmission of radio signals over irregular terrain. Some of the work involved the use of multiple antennas on both transmit and receive ends to improve the signal to noise ration and link performances [4, 5]. Figure 10.3 shows a typical MIMO block diagram.

Therefore, MIMO techniques are known to increase the performance of communication systems in irregular terrain [6]. In many MIMO systems, the channel capacity can be increased by using spatial multiplexing. However, to increase the reliability of the channel, space–time coding (STC) is used [7]. Many diversity and high data rate schemes for transmitting data via MIMO systems have been proposed [8–10].

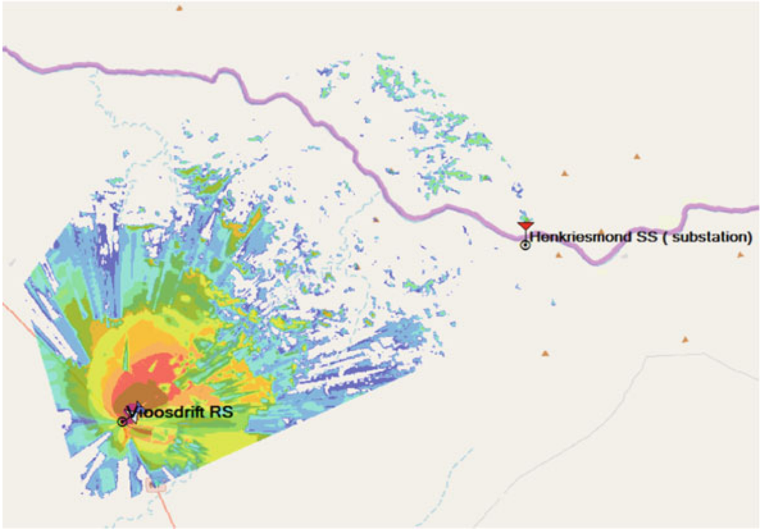


Fig. 10.2 Atoll wireless network coverage prediction diagram

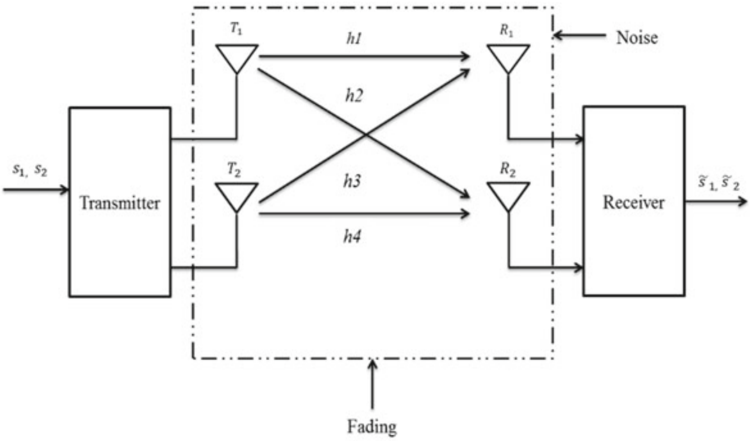


Fig. 10.3 MIMO block diagram

However, OSTBC is receiving much attention due to the simplicity of its signal decoding and good error probability characteristics.

Alamouti and his colleagues in [11] for instance designed a STBC code for a MIMO system with two antennas. Others took Alamouti's work further by designing STBC codes for more than two transmitting antennas [8]. On the hand, Tarokh grouped the STBC codes into two categories, namely "STBCs for real signals" and "STBCs for complex signals". A complex STBC could also be referred to as an

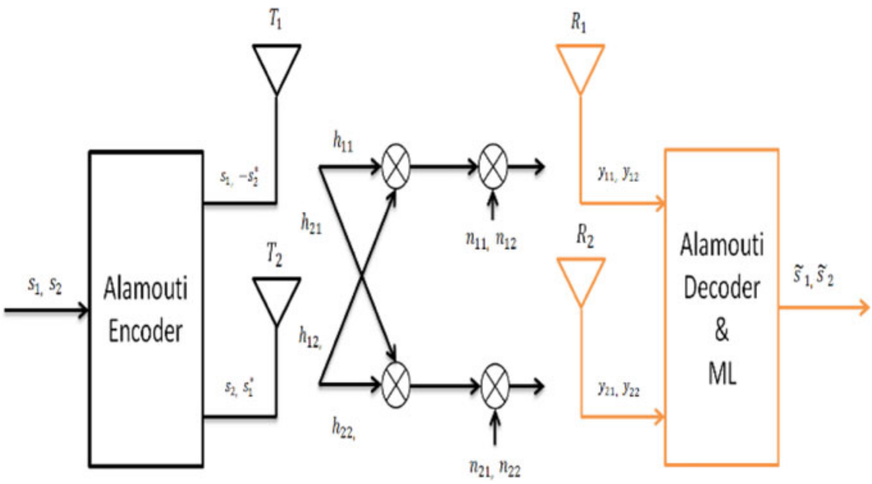


Fig. 10.4 MIMO 2 X 2 system with Alamouti code

orthogonal STBC (OSTBC). Figure 10.4 shows a 2 × 2 MIMO Alamouti code system diagram.

OSTBC can enhance performance by measuring the average number of channels used by each data transmitted. The maximum diversity a space–time system has is (Nt × Nr), which is the total number of channels between Nt transmitters and Nr receivers.

The use of multiple transmit antennas on OSTBC is to accomplish higher performance, less decoding complexity and less coding delay. The following equation is an example of Tarokh's complex transmission matrix which can achieve a full diversity for four transmit antennas [10]:

$$S_4^c\left(\frac{3}{4}\right) = \begin{bmatrix} S_1 & S_2 & \frac{S_3}{\sqrt{2}} & \frac{S_3}{\sqrt{2}} \\ -S_2^* & S_1^* & \frac{S_3}{\sqrt{2}} & \frac{-S_3}{\sqrt{2}} \\ \frac{S_3^*}{\sqrt{2}} & \frac{S_3^*}{\sqrt{2}} & \frac{-S_2-S_2^*+S_1-S_1^*}{2} & -\frac{S_1-S_1^*+S_2-S_2^*}{2} \\ \frac{S_3^*}{\sqrt{2}} & \frac{S_3^*}{\sqrt{2}} & \frac{-S_1-S_1^*+S_2-S_2^*}{2} & \frac{-S_1-S_2^*+S_1-S_1^*}{2} \end{bmatrix} \tag{10.1}$$

where

$(*)$ = the complex conjugate of the element.

$S_4^c\left(\frac{3}{4}\right)$ = the matrix code for complex transmission matrix for four transmit antennas.

$S_1, S_2, -S_2^*, S_1^*$ = Transmitted data symbols.

The signal power-to-noise power ratio of a system is essential to be able to predict its quality and signal reception capability and is measured in decibels [12]. In general, noise can be anything from atmospheric noise to circuit noise. BER is inversely proportional to SNR, i.e. low SNR causes high BER [13].

On the other hand, fading is prevalent in mountainous regions. This poses a challenge in radio wave propagation of these regions. With fading comes the weakening of signal strength at the receiver [14]. The most used fading model for irregular terrain simulations and used in this paper is the Rayleigh fading model.

The remainder of this article is organized as follows. In Sect. 10.3, we present the method and the proposed OSTBC MIMO transceiver model over a Rayleigh fading channel used to overcome the propagation challenges in irregular terrain of the Northern Cape. Section 10.4 presents and discusses the simulation results of the proposed model. Section 10.5 presents the conclusion and future work. The last section presents the list of references.

10.3 Methods

To achieve our objectives, the following methodology was applied:

- Literature Review: Existing literature related to factors affecting radio signal propagation and the unique characteristics of the low frequency band; the MIMO OSTBC technology and its application in irregular terrain environment was reviewed.
- System Model Development: A mathematical model for an OSTBC MIMO transceiver applied to a Rayleigh fading channel with added white Gaussian noise (AWGN) employing a 4×4 MIMO system was developed.
- Simulations: After developing the mathematical model, a low band VHF OSTBC MIMO model was developed and applied in MATLAB/Simulink software. The behaviour of the model was analysed, and the results of the radio link performance discussed.

The proposed OSTBC MIMO transceiver model over a Rayleigh fading channel is shown in Fig. 10.3. The model is simulated using MATLAB/Simulink software. The following are the different components blocks making up the system:

10.3.1 Random Binary Generator

The random binary generator block generates a binary sequence that is unpredictable. It uses Bernoulli binary generator to generate the random bits. It is used to simulate the information to be transmitted over the medium.

10.3.2 Quadrature Phase Shift Key (QPSK) Modulator

The binary data from the random bit generator is modulated in this block to a QPSK constellation. In this block, a splitter will divide the binary data stream into two data streams, $d_q(t)$ and $d_i(t)$. The symbol duration for both $d_q(t)$ and $d_i(t)$ streams is twice the bit duration ($T_{sym} = 2T_b$), thus increasing the data rate [10]. The QPSK modulator has been selected over other low order modulation schemes such as binary phase shift keying (BPSK) for this reason, that is, its ability to produce increased data rates, and it gives more efficient results than BPSK or QAM [15].

10.3.3 Orthogonal Space–Time Block Codes (OSTBC) Encoder

The output of the QPSK modulator which contains data symbols is presented to the input of the OSTBC encoder stage which encodes the data symbols [11] by using a complex orthogonal generalized code for three or four transmit antennas [8]. In this simulation, a 4×4 MIMO system with four antennas is implemented [16].

10.3.4 Multiple In, Multiple Out (MIMO) Rayleigh Fading Channel

The MIMO Rayleigh fading channel block models a baseband multipath Rayleigh fading propagation channel typical that of a non-line of sight, irregular, mountainous terrain. The block is configured as a 4×4 MIMO channel with transmit and receive antenna selection.

10.3.5 Added White Gaussian Noise (AWGN)

The AWGN channel block models symbol errors resulting from a noisy propagation channel.

10.3.5.1 Orthogonal Space–Time Block Codes (OSTBC) Combiner

The orthogonal space–time block codes (OSTBC) Combiner stage will combine the received signals from the receive antennas and the channel state information (CSI) in order to recover the information of the symbols encoded by the OSTBC encoder.

10.3.5.2 Quadrature Phase Shift Key (QPSK) Demodulator

The QPSK demodulator block demodulates the signal recovered from OSTBC combiner.

10.3.5.3 Frame Error Rate

The Frame Error Rate Calculator works out the system frame error rate by comparing the demodulated bits with the original bits sent per frame in order to detect errors. The block also determines the duration of the simulation. The simulation stops when the pre-determined number of error frames is reached.

10.4 Results and Discussion

The aim of this section is to investigate the validity of the proposed model, which is typical of the irregular terrain in the Northern Cape. This is done by analysing the bit error rate (BER) results of a single-in single-out (SISO) system, typical of what exists in the current Eskom network. Also, we analysed the OSTBC MIMO system with four antennas employed at both the transmitter and the receiver.

Also, in the simulations, the receive signal strength results of the OSTBC MIMO system, when used with two different VHF frequencies, were investigated.

Therefore, the following two scenarios in a Rayleigh fading channel were used to investigate

- Whether the BER could be reduced significantly by employing the OSTBC MIMO system.
- Whether the receive signal strength (RSS) could be improved significantly by employing the OSTBC MIMO system in the low VHF frequency band.

The first case scenario showed the BER simulation results of a single-in single-out (SISO) system employing only one antenna at both the transmitter and receiver. The second case scenario showed the BER simulation results of a 4×4 OSTBC MIMO system employing four antennas at both the transmitter and receiver. All the simulation scenarios were performed in a Rayleigh fading channel in the MATLAB/Simulink software.

10.4.1 The Current Eskom Radio Link Without MIMO System

First case scenario 1×1 SISO system in a Rayleigh channel.

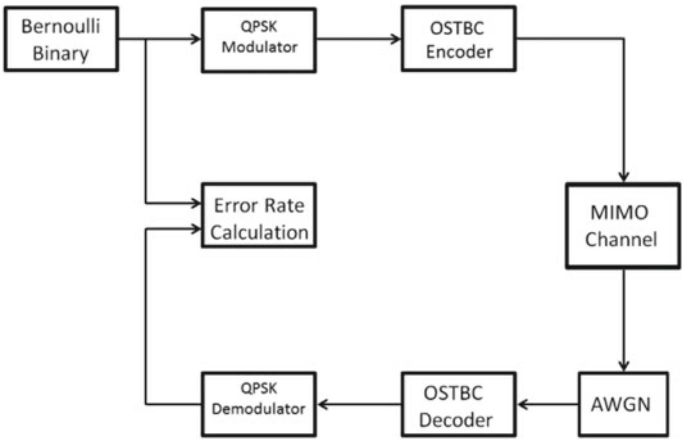


Fig. 10.5 Proposed model for OSTBC MIMO

The following simulation results show the performance of the existing Eskom radio link in the Northern Cape Province.

In the area under investigation, the Vioolsdrift Radio Site lies at 870 m above sea level, while the Henkriesmond Substation lies at 208 m above sea level. The two sites are 32.8 km apart, with uneven terrain in-between. Figure 10.4 is a line of sight diagram showing a no line of sight condition between Vioolsdrift Radio Site (TX) and Henkriesmond Substation (RX).

The high mountain peaks between the two sites make radio communication impossible. As seen in Fig. 10.5, the multipath radio signals are reflected off the mountainous terrain and lost in the medium, resulting in propagation loss terrain with diffraction loss [17]. It is expressed by the following equation

$$\mathrm{PL}_{\mathrm{Lee}} = \mathrm{PL}_0 + \gamma \log_{10}(d) + F_0 + L_{(v)} \tag{10.2}$$

where

PL_0 and γ are derived from empirical data, for example, $PL_0 = 89$ and $\gamma = 43:5$ for rural environmental. F_0 is the adjustment factor.

$L_{(\mathrm{v})}$ is the knife-edge diffraction loss.

Table 10.1 shows the performance results of the existing Eskom radio link. As seen in Table 10.1, the effects of the mountain can be observed at the 9 km mark towards the receiver (Fig. 10.5), where the path losses increased significantly due to reflection of the radio signals by the mountain, with Net path loss at 157.07 dBs.

According to the Eskom standard, the receiver level threshold for the UHF repeaters is between −60 and −90 dB [18]. The simulation results for the Eskom link show that the receive level is −120.08 dB. This is completely off the working levels, meaning this radio link will fail. Figure 10.6 shows the simulation results from MATLAB of BER versus E_b/N_0 curve based on the performance achieved across

Table 10.1 Link performance SISO results

Parameters	Vioolsdrift RS (TX)	Henkriesmond SS (RX)
Effective radiated power (W)	12.13	7.92
(dBm)	40.84	38.99
Free space loss (dB)	115.19	115.19
Polarization	Vertical	Vertical
Net path loss (dB)	160.74	157.07
RX signal (dBm)	−123.75	−120.08
Fade margin (dB)	−16.75	−13.08

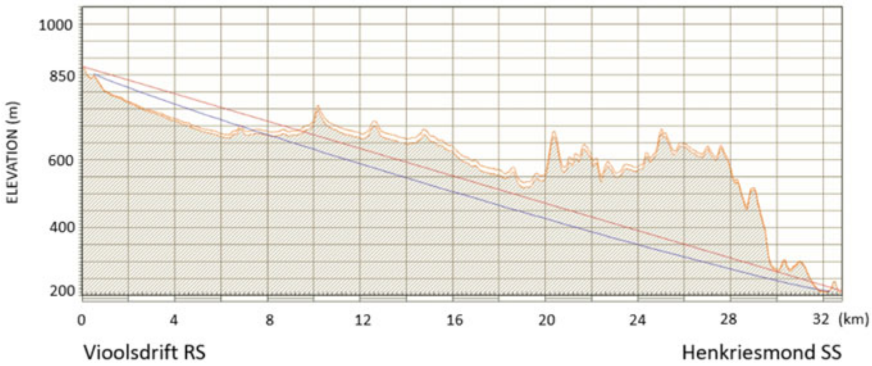


Fig. 10.6 No line of sight between radio site and substation

the 1 × 1 SISO System. The number of transmit antennas (Nt) and the number of receive antennas (Nr) used are 1, respectively. Based on (BER, E_b/N_0) curve shown in Fig. 10.6, as the signal (E_b/N_0) increases, the bit error rate (BER) gradually shows a very slight but noticeable decrease as seen in Table 10.2.

Based on the SISO system in Fig. 10.6, assume that our system is operating at 100 mb/s and selecting the closest to minimum BER reading in Fig. 10.6, at 15 dB(E_b/N_0) which is 10^{-2}. This means that 1-bit error occurs every 10^2 bits. Thus, to receive 10^2 bits, the time taken would be

$$\frac{\text{Number of Bits Sent}}{\text{Bit Rate}} = \frac{10^2}{10^8} = 1\,\mu s \tag{10.3}$$

1 μs is the average time it takes before an error can occur in this system. Therefore, we can see with the use of one antenna at both ends of the system, this radio link performance is very poor with errors occurring at every micro second.

1. Simulated results with the OSTBC MIMO system employed.
2. Second Case Scenario: 4 × 4 OSTBC MIMO system in a Rayleigh channel.

Table 10.2 BER values of a 2X2 OSTBC MIMO system (NT = 2) in a Rayleigh fading channel

E_b/N_0 (dB)	Bit Error Rate(BER)	Number of bits
1	0.0	0
2	0.0	0
3	0.21	3000
4	0.1835	3000
5	0.1581	3000
6	0.1338	3000
7	0.1109	3000
8	0.0881	3000
9	0.0681	3000
10	0.0517	3000
11	0.0386	3000
12	0.0288	3000
13	0.0206	3000
14	0.0145	3000
15	0.01	3000
16	0.0073	3000
17	0.0047	3000
18	0.0030	3000

The second case scenario as shown in Fig. 10.7, are simulation results of BER versus E_b/N_0 curve based on a 4 × 4 OSTBC MIMO System. The number of transmit antennas (Nt) used is 4, and the number of receive antennas (Nr) used is 4, respectively. As seen in Table 10.3, as the signal (E_b/N_0) increases, the bit error rate (BER) shows a more significant decrease to that shown in the system in Fig. 10.6.

Based on the (BER, E_b/N_0) curve shown in Fig. 10.7, again assuming that the system is operating at 100 mb/s, and choosing the minimum BER reading possible in Fig. 10.7, at 18 $dB(E_b/N_0)$ which is approximately 10^{-5}. This means that 1-bit error occurs every 10^5 bits [19]. Hence, to receive 10^5 bits, the time taken would be approximately

$$\frac{\text{Number of Bits Sent}}{\text{Bit Rate}} = \frac{10^5}{10^8} = 1\,\text{ms} \tag{10.4}$$

Therefore, 1 ms is the average time it takes before an error could occur in this system.

It can be seen that the use of four antennas at both ends greatly improves the quality of the performance of the system. In this simulation, the performance of scenario two is better than scenario one. The Receive Signal Strength results of the VHF OSTBC MIMO System uses two different VHF frequencies.

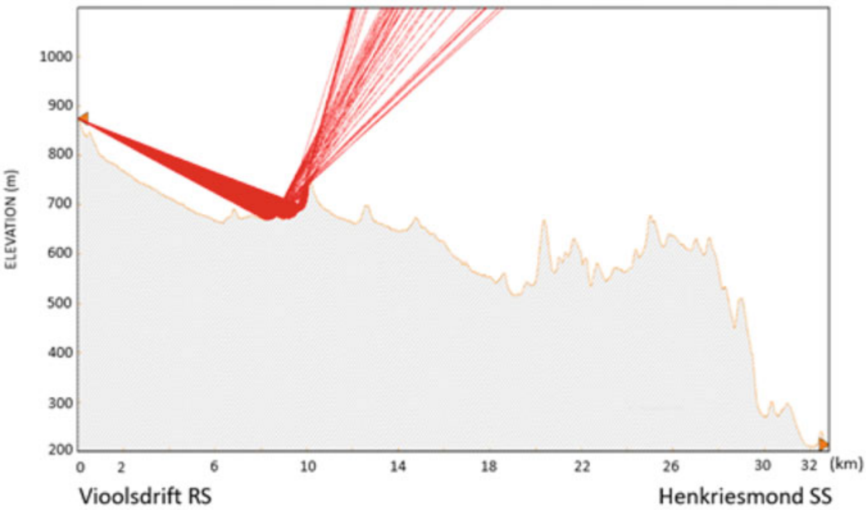


Fig. 10.7 Multipath propagation diagram

Table 10.3 BER values of a 4×4 OSTBC MIMO system in a Rayleigh channel

E_b/N_0(dB)	Bit error rate (BER)	Number of bits
1	0.0	0
2	0.0	0
3	0.1267	3000
4	0.1009	3000
5	0.0784	3000
6	0.0580	3000
7	0.0395	3000
8	0.0252	3000
9	0.0147	3000
10	0.0078	3000
11	0.0032	3000
12	0.0019	3000
13	0.000666	3000
14	0.000133	3000
15	0.0000366	3000
16	0.0000366	3000
17	0.0	3000
18	0.0	3000

Fig. 10.8 BER curve of a 1 × 1 SISO system

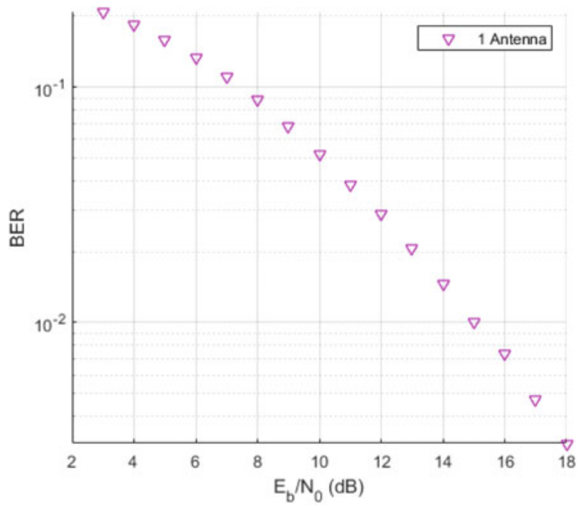


Figure 10.8 shows a proposed radio link path into Henkriesmond Substation. A microwave radio link between Vioolsdrift Radio Site and Doringwater Substation and a VHF repeater at Doringwater Substation propagating towards Henkriesmond Substation is proposed. The repeater at Doringwater Substation makes use of a 12 element Yagi antenna at a height of 60 m, and two different VHF frequency bands, namely 135 and 70 MHz, will be used. The distance between the transmitter and receiver is 11 km. The interfering obstacle (mountain) is 700 m above sea level. This part of the simulation will only focus on the radio signal propagation performance of the repeater at Doringwater Substation towards Henkriesmond Substation. The design of the microwave link from Vioolsdrift Radio Site to Doringwater Substation will not be dealt with in this study.

1. A 4 × 4 OSTBC MIMO System Operating at a VHF Frequency of 135 MHz.

Figure 10.9is a simulated MATLAB SiteViewer propagation pattern for a VHF OSTBC MIMO repeater situated at Doringwater Substation towards Henkriesmond Substation operating at a frequency of 135 MHz. As seen in Table 10.4, which contains the performance profile of this repeater, the receive (RX) signal strength of this system is −96.92 dBm. This reading means that the RTU at Henkriesmond Substation will fail at −96.92 dBm, because the receive signal strength threshold for RTU operation at Eskom is between −60 dBm and −90 dBm.

2. A 4 × 4 OSTBC MIMO System Operating at a VHF Frequency of 70 MHz.

Figure 10.10is also a simulated MATLAB SiteViewer propagation pattern similar to Fig. 10.9 except that its operating frequency is much lower at 70 MHz. The receive (RX) signal strength reading of this system has now improved beyond the threshold to −89.71 dBm, see Table 10.5. This reading means that the RTU at Henkriesmond Substation will operate perfectly at -89.71 dBm, because receive signal strength is

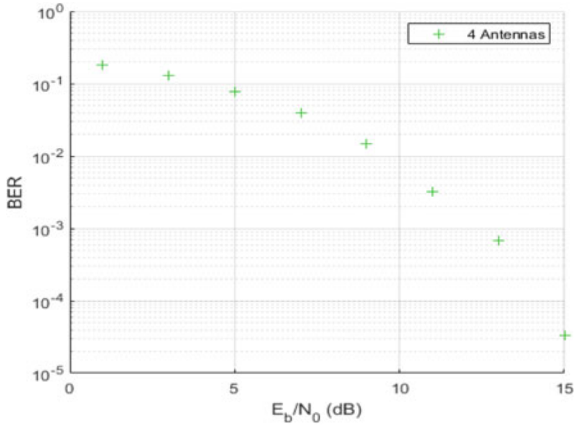


Fig. 10.9 BER curve of a 4 × 4 OSTBC MIMO system

Table 10.4 Propagation profile of the VHF repeater to Henkriesmond substation at 135MHz

Doringwater substation (TX)	Parameters
Latitude	−29.08333333
Longitude	17.94527778
Frequency (MHz)	135
TX power (watts)	10
Elevation (m)	700
Antenna height (m)	60
Antenna model	Yagi (12 El) Y425_12
Antenna gain (dBi)	16
MIMO array	4 × 4
Azimuth (°)	30
Henkriesmond Substation (RX)	
Latitude	−28.90111111
Longitude	18.13694444
Antenna model	Yagi (12 El) Y425_12
Antenna height (m)	60
RX sensitivity criteria (dBm)	−90
RX sensitivity (dBm)	−107
RX signal strength (dBm)	−96.92

within the RTU operating threshold standard for Eskom which is between −60 and −90 dBm [20] (Fig. 10.11).

The OSTBC MIMO technology is very effective in helping reduce the impact of time dispersion, path loss and interference [21].

The simulation results received in this section confirm that the OSTBC MIMO system operated at lower frequencies does overcome radio signal reflection, fading

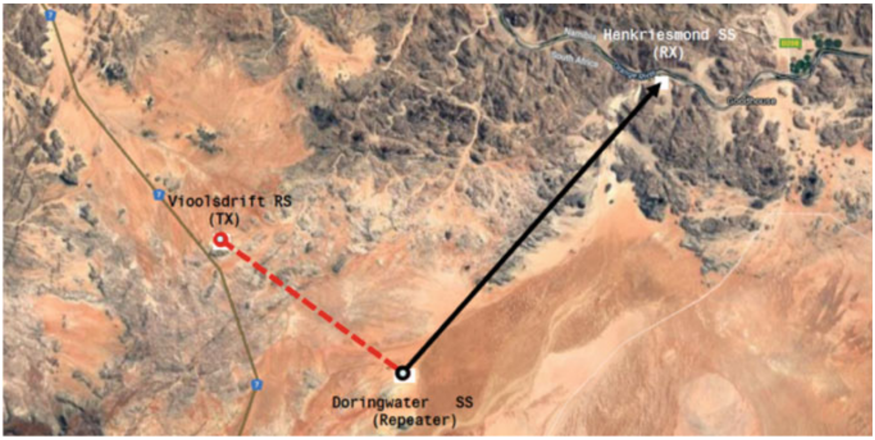


Fig. 10.10 Google earth image of the proposed radio link path into Henkriesmond Substation

Table 10.5 Propagation profile of the VHF repeater to Henkriesmond substation at 100MHz

Doringwater Substation (RX)	Parameters
Latitude	-29.08333333
Longitude	17.94527778
Frequency (MHz)	100
TX power (watts)	10
Elevation (m)	700
Antenna height (m)	60
Antenna model	Yagi (12 El) Y425_12
Antenna gain (dBi)	16
MIMO Array	4 x 4
Azimuth (°)	30
Henkriesmond Substation (TX)	
Latitude	-28.90111111
Longitude	18.13694444
Antenna model	Yagi (12 El) Y425_12
Antenna height (m)	60
RX Sensitivity Criteria (dBm)	-90
RX Sensitivity (dBm)	-107
Antenna height (m)	**-90.53**

The bold indicates the respective titles

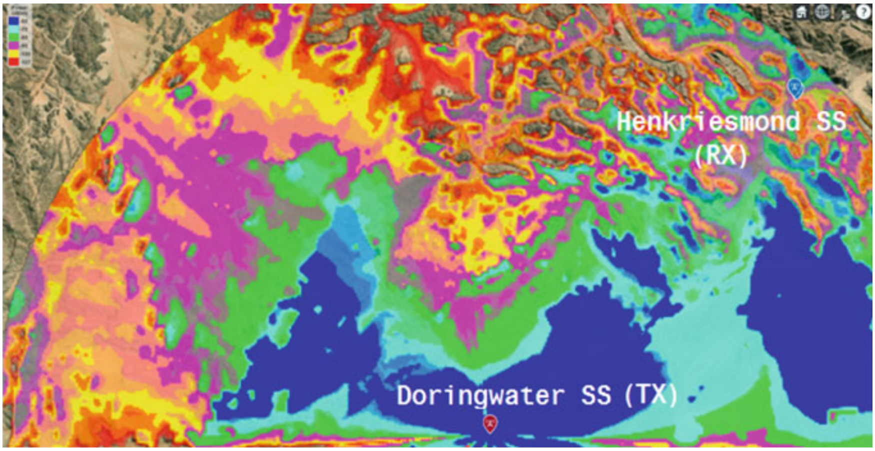


Fig. 10.11 MATLAB SiteViewer propagation diagram from Doringwater Substation to Henkriesmond Substation operating at 135 MHz

and scattering. The OSTBC MIMO system employed in a mountainous terrain can significantly reduce the bit error rate (BER), thus providing a quality communication channel than a single antenna transceiver [22].

10.5 Conclusion

This paper has demonstrated through various simulation results that the use of OSTBC MIMO system at low very high frequency (VHF) frequencies does greatly improve signal reception in irregular terrain.

Although the MIMO system has been exploited widely for many years especially in the cellular phone environment, its application in the low VHF band for SCADA purposes has not been widely investigated as has been in this paper.

Furthermore, the application of the OSTBC MIMO technology in the low VHF frequency band for improvement of signal reception in irregular terrain has certainly not been explored in the mountainous Northern Cape in South Africa.

Therefore, the information provided in this study not only presents useful information for radio engineers designing and installing radio systems in mountainous areas but also for the Eskom business to save costs, provide quality service and to improve safety. This study will also encourage more researchers to identify more research gaps to improve the system.

Further work involving testing this technique in the field using a real data is planned at Eskom in the future.

References

1. Wright, D.R., Theron, S., Lee: Birds of the Karoo: Ecology and Conservation, p. 8 (2019)
2. Atoll overview, p. 1 [Online] available at: https://www.forsk.com/atoll-overview. Accessed 04 Mar 2020
3. Maxama, X. B., Markus E.D.: VHF OSTBC MIMO system used to overcome the effects of irregular terrain on radio signals. In: Lecture Notes in Networks and Systems. Springer, Singapore, ISBN 978-981-336-172-0
4. Lai T.X., Ran T.A., Sessay A.B.: Performance analysis of space time block coded systems over frequency selective rayleigh fading channels. In: EEE Wireless Communication Network Conference **3**, 1480–1485 (2006)
5. Cortes-Pena L.: MIMO space-time block coding (STBC): simulation and results. ECE6604 Personal Mobile Communication pp. 1–8 (2006)
6. Coll J.F.: Channel Characterization and Wireless Communication Performance in Industrial Environments, Doctoral Thesis in Information and Communication Technology Stockholm Sweden (2014)
7. Al-Omary A.: End to end performance evaluation for MIMO-OSTBC adaptive modulation over rayleigh fading. Int. J. Eng. Technol. **3**, p. 8 (2018)
8. Tarokh, V., Jafarkhami, H., Calderbank, A.R.: Space-time block codes from orthogonal designs. IEEE Trans. Inf. Theory **45**(5), 1456–1467 (1999)
9. Foschini, G.J.: Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas. Bell Labs Tech. J. **1**(2), 41–59 (1998)
10. Hochwald, B.M., Marzetta, T.L.: Unitary space-time modulation for multiple-antenna communications in Rayleigh flat fading. IEEE Trans. Inform. Theory **46**(2), 543–564 (2000)
11. Alamouti, S.M.: A simple transmit diversity technique for wireless communications. IEEE J. Sel. Areas Commun. **16**(8), 1451–1458 (1998)
12. Delport P.: Optimisation of UHF Radio Systems for Electrical Distribution Networks, Master Thesis, University of Stellenbosch, p. 43, (2006)
13. Kumar, V., Dhyani, P.Dr., Anupma: Performance analysis of MIMO OFDM system using BPSK and QPSK modulation techniques under rayleigh fading channel. Int. J. Adv. Res. Comput. Eng. Technol. (IJARCET), **5**(1) (2016)
14. Abbas W., Abbas, N., Majeed, U., Khan, S.: Efficient STBC for the data rate of MIMO-OFDMA. Sci. Institute (Lahore), **28**, 247–255 (2016)
15. Ballal, B.R., Chadha, A., Satam, N.: Orthogonal frequency division multiplexing and its applications. Int. J. Sci. Res. **2**(1), 325–328 (2013)
16. Breed G.: Bit error rate: fundamental concepts and measurement issues. High Frequency Electron. p. 46 (2003)
17. Stec, B., Susek, W.: Theory and Measurement of Signal-to-Noise Ratio in Continuous-Wave Noise Radar, Faculty of Electronics, Military University of Technology, Gen. Urbanowicz St. No 2, 00-908 Warsaw, Poland, p. 1, 6 (2018)
18. Kumar, V., Dhyani, P.Dr., Anupma: Performance analysis of MIMO OFDM system using BPSK and QPSK modulation techniques under rayleigh fading channel. Int. J. A. Res. Comput. Eng. Technol. (IJARCET), **5**(1) (2016)
19. Rex, R.: Impact of BER on LAN throughput, p. 3 [Online] available at: https://www.brand-ex.com. Accessed 15 Jan 2020
20. Delport, P.: Optimisation of UHF Radio Scada Systems for Electrical Distribution Networks, Master's Thesis, p. 46 (2006)
21. Zakaria, Y., Hosek, J., Misurec, J.: Path loss measurements for wireless communication in urban and rural environments. Am. J. Eng. Appl. Sci. p. 1 (2015)
22. Maxama, X.B., Markus, E.D.: A survey on propagation challenges in wireless communication networks over irregular terrains. In: 2018 Open Innovations Conference (OI). IEEE 2018
23. Zang, J., Wang, X.: Measurements and modelling of path loss over irregular terrain for near-ground and short-range communications. Prog. Electromagn. Res. M **57**, 60 (2017)

24. Tarokh, V., Seshadri, N., Calderbank, A.R.: Space-time codes for high data rate wireless communication: performance criterion and code construction. IEEE Trans. Inform. Theory **44**(2), 744–765 (1998)

Chapter 11
Open Challenges of Communication Security in an IoT Environment—A Survey

Mahsa Mirlashari and Syed Afzal Murtaza Rizvi

Abstract Internet of Things (IoT) is a continuously growing concept in ICT concerning the progression of ubiquitous computing, machine-to-machine (M2M) communication, and wireless sensor networks. IoT is the network of physical objects such as watches and bicycles integrated with electronic equipment, sensors, and software. These devices gather and exchange data using network connectivity. However, with such an enormous number of IoT device connections on the network, the security of the data and communication become a major concern. This paper gives an overview of the IoT and the associated limitations and challenges of applying security in IoT devices and the classification of IoT attacks. Furthermore, the paper also discussed the security issues at different layers. Moreover, it demonstrates some identified solutions to the existing challenges and provides future direction to tackle the challenges present in IoT security.

11.1 Introduction

The IoT could mention the capability of connecting objects for transferring data over a network without H2H (human-to-human) or H2M (human-to-machine) participation. Due to this reason, a personal computer is not considered as an IoT device. A smartphone has lots of sensors, but it cannot be considered as an IoT device. Devices like a smartwatch, fitness band, etc., can be counted under the IoT devices.

In the IoT ecosystem, everything can be connected in the physical world seamlessly through the infrastructure of the existing Internet. Data is being collected and captured by the sensor devices when things interact with each other. IoT is capable enough to be applied to many domains ranging from automated home equipment to smart moving devices, etc. Furthermore, IoT can also be utilized for smart manufacture, agriculture, healthcare, and other social activities. We can convert physical

M. Mirlashari (✉) · S. A. M. Rizvi
Department of Computer Science, Jamia Millia Islamia, New Delhi, India

S. A. M. Rizvi
e-mail: sarizvi@jmi.ac.in

T. Senjyu et al. (eds.), *IOT with Smart Systems*, Smart Innovation, Systems and Technologies 251, https://doi.org/10.1007/978-981-16-3945-6_11

object (device) into an IoT device by simply connecting it to the Internet and control it. This holistic vision poses security attacks and vulnerability if the IoT gadget or system is not properly secure [1].

The present paper investigated the various challenges and shortcomings associated with securing IoT devices. The paper also discusses the feasible solutions suggested in the existing studies to solve the security issues The other section of this paper is organized as follows: Sect. 11.2 outlines the challenges in the existing IoT devices. Security issues in IoT are presented in Sect. 11.4. Section 11.3 provides different IoT attacks. The possible solutions to the discussed challenges have been explained in Sect.11.5. Section 11.6 discusses the conclusion.

11.2 Challenges in IoT Security

The steady rise in the growth of IoT-connected devices has introduced several unique security challenges, and therefore, it has gained the attention of the research community. We categorize the challenges into eight classes which are discussed in bellow:

11.2.1 Limitations of Applying Security

The IoT devices are highly dependent on the battery's power since they are largely deployed in the environment where there are no charging facilities. The IoT devices continuously need for replacing batteries. Therefore, this high reliance on the batteries makes the devices vulnerable to battery-exhausting attacks or battery-depletion attacks [2]. These attacks are sometimes considered as the special class of DoS attacks [3]. These attacks attempt to consume more energy of the devices to destroy their battery life fast.

11.2.2 Limitations of Computing Power

The limited memory space of IoT devices makes them incapable to handle and execute complex jobs, and thus, it does not support advanced cryptography techniques. Since computing is expensive, low-end devices cannot use encryption algorithms (DH, RSA, DSA) based on an asymmetric key. RFID tags cannot even use cryptographic algorithms (AES, DES, and 3DES) based on a symmetric key. Regarding authentication, it is worthwhile noting that authentication based on the digital signature is not suitable because of requiring the public–private key scheme. So, the only authentication based on the symmetrical key or other lighter methods applies to IoT devices [4, 5].

Having energy-efficient sensors and communication technology is the major challenge in designing the IoT devices. Therefore, security solutions are also expected to consume fewer energy resources such as CPU processing and memory.

11.2.3 Big IoT Data Analysis

In IoT, the ability to analyze huge amount of IoT based data collected from various sources such as sensors, healthcare applications, social media, and devices is a major challenge, for traditional database systems which are inefficient to store, retrieve, process, and analyze a huge amount of data for future access [6]. Not only does the high complexity of processing big IoT data prevent the simplified and smooth analysis of available data for the investigation, but also analytical algorithms scalability could have a major impact on the investigation [7].

11.2.4 Trusted Updates and Management

Management of IoT devices includes monitoring, administration, and recognition of replication problems and amendatory actions. The important function of IoT device management encompasses software updates, configuration, and overall control. As mentioned in [8], trusted updates and management of connected device software are big challenges in IoT. To grab these issues, it is advised to use blockchain technology, although blockchain technology having some problems such as efficiency, scalability, and key collision.

11.2.5 Quality of Service

There are various challenges while developing an IoT network, one of them being a quality of service (QoS) in terms of data drop, data quality sensing, resource consumption, etc. The QoS handles all the parameters that can influence the network's reliability, availability, and performance directly or indirectly. In this category, some considered points by the European Commission in 2014 are bandwidth, cost efficiency, capacity and throughput, latency, resource optimization, trustworthiness, and scalability [9].

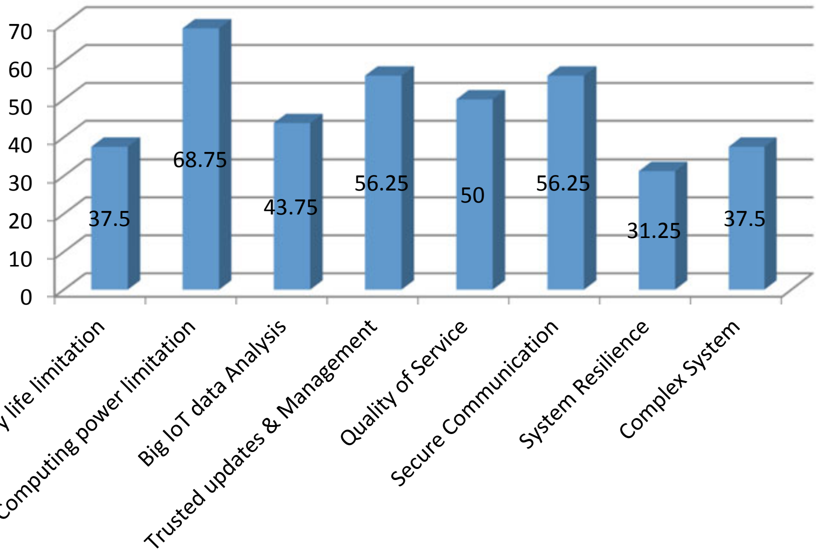


Fig. 11.1 Comparison between IoT security challenges

11.2.6 Secure Communication

The communication channel linking different communicating nodes, such as cloud services and IoT devices, must be secured against any attack. In most cases, data is sent by IoT devices without encryption is an easy target for different network attacks. Therefore, by using an appropriate encryption method or separate networks for separator devices and creating private communication channels, such attacks can be reduced [10].

Figure 11.1 highlights a clear trend of the IoT security challenges referenced in this paper and also demonstrated in [5, 11–17].

11.2.7 System Resilience

If smart devices are compromised, the ability of the system should be sufficiently efficient to defend other network nodes against any attack. Therefore, the significant role of the system resilience should not be overlooked in respond to unforeseen attacks/situations without receding [10].

11.2.8 Complex System

The numerous heterogeneous devices interact together in the IoT system, but with the increase of people, devices, interfaces, and interactions in the IoT system, security breaches risks are increasing as well. Consequently, it seems to achieve the high security required to manage the large-scale network which is not an easy task due to time, memory, and energy limitations.

11.3 Different IoT Attacks

The IoT attacks are classified into four major categories such as physical attack, encryption attack, software attack, and network attack [18, 19].

Physical attacks target the hardware components of the IoT network. For example, node tampering where the attacker tries to gain access over sensor node, and node jamming where attacker interferes with radio frequencies in to jam the signals of the sensor nodes and halt the communication. This attack also includes man in the middle attack, where the adversary deploys a malicious node between the legitimate nodes to watch and control the data flow. This attack typically requires the attackers to be in the proximity of the IoT system. *The encryption attack* attempts to break the encryption policy of the IoT system. Side-channel attacks and cryptanalysis are two main approaches used by the attacker to serve the purpose. In a side-channel attack, an attacker retrieves secret information from the unusual mechanisms such as by eavesdropping the power consumption or electromagnetic radiations of a cryptographic device, or by analyzing the running complexity of cryptographic operations. This little information can assist the attackers in producing the duplicate key of the encryption process. Side-channel attacks assume that the logic operations of the encryption show physical characteristics based on the input data. In cryptanalysis, the attacker aims to attack the encryption keys by exploiting the weaknesses in the encryption algorithm. These attacks include chosen plaintext attacks (CPAs), known-plaintext attacks (KPA), ciphertext-only attacks (COA), etc. *Software attack* exploits the vulnerabilities that exist in the implementation of IoT applications. This attack includes buffer overflows attacks, injecting malicious code, etc. The adversary can leverage the communication interface to installing malware into the IoT network which is later used to perform various unlawful operations such as distributing a virus, stealing data, and spying activities. The attacker can also exploit the software vulnerabilities to launch DDoS attacks. Lastly, wireless communications devices in the IoT environment are prone to *Network attacks*. An adversary can capture vital data by exploiting the RFID's wireless characteristics. The attacker can spoof RFID signals or create the clone of the RFID tag to gain access to the system [20]. The DDoS (distributed denial of service) attack is performed by an adversary with delivering more packets in the IoT network to make the server denies the client services.

Also, the routing information can also be spoofed or altered to create loops in the route, send false error messages, drop the traffic, etc. [21].

11.4 Security Issues in IoT

IoT security has become the most controversial issue in ICT. Three layers namely application layer (smart city, smart home, smart health, etc.), perception layer (sensor gateways, RFID nodes, sensor nodes, etc.), and network layer (mobile communication networks, Internet, cloud computing) present a reliable and safe environment in IoT, as shown in Fig. 4. All these issues at different layers are given in Table 11.1 and briefly discussed in the following.

11.4.1 Security Issues at Application Layer

In this layer, the main concern is about malicious attacks that make some problems in the normal execution of the application program code. Besides, due to the security challenges, the compromised applications in this layer cannot provide appropriate services such as authenticated services which are planned for it [22–24]. Common threats to application layer are briefly explained below:

Malicious code injection is an extreme threat and focuses on physically injecting malicious code into the IoT node to gain access to the IoT device [25]. Some of these malicious codes are computer viruses, worms, spyware, etc. It is noteworthy that some malicious programs are capable to act more than one type of malicious code. For instance, certain programs may be both a virus and a trojan horse. *Access control attack* is an authorization mechanism that gives data or account access to only legitimate users or processes. Access control attack is a crucial attack on IoT applications since the entire IoT application becomes vulnerable to attacks if the access is compromised [26]. *Smart meters* can get connected and intercommunicate with network devices inside homes, such as refrigerators and air conditioners. An attacker who accesses the meter also controls its software, and by checking the power consumption, the attacker knows whether the house is vacant or not. The security of the smart grid was a major concern, so an attack on the smart grid has catastrophic consequences and very high economic cost [27].

11.4.2 Security Issues at Perception Layer

The security issues in this layer, most threats originate from the outside objects concerning sensors and other data gathering utilities and node level security is a

Table 11.1 Recent security solution for IoT

Reference	Years	Contribution
[32]	2020	The authors investigated the main advantages and design obstacles for the incorporation of blockchain innovations for IoT applications
[4]	2020	In this paper, the authors defined IoT infrastructure, application, and protocol. Then, in the IoT model, security vulnerabilities are established. Besides, several new strategies identify to address IoT security problems. After a thorough overview, the authors realized that the current solution approach to solving the security problem in IoT is artificial intelligence, blockchain, and machine learning
[1]	2019	To differentiate network activity and to detect network attacks such as MITM, replay, and DoS and also to detect a multi-stage attack on IoT networks, the authors suggested a three-layer IDS using a supervised learning method of machine learning
[33]	2018	In this paper, the authors review the security model of IoT applications. Some of the problems in IoT systems, such as access control, authentication, and trust management, were stated in the paper. Then, some techniques were discussed as a solution
[34]	2018	The authors clarified that AI could make the system more realistic by computing unstructured data, huge quantities, and heterogeneous data in real-time
[25]	2020	The authors suggested security solutions for the safety of the smart city, such as approaches provide shared security monitoring, authentication, and analysis, as well as privacy and data integrity
[35]	2019	The authors in a study [35] address the privacy issues and suggested lightweight protocols and context-aware methods for privacy and most recent methods of virtualization are used to keep data integrity. Also, the SDN solution provides an implementation of IoT based lightweight cryptographic solutions through the conducted centralized routing at the SDN controller
[36]	2018	The authors in [36] presented legal adversaries (e.g., active, passive, and single-malicious active) and also explored smart blub and smart light as smart home devices. The results indicated that even a passive adversary can access the information, which determines the actions taken at a particular time interval

major concern, where sensors are the prime targets of the attacker as well [22–24]. Most common threats in this layer are as follow:

Eavesdropping attacks are easier and can be passive. In this attack, an attacker can access the received and transmitted data through sensors embedded in smart homes and healthcare areas by utilizing an unsecured network connection [28]. In a *sniffing attack*, a hacker can investigate the network and gain unencrypted data to crash or capturing the network traffic. It should also be noted that appropriate security measures are required to prevent unauthorized persons from accessing sensitive user data [26]. There is a probability that the data may include *Noise* during transfer through a wireless network that covers long distances [10]. Not only noisy data can arise from hardware failures, programming errors, and inaudible inputs from speech or optical character recognition (OCR) programs, but also the required storage space is significantly climbed by the noisy data and as well as can harm the conclusions of any data mining analysis.

11.4.3 Security Issues at Network Layer

In this case, the prominent security problems are malicious nodes and attacker which can expose the faulty connected devices in the network. Common threats in the network layer are briefly mentioned as under:

A *DoS attack* is a security attack that happens when a hacker sends a large volume of traffic at the same time and prevents legitimate users from accessing specific computer systems, devices, services, or other IT resources [4]. *End-to-End Security* delivers a safe mechanism for ensuring accurate transmission of data between sender and receiver. To achieving this aim, data privacy breaches will be the key security challenge for the IoT network layer [29]. Any access device that is expired, stolen, revoked, lost, obtained, or cancelled with intent to defraud is called *unauthorized access device* [30]. However, devices like pacemaker implants require precise timing to convey control signals at set times provided, if they left unattended, it might be very dangerous. Besides, some malicious nodes masquerade themselves as authenticated devices and could communicate with other connected devices. An attacker in the MITM (man in the middle) attack interferes with the communication between sensor nodes by putting a malicious node in between that allows monitoring and interception of all transmitted traffic and infracts sensor node privacy. This attack is carried out via the IoT system's network and does not demand that the attacker be physically present [31].

11.5 Existing Solutions

In Table 11.1, we are highlighting some of the well-known studies toward the solutions to different IoT related issues.

11.6 Conclusion

The Internet of Things is becoming a very promising paradigm; however, the related security and privacy issues hinder the adoption of IoT devices by the users. In this study, we have discussed the various security issues in IoT faced by the security experts along with the solutions proposed by the researchers over time. The IoT security challenges have been categorized into ten classes, namely limitations of applying security (limited battery), limited computing power, Big IoT data analysis, quality of service, secure communication, system resilience, complex system, trusted updates and management, different IoT attack classes, and issues at different IoT layers. All the classes are discussed in detail in the paper. The survey conducted in the paper will surely be helpful for the researchers of this domain to understand the various issues and challenges existing in IoT devices easily and conveniently.

References

1. Anthi, E., Williams, L., Słowińska, M., Theodorakopoulos, G., Burnap, P.: A supervised intrusion detection system for smart home IoT devices. IEEE Internet Things J. **6**(5), 9042–9053 (2019)
2. Kim, H., Shin, K.G., Pillai, P.: MODELZ: monitoring, detection, and analysis of energy-greedy anomalies in mobile handsets. IEEE Trans. Mob. Comput. **10**(7), 968–981 (2011)
3. Shakhov, V., Koo, I.: Depletion-of-battery attack: specificity. Modell. Analysis Sensors **18**(6), 1849 (2018)
4. Mohanta, B.K., Jena, D., Satapathy, U., Patnaik, S.: Survey on IoT security: challenges and solution using machine learning, artificial intelligence and blockchain technology. Internet of Things 100227 (2020)
5. Sha, K., Wei, W., Yang, T.A., Wang, Z., Shi, W.: On security challenges and open issues in Internet of Things. Futur. Gener. Comput. Syst. **83**, 326–337 (2018)
6. Yaqoob, I., Hashem, I.A.T., Ahmed, A., Kazmi, S.A., Hong, C.S.: Internet of things forensics: recent advances, taxonomy, requirements, and open challenges. Futur. Gener. Comput. Syst. **92**, 265–275 (2019)
7. Quick, D., Choo, K.K.R.: Quick analysis of digital forensic data. In: Big Digital Forensic Data, pp. 5–28. Springer, Singapore (2018)
8. Khan, M.A., Salah, K.: IoT security: review, blockchain solutions, and open challenges. Futur. Gener. Comput. Syst. **82**, 395–411 (2018)
9. Shukla, A., Tripathi, S.: Security challenges and issues of internet of things: possible Solutions. In: Proceedings of 3rd International Conference on Internet of Things and Connected Technologies (ICIoTCT), pp. 26–27 (2018)
10. Kumar, S.A., Vealey, T., Srivastava, H.: Security in internet of things: challenges, solutions and future directions. In: 2016 49th Hawaii International Conference on System Sciences (HICSS) IEEE, pp. 5772–5781 (2016)
11. Alam, S., Siddiqui, S.T., Ahmad, A., Ahmad, R., Shuaib, M.: Internet of things (IoT) enabling technologies, requirements, and security challenges. In: Advances in Data and Information Sciences,pp. 119–126. Springer, Singapore (2020)
12. Niraja, K.S., Rao, S.S.: Security challenges and counter measures in internet of things. In: 2020 International Conference on Computer Communication and Informatics (ICCCI) IEEE, pp. 1–3 (2020)
13. Patel, K., Vyas, S., Pandya, V.: IoT: leading challenges, issues and explication using latest technologies. In: 2019 3rd International Conference on Electronics, Communication and Aerospace Technology (ICECA) IEEE, pp. 757–762 (2019)
14. Balaji, S., Nathani, K., Santhakumar, R.: IoT technology, applications and challenges: a contemporary survey. Wireless Pers. Commun. **108**(1), 363–388 (2019)
15. Sfar, A.R., Natalizio, E., Challal, Y., Chtourou, Z.: A roadmap for security challenges in the Internet of Things. Digital Commun. Netw. **4**(2), 118–137 (2018)
16. Tabassum, K., Ibrahim, A., El Rahman, S.A.: Security issues and challenges in IoT. In 2019 International Conference on Computer and Information Sciences (ICCIS) IEEE, pp. 1–5 (2019)
17. Adat, V., Gupta, B.B.: Security in internet of things: issues, challenges, taxonomy, and architecture. Telecommun. Syst. **67**(3), 423–441 (2018)

18. Babar, S., Stango, A., Prasad, N., Sen, J., Prasad, R.: Proposed embedded security framework for internet of things (IoT). In: 2011 2nd International Conference on Wireless Communication, Vehicular Technology, Information Theory and Aerospace and Electronic Systems Technology (Wireless VITAE) IEEE, pp. 1–5 (2011)
19. 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)
20. Mitrokotsa, A., Rieback, M.R., Tanenbaum, A.S.: Classification of RFID attacks. Gen **15693**, 14443 (2010)
21. Wu, D., Hu, G., Ni, G.: Research and improve on secure routing protocols in wireless sensor networks. In: 2008 4th IEEE International Conference on Circuits and Systems for Communications, IEEE, pp. 853–856
22. Suo, H., Wan, J., Zou, C., Liu, J.: Security in the internet of things: a review. In: 2012 International Conference on Computer Science and Electronics Engineering IEEE, vol. 3, pp. 648–651 (2012)
23. Xiaohui, X: Study on security problems and key technologies of the internet of things. In: 2013 International conference on computational and information sciences IEEE, pp. 407–410 (2013)
24. Kozlov, D., Veijalainen, J., Ali, Y.: Security and privacy threats in IoT architectures. In: Proceedings of the 7th International Conference on Body Area Networks ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), pp. 256–262 (2012)
25. Atlam, H.F., Wills, G.B.: IoT security, privacy, safety and ethics. In: Digital Twin Technologies and Smart Cities, pp. 123–149. Springer, Cham (2020)
26. Hassija, V., Chamola, V., Saxena, V., Jain, D., Goyal, P., Sikdar, B.: A survey on IoT security: application areas, security threats, and solution architectures. IEEE Access **7**, 82721–82743 (2019)
27. Hackers could turn your smart meter into a bomb and blow your family to smithereens—new claim. https://www.theregister.co.uk/2017/01/04/smart_metres_ccc. Accessed 24 Oct 2020
28. What is an eavesdropping attack. https://www.investopedia.com/terms/e/eavesdropping-attack.asp. Accessed on: 24 October 2020
29. HaddadPajouh, H., Dehghantanha, A., Parizi, R.M., Aledhari, M., Karimipour, H.: A survey on internet of things security: requirements, challenges, and solutions. Internet of Things 100129
30. 1030.DEFINITION. https://www.justice.gov/jm/criminal-resource-manual-1030definitions. Accessed 24 Oct 2020
31. Orifama, D.G., Okoro, H.: Security challenges in IoT platforms and possible solutions. Comput. **8**(1), 1–7 (2020)
32. Dedeoglu, V., Jurdak, R., Dorri, A., Lunardi, R.C., Michelin, R.A., Zorzo, A.F., Kanhere, S.S.: Blockchain technologies for IoT. In: Advanced Applications of Blockchain Technology, pp. 55–89. Springer, Singapore (2020)
33. Das, A.K., Zeadally, S., He, D.: Taxonomy and analysis of security protocols for internet of things. Futur. Gener. Comput. Syst. **89**, 110–125 (2018)
34. Ghosh, A., Chakraborty, D., Law, A.: Artificial intelligence in internet of things. CAAI Trans. Intell. Technol. **3**(4), 208–218 (2018)
35. Hameed, S., Khan, F.I., Hameed, B.: Understanding security requirements and challenges in internet of things (IoT): a review. J. Comput. Netw. Commun. (2019)
36. Do, Q., Martini, B., Choo, K.K.R.: Cyber-physical systems information gathering: a smart home case study. Comput. Netw. **138**, 1–12 (2018)

Chapter 12
A Compact Ultra-Wideband (UWB) MIMO Antenna for K and Ka Band Applications

Amrees Pandey, Aditya Kumar Singh, Sweta Singh, and Rajeev Singh

Abstract A two port dumbbell shaped multiple input multiple output (MIMO) antenna of size ($30 \times 30 \times 1.6$ mm^3) is conceived, designed, and simulated. Ultra-wideband (24.95–31.31 GHz) at port-1 and a wideband (25.55–30.37 GHz) at port-2 with impedance bandwidths of 22.7% and 17.3%, respectively is observed. Maximum peak gain of 8.2 dBi (port-1) and 7.9 dBi (port-2) is observed. Simulated isolation is less than −15 dB, maximum envelope correlation coefficient is 0.0012, and diversity gain between 9.997 and 9.999 is obtained. Antenna is suitable for K and Ka band applications.

12.1 Introduction

Ultra-wideband (UWB) technology in recent times is preferred due to quality of service, high data transmission rate, and other inherent advantages [1, 2]. Modern 4G technology demands fast and higher data rate for transmission and reception for which higher diversity gain in broadside direction is desirable. MIMO microstrip patch antenna system is capable to provide higher diversity gain with trade-off between reliability and data rate [3]. Focus of researchers is intended to design compact MIMO microstrip patch antenna which also provide higher isolation between ports. Many techniques are reported to reduce mutual coupling effect in MIMO microstrip patch antenna system such as diversity technique, neutralization line, meta-surface, and symmetrical designs have been used to reduce mutual coupling in MIMO antenna systems [4]. Mutual coupling between radiating elements can also be minimized by using defected ground structures (DGS) [5, 6].

In UWB system operating in multipath environment, signal fading is an important parameter which can be fixed by UWB MIMO antenna as it provides and improved channel capacity. Small sized MIMO antenna, however, degrades the performance

A. Pandey (✉) · A. K. Singh · S. Singh · R. Singh
Department of Electronics and Communication, University of Allahabad, Prayagraj, India

R. Singh
e-mail: rsingh68@allduniv.ac.in

T. Senjyu et al. (eds.), *IOT with Smart Systems*, Smart Innovation, Systems and Technologies 251, https://doi.org/10.1007/978-981-16-3945-6_12

of antenna elements [7, 8]. A compact UWB MIMO microstrip patch antenna with pattern diversity and band rejection characteristics is reported [9].

In this paper, we propose a dumbbell shaped 2 × 2 MIMO patch antenna with defected ground structure on FR-4 as substrate for K and Ka band applications. Dumbbell shaped MIMO antenna is designed and simulated by means of electromagnetic tool, High Frequency Structure Simulator (HFSS) v13 by AnSoft. This tool utilizes the mathematical formulation of finite element method. Evolution of antenna design is presented in Sect. 12.2, result and discussion in Sect. 12.3 and conclusions in Sect. 12.4 is presented.

12.2 Evolution of Antenna Design

The evolution of the proposed antenna begins with Ant-1. The patch and ground structure of Ant-1 with two ports MIMO microstrip line feeding is presented in Fig. 12.1a, b respectively. Further Ant-2 was designed with two ports MIMO microstrip line feeding with the above same patch and defected ground structure as presented in Fig. 12.2a, b. Ant-3 is the proposed antenna designed with two port MIMO microstrip line feeding with the above same patch and defected ground structure as presented in Fig. 12.3a, b. Radiating patch geometry of all the three antennas is same. All these two port MIMO antennas were designed using dielectric material epoxy FR-4 with substrate height 1.6 mm and simulated with Ansoft HFSS software tools (Table 12.1).

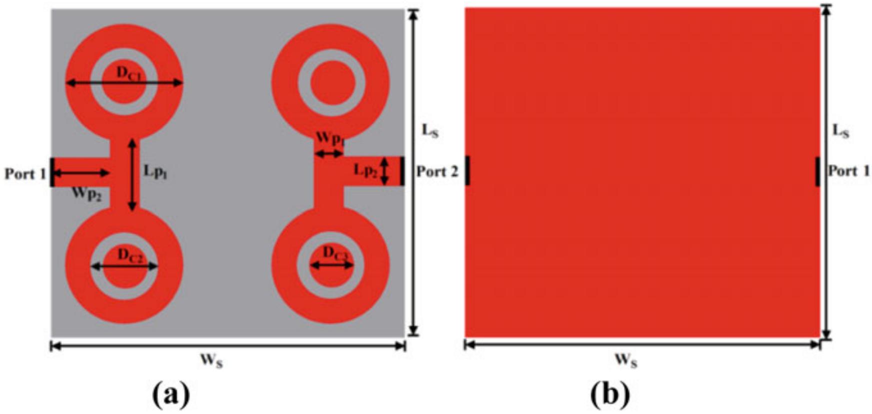


Fig. 12.1 **a** Upper and **b** lower view of ant-1

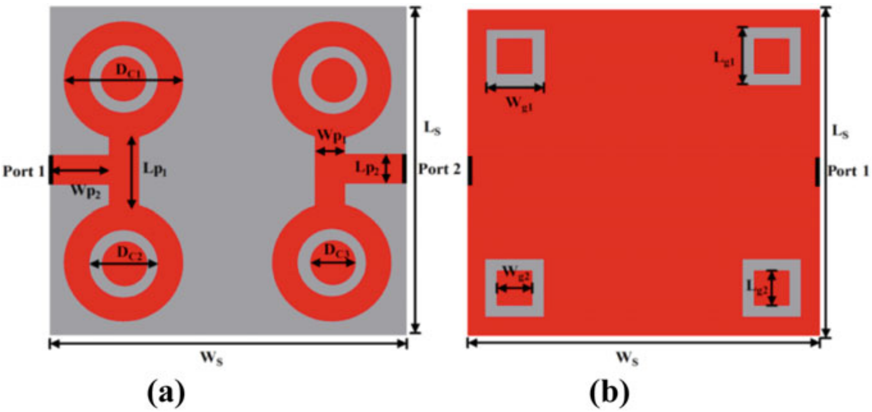


Fig. 12.2 **a** Upper and **b** lower view of ant-2

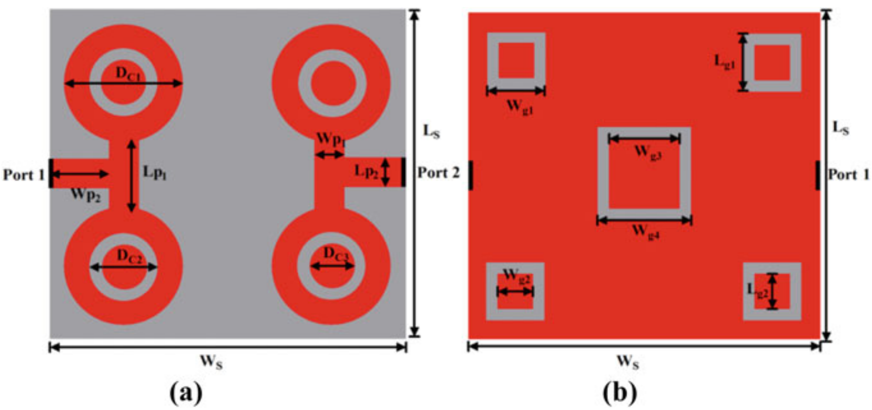


Fig. 12.3 **a** Upper and **b** lower view of ant-3 (proposed)

12.3 Results and Discussion

The showing of proposed antenna is studied in expression of surface current distribution, return loss, diversity gain, envelope correlation coefficient (ECC), and radiation pattern.

Ant-1 is without defected ground and its $|S_{11}|$ behavior is shown in Figs. 12.4 and 12.5. We observe simulated impedance bandwidths of 16.51% and 15.66%, resonance peaks at 26.38 GHz and 26.19 GHz (cf. Fig. 12.5) having peak gains of -1.2 dBi and -1.45 dBi (cf. Fig. 12.6) at port 1 and 2, respectively.

Impedance bandwidths, resonance peaks, and peak gains are depicted in Figs. 12.4 and 12.5 also for Ant-2 and Ant-3.

Table 12.1 Specifications of the proposed antenna (ant-3)

Geometrical parameters	Values
$L_S \times W_S$ (Substrate length × Substrate width)	30 mm × 30 mm
$L_{P1} \times W_{P1}$ (Patch length$_1$ × Patch width$_1$)	8 mm × 2 mm
$L_{P2} \times W_{P2}$ (Patch length$_2$ × Patch width$_2$)	2 mm × 6 mm
D_{C1} (Diameter of circle$_1$)	10 mm
D_{C2} (Diameter of circle$_2$)	6 mm
D_{C3} (Diameter of circle$_3$)	4 mm
$L_{g1} \times W_{g1}$ (Ground length$_1$ × Ground width$_1$)	5 mm × 5 mm
$L_{g2} \times W_{g2}$ (Ground length$_2$ × Ground width$_2$)	3 mm × 3 mm
W_{g3} (Ground width$_3$)	6 mm
W_{g4} (Ground width$_4$)	8 mm

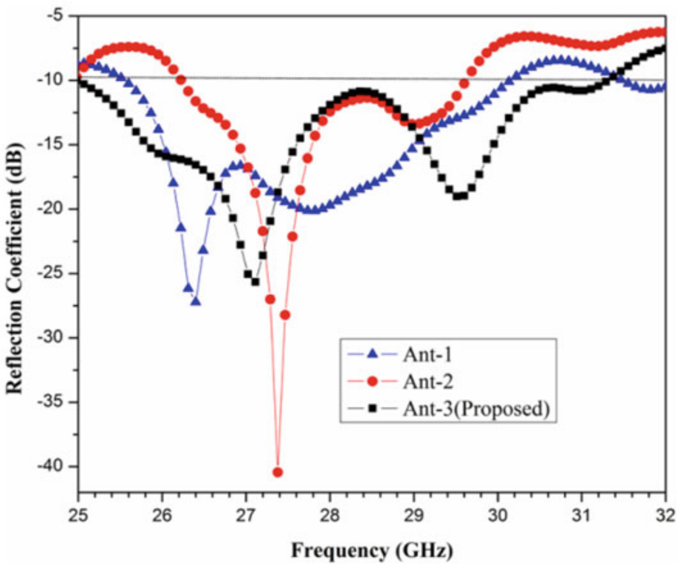


Fig 12.4 $|S_{11}|$ versus frequency of ant-1, ant-2, and ant-3 at port-1

For Ant-2 (IBW-15.2%, resonating peak at 27.18 GHz, peak gain 9.3 dBi port 1), we observe an impedance bandwidth of 7.5% at port 2 with same resonating peak and peak gain as obtained at port 1.

Table 12.2 depicts the values of impedance bandwidth, resonating bands, peak gains for Ant-3 (proposed) at port 1 and 2. An ultra-wideband having impedance

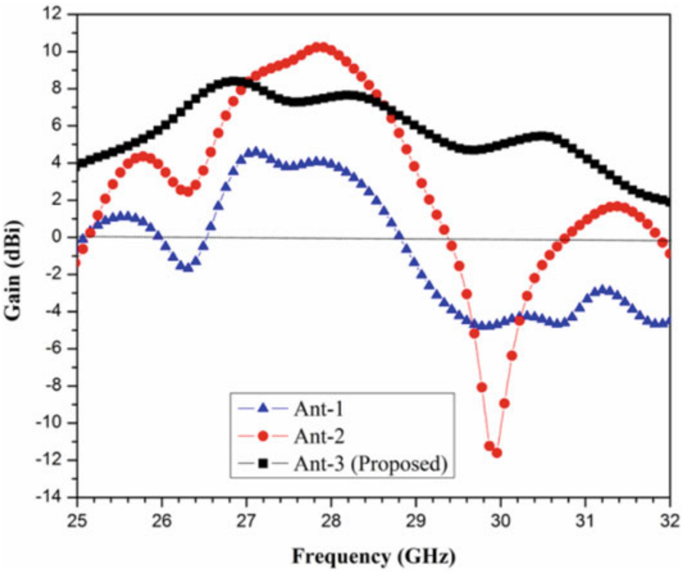


Fig 12.5 Simulated gain versus frequency of ant-1, ant-2, and ant-3

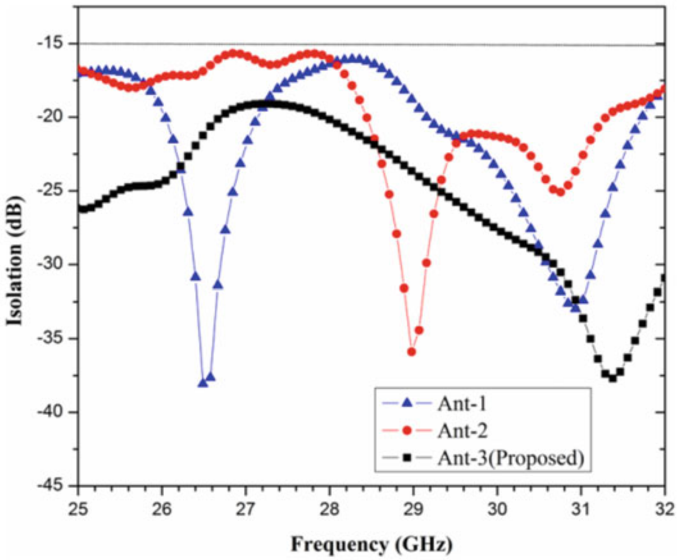


Fig 12.6 Simulated isolation (between ports) of ant-1, ant-2, and ant-3

Table 12.2 Port characteristics of proposed (ant-3) MIMO antenna

Port	Band frequency (GHz)	f_r (GHz)	Return loss (dB)	Gain (dBi)	Impedance BW (in %)
Port1	24.95–31.31	27.11	−25.6	8.2	22.7
		29.54	−18.9	4.8	
Port2	25.55–30.37	27.2	−27.15	7.9	17.3
		29.6	−22.32	4.8	

bandwidths of 22.7% and resonant frequency (f_r) at 27.11 GHz and 29.54 GHz along peak gains 8.2 dBi and 4.8 dBi, respectively, and isolation less than −15 dB is observed at port 1. A wideband with an impedance bandwidth of 17.3% and resonant frequency 27.2 GHz and 29.6 GHz peak gains of 7.9 dBi and 4.8 dBi, respectively, (cf. Figs. 12.4 and 12.5) and isolation less than −15 dB is observed at port 2 (cf. Figs. 12.6 and 12.7). However, we observe two resonating frequencies in a single band at both the ports (cf Figs. 12.4 and 12.7 and Table 12.2), which is termed as "ringing resonant frequencies" [10, 11]. As given in Table 12.2, we observe variations in gain (cf. Figure 12.7) between the first resonant peak and the second resonant peak at both ports. The ringing phenomenon causes variations in the antenna parameters, such as gain and return loss. Degradation in gain and return is observed (cf Table 12.2) due to "ringing effect."

At port 1, for the proposed antenna (Ant-3), we observe the current distribution as two different resonant frequencies is 27.11 GHz and 29.54 GHz as shown in Fig. 12.8a, b. Surface current densities at 27.11 GHz and 29.54 GHz are 35.25 A/m

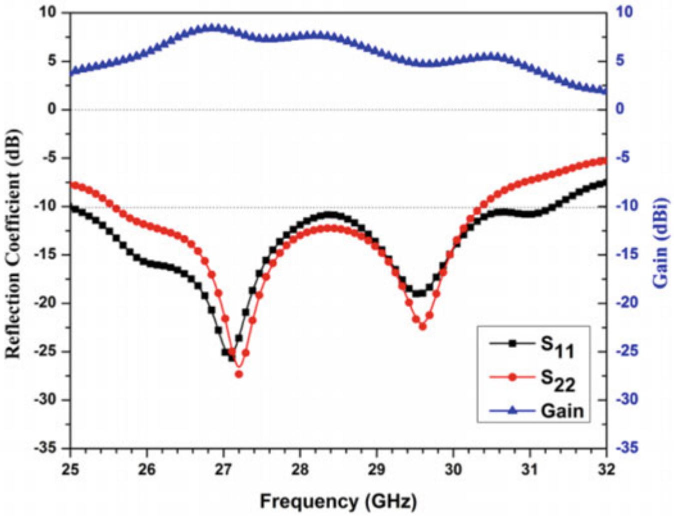


Fig 12.7 Simulated S_{11}, S_{22}, and gain versus frequency for the proposed antenna (ant-3)

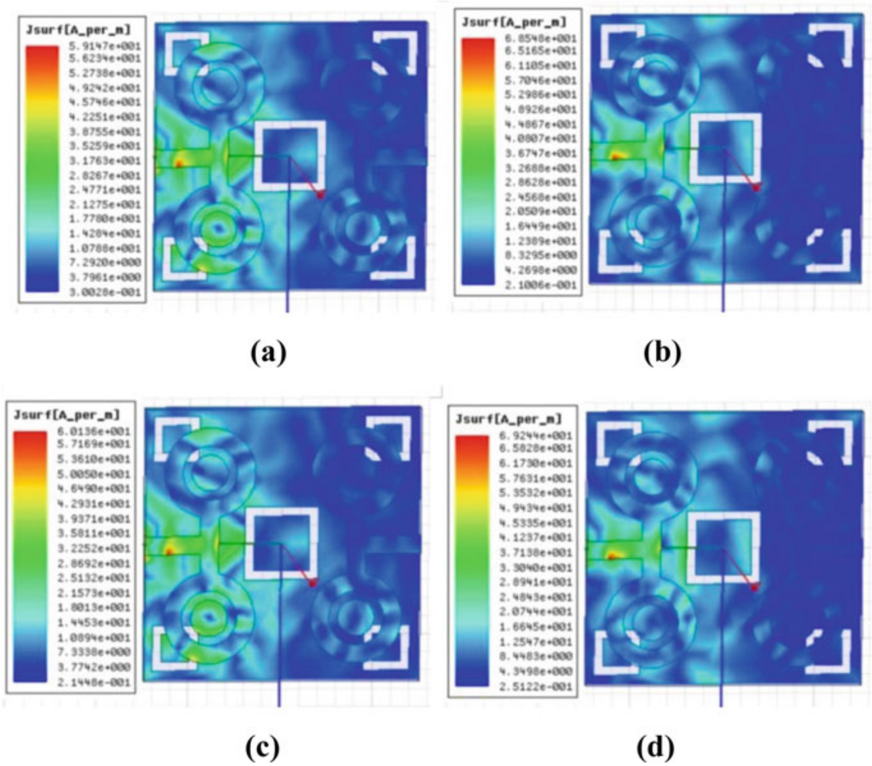


Fig 12.8 Current distribution at port 1 for **a** 27.11 GHz **b** 29.54 GHz at port 2 for **c** 27.2 GHz **d** 29.6 GHz

and 40.80 A/m, respectively. Similarly surface current distribution at resonating frequency 27.2 GHz and 29.6 GHz corresponding to port 2 is 39.37 A/m and 41.23 A/m, respectively, as presented in Fig. 12.8c, d.

ECC and diversity gain (DG) as a function of frequency is shown in Fig. 12.9. Both the antenna parameters are related as given by Eqs. 12.1 and 12.2 [8].

$$\text{Diversity Gain} = 10\sqrt{1 - \text{ECC}^2} \tag{12.1}$$

$$\text{ECC} = \frac{|S_{11} * S_{12} + S_{21} * S_{22}|}{\left(1 - |S_{11}^2| - |S_{21}^2|\right)\left(1 - |S_{22}^2| - |S_{12}^2|\right)} \tag{12.2}$$

From the Eqs. 12.1 and 12.2, it is clear that the lower the value of ECC, higher the diversity gain, and higher data transfer rate. Lower value of ECC in the range of 0–0.5 is considered to be appropriate for a MIMO antenna to operate efficiently, and when ECC value is close to zero, the efficiency of MIMO antenna is increased as the mutual coupling and signal interruptions between antenna elements are minimized.

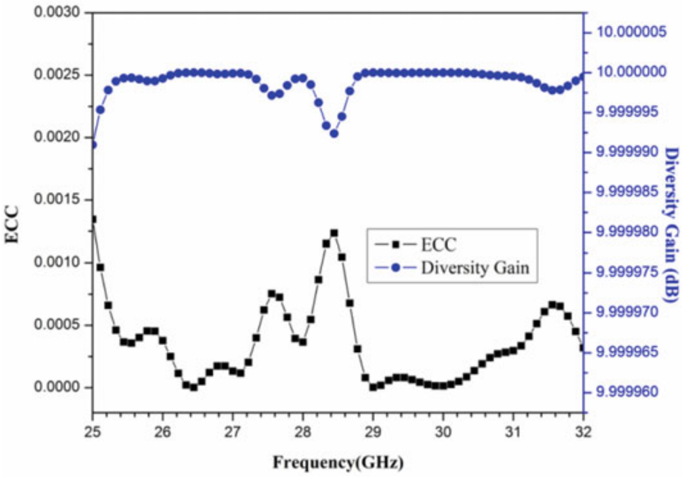


Fig. 12.9 Simulated ECC and DG of proposed antenna (ant-3)

ECC for proposed MIMO dumbbell shaped antenna shows a maximum value of 0.0012, which is near to zero which is indicative of minimum mutual coupling effect between antenna elements leading to higher data rates and diversity gain for the proposed antenna. The mutual coupling of effect between the ports of MIMO could be understood with the two scattering parameters S_{12} or S_{21}. The scattering parameter S_{12}/S_{21} provides isolation for the ports between MIMO antenna. The isolation value is less than −15 dB (cf Figs. 12.6 and 12.10) for the observed bandwidth. Radiation

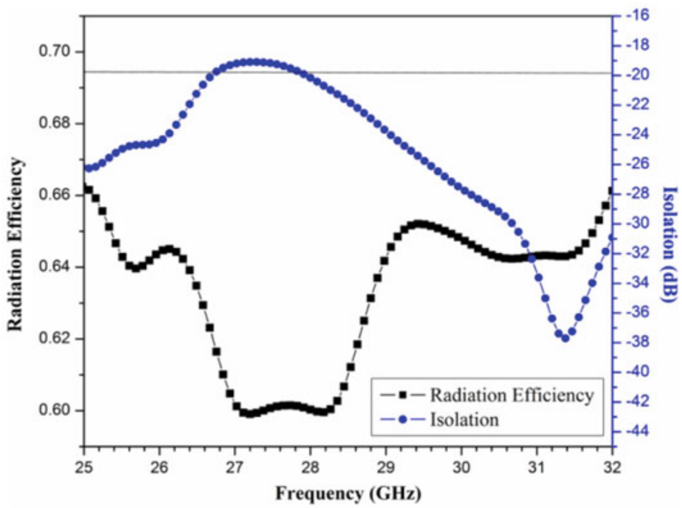


Fig. 12.10 Simulated isolation and radiation efficiency of proposed antenna (ant-3)

efficiency of around 70% and diversity gain response for proposed MIMO antenna range between 9.979 and 9.999 dB as presented in Fig. 12.9.

The simulated *E* plane and *H* plane for the far field radiation pattern of the Ant-3 as presented in Fig. 12.11a, b (at port 1 at 27.11 GHz and 29.54 GHz) and in Fig. 12.11c, d (at port 2 at 27.2 GHz and 29.6 GHz). The antenna for all resonating frequencies shows omnidirectional radiation pattern**.**

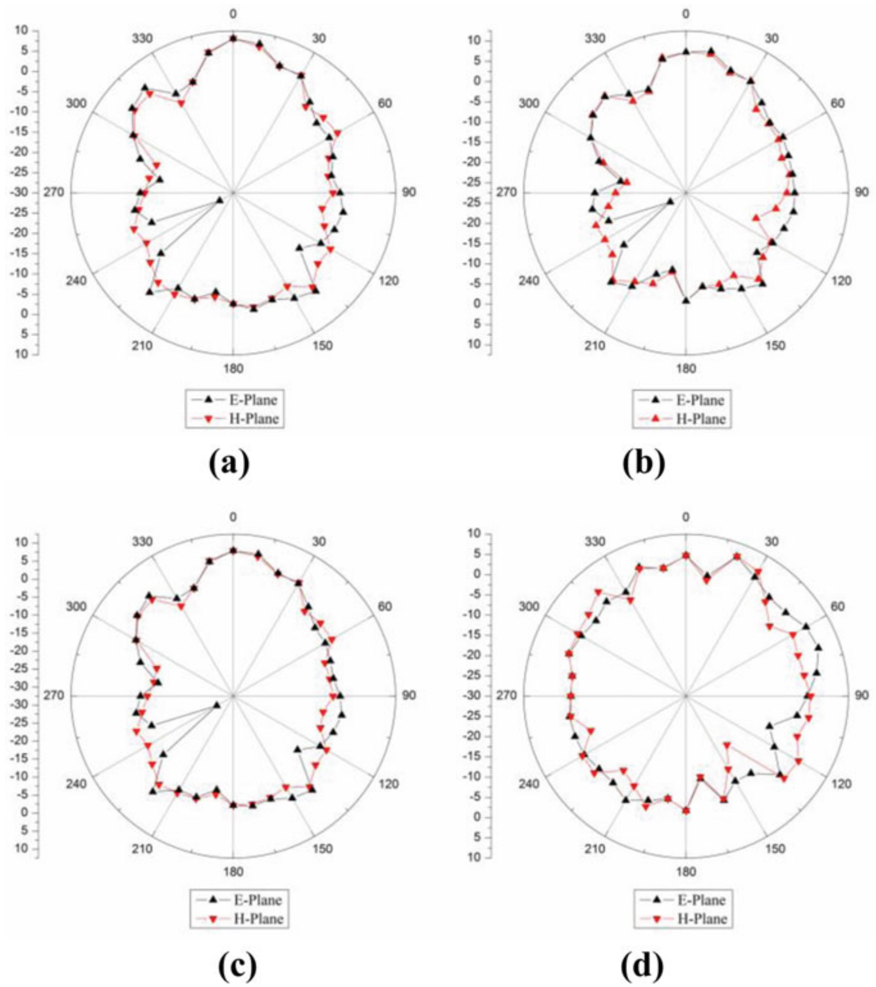


Fig. 12.11 Radiation pattern in *E* and *H* plane at port 1 at **a** 27.11 GHz **b** 29.54 GHz and at port 2 at **c** 27.2 GHz **d** 29.6 GHz

12.4 Conclusions

Dumbbell shaped MIMIO antenna (cf. Figs. 12.1, 12.2, and 12.3) is evolved through simulation of different stages and the characteristics are discussed in Sect. 12.3. Perusal of Fig. 12.4 depicts impedance bandwidth (22.7% at port 1 and 17.3% at port 2), resonant frequency (27.11 and 29.54 GHZ at port 1, 27.2 and 29.6 GHz at port 2). Figure 12.5 shows peak gain (8.2 dBi and 4.8 dBi at port 1, 7.9 dBi and 4.8 dBi at port 2) for MIMO antenna (Ant-3). Isolation (cf Fig. 12.6) less than −15 dB is observed at port 1 and 2 both. Simulated envelope correlation coefficient is 0.0012, and diversity gain between 9.997 and 9.999 is observed from the perusal of Fig. 12.9. However, the gain of the antenna varies and degrades as the frequency is shifted toward higher side. This is attributed due to ringing effect [10, 11]. The proposed MIMO patch antenna is suitable for K and Ka band applications and in particular for 4G/5G mobile applications and satellite communications.

References

1. Khan, M.S., Capobianco, A.D., Iftikhar, A., et al.: Ultra-compact dual polarized UWB MIMO antenna with meandered feeding lines. IET Microw. Antennas Propag. **11**, 997–1002 (2017)
2. Tang, X., Yao, Z., Li, Y., Zong, W., Liu, G., Shan, F.: A high performance UWB MIMO antenna with defected ground structure and U-shape branches. Int. J. RF Microw. Comput. Aided Eng e22270, 1–14 (2020)
3. Alnaiemy, Y., Elwi, T. A., Nagy, L.: Mutual coupling reduction in patch antenna array based on EBG structure for MIMO applications. Period. Polytech. Electr. Eng. Comput. Sci. 1–11 (2019)
4. Ibrahim, A., Machac., Shubair, R.: Compact UWB MIMO antenna with pattern diversity and band rejection characteristics. Microw. Opt. Technol. Lett. **59**, 1460–1464 (2017)
5. Nadeem, I., Choi, D.-Y.: Study on mutual coupling reduction technique for MIMO antennas. IEEE Access **7**, 563–586 (2019). https://doi.org/10.1109/ACCESS.2018.2885558
6. Khandelwal, M., Kanuajia, B. K., Santanu, D.: Bandwidth enhancement and cross-polarization suppression in ultra-wideband microstrip antenna with defected ground plane. Microw. Opt. Technol. Lett. **56**, 2141–2145 (2014)
7. Saif Al, H., Muhammad, U., et al.: Compact ultra-wide band MIMO antenna system for lower 5G bands. Wirel. Commun. Mobile Comput. **2018.** https://doi.org/10.1155/2018/2396873
8. Chandrasekhar, J.R., Venkateswara, N.R.: Compact MIMO antenna with WLAN band-notch characteristics for portable UWB systems. Prog. Electromagn. Res. C **88**, 1–12 (2018)
9. Malviya, L., Panigrahi, R., Kartikeyan, M.: MIMO antennas with diversity and mutual coupling reduction techniques, a review. Int. J. Microw. Wirel. Technol. **9**, 1763–1780 (2017)
10. Singh, V., Mishra, B., Singh, R.: Dual-wideband semi-circular patch antenna for Ku/K band applications. Microw Opt Technol Lett. **61**, 323–329 (2018)
11. Srivastava, K., Mishra, B., Patel, A.K., Singh, R.: Circularly polarized defected ground stub-matched triple-band microstrip antenna for C and X-band applications. Microw. Opt. Technol. Lett. pp 1–9 (2020). https://doi.org/10.1002/mop.32450

Chapter 13
An Improved Algorithm for Analyzing Students' Academic Performance

Anita Mahajan and Ilyas Khan

Abstract To keep track of educational information of students and keep them organized is a major task. This information is also useful in predicting the success rate and hence very important. Excellently working and efficient algorithms are used to do so. The main motive of this research paper is to propose an advanced version of C4.5 decision tree algorithm for assessing information easily and accurately. After carrying out various tests and comparing the accuracy of the existing algorithms with the advanced version, it was seen that this algorithm is more accurate while assessing test data and less accurate while assessing train data.

13.1 Introduction

The students' academic performance depends on various things like social, individual, mental, monetary and other different factors. There is a necessary need of prediction models including all these factors so that academic performance of the students can be predicted [1]. We can find out the students who are weak and have less academic performance by our prediction model. The students having low performance can be taken care of individually by mentors and teachers so that their performance can get improved later on. The performance measures like their CGPA, SGPA, test marks and other factors are included as variables to find out the effect of different courses or activities organized by educational organizations. The performance of students is the result of interactions among factors such as age, gender, capability of grasping, education support, motivation and enthusiasm of learning and achievement not just the simple cause and effect relationship [2, 3].

Our proposed research will be useful for taking decisions to improve academic performance of students as well as providing feedback to professors so that teaching learning methods can be improved. In addition to this, grouping of students as per

A. Mahajan (✉)
Acropolis Institute of Technology and Research, Indore, India

I. Khan
Acropolis Institute of Technology and Research, Indore, India

T. Senjyu et al. (eds.), *IOT with Smart Systems*, Smart Innovation, Systems and Technologies 251, https://doi.org/10.1007/978-981-16-3945-6_14

their interest and skills can be done using classification, so that they can be given a task/project of their interest which will be helpful in improving their performance.

Many research experts have been working on this topic from several years. In Spain, the most influencing factors of various extracurricular activities, be it related to study or sport, were noted by [4] J. A. Moriana, F. Alos, R. Alcala, M. J. Pino, J. Herruzo and R. Ruiz. The outcomes were diverse in forms, for example, the sport activities were both indoors and outdoors while the academic ones were both tutored and private. Data of 250 students from 12 different schools was taken into consideration. By the use of ANOVA, i.e., analysis of variance, it was observed that students taking part in activities outside of the school had better academic results [5].

The performance of the first and second semester's bachelor students of computer science course was predicted by [6] Fadhilah Ahmad, Azwa Abdul Aziz and Nur Hafieza Ismail, in 2013. They surveyed parameters, methods and tools, the three elements needed to make prediction on students' academic performances. The patterns were extracted using Naive Bayes classifier using the data mining tool, i.e., Weka. So, the abovementioned framework can have an application in system implementation and prediction of student's academic performance in higher learning institutions.

The main aim of [7] Umamaheswari. K. S. Niraimathi in 2013 was to prepare the students for interviews. The students were categorized in grade order in all their education studies. Different social and demographic variables like age, gender, name, lower-class grade, higher-class grade, degree proficiency and extra knowledge or skill, etc., were taken into consideration. An examination was conducted to find out the extent to which these factors help in categorizing the students in the order of ranks to proceed for the recruitment process [8]. Because of this process, the short listings of students reduced. Clustering, association rules, classification and outlier detection had been used for evaluation of the students' performance.

A student's future can be predicted by the application of data mining techniques as proposed by [9] Priyanka A. Patil, R. V. Mane in 2014. Data collected from students is used in sequential pattern mining algorithms, as it finds most frequent patterns. The results can be calculated based on the patterns that are more frequent, using the classification algorithms. Guidance can be provided to those students who are more prone to failure.

The aim of the paper proposed by [10] Jyoti Bansode in 2016 was to improve teaching–learning process. One of the major applications of EDM is to predict student's performance, and hence, decision tree can be easily used for prediction. Students who are found poor can be pre-warned and necessary actions can be taken by the administration to improve the performance of the student.

13.2 Proposed Work

We have collected data of undergraduate students of different streams of bachelor of engineering during the period (2015–2018). Initially, about 1000 students' records have been collected. Then data preprocessing is done, and duplicate entries have been removed. The objective of proposed research work is to suggest students for improvement in academics.

13.2.1 Proposed Algorithm

We have used the concept of the decision tree algorithm in our proposed algorithm. Any decision tree has a root node in the starting from which we take decisions and actions. Then as per the decision tree algorithm, further nodes are split from the first node recursively. In the resultant decision tree, a probable scenario of decision and its possible outcome is represented by each branch. There are many decision tree algorithms among which two are widely used: ID3 and C4.5, where C4.5 is an improved version of ID3 which is developed by Quinlan Ross [11]. C4.5 handles both continuous and categorical attributes. On the basis of the selected threshold, C4.5 does the splitting of attribute values into two partitions due to which the child nodes are decided. The values that are above the threshold value become one child, and the other remaining values become another child. Decision tree is made on the basis of decisions made as per attribute selection measure. In C4.5, gain ratio is used as an attribute selection measure for building a decision tree. Following are the three equations used for calculation of entropy, gain ratio and split info of the attributes:

$$\text{Entropy(Decision)} = \sum -p(I) \cdot \log 2p(I) \tag{13.1}$$

$$\text{Gain Ratio}(B) = \text{Gain}(B)/\text{Split Info}(B) \tag{13.2}$$

$$\text{Split Info}(B) = -\sum |Ej|/|E| \times \log 2|Ej|/|E| \tag{13.3}$$

On the basis of these values, attribute selection is done and decision tree is made. We have proposed an enhanced algorithm of C4.5 decision tree algorithm which overcomes few of the drawbacks of the decision tree algorithm C4.5. We have used mean, mode, median with fourfold cross-validation so that outliers and data impurity can be handled. We have first collected the data, and then we have done data preprocessing to eliminate the impurities in data. Then we have prepared a model using C4.5 and our proposed algorithm. The data is fed as input in the form of .csv file, and the model is applied on that data. As a result, we get an improvement in accuracy on test data but decrease in accuracy on train data.

13.3 Results

We have collected data through our Google form. We have applied data preprocessing for the collected data; then we have stored it as a .csv file for mining data-related work. The data preprocessing was a time taking process; an un-normalized database table was used for storing the original data with a unique row for student name, roll no., ID, semester, course, CGPA and grade, etc., for each course.

13.3.1 Target Audience

- The main point was to select which will be the target audience for this research work.
- So, after some analysis, we selected the engineering students of first and second year, respectively, for this purpose.

13.3.2 Attributes Selection

The second main point was to figure out the various parameters which may affect the performance of the students.

So, a list of attributes was selected, and the priority order was set for each parameter.

Following were the attributes which were being used to get the data from the students (Table 13.1).

13.3.2.1 Data Snapshot

See Fig. 13.1.

13.3.2.2 Practical Setup

Implementation Language Used: R

Language R is a reporting-based programming language mainly used for graphical representation and statistical analysis. Ross Ihaka and Robert Gentleman have developed R at the University of Auckland and being currently handled by R Development Core Team. It is a programming language under the GNU General Public License. This language supports various platforms like Linux, Windows and Mac.

The core of R says that it is an interpreted computer language giving functionality of branching, looping and modular programming using functions. R supports

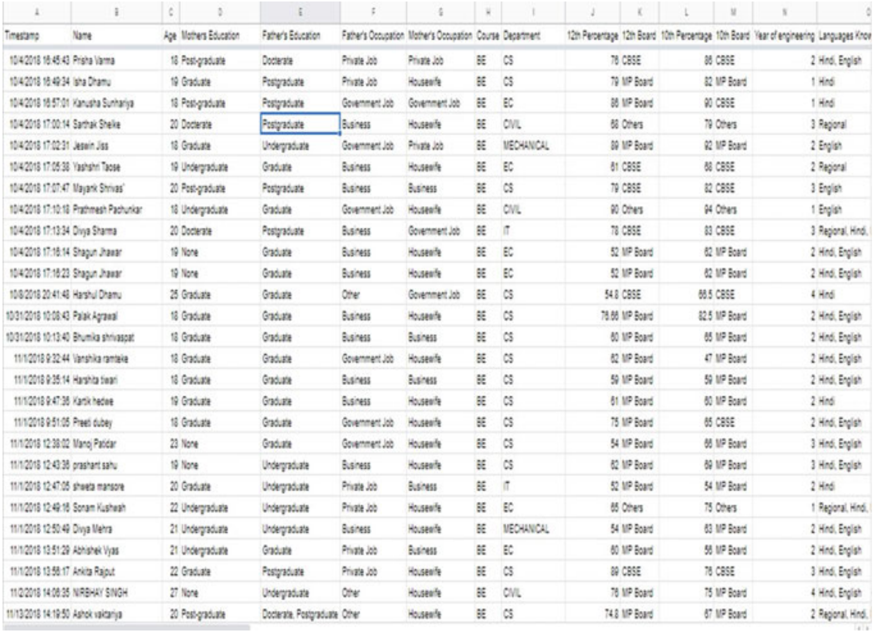
Table 13.1 Data table

Attribute name	Data type
Timestamp	Datetime
Name	Tinytext
Age	Double
Mothers education	Tinytext
Father's education	Tinytext
Father's occupation	Tinytext
Mother's occupation	Tinytext
Course	Tinytext
Department	Tinytext
12th percentage	Double
12th board	Tinytext
10th percentage	Double
10th board	Tinytext
Year of engineering	Double
Languages known	Tinytext
Coaching	Tinytext
Number of hours spent on study per day	Tinytext
Preparing for higher studies?	Tinytext
Projects under taken during semester	Tinytext
Time spent on social media in a day	Tinytext
Hostelite or local	Tinytext
Relationship status	Tinytext
Social media	Tinytext
Time of taking book before exam	Tinytext
Type of book	Tinytext
Topic clearance	Tinytext
First-year CGPA	Double
Gender	Tinytext

procedures and functions which can be written in different programming languages like FORTRAN, C, C++, .Net and Python.

R is distributed under a GNU style copy left as a free software, and it is an official part of the GNU project called GNU S.

We selected the C4.5 classification model and our proposed algorithm and compared their effectiveness on the basis of a prepared model for performance prediction of students. Based on the previous classes' results, the model was prepared by using performance as a class attribute. To finalize the best model, we have computed the average accuracy in predicting the performance for all students. For each level of



Timestamp	Name	Age	Mothers Education	Father's Education	Father's Occupation	Mother's Occupation	Course	Department	12th Percentage	12th Board	10th Percentage	10th Board	Year of engineering	Languages Know
10/4/2018 16:45:43	Prisha Varma	18	Post-graduate	Doctorate	Private Job	Private Job	BE	CS	78	CBSE	86	CBSE	2	Hindi, English
10/4/2018 16:49:34	Isha Dhamu	19	Graduate	Postgraduate	Private Job	Housewife	BE	CS	79	MP Board	82	MP Board	1	Hindi
10/4/2018 16:57:01	Kanusha Sunhariya	18	Post-graduate	Postgraduate	Government Job	Government Job	BE	EC	86	MP Board	90	CBSE	1	Hindi
10/4/2018 17:00:14	Sarthak Shelke	20	Doctorate	Postgraduate	Business	Housewife	BE	CIVIL	68	Others	79	Others	3	Regional
10/4/2018 17:02:31	Jeswin Jiss	18	Graduate	Undergraduate	Government Job	Private Job	BE	MECHANICAL	89	MP Board	92	MP Board	2	English
10/4/2018 17:05:38	Yashshri Taose	19	Undergraduate	Graduate	Business	Housewife	BE	EC	81	CBSE	68	CBSE	2	Regional
10/4/2018 17:07:47	Mayank Shrivas'	20	Post-graduate	Postgraduate	Business	Business	BE	CS	79	CBSE	82	CBSE	3	English
10/4/2018 17:10:18	Prathmesh Pachunkar	18	Undergraduate	Graduate	Government Job	Housewife	BE	CIVIL	90	Others	94	Others	1	English
10/4/2018 17:13:34	Divya Sharma	20	Doctorate	Postgraduate	Business	Government Job	BE	IT	78	CBSE	83	CBSE	3	Regional, Hindi,
10/4/2018 17:18:14	Shagun Jhawar	19	None	Graduate	Business	Housewife	BE	EC	52	MP Board	62	MP Board	2	Hindi, English
10/4/2018 17:18:23	Shagun Jhawar	19	None	Graduate	Business	Housewife	BE	EC	52	MP Board	62	MP Board	2	Hindi, English
10/8/2018 20:41:48	Harshul Dhamu	25	Graduate	Graduate	Other	Government Job	BE	CS	54.8	CBSE	66.5	CBSE	4	Hindi
10/31/2018 10:08:43	Palak Agrawal	18	Graduate	Graduate	Business	Housewife	BE	CS	78.66	MP Board	82.5	MP Board	2	Hindi, English
10/31/2018 10:13:40	Bhumika shrivaspat	18	Graduate	Graduate	Business	Business	BE	CS	80	MP Board	65	MP Board	2	Hindi, English
11/1/2018 9:32:44	Vanshika ramteke	18	Graduate	Graduate	Government Job	Housewife	BE	CS	82	MP Board	47	MP Board	2	Hindi, English
11/1/2018 9:35:14	Harshita tiwari	18	Graduate	Graduate	Business	Business	BE	CS	59	MP Board	59	MP Board	2	Hindi, English
11/1/2018 9:47:36	Kartik hedwe	19	Graduate	Graduate	Business	Housewife	BE	CS	61	MP Board	60	MP Board	2	Hindi
11/1/2018 9:51:05	Preeti dubey	18	Graduate	Graduate	Government Job	Housewife	BE	CS	75	MP Board	65	CBSE	2	Hindi, English
11/1/2018 12:38:02	Manoj Patidar	23	None	Graduate	Government Job	Housewife	BE	CS	54	MP Board	66	MP Board	3	Hindi, English
11/1/2018 12:43:38	prashant sahu	19	None	Undergraduate	Business	Housewife	BE	CS	62	MP Board	69	MP Board	3	Hindi, English
11/1/2018 12:47:05	shweta mansore	20	Graduate	Undergraduate	Private Job	Business	BE	IT	52	MP Board	54	MP Board	2	Hindi
11/1/2018 12:49:16	Sonam Kushwah	22	Undergraduate	Undergraduate	Private Job	Housewife	BE	EC	65	Others	75	Others	1	Regional, Hindi,
11/1/2018 12:50:49	Divya Mehra	21	Undergraduate	Undergraduate	Business	Housewife	BE	MECHANICAL	54	MP Board	63	MP Board	2	Hindi, English
11/1/2018 13:51:29	Abhishek Vyas	21	Undergraduate	Graduate	Private Job	Business	BE	EC	60	MP Board	56	MP Board	2	Hindi, English
11/1/2018 13:58:17	Ankita Rajput	22	Graduate	Postgraduate	Private Job	Housewife	BE	CS	89	CBSE	76	CBSE	3	Hindi, English
11/2/2018 14:06:35	NIRBHAY SINGH	27	None	Undergraduate	Other	Housewife	BE	CIVIL	76	MP Board	75	MP Board	4	Hindi, English
11/13/2018 14:19:50	Ashok vaktariya	20	Post-graduate	Doctorate, Postgraduate	Other	Housewife	BE	CS	74.8	MP Board	67	MP Board	2	Regional, Hindi,

Fig. 13.1 Snapshot of data

Table 13.2 Accuracy measures

Algorithm	Train data	Test data
C4.5	91.59	84.23
Proposed algorithm	75.63	92.93

performance, Table 13.2 gives the performance analysis of the proposed classification algorithms and C4.5 algorithm as well as the percentages of accuracy. The table shows that both the algorithms perform very good; however, higher accuracy was attained through our proposed model on test data and higher accuracy was attained through C4.5 for train data.

Based on the results, our proposed algorithm was applied for predicting the performances of students. Students are allowed to search for their performance prediction through the proposed small system. To find the prediction, we have built the classification models using training data and specific attributes and on the basis of the records of students. To construct the model, different attributes are used which includes the result, CGPA, grade, time spent per day, mother's education, father's education, family background, interests, etc. On the basis of the performance predicted, the recommendation of the best possibilities for each student is being provided by the system. The goal is to help the students to make them aware of improvements needed to enhance their academic attainment level.

Table 13.2 shows the accuracy measures for train data and test data.

13.4 Conclusion

Decision tree remains in the list of important algorithms used in data mining. There are different types of decision tree algorithms like ID3, CART and C4.5. Among this, most popular decision tree algorithm is C4.5 algorithm. The C4.5 algorithm gives a decision tree as an output. This tree is constructed by data preprocessing and attribute selection or parameters selection. After building the decision tree, pruning of the resultant tree is done. Following this, the algorithm then analyzes, estimates and completes the classified data mining. In this research work, data preprocessing and mean, mode, median are used for improving the classification accuracy for datasets including fourfold cross-validation. At first, the data is preprocessed, and then we analyzed the proposed approach algorithm with the C4.5 classifier and compared the performance of both. The results show that using mean, mode and median as threshold values can significantly improve the accuracy results. The accuracy increased for testing dataset but decreased for training dataset.

References

1. Mahajan, A., et al.: Data mining: elementary tools for analyzing data. Int. J. Emerg. Technol. Innov. Res. (JETIR) **5**(12) (2018)
2. Shahiria, A.M., Husaina, W., Rashida, N.A. (2015) A review on predicting student's performance using data mining techniques. In: The Third Information Systems International Conference. Procedia Comput. Sci. **72**, 414–422 (2015)
3. Kaur, P., Singh, M., Josan, G.S.: Classification and prediction based data mining algorithms to predict slow learners in education sector. In: 3rd International Conference on Recent Trends in Computing 2015 (ICRTC-2015) (2015)
4. Moriana J. A., Alos F., Alcala R., Pino M. J., Herruzo J., & Ruiz R., "Extra Curricular Activities and Academic Performance in Secondary Students", Electronic Journal of Research in Educational Psychology,Vol. 4, No. 1, 2006, pp. 35-46
5. García, E.P.I., Mora, P.M.: Model prediction of academic performance for first year students. In: 2011 10th Mexican International Conference on Artificial Intelligence, MICAI, pp. 169–174 (2011)
6. Aziz, A.A., Ismail, N.H., Ahmad, F. Mining student's academic performance. J. Theor. Appl. Inf. Technol. **53**(3), 2005–2013 (2013)
7. Umamaheswari & Niraimathi, K.S. (2013). " A Study on Student Data Analysis Using Data Mining Techniques" , India, International Journal of Advanced Research in Computer Science and Software Engineering, 3 (8)
8. Ramesh, V., Parkavi, P., Yasodha, P.: Performance analysis of data mining techniques for placement chance prediction. Int. J. Sci. Eng. Res. **2**(8) (2011). ISSN 2229-5518
9. Patil, P.A., Mane, R.V.: Student performance prediction by pattern mining technique. Int. J. Comput. Technol. Appl. **5**, 431–434 (2014). ISSN: 2229-6093
10. Bansode, J.: Student mining educational data to predict student's academic performance. Int. J. Recent Innov. Trends Comput. Commun. **4**(1) (2016). ISSN: 2321-8169
11. Han, J., Kamber, M.: Data Mining Concepts and Techniques, 2nd edn. Morgan Kaufmann Publishers (2008)

Chapter 14
Designing a Secure Integration of IoT Ecosystem and Intrusion Detection Using Machine Learning Approaches

Anshul Jain, Tanya Singh, and Satyendra K. Sharma

Abstract The Internet has become an indistinguishable part of human life, and the number of associated devices is also expanding exponentially. Specifically, the Internet of Things (IoT) devices has become a regular and indispensable part of human life which has now intruded in every corner, be it home, office or industry. However, security is still the primary concern, which is deemed to be solved by applying different machine learning techniques. Moreover, IoT and machine learning are a deadly combination and can help achieve many of these security issues. Machine learning techniques can help find real solutions to the most typical problems faced in the IoT ecosystem. It can also achieve the result with very high accuracy from the anonymous data provided to it. Besides, machine learning techniques have solid generalisability, so that they are additionally ready to distinguish obscure attacks. The authors here discuss the security issues in IoT ecosystems and how they can be solved using machine learning techniques. A modular architecture using machine learning and different intrusion detection techniques is also proposed in this paper.

14.1 Introduction

The end-to-end IoT network security is a critical issue for the fast development of IoT-enabled devices, business and their ecosystem. Some existing applications and solutions manage the security, but it has proved insufficient in current circumstances. Therefore, building an intelligent and reliable solution is the need of the hour. An intelligent solution that can educate itself parallel to the intruders building destructive problems could help cater to this challenging situation. All the studies and research

A. Jain (✉)
Amity Institute of Information Technology, Amity University, Noida, Uttar Pradesh, India

T. Singh
Amity School of Engineering and Technology, Amity University, Noida, Uttar Pradesh, India
e-mail: tsingh2@amity.edu

S. K. Sharma
Modern Institute of Technology & Research Center, Alwar, India

T. Senjyu et al. (eds.), *IOT with Smart Systems*, Smart Innovation,
Systems and Technologies 251, https://doi.org/10.1007/978-981-16-3945-6_15

done in academia and industry show that machine learning solutions can provide a viable solution to all these problems. Machine learning is a technique that can learn and unlearn different algorithms that can detect the network's unusual traffic. A substantial issue is to detect the interruptions without impacting the entire network and traffic. An intrusion detection system (IDS) is a methodology that detects the network's intrusion at different levels to identify the abnormality. Machine learning algorithms implemented with IDS can detect these elaborate scenarios and improve the IoT ecosystem security. However, it requires high-quality algorithms that can scan billions of devices and the data flowing through it and still consume low power, memory and processing.

Machine learning algorithms can be re-designed to use in IoT ecosystem to achieve the perfect result in decreasing recognition time. Machine learning for network intrusion detection is a precinct of progress and dynamic exploration. However, regardless of the vast amount of work done in IoT and machine learning, not many of these researches have received far-reaching results to achieve the desired goals. This paper proposed a modular architecture of the IoT ecosystem, which uses IDS in all the layers. This architecture further discusses and shows how machine learning techniques can be combined and used in an IoT ecosystem.

This paper's organisation is as follows. Section 14.2 discusses different security issues in the IoT ecosystem. It further explains different modules where IDS can be deployed in any IoT ecosystem. Section 14.3 explains different IDS methodologies that can be used within any network to detect intrusions. These IDS techniques can be combined with different machine learning techniques to make them robust and comprehensive. These machine learning techniques are discussed in Sect. 14.4, which are further used in the architecture proposed in this paper. Section 14.5 proposes a modular and decoupled architecture that shows how IDS can be deployed in different parts of the IoT ecosystem to enhance its security and utilise machine learning benefits. Finally, the conclusion is presented in Sect. 14.6.

14.2 Security Issues in IoT Ecosystem

The IoT has already captured a large part of our daily lives. It has built a massive ecosystem around us including homes, kitchens, cars, medical, sports, entertainment and our cities. IoT devices easily identify weak areas in real time such as water for crop, exercises for body, variations in a heartbeat and body temperature. IoT has already entered a stage where it carries more personal data than ever before; therefore, it becomes compulsory to ensure the security of the data processed in the IoT's ecosystem. However, this vast inclusion of IoT in human life and many positive impacts and benefits, it has another dark side. Weak cybersecurity implemented in IoT ecosystem [1] can cause a significant impact on the private data used and shared among different devices in IoT ecosystems. Figure 14.1 shows some of the security issues [2]. Security in IoT can be implemented in four types in different layers of its ecosystem [3]. These four types are defined below.

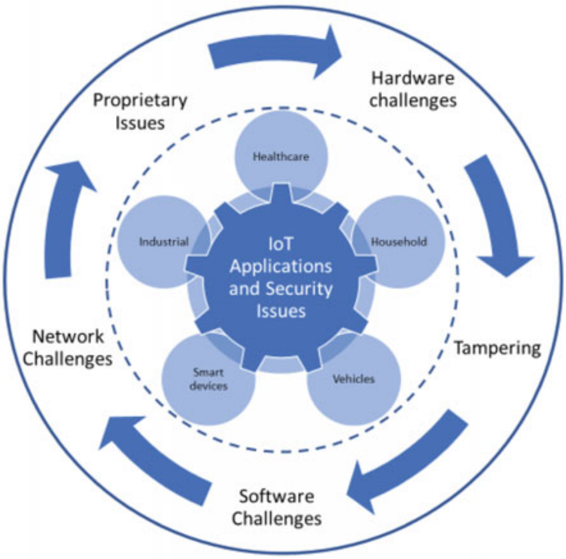


Fig. 14.1 Security issues in IoT

14.2.1 Edge Security

IoT is a truly distributed network that provides services to devices and their corresponding network around the globe. These devices can be sharing data for reporting, analysis or monitoring, which can be either used or misused if it gets into incorrect hands. Edge devices are used to connect such different networks for the transfer of data [4]. Therefore, securing such devices is much more important than any other network component because any impact on such devices can adversely affect the whole network. EDGE IDS is used to protect such a system.

14.2.2 Endpoint Security

Endpoint security solutions provide organisation-based protection against malware and botnets hazards. Additionally, the endpoint protection agreements guarantee that components joined to the network is accepted for interchange and validation. Endpoint security can also be related to hardware tampering [2], where a malicious attacker can tamper with hardware to gain access to the network or to push malicious data to any network.

14.2.3 Application Security

Application is managed and operated by the end-user at two points, in the front and backend. The backend application manages the device, whole application and data, whereas the frontend provides the user interface. Any vulnerability in the application at either end can be a significant risk to the whole ecosystem. It can expose the most critical data to hackers. Jain and Singh [2] discusses two of the most critical risks associated with application or software as challenges with embedded software and missing dynamic security patch updates.

14.2.4 Network Security

Securing networks and devices operating on the network layer are also as important as securing any other network components. This segment is very vulnerable to attacks like DoS/DDoS, and it can be secured by implementing solutions [5] such as encryption, VPN and dynamic routing protocols.

14.3 Intrusion Detection in IoT Ecosystem

An intrusion detection system (IDS) is a security mechanism that operates at all IoT ecosystem levels. We have discussed these levels in 14.2. An IDS designed for an IoT ecosystem should have the option of investigating packets of data and progressively reacting on the analysis, breaking down packets of data in different layers of IoT coordinating with different convention stacks and adjusting to various conditions applied in the IoT ecosystem. An IDS intended for IoT-based, critical situations should operate under authoritarian states designed to consume low power and memory, rapid reaction and high-volume data handling. Usual IDSs may not be entirely suitable for the IoT ecosystem as they are not designed for low power and CPU consumption. IoT security is a constant and vital issue; thus, an exceptional understanding of IoT frameworks security vulnerabilities and advancing relevant moderation approaches is needed. As shown in Fig. 14.5 of our proposed IoT and IDS ecosystem architecture diagram, different IDS mechanisms can be applied at different levels. The data collected at different levels can be processed on a cloud for detecting any attacks using machine learning techniques. Some of the IDS techniques are discussed in this section.

14.3.1 Signature-Based Detection

Signature-based recognition is achieved by comparing the identified data with the object database [6]. The signature-based detection technique is also called knowledge-based, where predefined variables/signatures are compared with their corresponding values stored in the database [7]. Article in [8] discusses how blockchain can enhance the security of an IoT ecosystem. The author in [9] has designed an IoT ecosystem for a 5G network where knowledge-based technique is used to detect any intrusion.

14.3.2 Anomaly-Based Detection

Technique anomaly-based detection technique thinks of current client experiments against client logs recently stacked. Anomaly is the best technique to detect new attacks by analysing abnormal activities. Therefore, it generates many false warnings because of erratic systems and customer behaviour [10]. The anomaly technique is a best-suited example of machine learning techniques because it learns from the logs and data generated. Additionally, it requires excellent collections of information to train the system for typical client profiles.

14.3.3 Hybrid Detection

The hybrid detection technique is the combination of the signature and anomaly intrusion discovery technique. The concept behind hybrid detection's use is to identify both the properties of executives with little recognised human affiliation and obscure attacks based on mark and identification systems for inconsistencies [11]. Hybrid is the best technique to be used in any ecosystem because it is more knowledgeable than the signature technique. It also enhances its knowledge in real time, and false-positive alerts generation is also less than the anomaly technique.

14.4 Machine Learning Techniques in IDS

IDS is used to detect abnormal data by scanning massive traffic flow in and out of any network. Unevenly distributed traffic flows through it with varied types, and it changes in real time. Therefore, reading real-time logs and detecting abnormalities is tedious, leading to many false-negative and false-positive alerts. This scenario makes the task manual and cumbersome for administrators managing it. To overcome these issues and to achieve IDS inclusion of machine learning techniques becomes

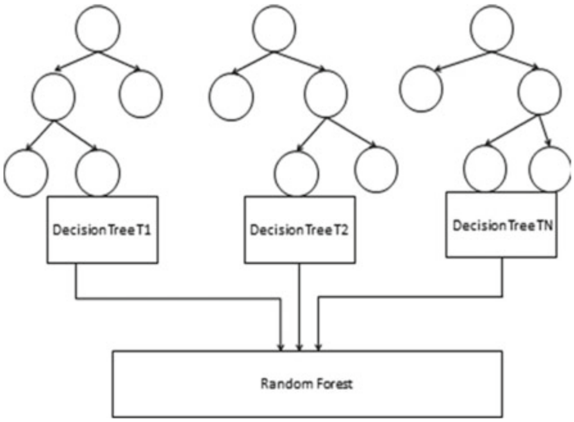


Fig. 14.2 Random forest and detection tree

a necessity. Section 14.3 discusses some of these techniques used to deploy in the improvement of IDS functioning. This section discusses three of the most common machine learning techniques and the algorithms used in IDS system development.

14.4.1 Random Forest and Detection Tree (DT)

Random forests consist of multiple single trees, each based on an arbitrary instance of data planning. They are generally more reliable than trees of one choice. There is no pruning of trees. Instead, trees are isolated from each other. Figure 14.2 shows each arbitrary wood tree's example and an arbitrary arrangement considered for separation at each centre. The two components make trees of assorted variety [12, 13]. The algorithm used in this technique is as follows:

1. Spot the best data set quality at the tree's foundation.
2. Divide the preparation into subsets. Subsets should be rendered in such a way that each subset includes data with a similar quality incentive.
3. Repeat stage 1 and stage 2 on each subset before finding a leaf node in each tree's pieces.

14.4.2 Bayesian Network

Bayesian networks are a kind of probabilistic graphic system that can be used to create data models and expert opinion. In addition, they can be used for a wide range of errors, including assumptions, detection, evaluation, mechanised intelligence, reasoning, time scheduling forecast and essential risk leadership [14]. Figure 14.3 shows the flow diagram of the Bayesian network algorithm.

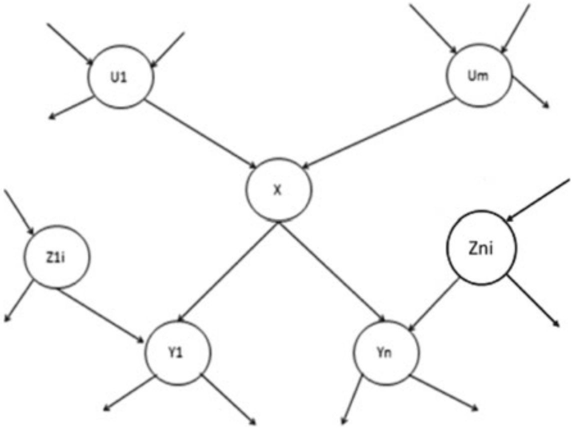


Fig. 14.3 Bayesian network algorithm

14.4.3 Neural Network

A neural network is an algorithm progression that attempts to perceive primary connections in many data through a procedure that emulates the way nature cerebrum works. Now, structures are alluding to neuronal constructs, either real or artificial. Figure 14.4 shows the architecture and functioning of the neural network.

Min et al. [15] propose a novel interruption identification framework called TR-IDS, exploiting both measurable highlights and highlights of payload. Word insertion and message convolutionary neural network <(text-CNN) is applied to extract from payloads sufficient information.

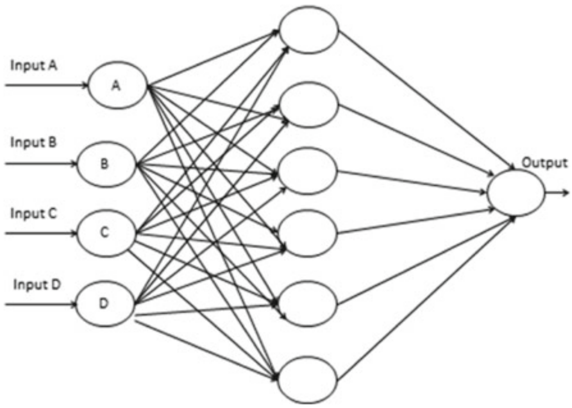


Fig. 14.4 Neural network algorithm

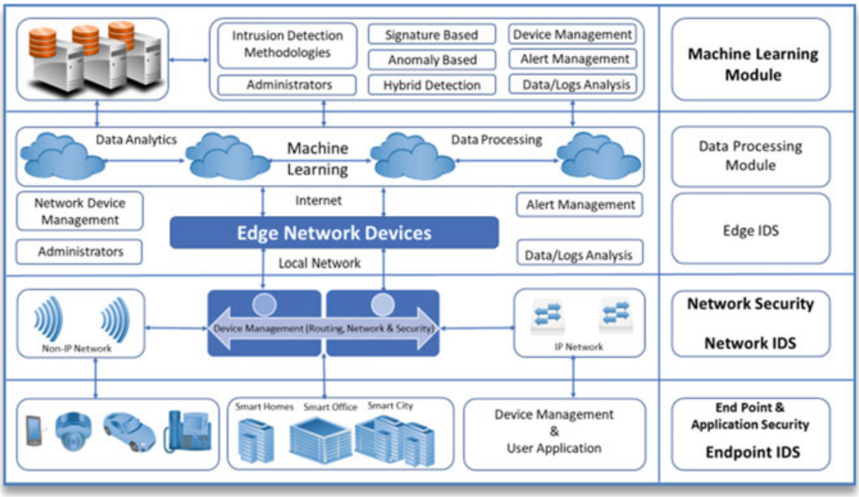


Fig. 14.5 Proposed IoT and IDS ecosystem using machine learning

14.5 Proposed Architecture of IoT and IDS Ecosystem Using Machine Learning Techniques

The architecture proposed in this paper is an extension and enhancement to our previous study done in [2, 8]. Figure 14.5 illustrates the proposed architecture. The modular architecture of the IoT ecosystem, as discussed in [8], introduces blockchain in a decoupled environment. This architecture is a joint venture of these two architectures that utilises both and introduces an architecture utilising machine learning techniques for intrusion detection.

This paper proposes an IoT and IDS architecture where IDS has been implemented on each layer to provide overall security to the IoT ecosystem. The bottom-most module has an endpoint IDS installed on all the individual devices and used to secure end devices and applications. Next, network IDS is installed above the endpoint and is used to secure local IP or non-IP networks. It is followed by edge IDS, which secures the devices connecting two different networks. This module also includes a cloud where all the IoT data is collected, stored and processed for further analysis. Edge is a unique module with an inbuilt machine learning module. After performing an analysis of stored data, the intrusion is detected based on the methodologies discussed in Sects. 14.3 and 14.4. Like our previous architectures, this is also a modular architecture that can be modified and used as per end-user requirements.

14.6 Conclusion

Machine learning algorithms may be feasible in constructing typical profiles and, subsequently, in IDS planning, depending on the inconsistent location approach. Therefore, machine learning and data mining algorithms assume significant components in planning IoT and IDS ecosystems. In this paper, we have briefly discussed the different security issues in the IoT ecosystem. We have also suggested the IDS techniques which can be used to mitigate these security issues. The paper further throws some light on the machine learning techniques that can be used to enhance the different IDS techniques discussed previously. A broad overview is provided in the proposed architecture where a modular approach is presented, combining all the three aspects of this paper, i.e. security issues, IDS techniques and machine learning methodologies. This architecture shows how IDS can be implemented in different IoT ecosystem modules combined with different machine learning techniques. Machine learning is the future of IoT and is a methodology that requires strict guidance. Unsupervised learning algorithms can "learn" and report abnormalities without a named data set. It can recognise new kinds of intrusions but is too prone to false-positive alarms. This paper is a small star in the IoT and machine learning universe. We are hopeful that machine learning is going to help IoT grow while ensuring its required security.

References

1. Zhao, K., Ge, L.: A survey on the Internet of Things security. In: 2013 Ninth International Conference on Computational Intelligence and Security, Leshan, 2013, pp. 663–667. https://doi.org/10.1109/CIS.2013.145
2. Jain, A., Singh, T.: Security challenges and solutions of IoT ecosystem. In: Tuba, M., Akashe, S., Joshi, A. (eds.) Information and Communication Technology for Sustainable Development. Advances in Intelligent Systems and Computing, vol. 933. Springer, Singapore (2020). https://doi.org/10.1007/978-981-13-7166-0_25
3. Jain, A., Singh, T., Sharma, S.K.: Threats paradigmin IoT ecosystem. In: 2018 7th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), pp. 1–7 (2018). https://doi.org/10.1109/ICRITO.2018.8748558
4. Netilion Blog: What-is-edge-device-iiot. https://netilion.endress.com/blog/what-is-edge-device-iiot/. Accessed 16 Jan 2021
5. Jain, A., Singh, T.: Securing communication in IoT ecosystem using cryptographic algorithms. Int. J. Eng. Adv. Technol. **9**(1), 7258–7268 (2019). https://doi.org/10.35940/ijeat.A1851.109119.
6. Canfora, G., Di Sorbo, A., Mercaldo, F., Visaggio, C.A.:Obfuscation techniques against signature-based detection: a case study. In: 2015 Mobile Systems Technologies Workshop (MST), Milan, 2015, pp. 21–26.https://doi.org/10.1109/MST.2015.8
7. Intrusion Detection Techniques, Methods & Best Practices. https://cybersecurity.att.com/blogs/security-essentials/intrusion-detection-techniques-methods-best-practices. Accessed 16 Jan 2021
8. Jain, A., Singh, T., Jain, N.: Framework for securing IoT ecosystem using blockchain: use cases suggesting theoretical architecture. In: Tuba, M., Akashe, S., Joshi, A. (eds.) ICT Systems and

Sustainability. Advances in Intelligent Systems and Computing, vol. 1270. Springer, Singapore (2021). https://doi.org/10.1007/978-981-15-8289-9_21
9. Jain, A., Singh, T., Sharma, S.K., Prajapati, V.: Implementing security in IoT ecosystem using 5G network slicing and pattern matched intrusion detection system: a simulation study. Interdiscipl. J. Inf. Knowl. Manag. **16**, 1–38 (2021). https://doi.org/10.28945/4675
10. Kruegel, C., Toth, T.: Using decision trees to improve signature-based intrusion detection. In: Vigna, G., Kruegel, C., Jonsson, E. (eds.) Recent Advances in Intrusion Detection. RAID 2003. Lecture Notes in Computer Science, vol. 2820. Springer, Berlin, Heidelberg (2003). https://doi.org/10.1007/978-3-540-45248-5_10
11. Zhang, J., Zulkernine, M.: A hybrid network intrusion detection technique using random forests. In: First International Conference on Availability, Reliability and Security (ARES'06), Vienna, Austria, 2006, pp. 8, 269. https://doi.org/10.1109/ARES.2006.7
12. Hasan, M., Islam, M.M., Zarif, M.I.I., Hashem, M.M.A.: Attack and anomaly detection in IoT sensors in IoT sites using machine learning approaches. Internet Things **7**, 100059 (2019). ISSN 2542-6605. https://doi.org/10.1016/j.iot.2019.100059
13. Singh, K., Malhotra, J.: IoT and cloud computing based automatic epileptic seizure detection using HOS features based random forest classification. J. Ambient Intell. Hum. Comput. (2019). https://doi.org/10.1007/s12652-019-01613-7
14. Harbi, Y., Aliouat, Z., Hammoudi, S.: Enhancement of IoT applications dependability using bayesian networks. In: Amine, A., Mouhoub, M., Ait Mohamed, O., Djebbar, B. (eds.) Computational Intelligence and Its Applications. CIIA 2018. IFIP Advances in Information and Communication Technology, vol. 522. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-89743-1_42
15. Min, E., Long, J., Liu, Q., Cui, J., Chen, W.: TR-IDS: anomaly-based intrusion detection through text-convolutional neural network and random forest. Secur. Commun. Netw. **2018**, Article ID 4943509, 9 pp. (2018). https://doi.org/10.1155/2018/4943509
16. Lin, J., Yu, W., Zhang, N., Yang, X., Zhang, H., Zhao, W.: A Survey on Internet of Things: architecture, enabling technologies, security and privacy, and applications. IEEE Internet Things J. **4**(5), 1125–1142 (2017). https://doi.org/10.1109/JIOT.2017.2683200

Chapter 15
Scaling of Rapid Tests During Pandemics Using Application: Conceptual Resolving of Irremediable COVID-19 Circumstances—PRATIDHI

Akshay A. Menon, Ananthu Vasudevan, Arun K. Nair, K. S. Krishna Das, and T. Anjali

Abstract The spread of COVID-19 contagion has to lead the world on a pause, because of its alarming rate. The only immediate solution was to practice social distancing and enforce lockdown. COVID-19 has declared an international emergency because it has been labeled pandemic, and different countries are still developing their vaccine. As far as now various countries with various governments have undertaken active plans and emergency measures to protect the public. The government and various health organizations are in charge of the survey. In fear of being affected, people tend to avoid hospital visits and circumvent the COVID-19 test. As of now, many Web applications are developed to avoid social interaction and community gathering. The COVID-19 test booking Web app provides the user to choose what, when, and where to take the COVID-19 test, it also helps the government health officials to maintain a safe database. This ensures the increase of COVID-19 test being taken and provides the individual proper safety and assurance.

15.1 Introduction

The current outbreak of COVID-19 has been the debate of the century. This outbreak has been reported in Nov 2019 in China, Wuhang. In the current scenario, consistent testing and isolation are the only way to prevent the escalation of COVID-19 furthermore. COVID-19 is a respiratory disease that severely affects the upper and lower tracts of the lungs. Also known as a "novel coronavirus," a new strain among the family of coronaviruses affecting more than 90.2 million persons and also resulting in 1.9 million as reported by World Health Organization (WHO) from the 223 countries around the world and reported SARS-CoV-2 as a pandemic virus [1, 2].

SARS-CoV-2 is a pathogenic virus that adversely affects the lungs. This virus can be passed on to another person through direct or indirect contact. When a virus

A. A. Menon (✉) · A. Vasudevan · A. K. Nair · K. S. Krishna Das · T. Anjali
Department of Computer Science and Engineering, Amrita School of Engineering, Amrita Vishwa Vidyapeetham, Amritapuri, India

T. Senjyu et al. (eds.), *IOT with Smart Systems*, Smart Innovation, Systems and Technologies 251, https://doi.org/10.1007/978-981-16-3945-6_16

passes on to another person through one's respiratory secretions, this is termed as direct contact [3, 4]. When a virus spreads onto a person through surfaces, it is termed as indirect contact. Since no vaccine has been developed for COVID-19, current precautions taken to prevent the widespread of COVID-19 are through isolation, social distancing, and regular sanitation [5]. The testing phase is a manual process where the individual sample is taking for testing in certified hospitals [6]. The nasopharyngeal swab is collected from the nasal cavity. This collection is done in hospitals where a huge amount of people gather for testing COVID-19. This exponentially increases the spread of COVID-19 dramatically. It is also scientifically proven that community gathering is the major reason for the widespread of COVID-19 [7]. Taking Kerala as an example, to date Kerala is one of the top states of India which has the highest number of active cases (64,447). Observing the increase in cases for COVID-19, the major cause of this rapid increase is a community gathering. Approximately 90% of the active cases are due to community gathering [8].

Social distancing is the best viable option to date since no vaccination has been developed which shows 100% succession. Some of the vaccines developed are COVISHIELD from UK, TOZINAMERAN from USA, Ad5-nCoV from China, etc.

Even though vaccines are developed still the virus will sustain in our environment as stated by WHO. This poses a huge chance for another pandemic to occur shortly. Even if one person has the virus, the characteristics of the virus may enable it to exponentially spread again.

Some people who have symptoms of COVID-19 may not approach for an antigen test or PCR test for COVID-19 [9, 10]. This is because of a simple cause, people show idleness to proceed for a COVID rapid test. Since these people do not show up, not all are quarantined. This causes the virus to spread even more dangerously [11]. This is one of the vital reasons for the widespread of COVID-19.

Another way implemented to reduce the spreading of the virus is by using mobile applications effectively [12, 13].

Papers mentioned in the related works are some of the conceptual papers we used to develop this idea [14, 15]. These papers are referred to understand the current existing technologies [16, 17]. Using those papers, we were able to find limitations in our proposed methods and could effectively tackle them by comparing them with previous developments in this particular area [18, 19]. This system uses a database for storing the user's data in a database by hosting the application in Amazon EC2 as a Software as a Service (SaaS) model [20]. To find nearby hospitals, we use Geo API in Google Maps [21]. We used tokens for patients affected with COVID, where tokens represent the availability of accommodation in a particular hospital which substantially decreases social gathering [22]. We also discuss how social gathering or community gathering effects or aids in the outspread of SARS-CoV-2 among the public. Since there has not been a previous method to this idea, this proposed method is derived from a library of existing methods for different COVID-related hindrances. The proposed method reduces social contact with one another, and also people will have no hesitation to apply for a rapid COVID test [23]. The proposed method provides full privacy of a person's health situation when undergone COVID health examination and can rapidly increase the scale of tests run each day [24, 25]. This will finally ensure the safety and privacy of patients who request treatment for COVID-19.

15.2 Related Works

Fahim Aslam has proposed the importance of social distancing, by analyzing how social distancing will limit the reproduction rate (R0) of the virus among communities, also the implementation of social distancing in previous pandemics and its effectiveness [1]. Mark J. Siedner and Guy Harling have developed a pretest–posttest result comparison by implementing social distancing statewide, also mathematical values of pre- and post-enactment of social distancing [2]. Noella Noronha and Alessia D'Elia surveyed the amount of COVID-19 applications developed and also how applications reduce the exponential propagation of the virus among the communities [3]. Patricia Biller Krauskopf has developed a COVID learning app where virtual sessions for caring patients are assisted by WHO itself [4]. Kannikar Intawong, Debra Olson, and Suwat Chariyalertsak developed an application that stores and interconnected databases between patients and hospitals. Implementing a self-screening to escalate health awareness among the public [5], Ravi Pratap Singh and Mohd Javaid represented IoT as a technological platform that provides automated and transparent treatment which substantially reduces the healthcare costs [6]. Abhuhammad S, Khabour OF, and Alzoubi KH proposed the method of contact-trace tracking, a method to determine the acceptability of COVID-19 contact-tracing technology, and a deep study of the ethical concerns for many people [7]. Ali Imran and Iryna Posokhova with their associates developed an AI-powered screening solution to scale up the amount of COVID tests, a cough-based diagnosis of respiratory disease testing through a smartphone application [8]. Richard I. Horowitz and Phyllis R. Freeman proposed a hypothesis for the controlled spreading of asymptomatic or minimally symptomatic infection with COVID-19 among a large number of individuals which increases global morbidity and mortality [9]. Anna Tsutsui and Yuko Ohno use Google Distance Matrix API to track healthcare access by implementing SAS and Web API's [10]. Changrong Xu and Feilun Chen's research on W3C Geolocation API tracking of client's position is based on client-side JavaScripts analysis which connects to a relational database containing user's location information [12]. Vidya K. Bhise and Ajit S. Mali published documentation in the fourth International Conference on Computing, Communications and Network Technologies (ICCCNT) comprising of the importance of storing user's data and essential database management using fewer resources on Amazon EC2 which significantly costs less for the developers of Web applications [13]. Parminder Singh, Pooja Gupta, Kiran Jyoti, and Anand Nayyar studied the significance of auto-scaling application set as the future for Web application in cloud platform [14]. Daniel R. Lanning, Gregory K. Harrell, and Jin Wang, in their study on Dijkstra's algorithm implemented in Google Map to find the shortest path between the source and destination

[15]. Heeket Mehta, Pratik Kanani, and Priya Lande worked on algorithms used in Google Maps to find the shortest distance between source and destination [16]. A Kumar, D Sathyapalan, A Ramachandran, K Subhash, L Biswas, K V Beena worked on how antibodies are formed with continuous contact with COVID-affected patients [17]. Unnikrishnan Payyappallimana, Kishor Patwardhan, Prasad Mangalath, Christian S. Kessler, Rama Jayasundar, Anupama Kizhakkeveettil, Antonio Morandi, and Ram Manohar P showed us how important is the usage of ayurvedic medicines for diagnosing a patient affected with a disease [18]. Bilha Baby, Aswathy R. Devan, Bhagyalakshmi Nair, and Lekshmi R. Nath proposed the different types of diagnostic and also about the types of treatment strategy for COVID-induced patients [19]. Shiva Prasad Kollur worked on how COVID-19 spread through community gathering and social contacts [20]. Priyadarsini, S Lakshmi, Dr. Suresh M, Huisingh, Donald talk about the necessary steps and effective methods learnt from previous pandemics to tackle this particular disease residing in the society [21]. Menon, J. C., Rakesh, P. S., John, D., Thachathodiyl, R. and Banerjee discuss about how Kerala has tackled the COVID pandemic and crucial measurements taken to avoid further outspread of the virus [22]. Sharma, P., Nair, V., and Jyotishi, A show the implementation of a grocery application which could be a generalized information regarding usage of applications and its importance in such tough times [23]. Arun, S. S. and Iyer, G. N. proposed the importance of data managements using machine learning, privacy of a user's data [24]. Nayak, A. R., Ayyar, S. C., Aiswarya, O., Mahitha, C. H., and Mohankumar, N. discuss about the importance of *Surveillance* to increase data collections and related security issues [25].

15.3 Proposed System

Users can book their medical test, and they will be allocated to the respective laboratory center based on slot availability and distance priority. Users will have flexibility in choosing the date and slot timing for their medical test. When the test is taken by the patient, the appointment slot will be empty, and results would be sent via e-mail or SMS. Figure 15.1 represents the basic principle behind the working of the application online.

15.3.1 Application Features

- In the built-notification system.
- Compatible with any network.
- Effective data isolation regarding the time slot of an appointment.
- Application size is lower a native app requires less storage.

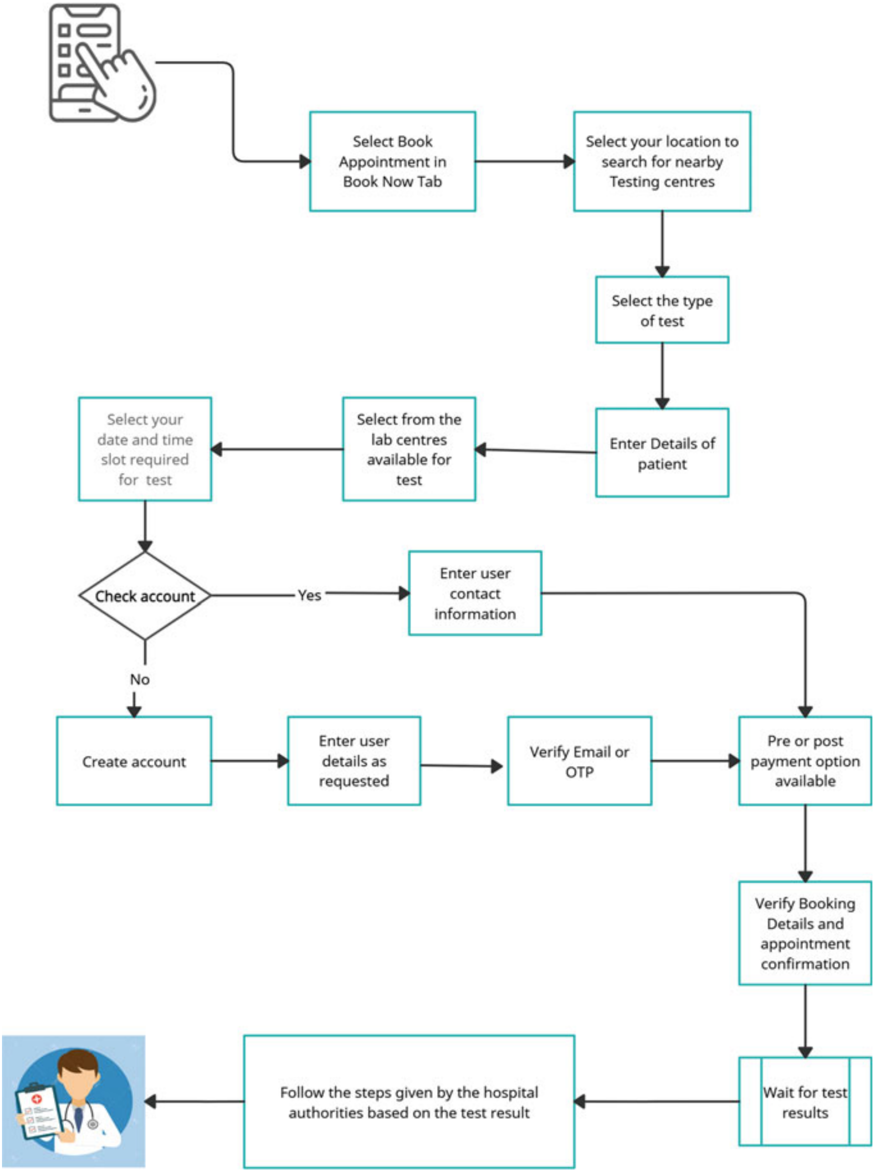


Fig. 15.1 Block diagram working

15.3.2 Application Configuration

The Web application is hosted in Amazon Web Services (AWS) and configured as an EC2 instance. Amazon AWS is a cloud platform to host and manage Web applications using fewer resources and high cost efficiency. For a better user experience, we offered 8 GB RAM and 80 GB SSD storage. This enables the application to run faster and smoothly in inadequate network data. The framework used here is PHP (Hypertext Preprocessor) as it is cost-efficient, easy to perform, better integration, and compatibility. Since PHP is highly flexible it was easy while building the project. For the storage of a user's information, we build a relational database in Amazon using MySQL. We connected it with the PHP framework. We created a database separately for tables and scripts. A user module and a booking module were implemented but configured separately. The booking module has submodules named hospitals, location, and customer/patient login module. To find the nearby testing center for the users, we implemented two methods: distance priority algorithm and Geo API using Google Maps. These both showed high efficiency during the testing phase. To connect the application server online, we used the Web view module in Android Studio by embedding a public URL.

15.3.3 Detection of Test Center-Algorithms

Users apply for a new appointment where to find the nearest testing center we use the simplest algorithm known as Dijkstra's algorithm which is implemented by Google Maps. Using the Geo API and with the help of this particular algorithm, we were able to find the nearest testing center concerning the positioning of the user. By Geo-mapping, the admin holder which is the respective testing center chosen by the user can effectively find the shortest path to the user's location. This is because Dijkstra's is a greedy algorithm, so the shortest path is given the highest priority. The algorithm has a source node and a destination node. The laboratory center location is the source node and the user's location is the destination node. Dijkstra's algorithm has a time complexity of O(|V2V2|) where V is a node that represents A house, city, or an intersection in real life. But this complexity can be reduced because we use a minimum-priority queue then the time complexity reduces to $O(|E| + |V|\log|V|)$. The above algorithm shows the working of Dijkstra's algorithm. Since Dijkstra's algorithm may fail to a huge amount of data, especially dealing with more number nodes when implemented in Google Maps, another algorithm is used to support Dijkstra's algorithm for better efficiency. A* Algorithm is also implemented as a cutting edge algorithm as it also finds the shortest distance between the source and destination nodes. A* is similar to Dijkstra's algorithm, which uses a heuristic function to navigate a better and more efficient path. A* algorithm assigns higher priority to the nodes which are supposed to be better (checks for parameters like time requirement, distance, and other such parameters) than the others, while Dijkstra explores all the nodes. Therefore, it is meant to be faster than Dijkstra's algorithm, even if the memory requirement and operations per node are more since it explores a lot fewer nodes and the gain is good in any case [16]. These algorithms were used to Geo map the nearest testing center.

15.3.4 Slot Allocation

The slot allocation algorithm implemented is as follows:

The slots availability is calculated as ($n >\ = 0$) where:

t = *total working hours.*
p = *maximum time for an individual test (15 min).*
n = *number of time slots available in a hospital (t/p).*

After the payment is completed, the allotted slot will be confirmed, thereby the slots will be reduced ($n - 1$). If a patient cancels the appointment, the slots would be incremented ($n + 1$).

A. *Procedure for Booking*

According to the user's choice, COVID-19 antigen test or COVID-19 PCR test can be selected, and these choices can be made as per the availability of hospitals or clinics nearby. Select the available laboratory centers and desired date and time slot for the test. If the desired slot being occupied, the person can put a reserve on it, in which case, if the latter gets canceled the reservation gets confirmed. In case if the reservation is not approved, the person should have to book a new time slot. The patient details must be entered, and the records and the present condition must be specified. Booking can be confirmed only if the person holds an account, and in such cases, the contact details and payment method must be entered. Otherwise, a new user account must be created with a proper e-mail id and verified phone number. The verification can be done either by OTP or verification code. The payment methods include offline and online modes. In any case, the verified appointment needs to have a cancelation, and the payment will be refunded accordingly. After a safe and secure testing procedure, the results will be mailed and massaged. After and before the test, there is an option to communicate with the concerned authorities so that recurring concerns can be rectified to an extent. If the results are positive, adequate measures will be provided by the hospital and more details have to be acquired. Home quarantine or hospital isolation is determined by the hospital by taking into consideration the patient's condition and the presence of family members. The patient's details, results, and condition are updated and displayed on the dashboard.

B. *Benefits of Admin Panel*

Control over the admin panel of the app is only given to authorized personals in the testing center as they could manage the appointment details and results of the test conducted on a patient. This ensures the privacy of the patient is affected by a disease. The information regarding the patient's test results would be handed over to the government health officials so they could get the number of persons affected with COVID-19, and they could take effective actions to prevent further outbreak of the virus, also this would create awareness among the public. The admin panel also could decline or accept the appointments concerning the slot availability.

III. *User Privileges*

Flexibility in choosing dates and time slots for the required medical tests is for a better user experience. Users can modify the timings for the test if encountered with any personal emergency. Users have better privacy as the test center will have the admin panel for the application. For better security and safety of the user and OTP verification via SMS or e-mail, criteria are implemented. Users will get the nearest testing center for user's easiness. A provision to see the results is embedded inside the application. Offline and online payment is available and if the test got declined the payment will be refunded to the user's account.

15.4 Result

Overall, many medical control measures helped us to contain the virus preventing it from further spreading it into the public. From the control measure, it consisted of washing hands, social distancing, etc., as some of the preventive measures. These rules were implemented after seeing the sudden spikes in the statistics of affected cases. Figure 15.2 shows the number of cases before and after implementing medical control measures. Figure 15.3 compares the results between pre- and post-implementation of lockdown.

As you can see, the variation in the COVID-19-affected cases before and after lockdown shows a huge difference. But even though lockdown is implemented, community gathering is not fully eradicated, even people do not apply for testing their health because of their hesitations and leak of privacy. This is where the application plays a major role in decreasing the trend of new cases furthermore. Even now community gathering is seen in hospitals and many medical centers because the ingathering of the active cases is important for the government so that they could take necessary measures. This application tackles this issue too as the information regarding the active cases can be retrieved from a database which reduces human effort by a large quantity. This application aids the effort taken by a person to apply for a medical test.

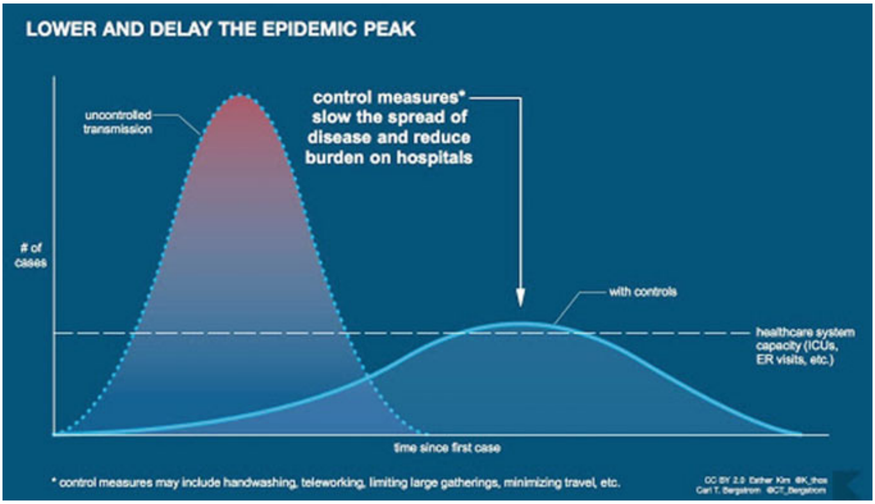


Fig. 15.2 COVID cases before and after implementing control measures

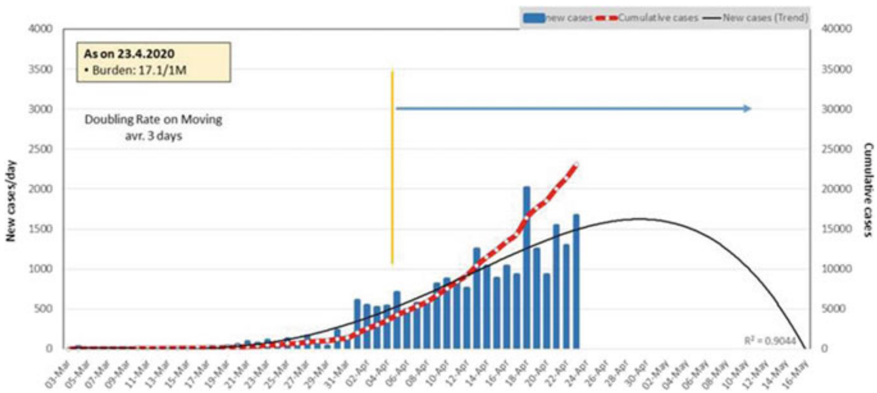


Fig. 15.3 Trend of COVID-19 cases before and after lockdown

15.5 Conclusion

This application provides the person a contactless social interaction and increase in the scale of the COVID-19 test. The person has the option to select the type of test and a suitable laboratory which ensures a safe and serene testing facility by avoiding the hospital visit. The person will also be provided with the option to select the desired slot according to the preferable date and time. Offline and online payment methods are provided for convenience. Government health officials collect data and maintain a database by direct human labor which will be solved by maintaining a database containing patient and hospital information. The privacy of the patient will be ensured. The patient details and test results are updated on the dashboard. The proper communication between the patient and hospital authorities will help in reducing their concerns. The main aim of this application is staying together by not standing together.

References

1. Aslam, F.: COVID-19 and Importance of Social Distancing (2020)
2. Siedner, M.J., Harling, G., Reynolds, Z., Gilbert, R.F., Haneuse, S., Venkataramani, A.S., Tsai, A.C.: Social distancing to slow the US COVID-19 epidemic: longitudinal pretest–posttest comparison group study. PLoS Med. **17**(8), e1003244 (2020)
3. Noronha, N., D'Elia, A., Coletta, G., Wagner, N., Archer, N., Navarro, T., Lokker, C.: Mobile Applications for COVID-19: A Scoping Review (2020)
4. Krauskopf, P.B.: World Health Organization Academy: COVID-19 learning and WHO info mobile apps. J. Nurse Pract. **16**(9), 704–705 (2020)
5. Intawong, K., Olson, D., Chariyalertsak, S.: Application technology to fight the COVID-19 pandemic: lessons learned in Thailand. Biochem. Biophys. Res. Commun. (2020)
6. Singh, R.P., Javaid, M., Haleem, A., Suman, R.: Internet of things (IoT) applications to fight against COVID-19 pandemic. Diabetes Metab. Syndr.: Clin. Res. Rev. (2020)
7. Abuhammad, S., Khabour, O.F., Alzoubi, K.H.: COVID-19 contact-tracing technology: acceptability and ethical issues of use. Patient Prefer. Adherence **14**, 1639 (2020)
8. Imran, A., Posokhova, I., Qureshi, H.N., Masood, U., Riaz, S., Ali, K., John, C.N., Nabeel, M.: AI4COVID-19: AI enabled preliminary diagnosis for COVID-19 from cough samples via an app. arXiv preprint arXiv:2004.01275 (2020)
9. Horowitz, R.I., Freeman, P.R.: Three novel prevention, diagnostic and treatment options for COVID-19 urgently necessitating controlled randomized trials. Med. Hypotheses 109851 (2020)
10. Tsutsui, A., Taniyama, Y., Ohno, Y.: Driving to childhood cancer hub hospitals: a study on hospital accessibility in Japan. Asian Pac. J. Cancer Prev.: APJCP **21**(6), 1725 (2020)
11. Nayak, A.R., Ayyar, S.C.S., Aiswarya, O., Mahitha, C. H., Mohankumar, N.: Security surveillance bot for remote observation during pandemics. In: 2020 5th International Conference on Communication and Electronics Systems (ICCES), pp. 635–640. IEEE (2020)
12. Xu, C., Chen, F.: Research on map-positioning technology based on W3C Geolocation API. In: 2012 2nd International Conference on Consumer Electronics, Communications and Networks (CECNet), pp. 3451–3454. IEEE (2012)
13. Bhise, V.K., Mali, A.S.: Cloud resource provisioning for Amazon EC2. In: 2013 Fourth International Conference on Computing, Communications and Networking Technologies (ICCCNT), pp. 1–7. IEEE (2013)
14. Singh, P., Gupta, P., Jyoti, K., Nayyar, A.: Research on auto-scaling of web applications in cloud: survey, trends and future directions. Scalable Comput.: Pract. Exp. **20**(2), 399–432 (2019)
15. Lanning, D.R., Harrell, G.K., Wang, J.: Dijkstra's algorithm and Google maps. In: Proceedings of the 2014 ACM Southeast Regional Conference, pp. 1–3 (2014)
16. Mehta, H., Kanani, P., Lande, P.: Google maps. Int. J. Comput. Appl. (2019)
17. Int. J. Comput. Appl. **178**(8) (2019)
18. Kumar, A., Sathyapalan, D., Ramachandran, A., Subhash, K., Biswas, L., Beena, K.V.: SARS-CoV-2 antibodies in healthcare workers in a large university hospital, Kerala, India. Clin. Microbiol. Infect. (2020)
19. Payyappallimana, U., Patwardhan, K., Mangalath, P., Kessler, C.S., Jayasundar, R., Kizhakkeveettil, A., Morandi, A., Puthiyedath, R.: The COVID-19 pandemic and the relevance of ayurveda's whole systems approach to health and disease management. J. Altern. Complement. Med. **26**(12), 1089–1092 (2020)

20. Baby, B., Devan, A.R., Nair, B., Nath, L.R.: The impetus of COVID-19 in multiple organ affliction apart from respiratory infection: pathogenesis, diagnostic measures and current treatment strategy. Infect. Disord. Drug Targets (2020)
21. Vinusha, H.M., Shivamallu, C., Prasad, S.K., Begum, M., Gopinath, S.M.: Int. J. Res. Pharm. Sci.
22. Priyadarsini, S.L., Suresh, M., Huisingh, D.: What can we learn from previous pandemics to reduce the frequency of emerging infectious diseases like COVID-19? Global Transit. **2**, 202–220 (2020)
23. Menon, J.C., Rakesh, P.S., John, D., Thachathodiyl, R., Banerjee, A.: What was right about Kerala's response to the COVID-19 pandemic? BMJ Global Health **5**(7), e003212 (2020)
24. Sharma, P., Nair, V., Jyotishi, A.: Patterns of online grocery shopping in India: an empirical study. In: Proceedings of the 2014 International Conference on Interdisciplinary Advances in Applied Computing, pp. 1–5 (2014)
25. Arun, S.S., Iyer, G.N.: On the analysis of COVID19-novel corona viral disease pandemic spread data using machine learning techniques. In: 2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS), pp. 1222–1227. IEEE (2020)

Chapter 16
Binary Duck Travel Optimization Algorithm for Feature Selection in Breast Cancer Dataset Problem

Krishnaveni Arumugam, Shankar Ramasamy, and Duraisamy Subramani

Abstract To predict if a tumor is malignant or benign is the challenge to all the researchers achieving efficient prediction of women in breast tumor. Unrestrained cell expansion in hankies of the breast is called breast cancer disease. Successful detection of breast cancer in earlier stage is required to secure women against high mortality. To select the best features among breast cancer classification of benign/malignant from the input mammogram images originated from Wisconsin dataset is satisfying with the new algorithm binary duck travel optimization algorithm (bDTO). Sigmoid activation function (logistic regression) is used for binary classification of proposed method. Food forage activation of ducks position is updated by sigmoid with maximum likelihood function (SMLE) of entire duck flock. Classifying the given mammogram is cancer or non-cancerous is based on the optimal feature selection by bDTO through (SMLE). Feature extraction of mammogram Corpus is done with the aid of sigmoid activation function (SAF) classifier. (bDTO-SMLE-SAF) is an intrinsic procedure to eliminate irrelevant scope and select the optimal highlights by using Wisconsin families normal nucleoli, single epithelial cell size, bare nuclei, uniformity of cell size, uniformity of cell shape, bland chromatin, mitoses, marginal adhesion, clump thickness features that are evaluated by the quality measures exactness, compassion, specificity, accuracy, evoke, and F-value clearly shows that bDTO classifier has the maximum accuracy 92% while compared with the DTO and SAF classifiers. The efficiency of an algorithm is proved by the promising results for selecting the best feature of malignancy classification through bDTO algorithm.

16.1 Introduction

Bio-inspired computing (BIC) is considered as an interdisciplinary field that connects computer sciences with natural sciences. It investigates models and computational techniques inspired by nature. The ultimate aim is to develop new optimization techniques all the way through natural behavior to solve laborious problems. Stochastic

K. Arumugam (✉) · S. Ramasamy · D. Subramani
Chikkanna Government Arts College, Bharathiar University, Coimbatore, India

T. Senjyu et al. (eds.), *IOT with Smart Systems*, Smart Innovation, Systems and Technologies 251, https://doi.org/10.1007/978-981-16-3945-6_17

method follows the trial and error method and classified as heuristic and meta-heuristic. Heuristic means guidelines/procedures. Metaheuristic consists of simple procedures and applied in a variety of problems. A metaheuristic strategy, then again, makes basically no earlier supposition about the issue, can coordinate a few heuristics inside, and is normally depicted as far as a set (usually known as a populace) of candidate solutions for the issue. A metaheuristic is officially characterized as an iterative generation measure which generates best solutions by using trial and error method of simple heuristics. Independent problems are solved by metaheuristic algorithms in a single objective/multiobjective, maximization/minimization, quantitative/systematic, static/dynamic, deterministic/stochastic, local/global, with memory/memory less, greedy/iterative, parallel/neutral, and naturalized/hybridized. Genetic algorithm is a most popular algorithm in metaheuristic which consists of mutation and crossover methods to create genomics.

Recently used metaheuristic algorithms are emperor penguins colony (EPC) and seagull optimization algorithm (SOA). Apart from that many new algorithms are introduced by various researchers are duck travel algorithm (DTO) in 2020, bear smell search algorithm (BSA) in 2020, black widow optimization (BWO) in 2020, group search optimizer (GSO) in 2020, red deer algorithm (RDA) in 2020, marine predators algorithm (MPA) in 2020, kernel partial least square (KPLS) algorithm in 2020, butterfly optimization algorithm (BOA) in 2019, Harris hawks optimization in 2019, artificial feeding birds (AFB), water wave optimization (WWO) in 2019, donkey and smuggler optimization (DSO) algorithm in 2019, and crow search algorithm (CSA) in 2018. Some swam-based optimizers are ant lion optimizer (ALO) 2015, Harris hawks optimizer (HHO) 2019, sail fish optimizer (SFO) 2019, and multiverse optimizer (MFO) 2016. Important metaheuristic algorithms which are aid to detect the breast tumor are ant and bee colony optimization (ACO & BCO), firefly (F) 2008, harmony search (HS) 2001, cuckoo search algorithm (CSA) 2009, bat (BA) 2010, krill herd optimization algorithm (KHOA) 2012, flower pollination algorithm (FPA) 2012, moth flame algorithm (MFA) 2015, dragonfly calculation (DC) 2015, glow worm algorithm (GWA) 2009, monkey search algorithm (MSA) 2007, elephant search algorithm (ESA) 2015, butterfly optimization algorithm (BOA) 2019, runner root algorithm (RRA) 2015, laying chicken algorithm (LSA) 2017, and killer whale algorithm (KWA) 2017.

16.2 Related Work

More or less than one in thirty nine ladies (3%) face death issues from breast cancer and roughly (13%) will be identified with infiltrating breast cancer that is the majority (81%) breast cancers are invasive. The glands or ducts broke the uncontrolled cells from nearby breast hankie. Metaheuristics are iterative search method for optimization problem solving by using two phases like exploration and exploitation to solve multidisciplinary and also actual time-independent problems. Many new researchers have introduced novel metaheuristic algorithm to solve problem easily and quickly.

A number of works have in recent times been measured applying bio-inspired optimization techniques to picture related optimization procedures such as segmentation and classification considered as nonlinear and expensive. Based on the input mammogram images, the objective of whether the breast having the tumor or not is classified according to the specified constraints. Primary objective of image classification is calculating the constraints using metaheuristic algorithm to reduce high computational cost. Hierarchy is the base term of software model for defining "usability" of resources. For dimensionality reduction, feature selection and feature extraction methods are required and also suggested by all the researchers in the last decades. Duck travel optimization algorithm is a new metaheuristic optimization algorithm based on the behavior of duck's imprinting and foraging activity. Endurance of all animals does hunting and preventing activity for their flock from predators. In this research work, feature selection is done by using binary duck travel algorithm (bDTO) to reduce the dimensionality of breast cancer dataset to classification.

The suitable method of low dimension for feature extraction is principal component analysis. But LDA handles easily if randomly generated test cases and uneven frequencies of class variance. Feature contribution is combined by independent component analysis. Frequently used classification method is LDA for dimensionality reduction. Between class variance and within class variance are maximized through LDA classifier to achieve the maximal separability. LDA classifier is used to select the optimal features, and level 4 decomposition features harr, db4, sym8, bior4.4, dmey are used to input images for feature extraction.

16.3 Formulation Using for Making BDTO Algorithm

Feature selection inherits the significant features from parent class to child class of the obtainable information. Metaheuristic with binary or chaotic performs role play to minimize the false positive rate, decrease data dimensionality, decline cost expensive, and maximize the accuracy for best optimal feature selection and extraction. Binary ABC algorithm for feature selection is introduced in 2014 by Subanya et al., binary GWO and Binary ALO in 2016 proposed by Emery et al., for extracting best features, binary PSO-GSA algorithm introduced in 2018 by Sarhani et al., S-shaped binary WOA by Hussian et al. in 2019, Binary BOA by Arora et al. in 2019, and Binary GOA by Mafarja et al. in 2019.

16.3.1 Feature Selection Using DTO Optimization Algorithm

Duck traveler optimization is inspired by the behavior of the duck traveler. DTO is used to select the breast features for classification system. To summarize the observations from ducks' foraging behavior, the subsequent tasks are offered.

Fig. 16.1 Ducks in a row

Task 1: A duck inhabitants comprise of more than a few clusters. Each cluster containing similar ducks optimize the victuals search activity using their stack of intelligence.
Task 2: Based on the height of the neck + head, the duck uses that information to select the hunting region.
Task 3: They traveling as a flock and follow their local guide which has fed on most food in the last location.
Task 4: After a number of tasks, ducks return on exterior to allocate with its local colleague, via announcement of exploitation, the position, and profusion of foodstuff.
Task 5: If the groceries support is less for the ducks of a given group to live on, part of the group migrates to another place via communication of exploration.
Task 6: Based on the satisfaction of end criteria, output the optimal solution. Otherwise go to Task 2 (Fig. 16.1).

16.3.2 Feature Selection Using bDTO Optimization Algorithm

Ducks stick together tiny clusters in static flock, which move through a confined area and attract other swimming ducks. The key features of a static swarm are spatial motions and sudden shifts in the moving direction. An enormous quantity of ducks in dynamic duck flock leads the swimming duck in identical path.

$$S(t)\infty\sqrt{t}$$

Movement of each duck in a flock takes diffusion with convection and advection. In bDTO for binary ducks, imprinting and foraging behavior makes hunting process efficiently and also quickly. Applying phase rule and Maxwell–Boltzmann rule with the constraints either the ducklings discontinue moving or the ducklings carry on moving

$$I + F = D + 1$$

$$P(F) = \frac{1}{DfF/n^C}$$

Probability (P) of a duck will have force (F) to reduce dimensionality (D) decreasing force (f) with (n) time's changes of climate (C). The bDTO by the diffusion is given as follows:

$$D_{t+1} = D_t + h \times \text{prob}()F_t$$

$$h = \sqrt{I}/_F$$

I specifies the action of identifying imprinting time period in seconds of a duck and F denotes the number of ducks involved in the foraging for the same flock in proportion to time.

$$F = 100 \times \mathrm{I}$$

$$F_t = \frac{1}{h\sqrt{2\pi}} \exp - \left(\frac{(\text{dimension} - \text{duck})^2}{2h^2} \right)$$

$$\mathrm{DV}(t+1) = \mathrm{C3DV}(t) + \mathrm{C4r2}(\mathrm{DPbest}(t) - \mathrm{DPnd}(t)) + \mathrm{C5r2}(\mathrm{DPGbest} - \mathrm{DPnd}(t))$$

$$\mathrm{DPnd}(t+1) = \mathrm{DPnd}(t) + \mathrm{DV}(t+1)$$

At each cycle, the classification exactness of each search duck in current populace is contrasted and the best fitness value in that emphasis. The best arrangements are put away and DT-pbest grid is framed utilizing DTO with Boltzmann development. Ducks are likewise made to screens best esteem got so far any duck in the area and is saved as DT-Gbest. It improves the misuse ability of DTO, and it likewise assists with getting away from nearby optima.

$$\Delta D^i t + 1 = h\Delta Dti + C1r1\,(\mathrm{DPpbest}_{ti} - Dti) + C2r2\,(\mathrm{DPGpbest}_{ti} - Dti)$$

$$D^i t + 1 = D^i_t + \Delta D^i_{t+1}$$

16.3.3 Feature Selection Using Sigmoid Activation Function

Sigmoid activation function is called as a logistic regression mainly used for binary classification to return a probability value which can then mapped to two discrete classes. In this research, the given input mammogram is benign or malignant is predicted by using sigmoid activation function. Clear predictions through gradient smoothening give the output values between zero and one. Vanishing gradient issue makes the slow process but significant results only achieved by using sigmoid activation function (Table 16.1)

$$S(af) = 1/1 + \mathrm{e}^{-\mathrm{x}}$$

$$D_{t+1} = D_t + P_t$$

16.4 Experimental Results

The Wisconsin bosom malignant growth determination (WBCD) database is the product of the endeavors made at the University of Wisconsin Hospital's attempts to reliably diagnose bosom masses in the sense of FNA examination. The reason for the informational index is to classify a tumor based on cell representations assembled by minuscule evaluation as either benign or threatening. The insightful collection incorporates ten ascribes like spherical structure, thin tissue, cytology preparation, cancer cell size, cancer cell shape, uniform texture, mitoses, tolerable sticking, and muster width. Six hundred and ninety nine patients information base is contained in which four hundred and fifty eight are considerate models and two hundred and forty one samples are threatening. Sixteen ascribes values are missed consequently;

Table 16.1 Attributes and its echelon obtained by gain ratio

Badge	*	Name of the variable	μ	σ	Echelon	Bigind
1-10	9	Spherical structure	2867	3054	1	0.399
1-10	6	Thin tissue	3216	2214	2	0.395
1-10	7	Cytology preparation	3464	3641	3	0.393
1-10	3	Cancer cell size	3134	3051	4	0.386
1-10	4	Cancer cell shape	3207	2972	5	0.314
1-10	8	Uniform texture	3438	2438	6	0.303
1-10	10	Mitoses	1589	1715	7	0.299
1-10	5	Tolerable sticking	2807	2855	8	0.271
1-10	2	Muster width	4418	2816	9	0.210

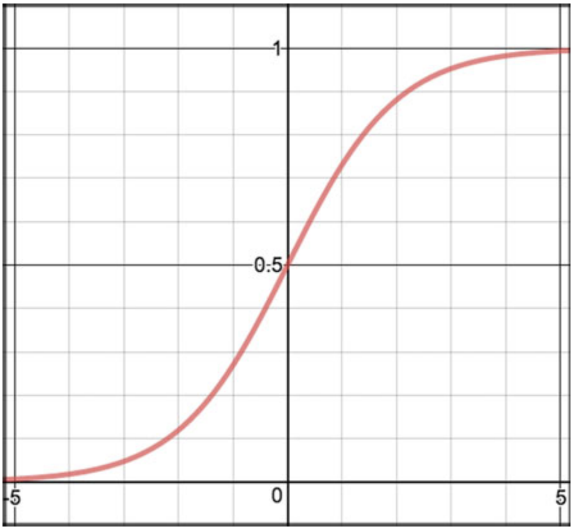


Fig. 16.2 Sigmoid activation function

we erased all from information base, and four hundred and forty four are utilized for preparing which contain two hundred and sixty benevolent and one hundred and eighty four threatening cases. Additionally two hundred and thirty nine utilized for testing which contains one hundred and eighty four benevolent and fifty five threatening cases (Fig. 16.2).

16.4.1 Performance Metrics

- **Accuracy**
  Correctly classified occurrences for all instances are measured by accuracy as below

$$\text{Accuracy} = \frac{\text{TOV} + \text{TPV}}{\text{TOV} + \text{TPV} + \text{FOV} + \text{FPV}}$$

- **Precision**
  Correctly classified occurrences for those instances are measured as optimistic value calculated by

$$\text{Precision} = \frac{\text{TOV}}{\text{TOV} + \text{FOV}}$$

- **Recall**
  Optimistic instances correctly classified are correctly measured by using recall ratio.

$$\text{Recall} = \frac{\text{TOV}}{\text{TOV} + \text{FPV}}$$

- **F-Measure**
  F-measure is the consolidated measurement of precision (P) and recall (R). It shows how exact the classifier is and furthermore how well the classifier is powerful.

$$F\text{-Measure} = \frac{2 * P * R}{P + R}$$

- **Sensitivity**
  Sensitivity means properly classified optimistic occurrences to a total number of optimistic instances which is measured by

$$\text{Sensitivity} = \frac{\text{TOV}}{\text{TOV} + \text{FPV}}$$

- **Specificity**
  Specificity means properly classified pessimistic occurrences to a total number of pessimistic instances which is measured by

$$\text{Specificity} = \frac{\text{TPV}}{\text{TPV} + \text{FOV}}$$

16.5 Conclusion

In this paper, a new binary model of the simple duck travel optimization algorithm known as bDTO to remedy the feature selection problem used to be proposed. To convert the native version of DTO to a binary version, sigmoid activation features are employed. In order to investigate the performance of the proposed two binary algorithms, the experiments are utilized on Wisconsin Breast datasets from UCI datasets and five assessment standards are performed. The experimental outcomes revealed that the proposed algorithms have executed foremost results versus the native algorithm. Furthermore, the results proved that bDTO has accomplished the smallest wide variety of features with better classification accuracy. For future work, the proposed algorithm added here will be used with extra common classifiers such as LDA, SVM, and ANN to affirm the performance (Tables 16.2 and 16.3; Fig. 16.3).

The computational complexity of the DTO algorithm is ($t_{\max} * n$) and bDTO is ($t_{\max} * n * d$) with d dimension. The final result shows that the mdb028 input image is in the malignant category.

Table 16.2 Confusion matrix (CM) for breast cancer

Analysis	Real	Forecasted
Pessimistic	1	2
Optimistic	3	4
Accuracy (ACC)	(1 + 4)/(1 + 2 + 3 + 4)	
True optimistic value (TOV)	4/(3 + 4)	
False pessimistic value (FPV)	3/(3 + 4)	
False optimistic value (FPV)	2/(1 + 2)	
True pessimistic value (TPV)	1/(1 + 2)	

Table 16.3 Metrics involved in classification

Metrics/methods	SAF	DTO	bDTO
Accuracy	0.846	0.886	0.92
Precision	0.804	0.832	0.876
Recall	0.822	0.921	0.939
F-measure	0.81	0.86	0.9
Sensitivity	0.727	0.76	0.84
Specificity	0.73	0.78	0.9

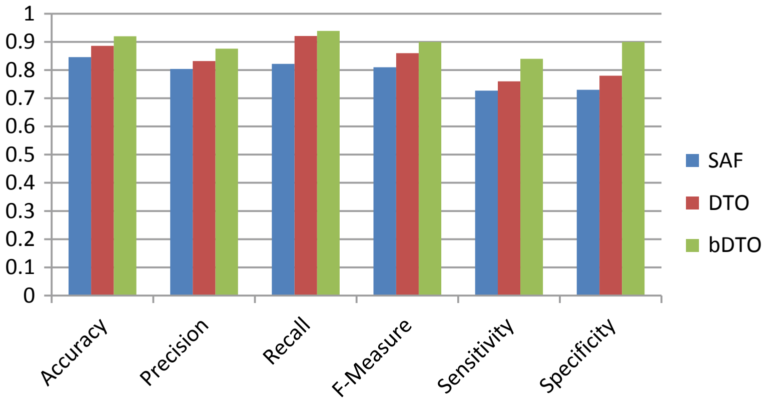
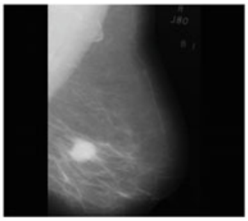
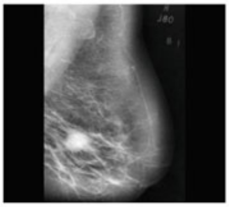


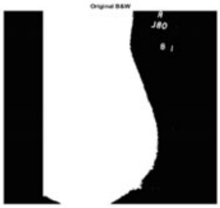
Fig. 16.3 Comparison between proposed and existing methods



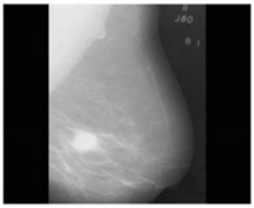
a) Input Image



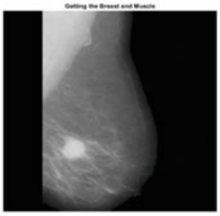
b) Preprocessed Image



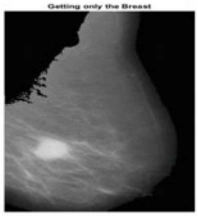
c) Black and White Image



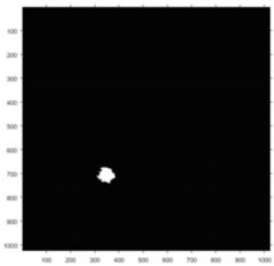
d) Gamma corrected Image



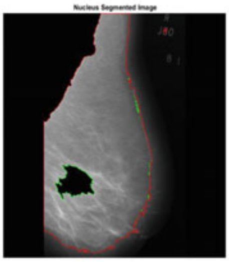
e) Getting the breast and muscle



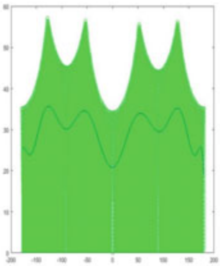
f) Getting only the breast



g) Segmented image



h) Region of Interest (ROI)



i) Graph

References

1. Abshouri, A.A., Alireza, B.: A new clustering method based on firefly and KHM. Int. J. Commun. Comput. **9**(4), 387 (2012)
2. Arora, S., Anand, P.: Binary butterfly optimization approaches for feature selection. Exp. Syst. Appl. **116**, 147–160 (2019)
3. Arora, S., Singh, S.: Butterfly optimization algorithm: a novel approach for global optimization. Soft Comput. **20**(3), 715–734 (2019)
4. Ashraf, M.H., Radwa, M., Hussein, M.E.: Multilevel thresholding for image segmentation using an improved electromagnetism optimization algorithm. Int. J. Interact. Multimed. Artif. Intell. **5**(4), 220–224 (2018)
5. Balakrishnan, S., Ganapathi, A.R.: LDA—a brief tutorial. Inst. Signal Inf. Process. **18**, 1–8 (1998)
6. Bayat, F.M.: The runner-root algorithm. J. Appl. Soft Comput. **33**, 292–303 (2015)
7. Bethiana, N.: Killer whale algorithm: an algorithm inspired by the life of killer whale. Procedia Comput. Sci. **124**, 151–157 (2017)

8. Casti, P., Mencattini, A., Salmeri, A.: Contour independent detection and classification of mammographic lesions. B.S.P.C **25**, 165–167 (2016)
9. Dhiman, G., Kumar, V.: Spotted Hyena Optimizer for solving complex and non-linear constrained engineering problems. HS NOA—Adv. Intell. Syst. Comput. **741**(4), 857–867 (2019)
10. Dorigo, M., Gianni, D.C.: Ant colony optimization: a new meta-heuristic. In: Proceedings of the 1999 Congress on Evolutionary Computation—CEC99, 2 Cat.No. 99th, 8406. IEEE (2019)
11. Gandomi, A.H., Alavi, A.H.: Krill herd: a new bio-inspired optimization algorithm. Commun. Nonlinear Sci. Numer. Simul. **17**(12), 4831–4845 (2012)
12. Haussien, A.G.: S-shaped binary whale optimization algorithm for feature selection. Recent Trends Signal Image Process. 79–87 (2019)
13. Hosseini, E.: Laying chicken algorithm: a new meta-heuristic approach to solve continuous programming problems. J. Appl. Comput. Math. **6**(344), 2 (2017)
14. Howlader, N., Noone, A.M., Krapcho, M. (eds.): SEER Cancer Statistics Review 1975–2016. National Cancer Institute, Bethesda, MD (2019)
15. Jain, R., Gupta, D., Khanna, A.: Usability feature optimization using MWOA. In: International Conference on Innovative Computing & Communication. Lecture Notes in Networks and Systems, vol. 56, pp. 382–397. Springer
16. Karaboga, D., Basturk, B.: A powerful and efficient algorithm for numerical function optimization: artificial bee colony (ABC) algorithm. J. Global Optim. **39**(3), 459–471 (2007)
17. Krishnanand, K.N., Ghose, D.: Glowworm swarm optimization: a new method for optimizing multi-modal functions. Int. J. Comput. Intell. Stud. **1**(1), 93–119 (2009)
18. Lamy, J.B.: Artificial feeding birds (AFB): a new metaheuristic inspired by the behavior of pigeons. Adv. Nat.-Inspir. Comput. Appl. **53**(3), 43–60 (2019)
19. Mirjalili, S.: Dragonfly algorithm: a new meta-heuristic optimization technique for solving single objective discrete and multi objective problems. J. Neural Comput. Appl. **27**(4), 1053–1073 (2015)
20. Mirjalili, S.: Moth flame optimization algorithm: a novel nature-inspired heuristic paradigm. Knowl. Based Syst. **89**, 228–249 (2015)
21. Mirjalili, S., Mirjalili, S.M., Lewis, A.: Grey wolf optimizer. Adv. Eng. Softw. **69**(2), 46–61 (2014)
22. Nag, K., Pal, N.R.: Genetic programming for classification and feature selection. Evol. Swarm Intell. Alg. 119–141 (2019)
23. Selva, B.K., Muneeswaran, K.: Firefly algorithm based feature selection for network intrusion detection. Comput. Secur. **81**, 148–155 (2019)
24. Shadravana, S., Najib, H.R., Bardsiri, V.K.: The Selfish Optimizer: a novel nature-inspired metaheuristic algorithm for solving constrained engineering optimization problems. Eng. Appl. AI **80**, 20–34 (2019)
25. Wolberg, W.H., Mangasarian, O.L.: Multisurface method of pattern separation for medical diagnosis applied to breast cytology. Appl. Math. **87**, 9193–9196 (1990)
26. Yang, X.S.: Flower pollination algorithm for global optimization. In: International Conference on Unconventional Computing and Natural Computation (UCNC), vol. 7445(3), pp. 240–249 (2012)
27. Zhang, J., Zhou, Y., Luo, Q.: Nature-inspired approach: a wind-driven water wave optimization algorithm. Appl. Int. **49**(1), 233–252 (2019)

Chapter 17
Low-Cost Health Monitoring Pedometer Using IoT

Ujwala Kshirsagar and Priti Shahane

Abstract The pedometer was initially used by fans of sport and exercise. But now, as a daily exercise measure, it is becoming popular. This paper discusses the idea of a pedometer as part of holistic approach designed to encourage a healthy lifestyle through regular physical activity. A paper deals with design and analysis of electromechanical pedometer which can measure steps in meter on a dial or counter. IoT-based pedometer using Raspberry Pi is proposed as a product in this paper. This pedometer is a fitness tracker system that tracks all motions and physical parameters with the aid of connected sensors. It links all the sensors data and display it to related metrics, such as steps count, distance covered, calories burned, heartbeat ratio, temperature of body, height and weight on one platform. The advantage of using such a system is, first and foremost, an increase in motivation to enhance physical activity. These devices are really important for fitness level monitoring. This wearable system helps us to offer a fitter and healthier experience, as all physical activities can be easily controlled by our fitness level. If we are getting physically more active in sports training, measuring our health is going to be easier than ever with proposed pedometer.

17.1 Introduction

A significant consideration is always well-being. However, numerous potentials yet risky variables, such as high blood pressure and irregular heartbeat rate, often interfere with the health of modern people. High blood pressure, also referred to as hypertension, raises the heart's strain of pumping blood supply into the arteries. One of the key risk factors for stroke, myocardial infarction (MI), cardiac failure, aneurysms, and peripheral arterial disease is hypertension, suggesting a wide variety of normal

U. Kshirsagar (✉) · P. Shahane
Symbiosis International (Deemed University), Pune, India
e-mail: ujwala.kshirsagar@sitpune.edu.in

P. Shahane
e-mail: pritis@sitpune.edu.in

T. Senjyu et al. (eds.), *IOT with Smart Systems*, Smart Innovation, Systems and Technologies 251, https://doi.org/10.1007/978-981-16-3945-6_18

heartbeat rates. The irregular rate of heartbeat can indicate the various underlying issues that are harmful to the body and harshly affected.

To avoid these problems, exercise (walking the simplest one) can be helpful hence we have to monitor the walking steps and calories burn. Maintaining such physical activity levels is necessary to avoid or postpone the onset of many medical conditions such as diabetes or mental health disorders. Traditional calorie estimating methods include wearing devices, such as pedometers, smart watches, or smart bracelets, which continuously track user behavior and estimate the energy expenditure [1]. In order to track health parameters regularly at home, we built a low-cost "IoT-based pedometer" health track device that can calculate walking steps, burning calories, heart rate, and body temperature. Typically used sensors for monitoring the walk are accelerometer or pedometer [2].

A pedometer is a device that counts each steps of a person by detecting the motion of the foot's step, usually portable and electronic or electromechanical [3]. As the distance of each person's move varies, an informal calibration performed by the user is required if presentation of the distance covered in a unit of length (such as in meters) is needed. For individuals looking to improve their physical activity, pedometers may be a motivational tool. Phase counters will help you to compete with yourself to get physically fit and reduce in weight. In clinical trials, pedometers have been shown to improve physical activity and decrease blood pressure levels and body mass index. The suggested pedometer measures the number of steps taken by the individual and multiplied it by the average length of step fed to the individual's distance walked. IoT software uses the Android application from a smartphone to watch walk of a person and also measures a person's heartbeat rate and displays the temperature and humidity using Raspberry Pi and communicates with the data from Raspberry Pi to the smartphone [4]. Mobile applications today can function as a pedometer or step counter with step sensor, heartbeat sensor, and temperature sensor to count the number of steps you take when walking, running, or step aerobics and also demonstrate the temperature and presence of humidity [5]. Figure 17.1 shows IoT-based pedometer using various sensors. IoT-based pedometer architecture consists of Raspberry Pi, cloud, two-stage sensor, heartbeat sensor, temperature sensor, power bank, LCD display, and Android application. The concept of IoT using Raspberry Pi for the implementation of pedometer is based on these health monitoring parameters to count the steps, calculate the burning of calories, the temperature, and humidity. The concept of IoT and the Raspberry Pi implementation is based on these health monitoring parameters.

The step sensor will calculate step counting and the burning of calories. The heartbeat sensor detects the rate of the person's heartbeat. The temperature sensor displays the humidity and the temperature. All this data is stored in a Raspberry Pi. The Raspberry Pi has uploaded the data to the Android app [5]. And the result is shown by the Android program. And it also appears on the LCD monitor. Figure 17.1 shows the architecture of IoT-based pedometer.

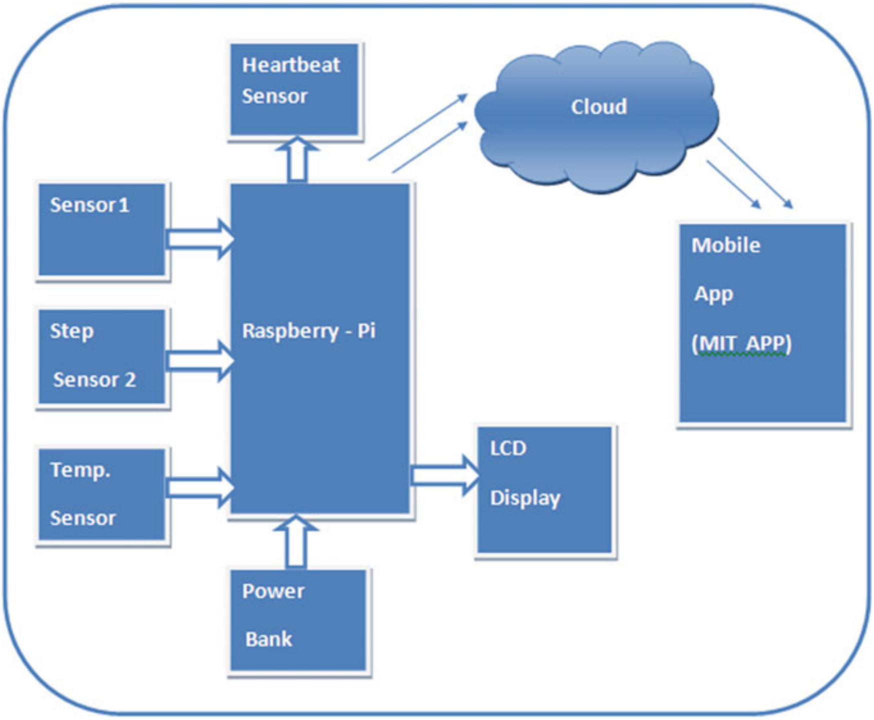


Fig. 17.1 Architecture of IoT-based pedometer

17.2 Related Work

IoT-based pedometer implementations have several frameworks and platforms to build. These concepts are also commonly adopted for use in designing the pedometer implementation Most of them, however, lack ease of setup and management of connected devices, problems with accuracy, user-friendly parameters for monitoring and control, safe network communication between connected sensors, devices, Raspberry Pi, and users. Pedometers are a common way for individuals to check whether they have completed the 10,000 daily steps recommended. Therefore, the purpose of this research was to determine the accuracy of four pedometer brands in measurement steps and to determine if the cost of the pedometer is related to the accuracy of the pedometer [6]. Chelsea G. Bender et al. (2017) presented the paper on Assessing the Health of Fitness Trackers. This paper compares the different parameters and experimental related to the fitness tracker in pursuit of the fitness tracker [7]. Ms Najme Zehra Naqvi et al. (2012) presented Counting of steps Using Smartphone Accelerometer [8]. Tang Meiyu et al. (2018) proposed the Concept of a Bluetooth 4.0.0 network pedometer creation that incorporates hotspot movement and Bluetooth 4.0 technology [9]. Wu, Shyi-Shiou et al. (2011) presented the specification of an

Intelligent-Pedometer using Android. In order to develop its application, Android is used. The walking motion of the user was detected by the acceleration sensor and the orientation sensor, and Bluetooth provided voice feedback [10]. Vivek P. Ardeshi et al. (2017) presented a study on Health Monitoring Systems using IoT Raspberry Pi. Raspberry Pi and IoT are a master computer in the proposed scheme for which all other devices such as different sensors are linked [11].

The literature survey shows that there are many issues related to the advancement of IoT-based technology for the pedometer device, and some of the research papers presented concentrate on effective basic management, incorporation of different protocols, monitoring and control of parameters of health monitoring, and communication range. For smart home systems however, not a single technical solution has been documented that solves all the problems mentioned above. Therefore, there is a need for an IoT-based pedometer to develop a technological solution to overcome all the problems.

17.3 Methodology

For the implementation of IoT-based pedometer, following equipment are being used.

17.3.1 Raspberry Pi Kit

The Raspberry Pi boards are compact and of low cost [9]. It serves as a server when the Internet is linked to the Raspberry Pi. The server sends data to the cloud automatically. Therefore, the health monitoring parameters are tracked.

17.3.2 DHT11

A temperature and humidity sensor integrated with a calibrated digital output features the DHT11 Temperature and Humidity Sensor. This sensor includes a humidity measurement component of the resistive kind and a temperature measurement component of the NTC and connects to a high-performance 8-bit microcontroller, offering excellent efficiency, quick response, ability to interfere and cost effectiveness.

17.3.3 Heartbeat Sensor

The heartbeat sensor is an electronic system used to measure heart rate, i.e., heartbeat velocity. The most important things that are required to keep us safe are the measurement of body temperature, heart rate, and blood pressure. There are many methods of measuring heart rate, and electrocardiography is the most accurate, but the best way to track heart rate is to use a heartbeat sensor.

17.3.4 Step Sensor

A phase sensor is used as the limit switch. A limit switch observes the actual movement of an object through direct contact with that object. A limit switch will be in its "natural state" when it is not in contact with anything. Of limit switch has its own "NO" (normally-open), "NC" (normally-closed), and "C" (common) screw terminal to link wires. Figure 17.2 shows the flowchart of the IoT-based pedometer. The embedded system with the Internet of things (IoT) has enabled for development of environments with highly heterogeneous entities and networks linked together. This can be seen in the notable growth over the last decade of wearable devices and applications. The proposed pedometer is focused to monitor and produce physical activities for health care. That is the case with personal fitness trackers, which provide new features that require a wide range with confidential user details to be collected in a single device.

17.4 Circuit Description and Experimental Results

The following experimental setup in Fig. 17.3 shows the output of IoT-based pedometer. Figure 17.3 indicates overall setup of the device. Due to small form factor, accuracy, efficiency and easy to use, continuous monitoring system came in form of wearable devices [12]. The main controlling unit of the device is Raspberry Pi which is supplied with the battery power. It indicates small heart rate sensor which can be mounted on the finger.

Two small limit switches will be fitted in the persons' shoes to count the no. of steps. DHT11 sensor is used to measure the temperature of the human body, the reading of heart rate, no. of steps completed, distance covered, and calories burned are showed with LCD as well as the same reading will be stored on the cloud.

Android app is developed with the help of MIT App Inventor. Here, the app developed for pedometer shows the reading of all the parameters like no. of steps, current temperature, presence of humidity, distance covered, and calories burns. These parameters are fetched from the cloud which is stored by our hardware circuit. The details output of each and every parameter is indicated by following testimonials.

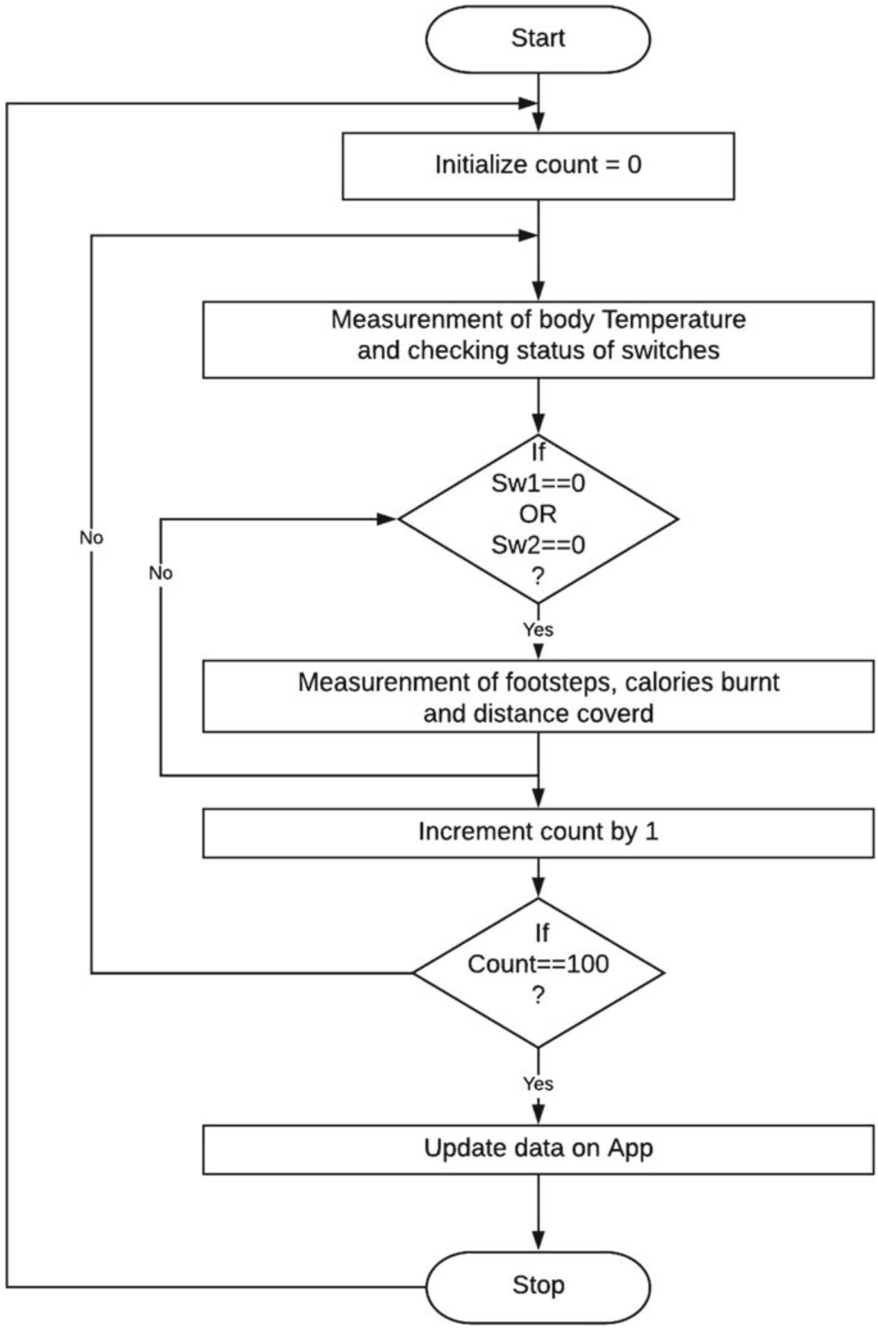


Fig. 17.2 Flowchart of IoT-based pedometer

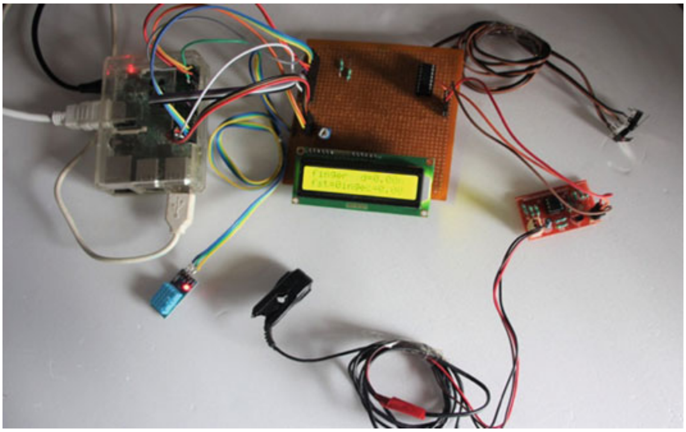


Fig. 17.3 Experimental setup

At the very first before performing the experiment, the circuit LCD display and app display are initialized as zero as shown in Fig. 17.4.

As shown in Fig. 17.5, for performing the pedometer action, firstly, we applied the inputs like weight in kg and height in centimeter to the program. After running the program, the step sensor limit switch is being activated for count of the steps and distance covered and simultaneously heartbeat sensor mounted on the finger for

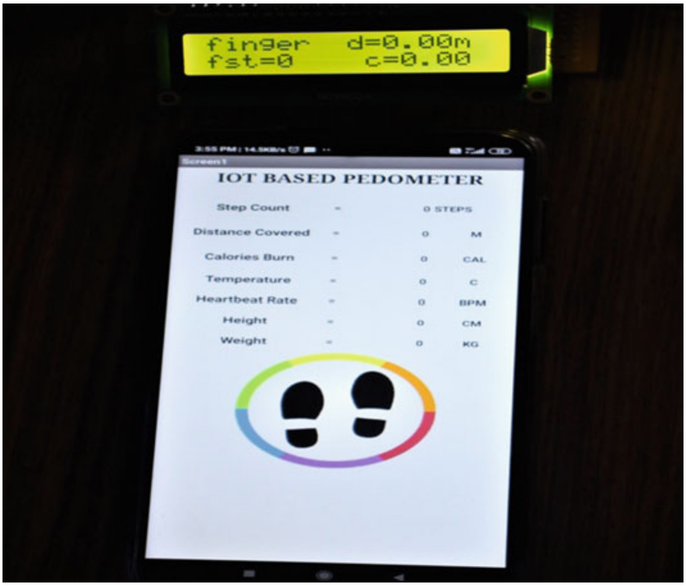


Fig. 17.4 LCD and pedometer MIT APP display

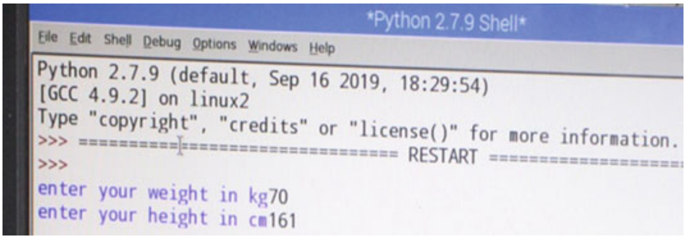


Fig. 17.5 Program display—weight in kg and height in centimeter

measuring the heartbeat rate in BPM. Using the DHT11 temperature sensor, after running the program for 100 times, it will show the temperature and presence of humidity will be displayed which we experimented as shown in Fig. 17.6.

Fig. 17.6 Temperature and humidity display

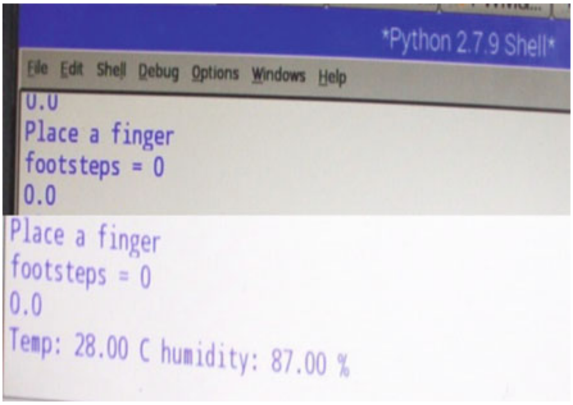


Fig. 17.7 Final output of IoT-based pedometer

Finally proposed pedometer shown in Fig. 17.7 shows all the expected parameters like no. of steps, distance covered in meter, temperature, presence of humidity, and calories as shown in the photograph.

17.4.1 Calculation for Calories Burn

The calculations of calories burned are shown below and same is encoded with Python coding to calculate total calories burned based on the weight of the person and distance covered.

Firstly, to measure how many calories you burn in 1 mile of casual walking, multiply your weight by 0.57. Calories burn per mile equals to 0.57 multiplied by weight of the person in pound. 1 kg equal to 2.2 pound; therefore, calories burn in meters' equals to calories burn per mile divided by 1609.

For example, let the person weight equal to 70 kg.

Calories burn in meters' equals to 0.57 * 154/1609 equals to 0.054.

Now, total calories burn equals to calories per mile multiplied by distance covered equal to 0.054 * 42 equals to 2.2 cal.

17.5 Conclusion

In this paper, we implemented pedometer for health monitoring. The novelty of the proposed pedometer is that the digital display of steps counts, distance covered, calories burned, heartbeat ratio, temperature of body, height and weight are brought on single platform with IoT. This technology helps to easily monitor the health of athletes, senior citizens, and ordinary people. For tracking and calculating the calories burned and heartbeat rate with a simple mathematical formula, we used a Python coding. We successfully build an Android application in the healthcare domain using the concept of Raspberry Pi and IoT device. This pedometer is often easily used by ordinary individuals at home. For various uses, doctors and trainers may make use of the collected data of IoT-based pedometer for a long duration of time.

References

1. Raheman, H.: Counting calories without wearables: device-free human energy expenditure estimation. In: International Conference on Wireless and Mobile Computing, Networking and Communications, 9253424 (2020)
2. Sathe, N.P., Hiwale, A.S.: Fitness mantra: theme for health analysis in rehabilitation process. In: International Conference on Big Data, IOT and Data Science, pp. 124–129 (2018)

3. Bajpai, A., Jilla, V., Tiwari, V.N., Venkatesan, S.M., Narayanan, R.: Quantifiable fitness tracking using wearable device. In: 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) (2015). https://doi.org/10.1109/embc.2015.7318688
4. Gunasegaram, M., Nallav, G.: Microcontroller based fitness analysis using IOT. Int. J. Phys. Educ. Sports Health **4**(3), 255–259 (2017)
5. Bhawarkar, N.B., Amaale, P.A., Chourey, S.R.: Modern health monitoring system using IOT and Raspberry Pi. In: 2nd National Conference Recent Innovations in Science and Engineering (NC-RISE 17), vol. 5(9) (2017). ISSN: 2321-8169
6. Husted, H.M., Llewellyn, T.L.: The accuracy of pedometers in measuring walking steps on a treadmill in college students. Int. J. Exerc. Sci. **10**(1), 146–153 (2017)
7. Bender, C.G., Hoff tot, J.C., Combs, B.T., Hooshangi, S., Cappos, J.: Measuring the fitness of fitness trackers. In: IEEE Sensors Applications Symposium (SAS) (2017). https://doi.org/10.1109/sas.2017.7894077
8. Naqvi, N.Z., Kumar, A., Chauhan, A., Sahni, K.: Step counting using smartphone-based accelerometer. Int. J. Comput. Sci. Eng. (IJCSE) **4**(05) (2012). ISSN: 0975-3397
9. Tang, M., Gao, T., Zhou, S., Zhao, K.: The design of network pedometer based on Bluetooth 4.0. In: International Conference on Intelligent Transportation, Big Data & Smart City (ICITBS) (2018). https://doi.org/10.1109/icitbs.2018.00063
10. Wu, S.-S., Wu, H.-Y.: The design of an intelligent pedometer using Android. In: 2011 Second International Conference on Innovations in Bio-Inspired Computing and Applications (2011). https://doi.org/10.1109/ibica.2011.83
11. Ardeshi, V.P., Sagar, S., Murmurwar, S., Hage, P.: Health monitoring systems using IoT and Raspberry Pi—a review. In: International Conference on Innovative Mechanisms for Industry Applications (ICIMIA) (2017). https://doi.org/10.1109/icimia.7975587
12. Mohanty, S.P., et al.: Smart-walk: an intelligent physiological monitoring system for smart families. In: IEEE International Conference on Consumer Electronics, pp. 1–4. IEEE (2018)

Chapter 18
Database Building, Recognition, and Categorization of Handwritten Kannada Words Using Convolution Neural Networks

Chandravva Hebbi, Y. M. Pradyumna, and H. R. Mamatha

Abstract This work focuses on building the database for handwritten Kannada words and makes it open source as there are no publicly available datasets online. The dataset can be used as a benchmark dataset to evaluate the performance of various recognition, feature extraction methods, and the results of various researchers. The dataset is rich in variations in writing styles because the data is collected from 370 different people. The words are extracted by the projection profiles. Experimentation is done to recognize the words using convolution neural networks (CNN) with an accuracy of 93.20%, and it is followed by category recognition with an accuracy of 91.90. An exhaustive dataset of 65,378 words and 11 categories is built with experimentation

18.1 Introduction

Recognition of handwritten characters and words from the scanned document images is the process of automatic reading of the characters and words from the document image. Recognition of words and characters is carried by either the online recognition method or by using the offline method. In the online recognition method, electronic devices are used for writing. The temporal information of the word or character like direction, the position of the pen, and pressure is used in the case of the online recognition system. In the case of offline word recognition, the inscribed words are first made available and then these words are given for the recognition system. The inscribed words can be printed or handwritten. Recognition of printed words is easier than handwritten. Recognition of handwritten words is challenging as these words come with a different orientation, size, style, pressure with which character is written, speed of writing, the physical and mental conditions of the writer while writing, and occlusion. Occlusion may be due to oil stains on the paper, water droplets, and noise in

C. Hebbi (✉) · Y. M. Pradyumna · H. R. Mamatha
Department of Computer Science and Engineering, PES University, Bengaluru 85, India

H. R. Mamatha
e-mail: mamathahr@pes.edu

T. Senjyu et al. (eds.), *IOT with Smart Systems*, Smart Innovation, Systems and Technologies 251, https://doi.org/10.1007/978-981-16-3945-6_19

the image. Analytical and holistic are the two main approaches to recognition words. In the analytical approach, the word image is split up into individual characters, and each character is recognized separately. The holistic approach is segmentation-free; i.e., the entire word image is used for the recognition. Each word is considered an indivisible entity, and the features of the entire world are considered for recognition [1].

The work on the development of OCR for Indic scripts is still in the infant stage. Researchers are working on different Indic scripts like Bangla, Gurumukhi, Devanagari, and also on some of the south Indian languages like Kannada, Tamil, and Telugu. This has motivated us to work on the handwritten Kannada words and build the dataset for the same. The aim and objective of the proposed method are to build a dataset for the handwritten Kannada simple words, recognize these words, and classify these words into name of the place, animals, and birds using machine learning techniques.

18.1.1 About Kannada Script

Kannada is the official language of the Government of Karnataka. The script is derived from the Brahmi script. The character set is large, and characters are similar and complex in structure. The character set consists of 49 alphabets (aksharas). The modern Kannada character set consists of 13 swaragalu (vowels) and 36 Vyanjanagalu (consonants). The combination of consonants and vowels or the consonant, consonant, vowel modifiers can be done, and these characters are called compound characters or vattaksharas. The number of combinations that can be done is 36 × 36 × 13. The combination of consonant with a consonant is referred to be vattakshara. Figure 18.1 shows some of the combinations of the characters with vowels.

When the number of combinations of the characters has increased, the complexity also increases. In the proposed model, we have considered the entire word for recognition instead of segmenting the words in characters. The rest of the paper is organized into existing work, proposed methodology, results, and conclusion. Rest of the paper is organized in related work, proposed methodology, and recognition model.

18.2 Related Work

The research work in the field of document image processing to build OCR for languages is still an ongoing process across the globe. The research toward word recognition that has been carried out for languages like English, Arabic, and some

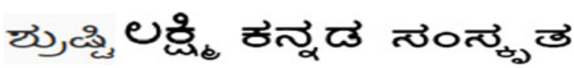


Fig. 18.1 Combination of Kannada character

of the Indian languages like Bangla, Devanagari, Kannada, Tamil, Telugu has been discussed below.

18.2.1 English, Bangla, and Devanagari Word Recognition

Recognition of handwritten English words with data augmentation and normalization using a convolutional neural network (CNN) and long short-term memory (LSTM) is presented in [1]. A comparative study with IAM, RIMES, the historical German database has been carried out. In [2], recognition of Arabic handwritten offline words using discrete cosine transforms (DCT) features support vector machines (SVM) with radial basis function (RBF) Kernel is presented. IFNI/ENIT database of Arabic dataset was used for the experimentation. The authors claim a good accuracy was obtained. Recognition of handwritten Arabic words using multi-level local phase quantization, Gabor features, and histogram of gradient using support vector machines, K-nearest neighbors, naïve Bayes is presented in [3]. The dataset consists of 1000 handwritten words collected from 100 users. An accuracy of 97.84% was achieved with the mentioned dataset. In [4], authors have presented the work on the recognition of Bangla handwritten words using the elliptical features. The extracted features are fed to classifiers like naive Bayes, dagging, bagging, support vector machines, and multilayer perceptron (MLP). The authors claim with the MLP method good recognition accuracy was achieved. In [5], the method to recognize using the histogram of gradient (HOG) features and neural network-based classifier has been discussed. The authors have achieved good accuracy with a small dataset. Recognition of handwritten Bangla words using the lexicon reduction technique was discussed in [6]. K-means clustering algorithm is applied to classify 35,700 words into 8 classes with an error rate of 3.8%.

In [7, 8], authors have described the methods to recognize handwritten Marathi characters and letters using geometrical features of the character images and two-letter words. The geometrical features of the words and characters extracted are given as input to the classifier neural network. The methods to recognize the handwritten Devanagari words using scale-invariant feature transform (SIFT) speed up robust feature (SURF) are discussed in [9]. These methods give scale- and rotation-invariant features. The experimentation is carried out on 1000 Devanagari words. The features obtained by the methods are compared with stored features to recognize the words.

Recognition of handwritten Devanagari words using the gradient, structural, statistical features, directional distance distribution, and gradient–structural–concavity features with classifiers Baum–Welch and Viterbi, SVM, and k-NN classifiers is discussed in [10–12]. Recognition of Marathi and Hindi handwritten words using the gradient, cavity, and structural method and binary vector matching in the first stage and vertical projection profile feature and dynamic time wrapping in stage 2 is described in [13]. The dataset consists of 26,720 words of the legal amount. The researcher claims very good accuracy for the proposed technique. In [14], recognition of offline handwritten Devanagari characters and words of using neural networks is discussed. An accuracy of 90% is achieved for characters and words.

Recognition of handwritten Telugu words using bidirectional LSTM is presented in [15]. The experiment was done on 120 K handwritten words collected from 11 users, and the words are labeled. In [16], recognition of Tamil handwritten words with Gabor, geometric features, an SVM classifier is discussed. The dataset consists of 4270 words of 217 country names. The words are recognized with good accuracy. Recognition of online handwritten Devanagari words using versions of recurrent neural network: long short-term memory and bidirectional long short-term memory is described in [17]. In [18], the recognition of handwritten Malayalam words using deep learning is presented. The dataset was built and experimented with deep learning architecture. Very good accuracy is obtained using the method on a large dataset. Segmentation-based handwritten Farsi word recognition with neural networks is presented in [19]. The dataset consists of 17,000 handwritten Farsi words. The rate of recognition is very high.

18.2.2 Recognition of Kannada Words

Recognition of handwritten district names of Karnataka State using the Euclidean classifier is presented in [21]. The data was collected from 60 people. Various preprocessing techniques were applied to remove the noise and enhance the images.

In [22], authors have proposed a method for handwritten Kannada words recognition using locality preserving projection (LPP), and this method is used for feature extraction and dimensionality reduction. Classification is done with support vector machine (SVM), and experimentation is done with names of districts and taluks of Karnataka State. In [23], the methods for words spotting and recognition using various deep learning models are presented. The feature extraction methods used are local binary pattern (LBP), HOG, and Gabor. An accuracy of 82% is claimed by the authors with a dataset of 11000 words written by 100 people using the CNN model with spatial transformation.

A literature survey reveals that the research work on the recognition of handwritten words for Indic scripts like Bangla, Marathi, Gurumukhi, Tamil, and Malayalam has been reported with good accuracy. The work on the handwritten word recognition for South Indian languages is still in the infant stage. Very few works have been reported to recognize the handwritten and printed Kannada words. This motivated us to take up the work on recognizing handwritten words and classify the words into different categories and publish the dataset online for the other researcher to work on this domain. This dataset can be used as a benchmark for their findings. The dataset can be used to build the mobile applications to learn the language or keyword spotting for the document categorization, scrape the information of the word from the web after the word is recognized. With these applications in mind, the dataset is built and will be published online.

18.3 Proposed Methodology

The proposed system architecture is divided into two phases: building the database and classification using CNN.

18.3.1 Database Building

The database building phase is divided into data collection and database creation.

Data Collection

There were not many images of handwritten Kannada words which are available publicly for the researchers to use directly. Hence, there was a need for the development of a dataset for the handwritten Kannada words. As a consequence of this data was collected from the people by handing over the form as shown in Fig. 18.2. Each

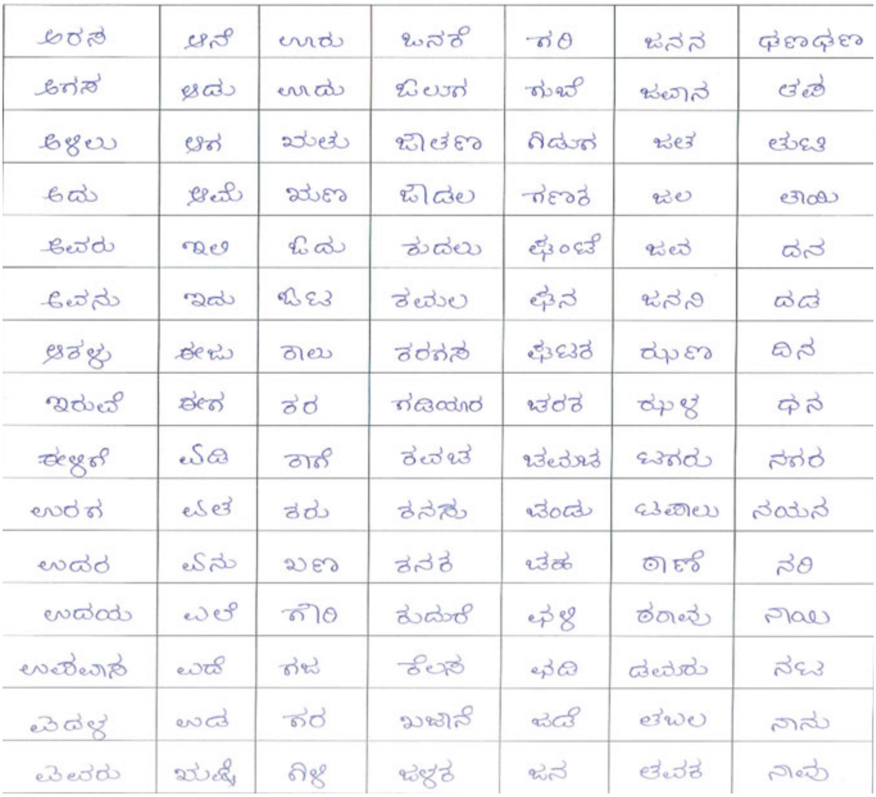


Fig. 18.2 Sample words written by an author

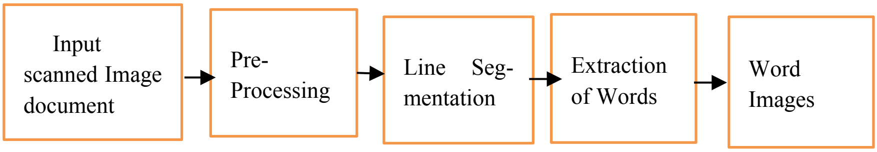


Fig. 18.3 System architecture for database building

author (writer) was asked to write a total of 210 specified simple (words without consonant conjugates) words. The words collected are the names of birds, animals, places, trees, body parts, things, nouns, verbs, and pronouns. Writer information is also collected by taking their consent in providing their handwritten data for the research.

Writer Information

The writer's details such as name, age, gender, profession, education qualification, and other details have been collected. Data are collected from people who know Kannada and also from people who do not know the script. Much of the data is collected from grade 10, grade 12, and degree students.

Dataset creation

All the documents have been scanned in the Kyocera Ecosys FS-6525 MFP scanner with 300 dpi as color images. The scanned documents were named with authors id in JPG image format. Figure 18.3 shows the steps involved in database creation.

Preprocessing

The documents are preprocessed by first eliminating the grid lines by using morphological operations. A median blur filter was used to remove fine noise in the document image. Binary thresholding and color inversion were performed on the images.

Text Line Segmentation and extraction of words

The text lines and words are segmented using horizontal and vertical projection profiles [20]. The thresholds were decided after observation of the nature of the projection profile. Words were segmented using contours after dilating each line to merge the characters of the words to avoid splitting of characters within the words. The bounding boxes for each of the contours were used to extract the regions of interest. We were able to successfully extract and move words to their respective folders with an accuracy of about 75%. This accuracy of extraction is affected due to the distortion of the projection profiles and also because of the inconsistency present in the position of the words in a line. The images are extracted and written to folders according to the index value. Extracted word images are labeled with file name.

The words are labeled after authors id as the dataset will be further used for the author identification or forensic analysis. The size of the dataset is 65,378. Figure 18.4 shows the sample dataset containing the word "Ame." Similarly, a dataset for the other words is also developed.

Fig. 18.4 Sample dataset for the word "Ame"

Sample Dataset

See Fig. 18.4.

Words not included in dataset

Some of the words would not be part of dataset as some of the words were wrongly written, words not segmented correctly due to the spacing between the characters or no spacing between the words, some of words had distortion and some of the words had noise in them. Figure 18.5 shows the words that are not included in the dataset.

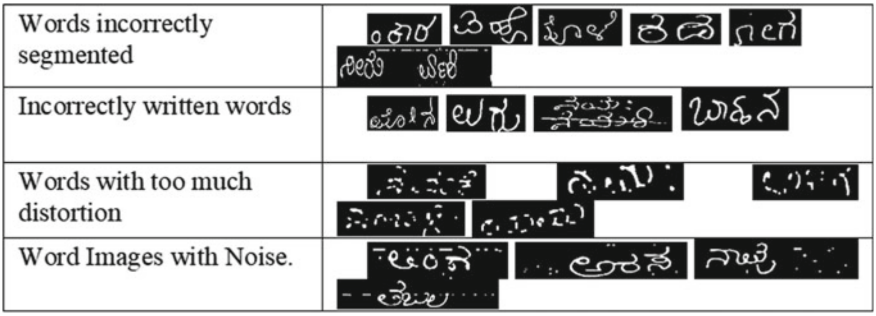


Fig. 18.5 Words with noise

18.4 Recognition Model

The recognition model consists of training and testing phases. The recognition model consists of training and test phases as shown in Fig. 18.6.

18.4.1 Training

All the images are normalized with the min–max normalization method before it is given to the model. A convolutional neural network was used to recognize the words. The neural network takes a 3 channel 300 × 300 RGB image as input. The model had six convolutional layers for extracting the features. Each of the convolutions was followed by a 2 × 2 max pooling layer of stride 1. Batch normalization was used for each of the convolutional layers to fit the training set well since it has a regularizing effect on small batch sizes. The extracted features are then flattened and passed to a fully connected network, with 1800 units for each of the two hidden layers with dropout regularization were used. (Dropout probability = 0.4). The output layer uses a softmax activation, and the hidden layers use ReLU/tanh activation functions. The final layer of the model has 210 units and a softmax activation function.

The model is trained using a categorical cross-entropy loss between the predicted probabilities and the one-hot label vectors. An Adam optimizer with a learning rate of 0.01 was used to rain the model. The raw images are resized 200 × 200 while training.

A total of about 65,378 images were used for this classification problem, 52,378 images were used to train the classifier, and 13,077 images were used for testing the trained model. An 80–20 split of the data was used. The splits were stratified to maintain similar distributions of classes in the train and test sets. This is essential while training a model on a large number of classes to make sure that some classes are not underrepresented in the train set.

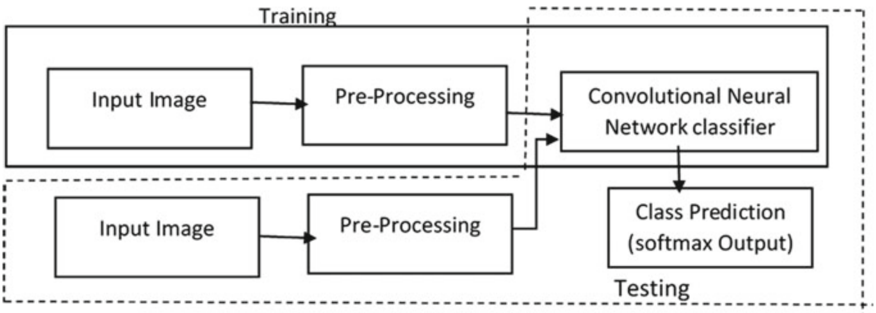


Fig. 18.6 Proposed recognition Model

a. Input image b. Output

Fig. 18.7 Recognized word "arasa"

18.4.2 Testing

The trained model was tested with a test set, and it performed satisfactorily with a test set accuracy of 93.32%. Figure 19.7a and b shows the sample input image and 19.7b shows the output obtained from the recognition model. After the word is recognized, its name in English is printed on the console.

18.5 Results and Discussions

The experiment was carried out on the dataset created with the proposed methodology using convolutional neural network.

The accuracy of sample 10 words is shown in Table 18.1. The accuracy of some of the words is 100%, and for the word "hogu," the accuracy is 45%; this was a huge drop in the accuracy of the word as the style of writing of this word was varying from each writer. The average accuracy of the proposed model is 93.32%.

The dataset was manually categorized into 11 classes, namely organs, events, people, birds, animals/reptiles/insects, feelings of people, drinks/time, instrument/device trees, things, places. Figure 18.8 shows the accuracy of the model.

From Table 18.2, it is very clear that authors have worked on their dataset, and in some cases, the size of the dataset is small. In [23] and proposed methodology, the classifier model used is CNN and the dataset is handwritten Kannada words. The

Table 18.1 Recognition of words and their accuracy

# Images	# images correctly recognized	# images in the test set	accuracy
325	60	65	92.31
320	58	64	90.63
354	63	71	88.73
295	54	59	91.53
328	65	66	98.48
314	62	63	98.41
330	63	66	95.45
286	52	57	91.23
289	54	58	93.10
324	64	65	98.46

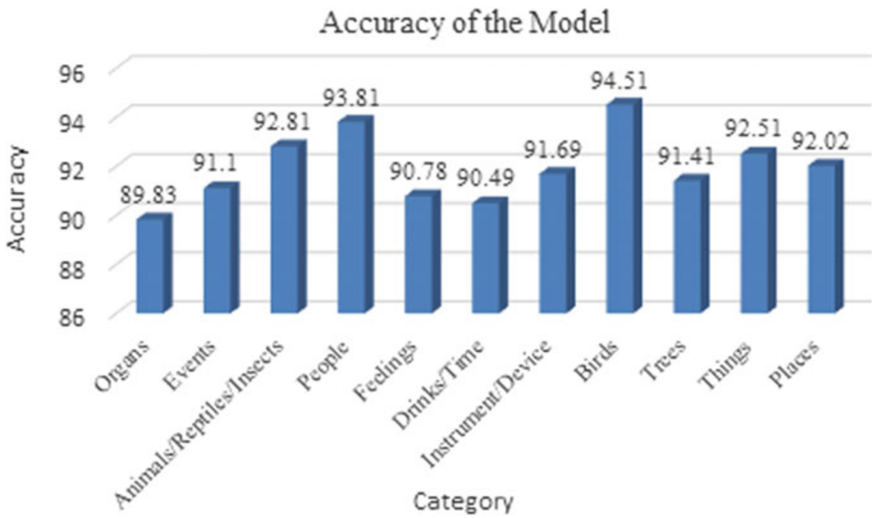


Fig. 18.8 Accuracy of the classifier model

Table 18.2 Comparative study of recognition models

Authors	Feature extraction method	Classifier	Dataset size	Accuracy (%)
M. S. Patel et al. [21]	Freeman's chain code (FCC)	Euclidean distance classifier, Dynamic Time Wrapping	1200	92
M. S. Patel et al. [22]	Locality Preserving Projections (LPP)	Support vector machine (SVM)	600	80
Tulika Sureka et al. [23]	KAZE, HoG, LBP, Gabor Wavelets	convolutional neural network (CNN)	11,000	82
Proposed method	–	convolutional neural network (CNN)	65,375	93.32 91.90

size of the dataset and images in the dataset is different. The proposed methodology gave us good results compared to others.

18.6 Conclusion

The paper aims to build the dataset for 210 handwritten Kannada words using projection profiles. These words are later fed to the convolution neural network to recognize these words. The average accuracy of recognition achieved is 93.32%. We have also classified these words into 11 categories, namely animals, birds, words-related devices/instruments, feelings of the people, words related to person and name of the person, body parts, events, places, things, plants, drinks, and time. An accuracy of 91.90% was achieved with the proposed method. The scope for further research is to display the information about the recognized word using web crawlers, translate the word to English; e.g., the word "arasa" can be translated as king, and voice output of the word recognized.

References

1. Wigington, C., Stewart, S., Davis, B., Barrett, B., Price, B., Cohen, S.: Data augmentation for recognition of handwritten words and lines using a CNN-LSTM network. In: 14th IAPR International Conference on Document Analysis and Recognition, pp. 639–645 (2017)
2. EI qacimy, B., Ait kerroum, M., Hammouch, A.: Word-based Arabic handwritten recognition using SVM classifier with a reject option. In: 15th International Conference on Intelligent Systems Design and Applications (ISDA), Marrakech, pp. 64–68 (2015)
3. Aiadi, O., Korich, A., Kherfi, M.L.: Recognizing Arabic handwritten words using multiple features and classifier selection. In: 2019 International Conference on Networking and Advanced Systems (ICNAS), pp. 1–5 (2019)
4. Bhowmik, S., Malakar, S., Sarkar, R., Nasipuri, M.: Handwritten Bangla word recognition using elliptical features. In: Sixth International Conference on Computational Intelligence and Communication Networks, pp. 257–261 (2014)
5. Bhowmik, S., Roushan, M.G., Ram, S., Polley, S., Malakar, S.: Handwritten Bangla word recognition using HOG descriptor. In: Fourth International Conference of Emerging Applications of Information Technology, pp. 193–197 (2014)
6. Bhowmik, T.K., Roy, U., Parui, S.K.: Lexicon reduction technique for Bangla handwritten word recognition. In: 10th IAPR International Workshop on Document Analysis Systems, pp. 195–199 (2012)
7. Oval, S.G., Shirawale, S.: Offline recognition of handwritten Devanagari words using recurrent neural network. Advances in Intelligent Systems and Computing, Proceedings of the 3rd International Conference on Frontiers of Intelligent Computing (FICTA), pp. 413–421 (2014)
8. Sawant, S., Baji, S.: Handwritten character and word recognition using their geometrical features through neural networks. Int. J. Appl. Innov. Eng. Manag. (IJAIEM) **5**(7), 077–085 (2016)
9. Guruprasad, P.: Handwritten Devanagari word recognition robust invariant feature transform. In: International Conference on Applied and Theoretical Computing and Communication Technology (iCATccT), pp. 327–330 (2015)
10. Shaikh, A., Dagade, R.: Offline recognition of handwritten Devanagari words using Hidden Markov Model. IJIRST—Int. J. Innov. Res. Sci. Technol. **1**(11), 526–531 (2015)
11. Singh, B., Mittal, A., Ansari, M.A., Ghosh, D.: Handwritten Devanagari word recognition: a curvelet transform based approach. Int. J. Comput. Sci. Eng. (IJCSE) **3**(4), 1658–1665 (2011)
12. Shaw, B., Bhattacharya, U., Parui, S.K.: Offline handwritten Devanagari word recognition: information fusion at feature and classifier levels. In: 3rd IAPR Asian Conference on Pattern Recognition (ACPR), pp. 720–724 (2015)
13. Jayadevan, R., Kolhe, S.R., Patil, P.M., Pal, U.: Database development and recognition of handwritten Devanagari legal amount words. In: International Conference on Document Analysis and Recognition, pp. 304–308 (2011)
14. Dedhe, V., Patil, S.: Handwritten Devanagari special characters and words recognition using neural network. Int. J. Eng. Sci. Res. Technol. **2**(9), 2521–2526 (2013)

15. Dutta, K., Krishnan, P., Mathew, M., Jawahar, C.V.: Towards spotting and recognition of handwritten words in Indic scripts. In: 16th International Conference on Frontiers in Handwriting Recognition (ICFHR), Niagara Falls, NY, pp. 32–37 (2018)
16. Subramaniam, T., Pal, U., Premaretne, H., Kodikara, N.: Holistic recognition of handwritten Tamil words. In: Third International Conference on Emerging Applications of Information Technology (EAIT), pp. 165–169 (2012)
17. Ghosh, R., Vamshi, C., Kumar, P.: RNN based online handwritten word recognition in Devanagari and Bengali scripts using horizontal zoning. J. Pattern Recogn. **92**, 203–218 (2019)
18. Jino, P.J., Balakrishnan, K., Bhattacharya, U.: Offline handwritten Malayalam word recognition using a deep architecture. Soft Computing for Problem Solving: SocProS, pp. 913–925 (2017)
19. Ghadikolaie, M.F.Y., Kabir, E., Razzazi, F.: Sub-word based offline handwritten farsi word recognition using recurrent neural network. Electron. Telecommun. Res. Inst. (ETRI) J. **38**(4), 703–713 (2016)
20. Mamatha, H.R., Srikantamurthy, K.: Morphological operations and projection profiles based segmentation of handwritten Kannada document. Int. J. Appl. Inf. Syst. (IJAIS) **4**(2249-0868), 13–19 (2012)
21. Patel, M.S., Reddy, S.L., Sandyal, K.S.: An improved handwritten word recognition rate of South Indian Kannada words using better feature extraction approach. In: Proceedings of the 3rd International Conference on Frontiers of Intelligent Computing: Theory and Applications (FICTA), pp. 553–561 (2015)
22. Patel, M.S., Rohith Kumar, S.C., Reddy, L.: Offline Kannada handwritten word recognition using locality preserving projection (LPP) for feature extraction. Int. J. Innov. Res. Sci. Eng. Technol. **4**, 5078–5086 (2015)
23. Sureka, T., Swetha, K.S.N., Arora, I., Mamatha, H.R.: Word recognition techniques for Kannada handwritten documents. In: 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT), Kanpur, India, pp. 1–7

Chapter 19
Future-Oriented Smart Village Model Using Advance IoT Sensors Based Technology

Dushyantsinh Rathod, Jaykumar Dave, Vijaykumar Gadhavi, Ramesh Prajapati, and Ghanshyam Parmar

Abstract The thought of Web of Things (WoT) is the rise of innovation. The method of reasoning of great working of web, versatile, data, and interactions innovation. It gives an empowers different gadgets in a system to impart and associated with one another to play out their work in a pleasing manner. The rising populace of the world makes it major to energize the urban areas and towns to work in a sharp manner. Hence, the idea of Canny urban areas appeared. It gives a far-reaching see concerning change inside the personal satisfaction in towns. It is also described the usage of different sensors for different purposes with its applications.

19.1 Introduction to IoT

The epic difficulties in acknowledgment of a common headway that screens and planning the entirety of the town establishment and organizations to utilize the aggregate bits of knowledge. The headway of an IoT-based adroit town joins cloud-based arrange which can give a virtual establishment to deal with and composed the examination instruments watching equipment, limit, and perception stage inside the IoT combining which expected to sharp charging and data investigation in imperativeness organization. Waste assortment structure updated with cloud-based IoT organizations which enable vigorous arranging and coordinating in a waste assortment system has all the earmarks of being a gainful framework. The capacities of the Web of Things (WoT) are evidently endless with potential outcomes. The visualized segments in each bunch are recorded beneath:

D. Rathod (✉) · J. Dave
Sankalchand Patel College of Engineering, Visnagar, Gujarat, India

V. Gadhavi
Swaminarayan Institute of Technology, Gandhinagar, Gujarat, India

R. Prajapati
L.J. Institute of Engineering and Technology, Ahmedabad, Gujarat, India

G. Parmar
Natubhai V. Patel College of Pure and Applied Sciences, Vallabh Vidhyanagar, Gujarat, India

T. Senjyu et al. (eds.), *IOT with Smart Systems*, Smart Innovation, Systems and Technologies 251, https://doi.org/10.1007/978-981-16-3945-6_20

- Mobile clinical administration services.
- Education framework.
- Sanitation system.
- Water executive framework.
- Road framework inter-town availability.
- Lighting control.
- Fleet the board.
- Digital literacy and people service centers.
- Economic action-based skill advancement program.
- Agro-processing, agri-services, storage, and advancement.
- Smart cultivating.

Figure 19.1 addresses the for the most part piece graph of the sharp town structure. The immense IT establishment is required by common improvement next to the epic budgetary back which is to be combined. Sensors, a large number of arranging equipment and registering devices are worked in this complex sort out. In the event of wise water framework organization system, each field must be fitted with a sensors and data control unit which are extremely capable and strong [1–3].

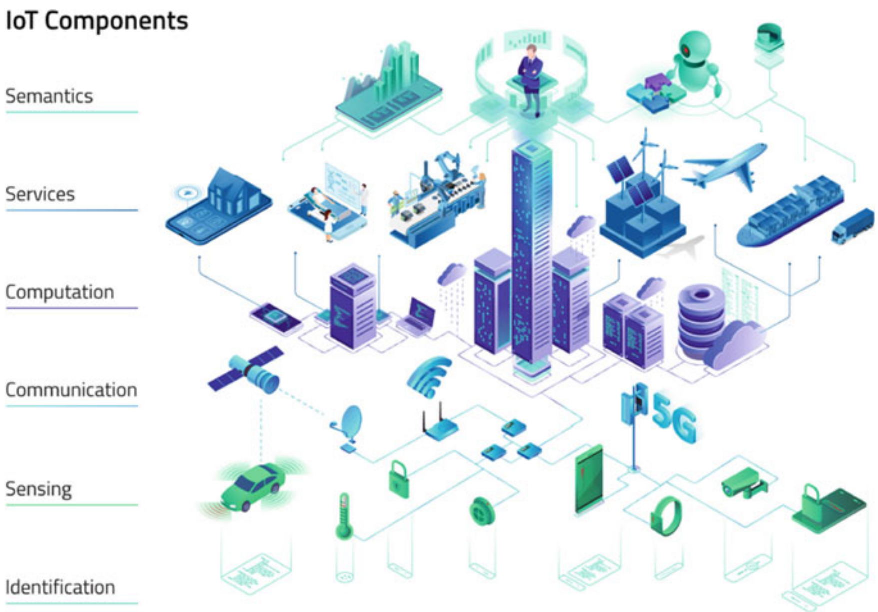


Fig. 19.1 IoT architectural components

Fig. 19.2 Smart village model

19.2 Innovations in Villages

A colossal pace of our masses lives in urban areas. Along these lines, the experts just as the legislatures focus the possible objective is to attain shrewd house, environment architecture, instruction, reconnaissance architecture, and keen farm among others. Figure 19.2 summarizes innovations uses and areas of intrigue in the village [4].

19.3 Types of IoT Sensors

See Fig. 19.3.

A. ***Accelerometer Sensor***
An accelerometer is a gadget that estimates changes in gravitational speeding up in a gadget and to detect movement in numerous ways, an accelerometer must be planned with multi-pivot sensors or different straight hub sensors.

B. ***Proximity Sensor***
This specific kind of light-sensing is generally applying in many applications which requesting guard with adequacy. Diverse applying regions such a sensing are fight disclosure, count many number of things, estimating aggregate insurgency, texture revelation, estimating development heading, stopping sensors, and so on [5].

C. ***Infrared Sensor***
Infrared light recognizes infrared waves so as to detect a couple of attributes of specific articles. They can additionally degree warm transmission.

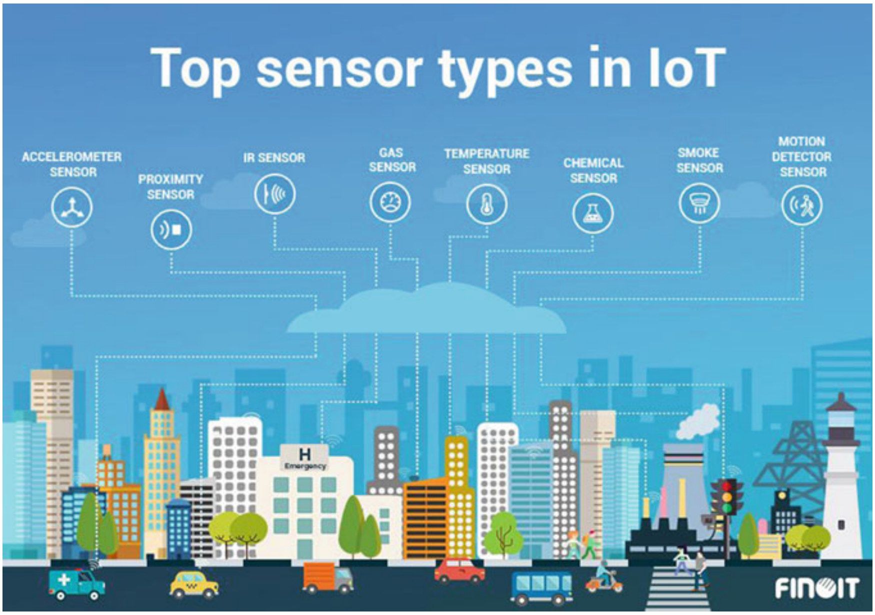


Fig. 19.3 Top sensors in IoT

D. ***Temperature Sensor***
Temperature lights are obliging in distinguishing estimating warm imperativeness. Makers in [6] used temperature lights for the checking of ecological states of nature [5].

E. ***Chemical Sensor***
A synthetic sensor is an informative contraption used to gauge the compound piece of nature. Examine quality checking should be possible utilizing a distant compound sensor organize by observing synthetic tufts inside the earth [5, 7].

F. ***Motion Detection Sensor***
A development finder might be a contraption used to detect all the active and physical improvement inside nature. An application for watching homes inside the nonappearance of property [5].

G. ***Gas Sensor***
Gas sensors are electronic gadgets that identify and recognize diverse sorts of gasses. They are commonly utilized to identify poisonous or hazardous gasses and degree gas concentration.

H. ***Smoke Sensor***
A smoke finder may detect smokes, fire, and flame or a few finders on the off chance that there are different smoke locators interlinked.

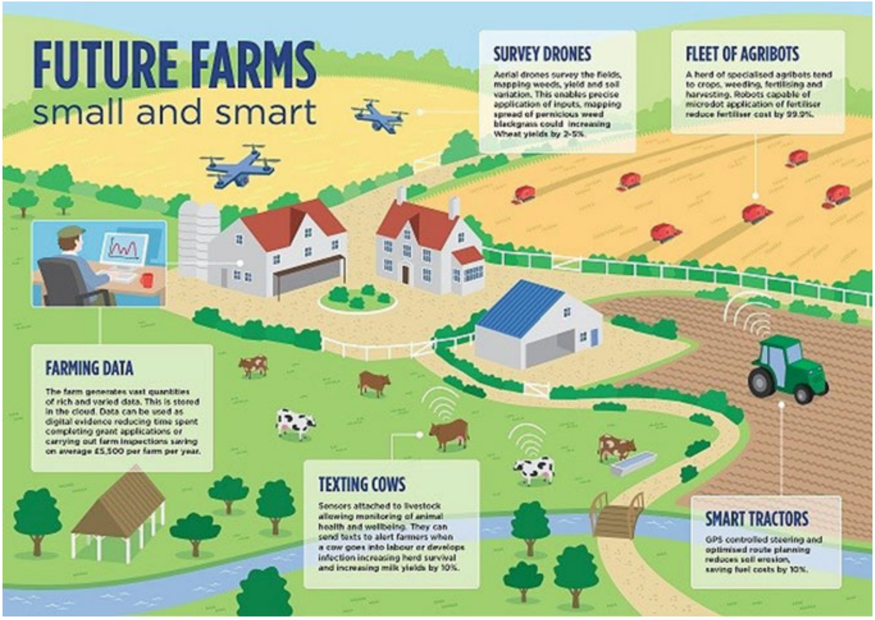


Fig. 19.4 Smart future farming

19.4 Applications of IoT in Smart Village Model

19.4.1 Smart Future Farms

As farming is spine everything being equal, the farmers should advantage the premier from the structure of IoT and Shrewd towns. Additionally, lights recognize developing fruits characteristic items then alarm vehicle benefit suppliers to maintain a strategic distance from any delays. From that point, appropriate courses of action can be made within the showcase to offer the deliver [4]. These days ranchers can use IoT to enhance their cultivated productivity such as water system, fertilization, collecting information, and climate estimate by checking with sensors to progress their choice making [8] (Fig. 19.4).

19.4.2 Smart Irrigation System

Keen water system frameworks can make utilize of sensors within the areas and farther adherent information to guarantee the ideal utilize of accessible water assets. Provincial water framework water is getting the chance to be scared not so to speak in completely dry and semidry areas yet in addition inside the tall precipitation districts.

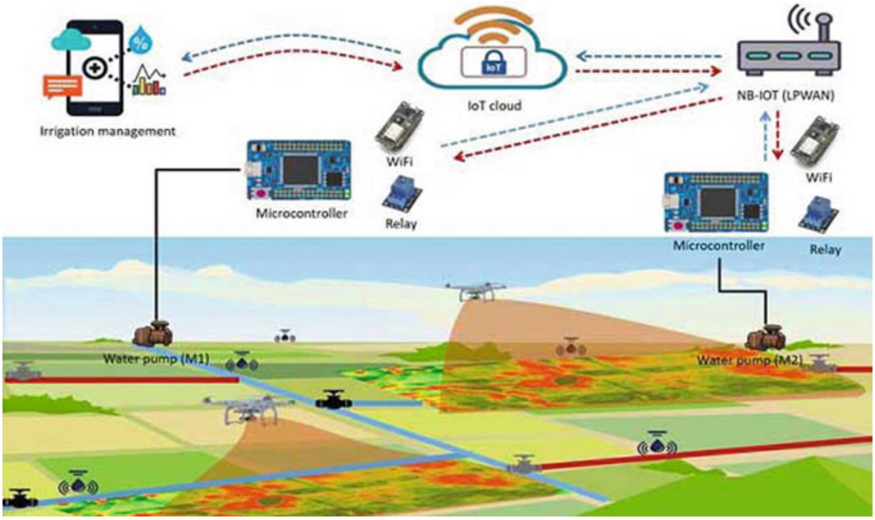


Fig. 19.5 Smart irrigation system

Since of the lopsided scattering of precipitation plan the vast majority harvests [9], current-day stream water framework (SDI) has a pivotal influence for insightful use of water according to the essential of the alter. In any case, this structure despite everything needs to keep up by the administrators [8] (Fig. 19.5).

19.4.3 Advance Education System

Guidance is the principal infers to execute all the degrees of progress throughout everyday life. Educating people approximately utilize the modern innovations motivate superior execution. Dealing with children and youngsters gets to be less requesting show. learning premium the greater part of the kids and can offer to help them learn in a keen route as opposed to scrutinizing the reading material inside the homerooms. Web of Things converts schools into Savvy cluster [10] (Fig. 19.6).

19.4.4 Smart Healthcare System

Keen wellbeing administrations are required to make strides the personal satisfaction inside towns. The town apothecary and clinics require progressed gadgets which are related to one another (Fig. 19.7).

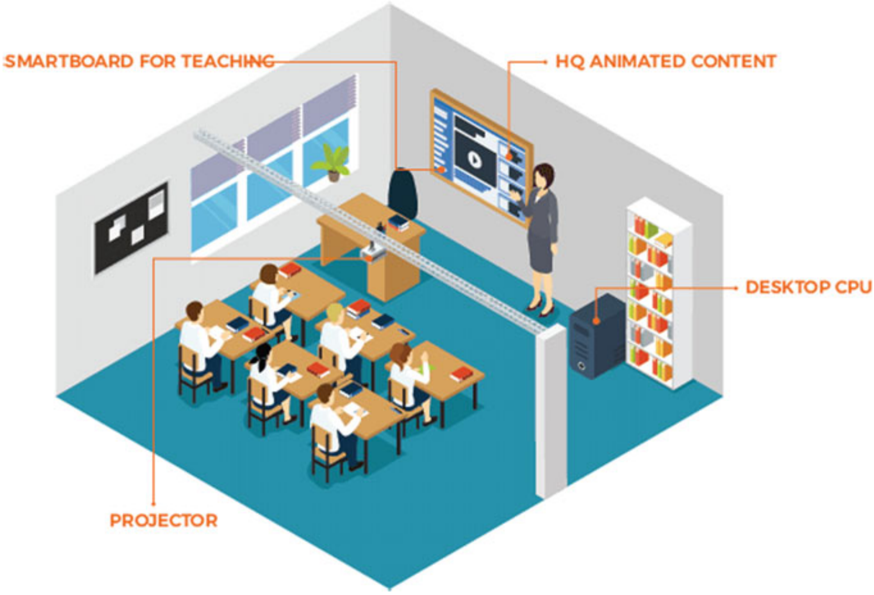


Fig. 19.6 Smart education system

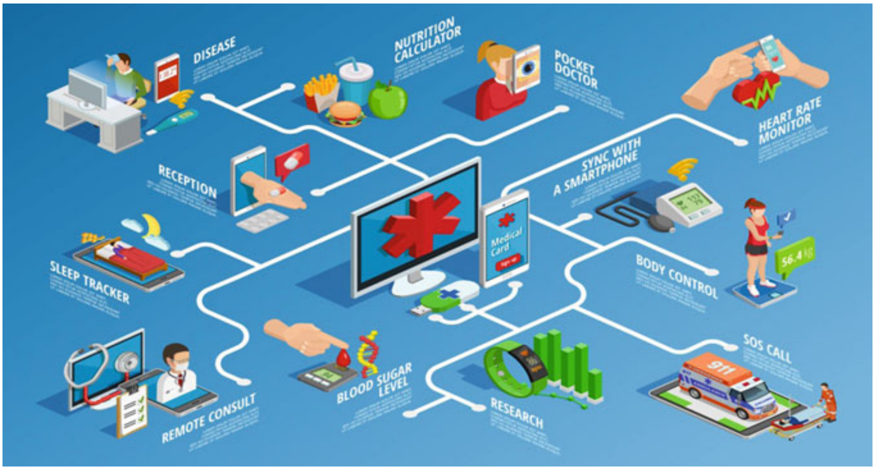


Fig. 19.7 Smart healthcare system

Another fruitful commitment of IoT in human services system is "Audemix gadget," This device catches and gathers the patient's data for review in an agreeable manner through powered by voice order for reducing specialist's work [11] (Fig. 19.8).

Fig. 19.8 Smart Skydio Drones

19.4.5 Advance Dairy System

The assistant control of an enormous many farmers are raising cows dairy things. The use of sensors inside the territories can abstain from crafted by management by a human, and it tends to be done distantly by the agriculturists. It is the strategy of recognizing the most excellent classifier from each unlabeled and labeled data. By utilizing unlabeled information, it exchanges tall execution of classification (Fig. 19.9).

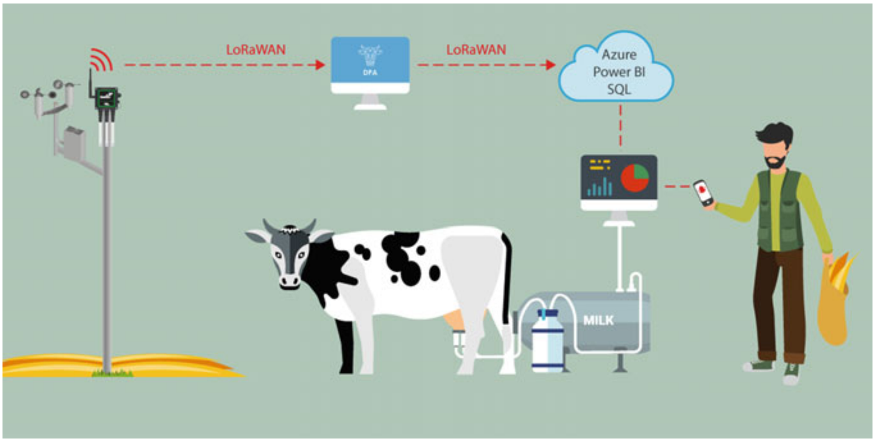


Fig. 19.9 Smart dairy system

Fig. 19.10 Smart home and surveillance system

19.4.6 Smart Home and Security System

Protection could be a significant worry in towns. Because of these factors, keen perception systems are required in towns. These will take a shot at the premise created by sensing nearby crisis catches found in a few pieces of the town. The information can too be analyzed to maintain a strategic distance from such episodes within the future (Fig. 19.10).

19.4.7 Smart Drinking Water and Canal Management System

In spite of monitor water steam, there is prompt necessity of quality checking as well as affirmation of spillage (in case any). For this purpose, a remote framework has to be built up to guarantee client inviting framework; thus, Web of Things (WoT) shows up as a boon to the framework.

Presently days, in canal water conveyance framework, there is a parcel of debasement at water conveyance focuses. To dodge such debasement, we must be creating a computerized system, which is able to grant the correct conveyance of water to the agriculturists and dodge the debasement. As per the necessity indicated by ranchers, the computerized framework will open the valve for given time span as per rancher ask, and after the time bound, the valve will be near consequently. It ought to be done by keen IOT technology [12] (Figs. 19.11 and 19.12).

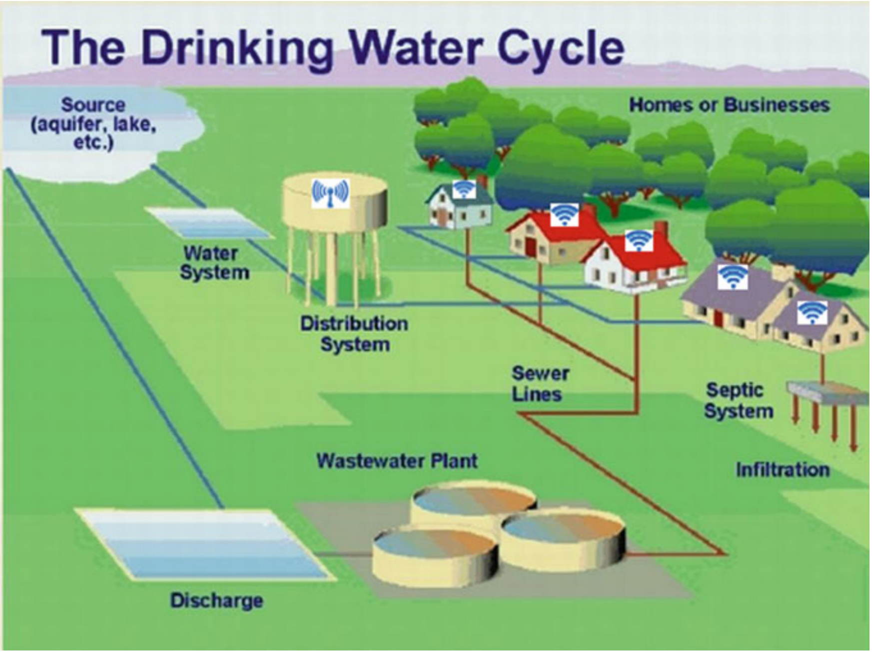


Fig. 19.11 Smart drinking system

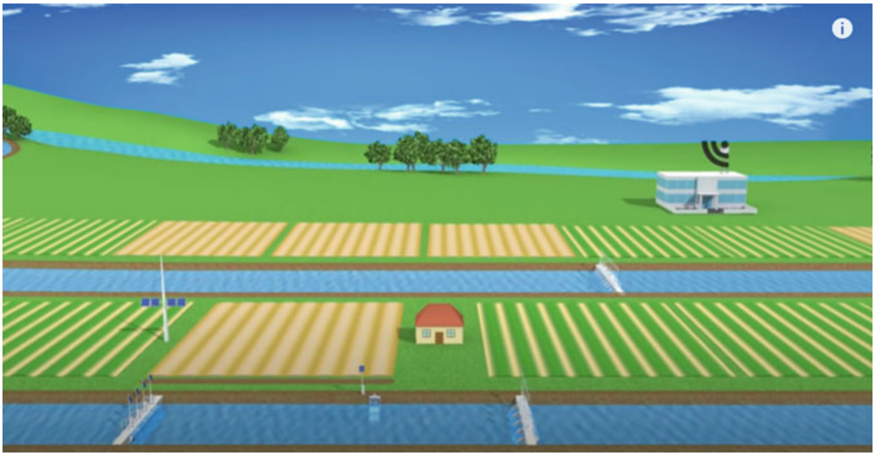


Fig. 19.12 Smart canal system

Table 19.1 Smart IoT real-time applications and light sensors

IoT applications	Type of sensors
Smart city	Velocity, light, accelerometer, position, temperature, proximity, humidity, pressure, infrared
Smart environment	Light, temperature, humidity, chemical, gyroscope, bio-sensors, chemicals, accelerometer, optical
Smart water	Temperature, humidity, occupancy, water quality
Smart building	Light, accelerometer, chemical, gyroscope, magneto
Smart health	Light, gyroscope, biosensors, chemicals, magneto, accelerometer, pressure
Smart home	Light, gyroscope, biosensors, chemicals, magneto, accelerometer, temperature, proximity, position, infrared
Smart transport	Gyroscope, pressure, chemicals, magneto, accelerometer, temperature, motion, infrared
Smart security	Light, gyroscope, chemical, magneto, accelerometer, temperature, infrared
Smart agriculture	Temperature, humidity, water quality, chemical, proximity, position

19.5 IoT Real-Time Applications and Light Sensors

Sensor lights are used in full IoT real-time applications. Subsequent to breaking down various kinds of lights sagacious uses the IoT [5] (Table 19.1).

19.6 Conclusion and Future Scope

Internet of things is the concept in which the virtual world of information technology connected to the real world of things. The technologies of Internet of things such as RFID and sensor make our life become better and more comfortable in village, so that in the future the village will become the smart village using IoT-based sensors technology. The limitation is only the expenses and maintenance.

References

1. Davies, A.: IOT, smart technologies, smart policing: the impact for rural communities. Smart Village Technology. Springer (2020). ISBN: 978-3-03-037794-6. https://doi.org/10.1007/978-3-030-37794-6
2. Soorjani, M., Bhagyeshree, M., Pragya, N., Siddhartha, S., Shrikant, P.: Smart village initiative: an overview. Smart Village Technology. Springer (2020). ISBN: 978-3-03-037794-6. https://doi.org/10.1007/978-3-030-37794-6
3. Gayatri, N., Kumar, L.: Implementation of IOT based smart village for the rural development. IJMET, SCOPUS INDEX (2017). ISSN: 0976-6359

4. Kaur. K.: The idea of smart villages based on IOT. IRJET (2016). ISSN: 2395-0056
5. Deepti, S., Nasig, G.: Smart sensors: analysis of different types of IOT sensors. In: IEEE ICOEI-2019 (2019). ISBN-978-1-5386-9439-8
6. Zafar, S., Miraj, G., Baloch, R., Murtaza, D., Arshad, K.: An IoT based real-time environmental monitoring system using Arduino and cloud service. Eng. Technol. Appl. Sci. Res. **8**(4), 3238–3242 (2018)
7. Tapashetti, A., Vegiraju, D., Ogunfunmi, T.: IoT-enabled air quality monitoring device: a low cost smart health solution. In: Proceedings of 2016 IEEE Global Humanitarian Technology Conference (GHTC), pp. 682–685. IEEE (2016). Available Online: https://doi.org/10.1109/GHTC.2016.7857352
8. Ahmad, V., Mehmood, N.: Smart farming: an overview. Smart Village Technology. Springer (2020). ISBN: 978-3-03-037794-6. https://doi.org/10.1007/978-3-030-37794-6
9. Pereira, L.S., Oweis, T., Zairi, A.: Irrigation management under water scarcity. Agric. Water Manag. **57**, 175–206 (2002). https://doi.org/10.1016/S0378
10. Giovannella, C., et al.: Scenarios for active learning in smart territories. IxD&A **16**, 7–16 (2013)
11. Srabanti, C., Pransenjit, D., Sauvik, P.: IoT foundations and it's application. IoT and Analytics for Agriculture. Springer (2020). ISBN: 978-981-13-9177-4. https://doi.org/10.1007/978-981-13-9177-4
12. Antonio, S., Karla, P.: Water management through smart water management system. Smart Village Technology. Springer (2020). ISBN: 978-3-03-037794-6. https://doi.org/10.1007/978-3-030-37794-6

Chapter 20
Impact of HR Matrices on HR Analytics and Decision Making

Maria Afzal and Amirul Hasan Ansari

Abstract For several years now, HR analytics has been a widely researched field. Researchers around the world have researched and evaluated various variables that impact HR analytics. This is important since this approach assists in HR decision making and better selection. We identify basic parameters in this paper, namely organizational nature and organizational tenure, in order to research their effect on HR analytics. The parameters are further divided into sub-parameters, such as talent management, quantitative literacy and adaptation of innovation. Data is collected via a questionnaire from the Delhi/ NCR, India, IT market. To conduct data analysis, the paper utilizes the SEM model. The suggested structure satisfies research hypotheses and research questions after evaluating the data.

20.1 Introduction

In a number of ways, HR analytics has been named and identified. Alternatives such as talent analytics, predictive analytics, people analytics and workforce intelligence are used in its labeling. First, the concepts suggest that HR analytics is not simply about metrics intended to measure the human resource performance or effectiveness of an enterprise. Second, the incorporation of information from both inside and outside the company consists of HR analytics. Third, the use of information technology to collect, manipulate and publish information is involved. Finally, it includes connecting human resource decisions to the organization's overall priorities and results. The presence of vast volumes of data due to the nature of the information era is among the characteristics of the current business climate. In particular, since it is typically raw, the organization and analysis of this information are a major challenge. The need to turn this knowledge into usable data has contributed to the advent and rapid growth of HR analytics. In an enterprise, the capacity and resourcefulness of HR have become major components of business performance. Competing firms are therefore engaged in a big fight for HR expertise. Different techniques and

M. Afzal (✉) · A. H. Ansari
Jamia Millia Islamia, New Delhi, India

T. Senjyu et al. (eds.), *IOT with Smart Systems*, Smart Innovation, Systems and Technologies 251, https://doi.org/10.1007/978-981-16-3945-6_21

methods of assessing HR have arisen to recognize the best HR talent. This principle of HR evaluation, ranking and aggregation of intelligence is known as HR analytics [1–5]. The effect of organizational parameters on HR analytics is studied in this paper. Contributions from the author are as follows:

- Paper studies the impact of parameters such as organization's nature and organization tenure on HR analytics.
- Detailed data is collected through well-defined questionnaires from IT sector of Delhi/NCR, India.
- The paper uses SEM model for data analysis, and we finally draw a conclusion on the impact of organization's parameters on HR analytics.

The chapter is subdivided into five parts. Introduction is explained in Sect. 20.1. The research hypothesis is explained in Sect. 20.2, research methodology is explained in Sect. 20.3, data analysis is carried out in Sect. 20.4, and the paper is concluded in Sect. 20.5.

20.2 Research Objectives and Hypotheses

The aim of the study is to establish the relationship between organizational features and HR analytics, while exploring the dimensions of the use of HR analytics in organizational features. Furthermore, the relationship between the implementation of HR analytics and the pattern of features of the company is also studied. It analyzed the dimensions of the function pattern of the company, i.e., talent management, quantitative literacy and adaptation of creativity to HR analytics implementation.

To give answer to the research questions mentioned in Sec 20.2, below mentioned objectives and hypotheses have been formulated:

1. To examine the relationship of organization feature on application of HR analytics and its dimensions among the different level of managers at organization separately.

   Corresponding Hypotheses:
   Hypothesis 1: Talent management has a positive relationship on application of HR analytics.
   Hypothesis 2: Quantitatively literacy values have a positive relationship on application HR analytics.
   Hypothesis 3: Innovation adaption has a positive relationship on application of HR analytics.

2. To examine the relationship decision-making pattern on application of HR analytics and its dimensions among the different level of managers at organization separately.

Corresponding Hypotheses:
Hypothesis 4: Effort expectancy has positive effect on application of HR analytics.
Hypothesis 5: Fear appeal has a negative relationship on application of HR analytics.
Hypothesis 6: Self-efficacy has positive relationship on application of HR analytics.

3. To determine the relationship of organizational tenure on application of HR analytics

   Corresponding Hypothesis:
   Hypothesis 7: Older the organizational tenure, greater the application of HR analytics.

4. To determine the relationship of organization features on application of HR analytics.

   Corresponding Hypothesis:
   Hypothesis 8: Better the organization features, greater application of HR.

5. To study the relationship of organization's size on application of HR analytics.

20.2.1 Proposed Research Question

1. What different factors affect use of HR analytics in an organization and up to what extent the human resource analytics has been used by professionals?
2. Do the 'talent management', 'quantitatively literacy' and 'innovation adaption' related to the 'use of statistical methods', 'functions dedicated to HR', 'data collection' and 'degree of application' of HR analytics?
3. Do the HR analytics, 'use of statistical methods', 'functions dedicated to HR', 'data collection' and 'degree of application' anyway related to the decision-making pattern of managers varied among the different levels of organization?
4. Does the organization tenure help in inculcating the 'attributes' of authentic leadership?
5. Do the organization tenure such as 'effort expectancy', 'fear appeal' and 'self-efficacy' affect the relationship between HR analytics and decision making? A research model is shown in Fig. 20.1.

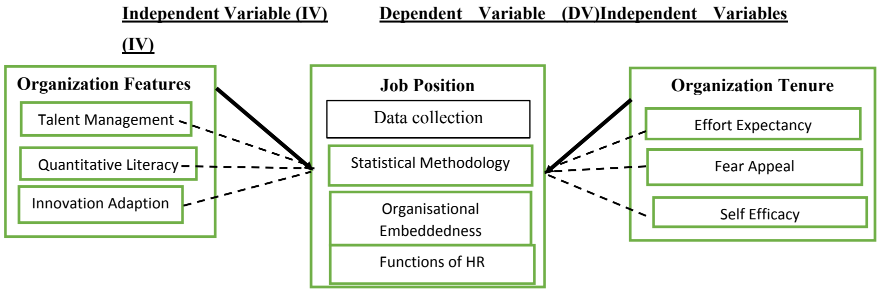


Fig. 20.1 Proposed research model

20.3 Research Methodology

20.3.1 Data

To gather more data on the results from the review of records, a questionnaire was circulated. The questionnaire includes questions that reflect on the effect on HR analytics of genuine leadership and decision-making trends.

20.3.2 Questionnaire Design, Analysis and Measurement of Variables

Although there are several intangible HRM findings, some scholars have established the key ones (Huang, H. C., and Chao, R. R. 2009). The same methodology has been taken by this report. Data collection is done by using questionnaire with targeted professionals created and implemented in accordance with the nature of the data needed for this analysis. The target population for the study is IT sector HR managers in Delhi, NCR, India. For data analysis, the thesis utilizes snow ball sampling. In this report, the sample size was chosen based on the criteria of SEM, which is the primary form of data analysis used in current research. Based on the complexity of the proposed conceptual model, a sample size of 250 is therefore considered acceptable for this analysis, which was drawn from the actual population of IT sector HR managers. In each of Delhi's IT sector organizations, NCR in India (the target population), the researcher explained the intent of the study to the human resources director and gave them questionnaires to be distributed among middle/line managers and other staff. After ten working days, the questionnaires were collected.

20.3.3 Data Analysis Tools

Exploratory Factor Analysis (EFA) Exploratory factor analysis (EFA) puts together and categorizes inter-correlated variables under a common factor. The goal of the factor analysis is the reduction of dimensions, which reduces the dimensionality of the original space and produces a new space with less dimensions. Factor analysis not only provides the possibility of obtaining a comprehensible view of the results, but also the probability of using the output in consequential studies. In this analysis, by calculating the number of factors needed to explain correlations between the products, the EFA is used to assess the uniformity of all factors [6–10].

Confirmatory Factor Analysis (CFA) CFA tests and develops proof of adequate validity of the construct. It also checks whether the research construct's measurements are compatible or suit the validity of the research model. CFA is an initial step for a full test of a structural model as a special kind of factor analysis. It consists of a sequence of different research steps. The establishment of construct validity, convergent and discriminant validity and goodness of fit include these levels. The validity of the construct guarantees the suitability of inferences taken on the basis of observations to verify whether the expected construct is assessed by the test. To test the validity of the construct, there are two aspects: convergent validity and discriminant validity. The things that are indicators of a specific construct should converge or share a high share of variance in general under convergent validity. The construct reliability and the average variance extracted will be determined by (AVE). Discriminant validity is the degree to which one concept is truly distinct from others. By testing if the square root of AVE is greater than the coefficient of correlation between the construct discussed and other constructs, it can be assessed [11–13].

20.4 Data Analysis and Outcome

20.4.1 Results of Descriptive Statistics

The results of the data analysis are developed in detail in this chapter. In order to test the construct's validity, discriminant and convergent validity are evaluated. It also describes the study of data by structural equation modeling (SEM) using SPSS17/Amos18. In addition, to determine if the proposed models and data match well together, the proposed model is evaluated by the fit indices. Last, it also addresses an alternative model (Tables 20.1, 20.2, 20.3 and 20.4).

From all respondents, the highest mean score was statistical methodology with lowest, followed by data collection and effort expectancy. This standard deviation (S.D. = 1.549) indicated that felt obligation was one of the most important factors of employee attitude and behavior.

Table 20.1 Descriptive statistics of observed variables

Construct	Observed variables	Min	Max	Mean	S.D.
Organization features (OF)	Talent management	1	6	3.22	1.234
	Quantitative literacy	1	6	4.22	1.543
	Innovation adaption	1	6	3.11	1.096
Job position (JP)	Data collection	1	6	3.44	1.342
	Statistical methodology	1	6	3.77	1.432
	Organizational embeddedness	1	6	5.11	1.343
	Functions of HR	1	6	5.65	1.567
Organization tenure (OT)	Effort expectancy	1	6	4.41	2.231
	Fear appeal	1	6	4.17	1.745
	Self-efficacy	1	6	4.76	1.876

Table 20.2 Summary of the mean of all constructs

Construct	Mean	S.D.
Talent management (TM)	4.54	1.231
Quantitative literacy (QL)	3.74	1.562
Innovation adaption (IA)	4.12	1.634
Data collection (DC)	3.14	1.540
Statistical methodology (SM)	4.87	1.547
Organizational embeddedness (QE)	3.32	1.648
Functions of HR (FHR)	4.16	1.549
Effort expectancy (EE)	3.09	2.451
Fear appeal (FA)	4.65	1.387
Self-efficacy (SE)	4.43	1.368

Table 20.3 Constructs and reliability analysis (Cronbach's alpha)

Constructs	Cronbach's alpha
Talent management (TM)	0.878
Quantitative literacy (QL)	0.812
Innovation adaption (IA)	0.904
Data collection (DC)	0.716
Statistical methodology (SM)	0.808
Organizational embeddedness (QE)	0.798
Functions of HR (FHR)	0.865
Effort expectancy (EE)	0.834
Fear appeal (FA)	0.756

Table 20.4 Construct measurement model

Factor	Standardized loading[a]
Talent management (TM)	0.542 (5.104)
Quantitative literacy (QL)	0.645 (5.657)
Innovation adaption (IA)	0.456[b]
Data collection (DC)	0.867 (9.341)
Statistical methodology (SM)	0.796[b]
Organizational embeddedness (QE)	−0.723 (−5.451)
Functions of HR (FHR)	0.434 (7.657)
Effort expectancy (EE)	−0.756 (−5.154)
Fear appeal (FA)	−0.638 (−5.965)
Self-efficacy (SE)	−0.553 (−4.659)

'a' signifies composite reliability
'b' is average variance extracted (AVE)

20.4.2 *Reliability Analysis*

The Cronbach's alpha of these constructs suggested that since all constructs were higher than 0.7, ranging from the lowest reliability of 0.716 (performance) to the highest reliable construct of 0.878, all constructs were highly reliable.

20.4.3 *Convergent Validity and Discriminant Validity*

20.4.3.1 Convergent Validity

Factor loadings in the range of 0.3–0.4 are deemed to reach the minimum threshold for structure understanding, according to Hair [14]. The association between an initial variable and its factor is expressed by a factor loading. The sample size needed for significance will be 250 with a loading factor of 0.35. The loading factor value must equal or exceed 0.30, according to Wanichabuncha [15].

Reliability is also a measure of convergent validity, as Hair [14] has suggested. In combination with the SEM model, build reliability (CR) value is also used. The calculation of reliability is that 0.7 or higher implies good reliability. Reliability between 0.6 and 0.7 might be appropriate given that other construct validity indicators of the model are fine.

20.4.3.2 Measurement Model 1 (CFA)

The author makes use of a reflective CFA model with constructs. Self-awareness, moral and ethics, openness, balanced processing, diligence, hyper-vigilance, procrastination, buck passing, data collection, statistical methodology, organizational integration and HR functions were observed in the following constructs (Fig. 20.2).

NFI = 0.903, TLI = 0.945, RMSEA = 0.045.

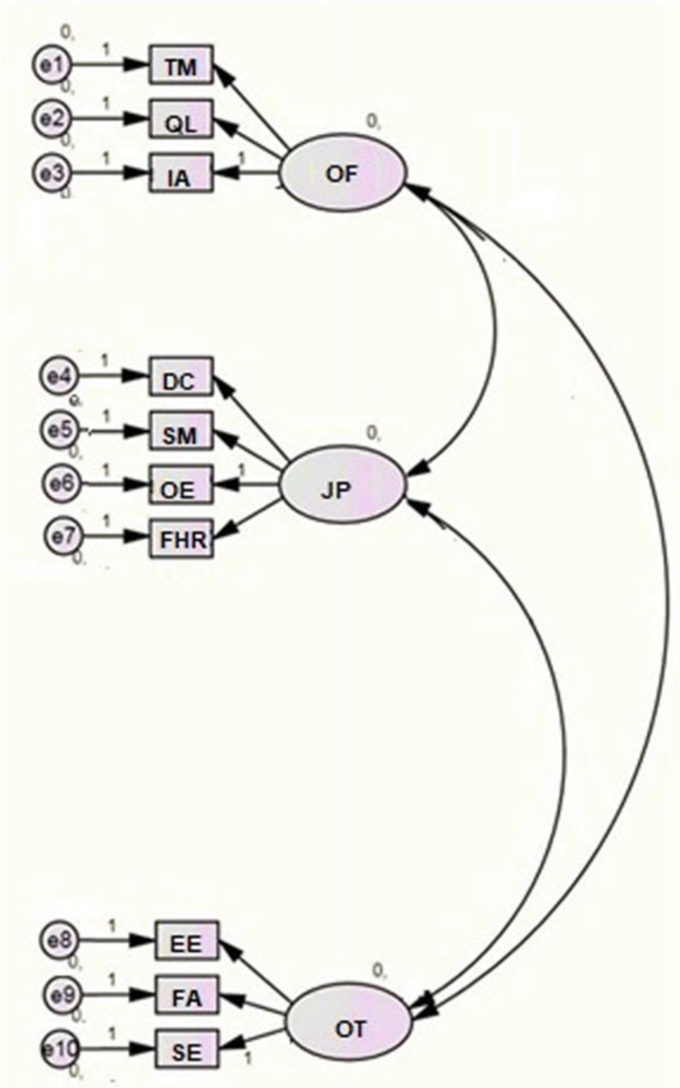


Fig. 20.2 Construct measurement model

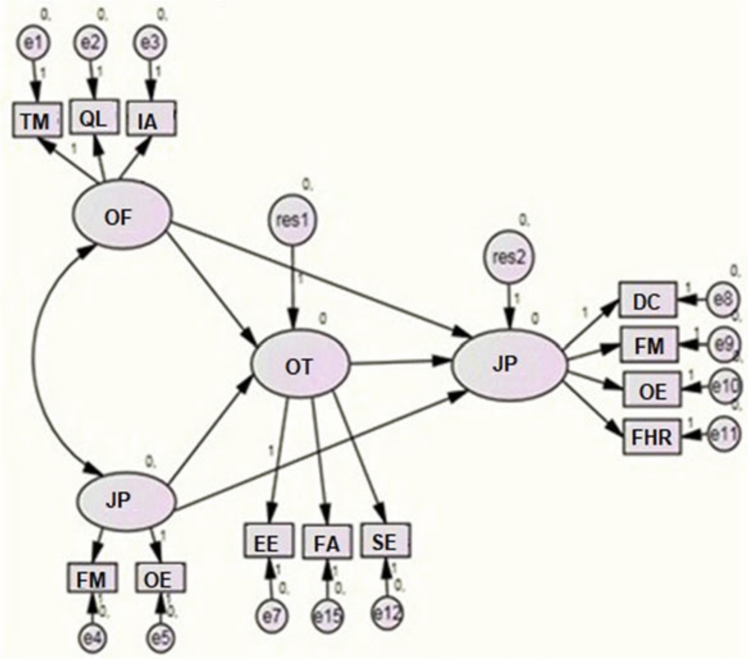


Fig. 20.3 Structural model

First-ordered frameworks have been tested for self-awareness, moral and ethics, openness, balanced processing, vigilance, hyper-vigilance, procrastination, buck passing, data collecting, statistical methods, HR functions and organizational embeddedness (Fig. 20.3).

Hypothesis testing and approval are shown in Table 20.5.

20.5 Conclusion

This paper analyzed the factors or structures influencing HR analytics and described them. To explain how the knowledge matches with the proposed models, the hypotheses have been checked with empirical evidence. According to the results of this analysis, both hypotheses were endorsed by the empirical test one, based on the data on the respondents. Effort expectancy and self-effectiveness affected talent management, creativity adaptation and organizational embeddedness. Moreover, the control variables of demographics were also included in this research, i.e., age, gender and teaching. The author used the structural equation modeling methodology in order to evaluate the model, where all relationships in the model were evaluated simultaneously. Therefore, the findings of the study fulfilled this aim.

Table 20.5 Hypothesis acceptance

H1: Talent management has a positive relationship on application of HR analytics	A significant path coefficient = 0.663 ($p < 0.06$)	Accepted
H2: Quantitively literacy values have a positive relationship on application HR analytics	A significant path coefficient = 0.660 ($p < 0.06$)	Accepted
H3: Innovation adaption has a positive relationship on application of HR analytics	A significant path coefficient = 0.623 ($p < 0.06$)	Accepted
H4: Effort expectancy has positive effect on application of HR analytics	A significant path coefficient = 0.658 ($p < 0.06$)	Accepted
H5: Fear appeal has a negative relationship on application of HR analytics	A significant path coefficient = 0.656 ($p < 0.06$)	Accepted
H1: Self-efficacy has positive relationship on application of HR analytics	A significant path coefficient = 0.642 ($p < 0.06$)	Accepted
H7: Older the organizational tenure, greater the application of HR analytics	A significant path coefficient = 0.622 ($p < 0.06$)	Accepted
H8: To study the relationship of organization's size on application of HR analytics	A significant path coefficient = 0.644 ($p < 0.06$)	Accepted

References

1. Bassi, L., Carpenter, R., McMurrer, D.: HR analytics Handbook, 1st edn. McBassi and Company, Amsterdam (2012)
2. Bradley, S.L.: An exploratory study of the role of the human resource information system professional. ProQuest Dissertations (Doctoral dissertation), 10635957 (2017)
3. Chaturvedi, V.: Talent analytics as an indispensable tool and an emerging facet of HR for organization. FIIB Business Rev. **5**(3), 13–20 (2016)
4. Falletta, S.: In search of HR intelligence: evidence-based HR analytics practices in high performing companies. People Strategy **36**(4), 28–37 (2014)
5. Fitz-enz, J., Mattox, J.R.: Predictive Analytics for Human Resources. Wiley, Hoboken, NJ (2014). https://doi.org/10.1002/9781118915042
6. Giuffrida, M.: HR can't ignore big data. Talent Manag. Mag. **46**, 16–19 (2013)
7. Miles, A.: Measuring customer behavior and profitability: using marketing analytics to examine customer and marketing behavioral patterns in business ventures. Acad. Mark. Stud. J. **18**(1), 141–165 (2014)
8. Mihalcea, A.D.: Employer branding and talent management in the digital age. Manag. Dynamics Knowl. Econ. **5**(2), 289–306 (2017). https://doi.org/10.25019/MDKE/5.2.07
9. Rafter, M.V.: Companies making HR big data work. Workforce Magazine (2013)
10. Spahic, J.: Exploring HR intelligence practices in fortune 1000 and select global organizations. ProQuest Dissertations (Doctoral dissertation), 10010853 (2015)
11. Sharma, A., Sharma, T.: HR analytics and performance appraisal system. Manag. Res. Rev. **40**(6), 684–697 (2017). https://doi.org/10.1108/MRR-04-2016-0084
12. Van den Heuvel, S., Bondarouk, T.: The rise (and fall?) of HR analytics: a study into the future application, value, structure, and system support. J. Organizational Effectiveness: People Perform. **4**(2), 157–178 (2015). https://doi.org/10.1108/JOEPP-03-2017-0022

13. Vargas, R.: Adoption factors impacting human resource analytics among human resource professionals. ProQuest Dissertations (Doctoral dissertation), 3716458 (2015)
14. Hair, J., et. al.: A primer on partial least squares structural equation modeling (PLS-SEM) (2010)
15. Wanichbuncha, K.: Advanced statistic for SPSS. Chulalongkorn University Publishing (2003)

Chapter 21
A Neural Network Based Customer Churn Prediction Algorithm for Telecom Sector

Kamsali Mani Teja Achari, Sumitra Binu, and K. T. Thomas

Abstract For telecommunication service providers, a key method for decreasing costs and making revenue is to focus on retaining existing subscribers rather than obtaining new customers. To support this strategy, it is significant to understand customer concerns as early as possible to avoid churn. When customers switch to another competitive service provider, it results in the instant loss of business. This work focuses on building a classification model for predicting customer churn. Four different deep learning models are designed by applying different activation functions on different layers for classifying the customers into two different categories. A comparison of the performance of the different models is done by using various performance measures such as accuracy, precision, recall, and area under the curve (AUC) to determine the best activation function for the model among tan*h*, ReLU, ELU, and SELU.

21.1 Introduction

Customers are very important for an organization to achieve profitable outcomes. Churn obstructs the profitable growth of a business. In the telecommunication industry, customer churn is a major problem and it might lead to bankruptcy of the industry. Retaining existing customers for business growth is imperative because the cost of acquiring new customers is much more than maintaining existing ones. When customers switch to another competitive service provider, it results in the instant loss of business for the service provider. Customers classically choose to switch providers when they are no longer satisfied with their existing service provider [1]. Hence, it is

K. M. T. Achari (✉) · S. Binu · K. T. Thomas
Christ (Deemed to be University), Lavasa, India
e-mail: kamsali.achari@science.christuniversity.in

S. Binu
e-mail: sumitra.binu@christuniversity.in

K. T. Thomas
e-mail: thomas.kt@christuniversity.in

T. Senjyu et al. (eds.), *IOT with Smart Systems*, Smart Innovation, Systems and Technologies 251, https://doi.org/10.1007/978-981-16-3945-6_22

crucial for the telecom sector to develop some means for predicting customers who are likely to churn and take measures to retain them. So to retain the customers, it is necessary to build a churn prediction model. In this paper, four artificial neural network models have been proposed.

The paper is organized as follows: The first section explains the notion of churn prediction model; the second part describes the literature survey on churn prediction models. The data preprocessing steps have been explained in the third section. The methodology has been explained in the fourth section. In the fifth section, the experimental results are discussed and finally the conclusion and future work.

21.2 Related Work

Sundarajan et al. [1] have compared the classification models on the telecom dataset. The data reveals that attributes and features such as tenure group, contract, paperless billing, monthly charges, and Internet service seem to play a role in consumer churn. Also, there seems to be no affiliation between the gender and the churn rate. The data also says that the customers having a service plan of month-to-month contracts, with paperless billing, and within 12-month tenure are more likely to churn. In the work [1], random forest classifier achieved 93.5% accuracy and SVM achieved 81% accuracy.

Bharadwaj et al. [2] have proposed two nonlinear models, viz., a logistic regression model and an artificial neural network model to predict churn in the telecommunication industry. As part of preprocessing, feature extraction is done to eliminate irrelevant features and feature scaling is done to normalize the data. The models are trained on telecom churn prediction dataset available in Kaggle repository. The logistic regression model when applied to the test data achieved an accuracy of 87.52%, and the artificial neural network model obtained an accuracy of 94.19%.

The study done by Spiteri et al. [3] reveals that random forest algorithm is an effective technique to foresee customer churn for a motor insurance company, reaching an accuracy of 91.18%. Feature analysis was conducted to give perception into the most relevant features for this application. They used the ROSE algorithm for oversampling the dataset, and they have selected features. Survival analysis was employed to model time until churn, and it was found that approximately 90% of the policyholders remain with the company for up to five years, whereas the majority of the policyholders do not dismiss the policy before the expiry date.

Sagir et al. [4] evaluated different versions of neural network-based individual classifiers for churn prediction. These classifiers include neural network (NN), multilayer perceptron (AutoMLP), and deep learning (DL). Results are computed and evaluated using various performance measures such as accuracy, precision, recall, and f-measure and statistical measures such as kappa, absolute error, relative error, and classification error. When applied on the dataset, DL achieved 91.42% accuracy, NN achieved 93.94% accuracy, AutoMLP achieved 93.91% accuracy, bagging DL achieved 91.51% accuracy, AdaBoost DL achieved 91.09% accuracy, bagging NN obtained 94% accuracy, AdaBoost NN obtained 93.07% accuracy, bagging MLP achieved 94.15% accuracy, and AdaBoost MLP achieved 93.88% accuracy level.

Patil et al. [5] focused on the area of retail business, and the data they used is available in the University of California Irvine (UCI) machine learning repository. The dataset is preprocessed by removing values that are not available, validating numerical values, and removing erroneous data points, and variable churn is attached to each data point. The models trained on the dataset are RF, SVM, and gradient boosting, and they performed cross-validation on the data. When the models were applied to test data, RF achieved accuracy of 65%, SVM achieved 70% accuracy, and extreme gradient boosting (XGB) achieved 71% accuracy.

Mishra et al. [6] used dataset from the University of California Irvine (UCI) Web site which has machine learning (ML) repository, and dataset consists of 3333 records. They compared different algorithms, viz., bagging, boosting, random forest, decision tree, SVM, and Naïve Bayes algorithms. The accuracy for bagging is 90.83%, boosting is 90.32%, random forest is 91.66%, decision tree is 90.97%, SVM is 90.12%, and Naïve Bayes is 86.53%. The performance metrics they have used is accuracy, sensitivity, and specificity. Random forest got specificity of 53.54% and sensitivity of 98.89%.

Laila et al. [7] used a dataset comprising 24 attributes including churn and 3334 records. The testing-to-training ratio used for the research was 60:40. The models were trained using a decision tree, Naïve Bayes algorithm, and rule induction. However, the decision tree failed to classify churners accurately but the non-churners were classified with commendable accuracy whereas Naïve Bayes algorithm has classified churners accurately but failed to classify non-churners accurately.

Semrl et al. [8] in their work focused on predicting customer churn in the gym, and dataset used for the study has approximately 30,000 observations. They used two factors whether the gym contract ends or customer behavior changes to end the contract. They used a binary classification task based on gym attendance from the first month of contract. Azure ML's best performing calculation is the logistic relapse with the AUC of 0.734. Neural network achieved an area under the curve (AUC) of 0.732, and the decision-based random forest algorithm yields a fundamentally lower area under the curve (AUC) of 0.675. The best performing was the neural network with the accuracy of 74.6%, and the most noticeably awful performing was the choice tree with the precision of 70.8%. Big ML offers a hyperparameter improvement choice for choice forests, called sequential model-based algorithm configuration (SMAC) down concept.

It has been proven that, in the case of machine learning models, feature engineering and feature scaling are necessary for improving accuracy and performance. However, the works discussed in the reviewed literature fail to consider this aspect. Also, a few datasets used for churn prediction have more non-churners compared to the churners which makes the dataset imbalanced. This may lead to under-sampling of data, which is not addressed by the works which are making predictions based on such imbalanced datasets. Activation functions which take into account the interaction effects play a major role in deep learning neural networks. So, we need to use efficient activation function in hidden layers and output layer to maximize accuracy which some of the papers did not focus upon.

This paper attempts to address the above-mentioned gaps identified in the literature. The work attempts to build a neural network model for customer churn prediction that takes into consideration the interaction effects of various parameters by using different activation functions which contributes to better classification accuracy.

Figure 21.1 shows the steps of process for all neural network models. At the end of step, performance is evaluated using evaluation metrics such as AUC score, recall, precision, and accuracy.

21.3 Data Preprocessing

The telecom customer churn dataset used in this model is fetched from Kaggle. It contains attributes pertaining to customers of the telecom industry and contains 21 attributes and 7043 rows.

Figure 21.2 shows the data types of the columns present in the dataset. Figure 21.3 shows the summary of column maximum of monthly charges is 118.75 and minimum is 18.25. The maximum of total charges is 8684, and minimum is 118.75.

Data preprocessing phase of this work included removing null and missing values, converting data into a format suitable for process and standardization of the data. There are null values in the column "total charges" in the dataset, and there are 11 missing values in the same column. The "total charges" column is numeric, and so the missing values are replaced using the mean value of the column using the "Fill na" function from Pandas library.

There are many columns with different categorical values represented in different formats. These values are converted into a standard format by replacing the values with Yes or No. The attributes such as "Online Security" and "Online Backup" which are assuming categorical values are replaced with a value "No." Replace function is used from Pandas library to replace different categorical values.

The data is standardized using a minmax scaler from the sklearn library so that the data will remain on the same scale. Then, the scaled data may converge faster.

During the training phase, the data is split into training set and test set as per the 80:20 split ratio. Thus, 80% of the data is chosen for training and remaining 20% is chosen for testing the new observations and for determining the class. A higher percentage of training data makes the model learn better.

As the target column of telecom customer churn dataset has imbalanced classes of churners and non-churners, the data needs to be balanced. In this work, the dataset is balanced using Synthetic Minority Over-Sampling Technique (SMOTE) algorithm and thus the minority class is oversampled. Otherwise, model will learn more from majority class, viz., non-churn.

Feature selection is the process, wherein you automatically or manually select those features which contribute most to your prediction variable or output. It reduces the complexity of a model and makes it easier to interpret the output for the model. It

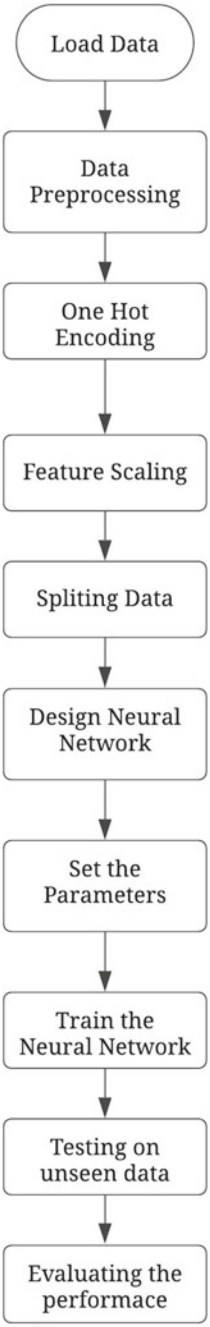


Fig. 21.1 Process flow of all the models

```
gender              object
SeniorCitizen       object
Partner             object
Dependents          object
tenure               int64
PhoneService        object
MultipleLines       object
InternetService     object
OnlineSecurity      object
OnlineBackup        object
DeviceProtection    object
TechSupport         object
StreamingTV         object
StreamingMovies     object
Contract            object
PaperlessBilling    object
PaymentMethod       object
MonthlyCharges     float64
TotalCharges       float64
Churn                int64
tenure_group        object
dtype: object
```

Fig. 21.2 Data types of columns

	tenure	MonthlyCharges	TotalCharges	Churn
count	7043.000000	7043.000000	7032.000000	7043.000000
mean	32.371149	64.761692	2283.300441	0.265370
std	24.559481	30.090047	2266.771362	0.441561
min	0.000000	18.250000	18.800000	0.000000
25%	9.000000	35.500000	401.450000	0.000000
50%	29.000000	70.350000	1397.475000	0.000000
75%	55.000000	89.850000	3794.737500	1.000000
max	72.000000	118.750000	8684.800000	1.000000

Fig. 21.3 Summary of columns

improves the accuracy of a model if the right subset is chosen. Mutual_info_classif function is used to select the features for the model. This method basically utilizes the mutual information [9]. It calculates mutual information value for each of the independent variables with respect to dependent variables and selects the ones which have the most information gain. It basically measures the dependency of features with

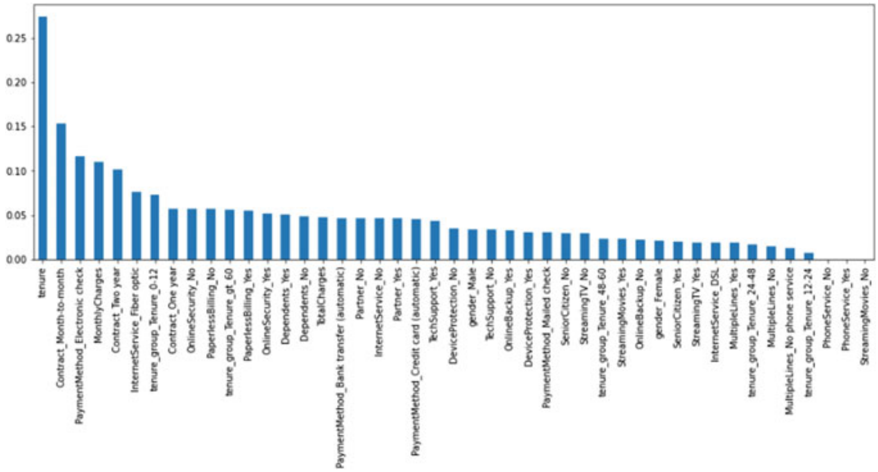


Fig. 21.4 Bar plot of important features

the target value. The higher score means more dependent variables and the percentage of important features is selected by applying a select percentile function.

As shown in Fig. 21.4, the important feature for classification of churner and non-churner is tenure column which has maximum importance followed by contract_from month_to_month and so on. The least important feature is streaming_movies_no column.

21.4 Methodology

During the model building phase, hyperparameters such as number of hidden layers, batch size, dropout rate, and learning rate are tuned to reach optimal values on different architectures. The optimal values are found using random search and grid search from scikit-learn library.

21.4.1 Model 1

The neural network model designed for customer churn prediction is a fully connected model containing 3 hidden layers. The number of neurons for first hidden layer is 40 and second hidden layer is 20. The tan$h$ activation function is shown in Eq. (21.1) and having the gradient shown in Eq. (21.2), and it is used for both the hidden layers. The final output layer has sigmoid activation function because it is a binary classification process. The weights are initialized using Glorot normal initializer.

To reduce overfitting, dropout is added between hidden layers with a ratio of 0.3. The learning rate is set to 0.005 using the optimizer ADAM. The model has "binary cross-entropy" as loss function and 20% of training data split for validation.

$$\mathrm{Tanh}(x) = \frac{e^z - e^{-z}}{e^z + e^{-z}} \tag{21.1}$$

which has the gradient:

$$\mathrm{Tanh}(x) = 1 = \tan(x)^2 \tag{21.2}$$

21.4.2 Model 2

For the second architecture, there are 3 hidden layers and 1 output layer in the architecture. The number of neurons for first hidden layer is 40, second hidden layer has 20 neurons, and third hidden layer has 10 neurons. All the hidden layers have the activation function ReLU, whose equation and gradient are shown in Eqs. (21.3) and (21.4), respectively. The final output layer has sigmoid activation function. The weights are initialized using he_uniform initializer. To reduce overfitting, dropout is added between hidden layers with a ratio of 0.2. The learning rate is set to 0.001 using the optimizer ADAM. The model has "binary cross-entropy" as loss function and 20% of training data split for validation.

$$\mathrm{relu}(x) = \max(0, x) \tag{21.3}$$

which has the gradient:

$$\frac{\mathrm{d}}{\mathrm{d}x}\,\mathrm{relu}\,(x) = \begin{cases} 0 \text{ if } x \le 0 \\ 1 \text{ if } x > 0 \end{cases} \tag{21.4}$$

21.4.3 Model 3

For the third architecture, there are 3 hidden layers and 1 output layer in the architecture. The number of neurons for first hidden layer is 40, second hidden layer has 20 neurons, and third hidden layer has 10 neurons. All the hidden layers have the activation function ELU, whose equation and gradient are shown in Eqs. (21.5) and (21.6), and the final output layer has sigmoid activation function. The weights are initialized using Glorot normal initializer. To reduce overfitting, dropout is added between hidden layers with a ratio of 0.3. The learning rate is set to 0.005 using the optimizer ADAM. The model has “binary cross-entropy” as loss function and 20% of training data split for validation.

$$\mathrm{Elu}(x) = \begin{cases} \alpha(\exp(x) - 1) & \text{if}\, x \leq 0 \\ x & \text{if}\, x > 0 \end{cases} \tag{21.5}$$

which has the gradient:

$$\frac{\mathrm{d}}{\mathrm{d}x}\,\mathrm{Elu}\;(x) = \begin{cases} \mathrm{elu}(x) + \propto & \text{if}\, x \leq 0 \\ 1 & \text{if}\, x > 0 \end{cases} \tag{21.6}$$

21.4.4 Model 4

For the fourth architecture, there are 2 hidden layers and 1 output layer in the architecture. The number of neurons for first hidden layer is 40 and second hidden layer is 20. The equation and gradient are shown in Eqs. (21.7) and (21.8), SELU is used for both the hidden layers, and the final output layer has sigmoid activation function. The weights are initialized using LeCun normal initializer. To reduce overfitting, dropout is added between hidden layers with a ratio of 0.3. The learning rate is set to 0.001 using the optimizer ADAM. The model has “binary cross-entropy” as loss function and 20% of training data split for validation.

$$\mathrm{selu}(x) = \lambda \begin{cases} \propto\, (\exp(x) - 1) & \text{if}\, x \leq 0 \\ x & \text{if}\, x > 0 \end{cases} \tag{21.7}$$

which has the gradient:

$$\frac{\mathrm{d}}{\mathrm{d}x}\,\mathrm{selu}(x) = \begin{cases} \mathrm{selu}(x) + \lambda \propto & \text{if}\, x \leq 0 \\ x & \text{if}\, x > 0 \end{cases} \tag{21.8}$$

21.5 Results and Interpretation

The models are trained on the training data with 100 epochs, validated to ensure the goodness of fit and the trained models are tested on unseen data to evaluate the best performing model from four architectures.

The various performance measures used to evaluate the model are as follows:

21.5.1 Confusion Matrix

Table 21.1 shows the confusion matrix of model 1. 864 non-churners and 229 churners are correctly predicted by the model. 171 are false positives which means that they are non-churners, but the model predicted them as churners. 145 are false negatives which means that they are churners, but the model predicted them as non-churners.

Table 21.2 shows the confusion matrix of model 2. 893 non-churners and 209 churners are predicted correctly by the model. 142 are false positives which means that they are non-churners, but the model predicted them as churners 0.165 that are false negatives which means that they are churners but the model predicted them as non-churners. Model 2 has less false negatives which is better than other models in this case.

Table 21.3 shows the confusion matrix of model 3. 828 non-churners and 245 churners are predicted correctly by the model. 207 are false positives, which means that they are non-churners, but the model predicted them as churners 0.129 that are false negatives which means that they are churners but the model predicted them as non-churners.

Table 21.4 shows the confusion matrix of model 4. 863 non-churners and 241 churners are predicted correctly by the model. 172 are false positives which means that they are non-churners, but the model predicted them as churners 0.133 that are false negatives which means that they are churners but the model predicted them as non-churners.

Table 21.1 Confusion matrix of model 1

	Actual (0)	Actual (1)	
Predicted (0)	TN = 864	FN = 171	1035
Predicted (1)	FP = 145	TP = 229	374
	1009	400	1409

Table 21.2 Confusion matrix of model 2

	Actual (0)	Actual (1)	
Predicted (0)	TN = 893	FN = 142	1035
Predicted (1)	FP = 165	TP = 209	374
	1058	351	1409

Table 21.3 Confusion matrix of model 3

	Actual (0)	Actual (1)	
Predicted (0)	TN = 828	FN = 207	1035
Predicted (1)	FP = 129	TP = 245	374
	957	452	1409

Table 21.4 Confusion matrix of model 4

	Actual (0)	Actual (1)	
Predicted (0)	TN = 863	FN = 172	1035
Predicted (1)	FP = 133	TP = 241	374
	996	413	1409

21.5.2 Evaluation Metrics

As shown in Table 21.5, the precision for model for class 0 is better compared to model 2 and precision for class 1 is lower compared to model 2. Compared to models 3 and 4, precision for model 1 for class 0 is the least. The precision for model 2 of class 0 is lower compared to all models, and for class 1, it is higher compared to other models. The precision for model 3 of class 0 is higher than model 1 and model 2, and for class 1, it is lower compared to all other models. The precision for model 4 for class 0 is higher compared to model 1 and model 2, and for class 1, it is higher than model 3. Model 4 can precisely classify non-churners from churners and reduce false positives.

As shown in Table 21.5, the recall for model 1 of class 0 is higher compared to model 3 and lower compared to all other models. The recall of class 1 for model 1 is higher compared to model 2 and lower compared to all other models. The recall of model 2 for class 0 is higher compared to all other models, and class 1 is lower compared to all other models. The recall of model 3 for class 0 is lower compared to other models, and for class 1, it is higher compared to other models. The recall of model 4 for class 0 is lower compared to other models, and for class 1, it is higher compared to model 1 and model 2. Model 2 has higher recall, so it can reduce false negatives.

As shown in Table 21.5, class 0 has the highest f1 score for model 1 compared to class 1. Model 2 has the same f1 score as model 1 for class 0 and class 1. Model 3 has less f1 score compared to model 1, model 2, and model 4. Model 4 has same f1 score for class 0 and has higher f1 score for class 1 compared to other models.

Table 21.5 Evaluation metrics for all models

Model name	Class	Precision	Recall	F1 score	AUC score
Model 1	0	0.86	0.83	0.85	0.72
	1	0.57	0.61	0.59	
Model 2	0	0.84	0.86	0.85	0.71
	1	0.60	0.56	0.58	
Model 3	0	0.87	0.80	0.83	0.727
	1	0.54	0.66	0.59	
Model 4	0	0.87	0.83	0.85	0.739

Table 21.6 Comparison of training and testing accuracy

S. no.	Model name	Training accuracy	Testing accuracy
1	Model 1	83.3	77.6
2	Model 2	86.6	78.2
3	Model 3	85.4	76.2
4	Model 4	79.5	78.4

As shown in Table 21.5, the model 1 has better AUC score compared to model 2. The model 2 has lower AUC score compared to all models. The model 4 has highest AUC score which indicates that it can classify churners and non-churners more accurately as compared to other models.

21.5.3 Training and Testing Accuracy

As shown in Table 21.6, the model1 has training accuracy of 83.3 but the testing accuracy of 77.6 is low compared to model 2 and model 4. There is high generalization error in model 1. Model 2 has training accuracy of 86.6, but the testing accuracy is 78.2 which is slightly high compared to model 1 and model 3. There is high generalization error in model 2. Model 3 has training accuracy of 85.4, but the testing accuracy of 76.2 is very low compared to all 3 models. Model 4 has training accuracy of 79.5 which is low compared to all 3 models and testing accuracy of 78.4 which is high compared to all 3 models. Model 4 has very low generalization error compared to all 3 models.

21.6 Conclusion and Future Work

Customer churn prediction which aids in devising strategies for retaining existing customers is very much imperative for the service providers in the telecom sector, to avoid revenue less. This research has implemented four different neural network models each using different activation functions and optimizers to predict whether a customer will churn or not. The first model using tanh activation function has better training accuracy but lesser testing accuracy. The training accuracy for first model is 83.3%, and testing accuracy is 77.6%. For the second model using ReLU activation function, training accuracy increased slightly but has lesser testing accuracy. The training accuracy for second model is 86.6%, and testing accuracy is 78.8%. For the third model using ELU activation function, training accuracy decreased slightly but has lesser testing accuracy. The training accuracy for third model is 85.4%, and testing accuracy is 76.2%. For the fourth model using SELU activation function, training accuracy decreased but has lesser error rate compared to other models. The training accuracy for the fourth model using SELU activation function is 79.5%, and testing accuracy is 78.4%. Thus, compared to other models SELU has less generalization error. The model can be trained with huge amount of data, and model architectures can be built with different layers, activation functions, epochs, and regularization techniques for achieving better performance.

References

1. Sundararajan, A., Gursoy, K.: Telecom customer churn prediction. https://doi.org/10.7282/t3-76xm-de75
2. Bharadwaj, S., Pahargarh, A., Anil, B.S., Gowra, P.S., Pahargarh, A., Kumar: Customer churn prediction in mobile networks using logistic regression and multilayer perceptron (MLP). 2nd International Conference on Green Computing and Internet of Things (ICGCIoT), Bangalore, India, pp. 436–438 (2018). https://doi.org/10.1109/ICGCIoT.2018.8752982
3. Spiteri, M., Azzopardi, G.: Customer churn prediction for a motor insurance company. 13th International Conference on Digital Information Management (ICDIM), Berlin, Germany, pp. 173–178 (2018). https://doi.org/10.1109/ICDIM.2018.8847066
4. Saghir, M., Bibi, Z., Bashir S., Khan, F.H.: Churn prediction using neural network based individual and ensemble models. 16th International Bhurban Conference on Applied Sciences and Technology (IBCAST), Islamabad, Pakistan, pp. 634–639 (2019). https://doi.org/10.1109/IBCAST.2019.8667113
5. Patil, A.P., Deepshika, M.P, Mittal, S., Shetty, S., Hiremath, S., Patil Y.E.: Customer churn prediction for retail business. International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS), pp. 845–851, Chennai (2017). https://doi.org/10.1109/ICECDS.2017.8389557
6. Mishra, A., Reddy, U.S.: A comparative study of customer churn prediction in telecom industry using ensemble based classifiers. International Conference on Inventive Computing and Informatics (ICICI), Coimbatore, pp. 721–725 (2017). https://doi.org/10.1109/ICICI.2017.8365230
7. Qaisi, L.M., Rodan A., Qaddoum, K., Al-Sayyed, R.: Customer churn prediction using data mining approach. 50th HCT Information Technology Trends (ITT), Dubai, United Arab Emirates, pp. 348-352 (2018). https://doi.org/10.1109/CTIT.2018.8649494
8. Semrl, J., Matei, A.: Churn prediction model for effective gym customer retention. International Conference on Behavioral, Economic, Socio-cultural Computing (BESC), Krakow, pp. 1–3 (2017). https://doi.org/10.1109/BESC.2017.8256385
9. SKlearn, https://scikitlearn.org/stable/modules/generated/sklearn.feature_selection.mutual_info_classif.html#sklearn.feature_selection.mutual_info_classif, Last Accessed 25 Oct 2020

Chapter 22
Wavelets and Convolutional Neural Networks-Based Automatic Segmentation and Prediction of MRI Brain Images

P. Muthu Krishnammal, L. Magthelin Therase, E. Anna Devi, and R. M. Joany

Abstract In clinical research, precise image segmentation and image prediction (classification) of MRI brain images has become a very crucial and challenging task as the pathological study and analysis of anatomical structures of brain is very complex. Moreover, processing of data amounts of dataset manually will also be the most challenging and time-consuming and produces inaccurate delineation of images, and making decisions on treatment planning will become difficult. Hence an automatic segmentation and classification algorithms based on convolutional neural network (CNN) is developed to predict brain tumors in the dataset if it is present. In MR images, the structural information varies as that of a non-periodic signal, and an efficient tool like wavelets which is a multi-resolution technique can be used for the characterization of complex signals. The isotropic features that may occur in spatial locations and scales can be detected by using wavelets. This paper presents a three-phased work wherein the significant features extracted by Wavelet transform, the image classification or the prediction of tumor classes is done by CNN and images are segmented by Spatial FCM and watershed segmentation algorithms.

22.1 Introduction

MRI is one of the most widely used methods in neurology and neurosurgery procedures and it provides excellent brain, spinal cord and vascular anatomical information. In addition to that, anatomy of the internal organs can be perceived in all three planes namely axial, coronal and sagittal [1]. The MRI scans can be used throughout the body as an extremely accurate means for detection of diseases and provide sufficient information to begin treatment for the patient. Other commonly found abnormalities include brain aneurysms, brain tumors, stroke, or spinal inflammation.

P. Muthu Krishnammal (✉)
Department of ECE, Narayana Engineering College-Gudur, Andhra Pradesh, India

L. Magthelin Therase · E. Anna Devi · R. M. Joany
Department of ECE, Sathyabama Institute of Science and Technology, Chennai, India

T. Senjyu et al. (eds.), *IOT with Smart Systems*, Smart Innovation, Systems and Technologies 251, https://doi.org/10.1007/978-981-16-3945-6_23

A brain tumor is a cluster or development of abnormal cells in the brain. It is also known as an intracranial tumor in which cells expand and 8 spread in uncontrollable manner, apparently uncontrolled by the regulation mechanisms of the normal cells. The two major brain tumor groups are: primary and metastatic (secondary) [2].

A multi-resolution transform, wavelet transform is used for the effective noise removal as well as feature extraction. The statistical features represent the intensity variations [3] and homogeneity property of images can be described by the texture elements with specific scale, directions and the analysis of local spatial variations are helpful for the problems in the pattern recognition [4]. For the better description of textural characteristics, GLCM, gray-level co-occurrence matrix is one among the best statistical approaches which are extensively used in the different phases of image processing such as image recognition, feature extraction, and classification [5, 6].

The purpose of image classification is to group of classes that have similar features. Computer vision and machine learning are the recent hot topics of research. Identification of cars exceeding the speed limits, tagging features used in Facebook are some of the realistic examples of image classification procedures where it strongly relies on the creation of labeled database. This image classification strategy using deep learning algorithms can also be extended for the classification of clinical images to classify them based on their anatomical structures. The tradition methods of classifiers fall into two categories: Supervised and unsupervised methods. In this paper, deep learning-based classification methods are presented and their performance is analyzed.

The image segmentation in clinical images is another daunting task in computer-assisted diagnostic processes which involves pattern recognition. The image segmentation is very important strategy which can be used for the study of pathological study, volume estimation for the detected objects, visualization by parts and surgical planning. Automatic segmentation is preferred because manual segmentation is a complicated and time-consuming operation.

22.2 Materials and Methods

22.2.1 Phase I: Feature Extraction

In general, the image features usually represent the distinctive properties of an image or object that has to be segmented. The shape descriptors and quantitative analysis of features are helpful to discriminate the interested structures and their relevance. The performance of image segmentation significantly depends on proper selection of features and precise extraction of features [7]. In MR imaging, feature extraction is a quantitative method used for the detection of structural anomalies present in the tissues of the brain. Since the brain tissues are very difficult to classify using the gray-level intensities or shape, textural feature extraction is very significant for further segmentation and classification. The visual pattern which contains the important information is defined by the texture that is represented by the spatial distribution of neighborhood gray levels [8]. The statistical approach is generally used, where the texture in a multidimensional feature space is represented as a vector. Statistical features are derived using the first-order and second-order gray-level intensities. The first-order features such as mean, median, and intensity are derived from the histogram [7]. Since any information regarding the pixel's spatial distribution is not included in the first-order features, these are generally used in combination with the second-order features which are computed using gray-level co-occurrence matrix (GLCM).

22.2.2 Extraction of Features by Discrete Wavelet Transform (DWT)

The major drawback of FT is that it can be used only for the stationary signals. Another transform, called the short-term Fourier transform (STFT), is used to find the frequency spectrum over a short period of time, but perhaps the size of the window limits precision. To analyze the non-stationary biomedical signals such as ECG and EMG, it is required to represent both time and frequency of the signals simultaneously [9, 10]. The wavelet transform adopts variable windowing method that provides time–frequency information which also adopts 'scale,' thereby provides the view of the signal in timescale and it is a multi-resolution technique [11].

The discrete wavelet transform (DWT) that makes use of positions and dyadic scales can be defined in the following manner: If $x(t)$ is a square-integrable function, the continuous wavelet transform of $x(t)$ with reference to a given mother wavelet $\psi(t)$ can be defined mathematically as,

$$W_\psi(a, b) = \int_{-\infty}^{\infty} x(t)\psi_{a,b}(t)\mathrm{d}t \tag{22.1}$$

where

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}}\left(\frac{t-a}{b}\right) \tag{22.2}$$

Here, $\psi(t)$ is the mother wavelet and $\psi_{a,b}(t)$, the wavelet is calculated from the given mother wavelet by using dilation factor (a) and translation parameter (b), both positive real numbers. Discretizing the equation by keeping under control of 'a' and 'b' to a discrete lattice $\left(a = 2^b \text{ and } a > 0\right)$, the resulting DWT is defined as,

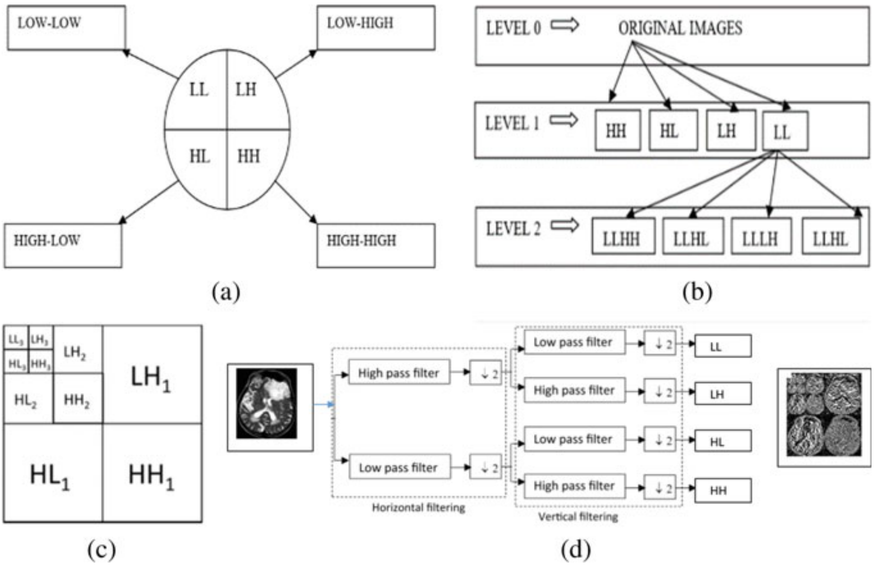


Fig. 22.1 **a** Level-1 decomposition. **b** Level 2 decomposition. **c** Level 3 decomposition. **d** Decomposition of MRI image using 2D DWT as an example

$$Ca_{j,k}(n) = \downarrow \left[\sum_n x(n) g_j^*\left(n - 2^j k\right) \right] \quad (22.3)$$

$$Cd_{j,k}(n) = \downarrow \left[\sum_n x(n) h_j^*\left(n - 2^j k\right) \right] \quad (22.4)$$

where $Ca_{j,k}$ denotes the approximation coefficients and $Cd_{j,k}$ denotes coefficients of detail components; $g(n)$ represents the LPF and $h(n)$ represent the HPF; $j \rightarrow$ the wavelet scale and $k \rightarrow$ translation factor and $\downarrow$ denotes the down sampling process. This decomposition process can be iteratively decomposed with successive approximations and by this kind of repeated decomposition, wavelet decomposition tree results [9, 10, 12] which is shown in Fig. 22.1.

22.2.3 GLCM Feature Extraction

GLCM is the statistical method wherein the spatial relationship among the pixels is considered. The GLCM functions determine the image texture by computing the frequency of occurrence of a pixels pair with specific intensities in a predetermined spatial closeness. In general, the spatial relationship is defined between the specific pixel of interest and the immediate horizontally adjacent pixel at the right. Feature extraction is helpful in identifying where the brain tumor is exactly located and helps to predict the next stage. The features can be extracted by transforming the input data into the set of feature vectors [13]. With GLCM, the most commonly used features known as Haralick features can be extracted. By using the criteria such as joint probability distribution and local correlation of pixels, the feature extraction using GLCM can be done and the value of the texture features is calculated. The different Haralick features used in this work are described in the results section.

22.2.4 Phase II: Image Classification by Convolutional Neural Networks

In deep learning, a CNN or *ConvNet* is the most commonly used deep neural network which finds its applications in various domains such as image and video recognition systems, image classification problems, natural language processing, medical imaging, and time series applications [14]. It is an algorithm in deep learning where the input image is taken, and the learnable weights and biases are assigned to the different objects in the image. Comparing to other traditional classification algorithms, CNN requires less pre-processing because of its self-learning capabilities. CNNs are generally composed of many convolutional layers. The feature map which represents the higher-level abstraction of the input data is generated by each of the layers in a CNN. By having very deep hierarchy of layers, it is possible to achieve superior performance [15].

22.2.5 Typical CNN Architecture

In the generalized CNN architecture, there are a series of feedforward layers, convolutional layers, and pooling layers. After the final pooling layer, there are many fully connected (FC) or dense layers that convert the two-dimensional features from the preceding layers into one-dimensional vector (flattening) for classification. The deep learning CNN models are used to train and test, where each input is fed through a series of kernels, i.e., convolutional layer with filters, pooling layers which together performs the feature extraction, fully connected layer (FC layer) that maps the extracted feature maps to the output 'y' to classify the images by converting the two-dimensional feature maps into one-dimensional map. Finally, a softmax function is applied at the end to classify the objects with probabilities between 0 and 1 [16, 17] (Fig. 22.2).

(a) **Convolutional Layer**: CNN is a unique neural network architecture which is designed for 2D or 3D data. Convolutional layer is the central to the CNN, and hence, it has got this name and the convolutional layer performs the stack of mathematical operation called 'convolution.' Performing convolution on the

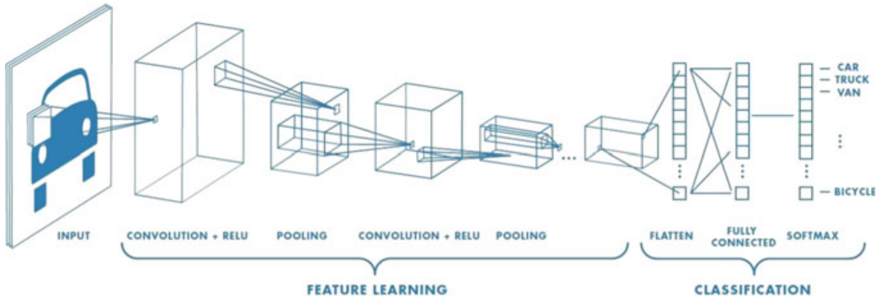


Fig. 22.2 Typical CNN architecture

input image (whose pixel values got stored in a 2D array) with a kernel (convolutional filter, an optimizable feature extractor), a feature map is produced [16–18]. It performs an element-wise (point-wise multiplication) multiplication between input vector and kernel's elements at every location and then summing up, results in the output tensor, the feature map. This convolution operation generally reduces the dimensions of the feature map at the output, and hence, zero padding is adopted in order to retain the in-plane dimensions.

(b) **Strides** decide the number of pixel shifts over an input matrix. For example, if the stride = 3, then filters are moved to three pixels at a time.

(c) **Padding** is used to preserve the dimensions of the input. There may be two possible results-either the dimensionality of the convolved feature is reduced in comparison with the input (valid padding) or the dimensionality is increased or remains the same (zero padding or same padding) where zeros are added to outside of the input and number of zeros padded will depend on the size of the kernel.

(d) **Activation functions**: The output of the convolutional layer is passed through the nonlinear activation unit which is used to introduce nonlinearity into the deep neural networks [18]. The conventional activation functions used are: rectified linear unit (ReLU), hyperbolic tangent, and sigmoid and functions. The activation unit gives the maximum value as the output, and it is computed as,

$$f(y) = \max(0, y). \tag{22.5}$$

(e) **Pooling layer** involves in dimensionality reduction in order to minimize the computational complexity involved of data processing. It also extracts the dominant features. The two types of pooling are: max pooling and average pooling. Max pooling produces the maximum value and average value is produced as a result of average pooling.

(f) **Fully Connected (FC) Layer**: Flattening of the 2D feature maps to 1D vector is performed by densely connected layers in order to map the final output of the network. The flattened output is now fed to a feedforward network and applying backpropagation to every iteration of training phase. Now, the model is able to distinguish the dominant features.

(g) The network's final layer is typically a softmax layer for the classification of the image, and it is activation function is given mathematically as,

$$y_r = \frac{\exp(a_{r)}}{\sum_{j=1}^{k} \exp(a_j)} \tag{22.6}$$

Training the network, using the regularization constraints, focuses on minimizing the cost function (loss function), i.e., difference between the ground truth labels and predictions. For multiclass problems, **cross-entropy** is typically the last layer of CNN. By using Eq. (22.7), we can calculate the consistency between the ground truth labels and predictions.

$$E(\theta) = -\sum_{i=1}^{n} \sum_{j=1}^{k} t_{ij} \ln y_j(x_i, \theta) \tag{22.7}$$

22.2.6 Phase III: Image Segmentation

The image segmentation methods work with an objective to divide the image into semantically meaningful information, homogeneous regions, and non-overlapping regions containing similar characteristics such as pixel color, intensity, or textural features [7]. The result of segmentation is either a labeled image to identify the homogeneous regions or a set of contours representing regional boundaries. These results are further used in the analysis of anatomical structures, pathological studies, visualization, and the surgical planning.

22.3 Segmentation Algorithms

22.3.1 Fuzzy C-*Means Clustering*

The FCM method is an improved variant of k-means segmentation algorithm, proposed by Bezdek. Assigning a membership function to the data points or pixels depending on the similarity to the specific class with respect to the other classes, the objective function is modified as:

$$J_m(\mu, v) = \sum_{i=1}^{c} \sum_{j=1}^{n} \mu_{ij}^{m} d^2(x_j, v_i) \tag{22.8}$$

where

$$\sum_{i=1}^{c} \mu_{ij} = 1$$

where $X = \left(x_1, x_2, \ldots, x_j, \ldots, x_n\right)$ describes a data matrix ($p \times n$), where p holds the dimension of x_j feature vector and n is the number of pixels in the image pixels. The distance 'd' is calculated using Euclidean norm as described in Eq. (22.9), and it the gives the similarity between v_i and x_j in the feature space.

$$d^2\left(x_j, v_i\right) = \left\|x_j - v_i\right\|^2 \tag{22.9}$$

A high membership value is assigned to the points which are closer to the centroid and low membership values for the points that are farther from the centroid. The objective function can be minimized and this process can be continued in an iterative manner until a solution is reached [10]. The major drawbacks of FCM algorithm are: There is no prior knowledge on the number of clusters, and spatial information is not accounted and highly sensitive to noise and other artifacts due to the imaging environment.

22.3.2 Spatial Constrained FCM (SFCM)

Modifying the FCM by incorporating the spatial constraints, Ahamed et al. [19], the objective function can be reformulated as,

$$J_{\mathrm{FCM\_}S} = \sum_{i=1}^{c}\sum_{j=1}^{n} \mu_{ij}^{m}\left\|x_j - v_i\right\|^2 + \frac{\propto}{N_R}\sum_{i=1}^{c}\sum_{j=1}^{n} \mu_{ij}^{m}\sum_{k\in N_i}\left\|x_k - v_j\right\|^2$$

where N_R is the number of pixels in every window and N_i is the neighborhood window of ith pixel. For smoothing being better, α is bigger, and smaller α emphasizes more detailed information of the given image. The major drawback is that calculation of the term $\sum_{k\in N_i}\left\|x_k - v_j\right\|^2$ will be time-consuming and it has to be carried out for every pixel in every iteration [20].

22.3.3 Watershed Algorithm

Vincent and Soille [21] introduced the concept of watershed segmentation, and it is a morphological gradient-based segmentation technique [22]. Watershed is built on the terms of watershed lines that is connected to topography and water source basin.

The gray values and borders of the objects are defined by taking advantage of the topological structure of the object in an image [23]. The rainfall simulation procedure is repeated for all the unlabeled pixels that are along the path of steepest descent. The paths reaching a common minimum adopt the same label as that minimum, and thereby a catchment basin is constituted, which refers to a partition in the image and thereby the image is segmented. It is a fast algorithm but suffers from severe over-segmentation problems.

22.4 Simulation Results

Dataset: The 150 digitized images of 256 × 256 DICOM MRI images were taken from the publicly available sources, and 60% of images were used for training and remaining for testing. Table 22.1 gives the Haralick features for the images of different classes of tumors such as normal, malignant, and benign cases that have been extracted using wavelet (DB wavelet with 4-level decomposition) and GLCM.

22.4.1 Feature Extraction

22.4.2 Classification

In this paper, a pre-trained CNN architecture with 25 layers is used. Figures 22.3 and 22.4 depict the performance in terms of batch accuracy, batch loss, and the time elapsed for the CNN at training and testing phases. The maximum number of epochs taken for the analysis is 100. It is observed that not only the CNN is getting trained well but also accuracy of CNN increases when the number of epochs is increased, and hence, the loss error function decreases. The network learns at the rate of 0.0004.

Table 22.1 Features extracted

Image	Feature (1.0e+04)	Energy	Contrast	Correlation	Homogeneity	Entropy
Normal image	Max	0.0000	0.4906	0.0000	0.0000	0.0006
	Min	0.0000	1.1106	0.0000	0.0000	0.0006
Benign image	Max	0.0000	0.7206	0.0000	0.0000	0.0005
	Min	0.0000	1.0945	0.0000	0.0000	0.0005
Malignant image	Max	0.0000	1.1890	0.0000	0.0000	0.0005
	Min	0.0000	1.4593	0.0000	0.0000	0.0005

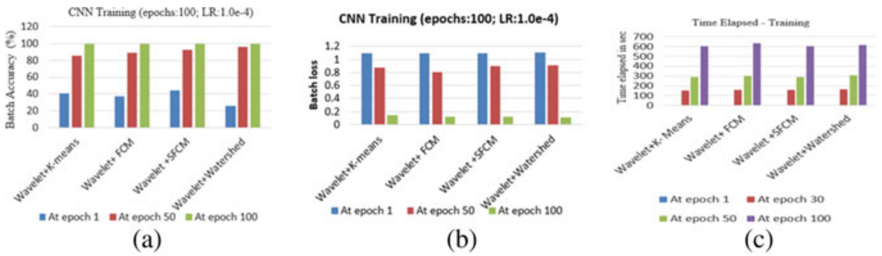


Fig. 22.3 CNN training stage: **a** Minimum batch accuracy. **b** Minimum batch loss. **c** Time elapsed

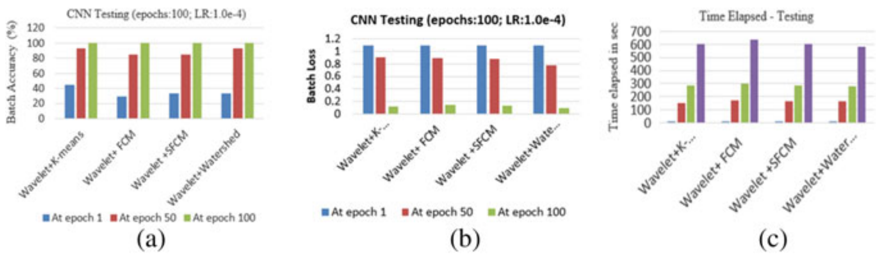


Fig. 22.4 CNN testing stage: **a** Minimum batch accuracy. **b** Minimum batch loss. **c** Time elapsed

22.4.3 Segmentation Stage

After the classification of tumor images, localization of tumors and segregation of tumor were done which was aided with the segmentation algorithms, namely *K*-means, FCM, spatial FCM and watershed segmentation algorithms and the outputs are depicted for the various classes of tumor such as normal, malignant and benign and cases. Figures 22.5, 22.6, and 22.7 show the simulated results for the different categories of tumor classes for the different methods of segment methodologies used in this paper.

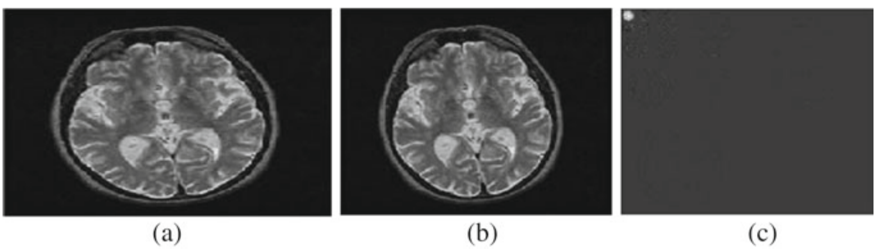


Fig. 22.5 Experimental results for normal image: **a** Input image **b** Test image **c** 4-level wavelet decomposition

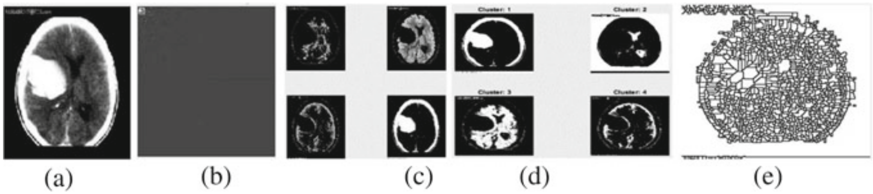


Fig. 22.6 Segmentation results of benign image: **a** Input image **b** 4-level wavelet decomposition **c** FCM. **d** Spatial FCM **e** Watershed

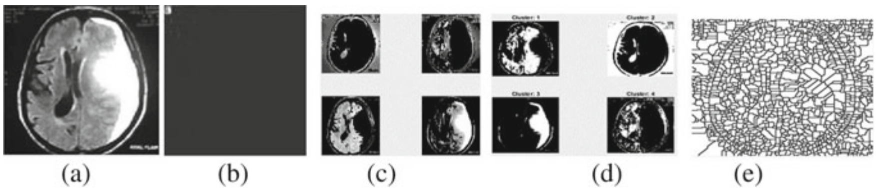


Fig. 22.7 Segmentation results of malignant image: **a** Input image **b** 4-level wavelet decomposition **c** FCM. **d** spatial FCM. **e** Watershed

22.5 Conclusion

In this paper, a 4-level Daubechies wavelet is chosen as it has less complexity in computation, and feature extraction is done in combination with GLCM in order to extract the features accurately. The CNN, which is a pre-trained 25 layered AlexNet architecture, is used for the process of prediction of tumor classes, i.e., classification problem, and it provides impressive outcomes by processing the feature set obtained using DB-DWT and GLCM; it is observed from the simulation results that 100% batch accuracy can be achieved during both training and testing phases of CNN classification and batch loss decreases when it is trained with large training dataset and also at the increased number of epochs. The self-learning and generalization capabilities of CNN help to achieve greater accuracy. Next, the traditional segmentation algorithms used in this paper also produce better results and as SFCM produces optimal results by having overhead on processing time in comparison with other clustering-based segmentation algorithms as it includes spatial constraints and the drawback of watershed algorithm is that it suffers from over-segmentation. From the results, it is perceived that the CNN offers superior performance for classification. Working with GPUs, the computational time efficiency and high accuracy can be maintained as it is very helpful in handling larger datasets.

References

1. Ortiz, A., Palacio, A.A., et al.: Segmentation of brain MRI using SOM-FCM-based method and 3D statistical descriptors. In: Computational and Mathematical Methods in Medicine, vol. 2013, pp. 1–12. Hindawi Publishing Corporation (2013)
2. The National Cancer Institute: https://www.cancer.gov/about-cancer/understanding/what-is-cancer
3. Shahnazar, A., et al.: A new developed approach for the prediction of ground vibration using a hybrid PSO-optimized ANFIS-based model. Environ. Earth Sci. **76**, 527 (2017)
4. Bharathi, D., Govindan, S.M.: A new hybrid approach for denoising medical images. In: Meghanathan, N., Nagamalai, D., Chaki, N. (eds.) Advances in Computing and Information Technology. Advances in Intelligent Systems and Computing, vol. 177. Springer, Berlin, Heidelberg (2013)
5. Buragohain, M. (2008) Adaptive network based fuzzy inference system (ANFIS) as a tool for system identification with special emphasis on training data minimization. Ph.D. thesis. Department of Electronics and Communication Engineering, Indian Institute of Technology Guwahati, Guwahati, India
6. Maji, P., Roy, S.: Rough-fuzzy clustering and unsupervised feature selection for wavelet based MR image segmentation. PLoS ONE **10**(4), e0123677. https://doi.org/10.1371/journal.pone.0123677
7. Gimel'farb, G., Delmas, P.: Lecture notes on "image processing-image filtering"
8. Muthu Krishnammal, P., Selvakumar Raja, S.: Automated brain image classification using neural network approach and abnormality analysis. Int. J. Eng. Technol. (IJET), **3**(7), 876–886 (2015). ISSN: 0975-4024
9. Mohsen, H., et al.: A machine learning technique for MRI brain images. In: The 8th International Conference on Informatics and Systems (INFOS2012), 14–16 May 2012. Bio-Inspired Optimization Algorithms and Their Applications Track, BIO, pp. 161–165
10. Mohsen, H., et al.: Classification using deep learning neural networks for brain tumors. Future Comput. Inf. J. **3**, 68–71 (2018)
11. Williams, T., Li, R.: Advanced image classification using wavelets and convolutional neural networks. In: 2016 15th IEEE International Conference on Machine Learning and Applications, pp. 233–239
12. Natarajan, P., et al., Tumor detection using threshold operation in MRI brain images. In: IEEE International Conference on Computational Intelligence and Computing Research (2012)
13. Haralick, R.M., Shanmugam, K., Dinstein, I.: Textural features for image classification. IEEE Trans. Syst. Man Cybern. **3**(6), 610–621 (1973)
14. Russell, S.: Artificial intelligence—a modern approach, 3rd edn
15. Yamashita, R., et al.: Convolutional neural networks: an overview and application in radiology. Insights Imaging **9**, 611–629 (2018)
16. Liu, J., Guo, L.: A new brain MRI image segmentation strategy based on wavelet transform and *K*-means clustering. In: ICSPCC 2015
17. Kass, M., Witkin, A., Terzopoulos, D.: Snakes: active contour Models. Int. J. Comput. Vision **1**, 321–331 (1988)
18. Muthu Krishammal, P., & Raja, S.S.: Deep learning based image classification and abnormalities analysis of MRI brain images. In: 2019 TEQIP III Sponsored International Conference on Microwave Integrated Circuits, Photonics and Wireless Networks (IMICPW), pp. 427–431. Tiruchirappalli, India (2019). https://doi.org/10.1109/IMICPW.2019.8933239
19. El-Dahshan, E.-S.A., et al.: Hybrid intelligent techniques for MRI brain images classification. Digit. Sig. Process. **20**, 433–441 (2010)
20. Krishnammal, P.M., Raja, S.S.: Medical image segmentation using fast discrete curvelet transform and classification methods for MRI brain images. Multimed. Tools Appl. **79**, 10099–10122 (2020)
21. Mustafa, Z.A., Abrahim, B.A., Kadah, Y.M.: K11. Modified hybrid median filter for image denoising. In: 29th National Radio Science Conference, pp. 705–712. 978-1-4673-1887-7/12

22. Zhanga, Y., et al.: A hybrid method for MRI brain image classification. Expert Syst. Appl. **38**, 10049–10053 (2011). https://doi.org/10.1016/j.eswa.2011.02.012
23. Zhang, X., Cui, J., Wang, W., Lin, C.: A study for texture feature extraction of high-resolution satellite images based on a direction measure and gray level co-occurrence matrix fusion algorithm. Sensors **17**, 1474:1–15 (2017)

Chapter 23
Effect of Meditation on Human Emotion Based on EEG Signal

Dinesh Datar and R. N. Khobragade

Abstract Mindfulness and wireless communications in various sectors such as education, self-regulation, manufacturing, marketing, internal security, and also interactive games and entertainment have their specific fingerprint. This paper discusses the natural style of meditation with the neurophysiological effect of the brain. The main aim of this study is to review the existing research and challenges of EEG signals on meditation effects. Although current studies demonstrated the effectiveness of mindfulness in relaxing stress and anxiety, generating mental, emotional well-being, more comprehensive investigation is needed for better design, consider user attributes like physical features to minimize bad impacts, randomized constrained measures, and broad sample sizes.

23.1 Introduction

Brain–computer interface is a highly growing field of research with a variety of medical application systems. Its medical achievements range from the prevention of traumatic injury to neuronal regeneration. Patients with serious neuromuscular conditions such as stroke, amyotrophic lateral sclerosis, or spinal injuries may regain control of their environment with BCI devices. Brain signals record the EEG or from inside the brain, ECoG from the skull was used as guidance for contact, movement, or controlling of connected devices [1].

Two key components are used in an entire BCI system: The first system to obtain and interpret the input signal, then create a command to operate the devices; and the user who generates the brain signals to send to the system. BCI technology developments concentrate primarily on these two components, improving higher output devices, such as higher data acquisition rates or advanced signal processing methods, and improving user training methods to generate steadier and also more reliable brain signals. Existing studies have indicated that factors can influence the

D. Datar (✉) · R. N. Khobragade
Sant Gadge Baba Amravati University, Amravati, India

T. Senjyu et al. (eds.), *IOT with Smart Systems*, Smart Innovation, Systems and Technologies 251, https://doi.org/10.1007/978-981-16-3945-6_24

output of the BCI, such as exhaustion, anger, focus as well as other behavioral and emotional variations [2].

People may build meta-cognitive reasoning to foresee the difference between what should be thought and what is being thought. In [3], different neurophysiologic improvements have been theorized during meditation on cognitive, hormonal, and autonomic systems. Evidence demonstrates the beneficial effects of mindfulness as it raises the brain regions' level of monoamines, parasympathetic involvement, and gray matter density (reflecting regulating emotion) and decreases negative impacts. The removal of stress-mediated depression results in these symptoms. In investigating these modifications, more study needs to be conducted with large samples in mind. EEG's spectral power and coherence describe bands of delta, theta, and alpha frequencies to categories various states of meditation.

The several investigations have been carried out to clinically examine that various neurological activity in the brain during mindfulness and meditation. Quite limited number of mindfulness medical procedures has been reported, and also, there are several missing in the control and synchronous group of antipsychotic medications including shortened time [4, 5].

The organization of this paper is as follows. Section 23.2 discussed the quantitative analysis of brain signal activity and its effects on various meditation styles. Section 23.3 describes the various brain wave frequency differences corresponds to different behavior and mental states of the brain. Section 23.4 presents the recognition of human emotion based on EEG signals needs more advanced and powerful machine learning techniques/algorithm that shows the high level of abstraction. Section 23.5 describes major challenges of EEG are associated with alpha, and theta power gives the maximum proficiency. Finally, Sects. 23.6 and 23.7 outline the future work and conclude this paper, respectively.

23.2 Electroencephalogram in Meditation

It is hard to make inferences from the different research looking at reflection and EEG signals, for the most part because of contrasts in the kind of meditation conjured, disparate practical conditions, and unmistakable investigation users. The general result of these investigations is that various parts of EEG are influenced diversely by various meditation types. This implies every meditation influences mind movement states that brain can adopt the stable changes. In this section discussed a quantitative analysis of brain signal activity and effect of meditation in various meditation positions.

23.2.1 Mindfulness Meditation

Meditation is the process of reducing physiological stress and improving psychological capabilities. It can be used for examining the BCI performance based on the short-term meditation program [6]. In self-referential development, mindfulness meditation can cause neural activity in the initial phase and maintain attention to physical stimuli. Kozasa et al. [6] measure the attention abilities between the meditators and non-meditators based on his/her age, year, and gender using fMRI adapted Stroop word-colored task. The main limitation of this study was the author cannot be suggested any kind of way to activate the brain region for non-meditators. Chow et al. [7] improved the meditation research on mindfulness meditation and EEG alpha neurofeedback (NFB). They show that improvement in attentional performance and observed that full EEG alpha amplitude with sub-band amplitude in 15 min session was raised. The study was limited to the upper sub-band of EEG alpha rhythm and neurocognitive. Another mindfulness meditation technique of integrative body–mind training (IBMT) and triarchic body pathway relaxation (TBRT) have been proposed but have limited information and need expert meditators for considering these techniques on large scale [9, 10].

23.2.2 Buddhist Meditation

Buddhist meditation practices become accompanied by high amplitude gamma oscillations based on active analysis [11–13]. The study is based on Samatha and Vipassana meditation is a core concept of long-term Buddhist meditation practice to develop attention and concentration based on neurofeedback technique. Samatha technique has the potential to increase the degree of resilience to seizures in epilepsy sufferers. B. Rael Cahn et al. found that Vipassana meditation increased gamma power of meditator with daily practices of above 10 years [13].

23.2.3 Mantra Meditation

Usually, mantra meditation consists of a repeated specific word, phrase, or sequence of vowel sounds while actively completely ignoring any internal and external stimuli. Over recent decades, the number of research investigating the effect of mantra therapy on psychological health metrics has increased; with significant reduction, the effects of fatigue, tension, depression, anxiety, and mental distress are frequently reported [14]. Julie Lynch et al. study about mantra meditation and suggested regular practices of mantra meditation relieve the stress, exciting educational, occupational, and psychotherapeutic potential, providing an individual solution to the general public that avoids the side effects of drugs, the stigma of care, and obstacles to cost and accessibility issues. Susan Thomas et al. studied the effect of Gayatri mantra meditation on the brain region based on EEG and fMRI. They found that the maximum brain region was activated superior temporal, lobe, right insula, left inferior lobule, culmen of the cerebellum, etc., after continuously listening to the Gayatri mantra [15].

23.2.4 *Qi Gong*

After the Qigong meditation, EEG brain signal action can be significantly increases and also to increases the frontal theta and posterior alpha. Using the Qigong technique to increase the fronto-central midline theta activity and the brain of alpha power increases with the same intensity level [16].

23.2.5 *Tai Chi and Yoga*

The Tai Chi meditation improves the brain activity and function in adult person [17]. The existing study shows that the tai chi meditation improves the theta activity in brain and decreases the nervousness of person and person get to relax after the tai chi and yoga [18]. The yoga is legitimately affecting our state of mind by making a few vibrations in breathing which changes the EEG signals recurrence in positive manner [19] (Table 23.1).

23.3 EEG Signal Characteristics

Primarily EEG signals can be classified with the help of their frequency, the amplitude of the signal and shape as well as where electrodes are placed on the scalp. Hertz (Hz) is the basic unit of frequency to determine normal and abnormal rhythms. The EEG waveform can be classified into alpha, beta, gamma, theta, and delta which is based on their frequency. The continuous rhythms of the brain signal are categorized by frequency bands. Brain wave frequency differs corresponds to different behavior and mental states of the brain (Table 23.2).

23.4 EEG-Based Emotion Recognition

Human emotion analysis is one of the most challenging tasks. To recognize the human emotion based on EEG signals needs more advanced and powerful machine

Table 23.1 Summary of effect of meditation style

Authors	Style of meditation	No. of participants	Techniques	Results
Yin-Qing Tan et al.	Mindfulness	5	–	Improvement in BCI performance
Elisa H. Kozasa et al.	Mindfulness	38	fMRI adapted Stroop Word-Color Task	Activate more brain region and increase the efficiency of brain of regular meditator as compared to non-meditator
Theodore Chow et al.	Mindfulness and Neurofeedback	74	–	Increases EEG alpha amplitude
Paul Dennison	Samatha (individual study)	–	Self-Directed Neurofeedback	Strong alpha and delta Decrease the beta activities
B. Rael Cahn et al.	Vipassana	16	Independent component analysis	Power: • Increases alpha • No effect in theta • Gamma increases frontal or parietal region • Decreases bilateral frontal delta
Rahul Ingle et al.	Vipassana	50	Time–Frequency analysis Support vector machine	Increase alpha and theta power in occipital and right temporal
Ankita Tiwari et al.	Yoga	37	–	Increase alpha after yoga
Susan Thomas et al.	Gayatri Mantra	20	fMRI, fast Fourier transform	Gamma and Beta increase after listening Gayatri mantra
Diana Henz et al.	Qigong	75	Discrete fast Fourier transform	• Increased midline fronto-central theta • Increased posterior alpha-1 and alpha-2 activity

Table 23.2 Summary of brain wave characteristics

Brain wave with frequencies	Frequency classification (Hz)	Characteristics	Summary
Delta (0.5–4 Hz)	0.5	Total relief of headache and relaxation	(1) Focal delta activity: abnormal indicating lesions (2) Lower delta power: complex depression, continual migraine, closed posterior head/neck damage (3) Excess delta activity: learning disability, Alzheimer's disease, edema
	0.5–1.5	pain relief through endorphin release	
	0.9	Euphoric condition	
	1	Feeling of well-being; stimulation of pituitary glands to release growth hormones	
	2	Nerve regeneration	
	2.5	Further pain and migraine relief from the production of endogenous opiates. Natural sedative effect	
	1–3	Restorative sleep and profound relaxation	
	3.4	Restful sleep	
	3.5	Feeling of calmness, reducing anger and irritability. Retention of languages	
	4	Enkephalin release for natural stress and pain reduction. Improved memory, subconscious learning, and problem-solving	
Theta (4–7 Hz)	4.5–4.9	It is also called as Tibetan state of meditation Induced relaxation, meditation, introspection, and a deeper sleep	In normal awaken adults, very small theta activity but a high value accounts for pathological conditions
	5–5.35	Solve unconscious problem, release endorphin to relief from pain, deep breathing, lung relaxation	
	5.5–5.8	Develop feeling of intuition, reduction of fear	
	6–6.5	Improvement in long-term memory, reduces unwillingness, and waking dreaming	

(continued)

Table 23.2 (continued)

Brain wave with frequencies	Frequency classification (Hz)	Characteristics	Summary
	6.5–7	Activate frontal lobe, reduces fatigue, increase peace	
Alpha (7–13 Hz)	7.5–8	Increase the interest in music, art, research, easily finding the solution Increases addition level	Alpha coma, a diffused alpha in EEG occurs in coma which does not respond to external stimuli
	8–14	Qi Gong meditation	
	8–10	Increase learning power, memorize	
	10–12	Improves the mind, body connection	
Beta (13–39 Hz)	13–27	Improve the attention, alertness	(1) Decreased beta activity: focal lesions, stroke, or tumor diffused encephalopathy such as anoxia (2) Excess beta activity during alcoholism along with decreased alpha and theta
	13–30	Easy to solve problem, active conscious	
	15–18	Increased mental ability, alertness, focused	
	18–24	Released high serotonin	
	20–40	Released ideal meditation frequency for stress	
Gamma 40 Hz+	40	Creates a good memory, problem-solving, high-level information processing	Sometimes of no clinical interest and filtered out in EEG recordings
	55	Tantric yoga which stimulates the "kundalini"	
	40–60	Release of beta-endorphins and give anxiolytic effects	
	62	Feeling of physical vigor	

learning techniques/algorithm that shows the high level of abstraction [20] (Table 23.3).

23.5 EEG Signal Processing Challenges

This section can add the signal processing part to this study. The scalp electrode can be placed on the human head issues the brain activity with the help of electric potential, give brain signal can be processed using various signal processing techniques and

Table 23.3 Summary of brain wave characteristics

Authors	Participants	Stimulation	Emotion	EEG feature	Classification technique /algorithms	Accuracy (%)
Jirayucharoensak et al. [20]	32	EEG signals	Negative–positive Passive–Active Dominated–Dominant Like–dislike	CSP	KNN, NB, SVM	53.42
Hassouneh et al. [21]	55	Video Clip	Happy, fear, angry, surprise, disgust	Infinite Impulse Response (IIR)	CNN, LSTM	87.25
Kam et al. [22]	9	Images	Mental practice of action	ERD/ERS, CSP	LDR, SVM	60
Asghar et al. [23]	32	Video	Arousal, Dominance, Valence, Liking	Combined Feature Vector	SVM KNN RF	SVM-97 KNN-78.9 RF-80.6
Zangeneh Soroush et al. [24]	40	Video	Arousal, Valence	CD, FD, LLE, ED, ATL, RR, DET, LAM, MLV, RT, L	MLP KNN SVM	90
Al-Nafjan et al. [25]	–	Video	Arousal, Valence	PSD	DNN	82.5
Guo et al. [26]	32	Video, Music	Arousal, Valence	CD, FD, LLE, ED, ATL, RR, DET, LAM, MLV, RT	SVM, HMM	62.5

methods to research on various brain activities and brain cognitive tasks. The EEG can be used to understand the internal activities of the brain based on electromagnetic activity [27, 28]. The EEG patterns contain the signal which is associated with encouragement, feeling, emotion, awareness, and other cognitive activity [28]. The characterize of the EEG signal can be described in the above section. The complete description of the delta, alpha, beta, theta, and gamma is in Table 23.2. The major challenges of EEG are associated with alpha, and theta power gives the maximum proficiency. However the dynamically updating of these brain waves in continuous meditation practices is depends on the specific brain region. To classify the EEG patterns based on various meditation techniques, advanced mindfulness studies need to be done. It's a very complex job to effectively observe nonlinear and non-stationary EEG raw information in the time–frequency domain. As such, different methods of linear and nonlinear signal processing and their correlation with physiology have been suggested [29]. For preprocessing the EEG raw signal require powerful algorithms for feature extraction. Observation of time frequencies is suitable for investigating rhythmic relevant data in EEG signals. In each frequency spectrum, spectral time series or cohesiveness includes sequence consistency evaluation among both signal sequences. Preprocessing is a non-trivial operation, as the removal of any unnecessary components integrated with the EEG signal is performed. Good preprocessing leads to an improvement in spectral efficiency, which in effect results in improved distinction of features and efficiency of classification. The key efforts to modulate samples is determined the EEG are basic low, high, and band-pass filtering [30].

23.6 Future Work

In this study, it is clearly observed that the specific EEG system component needs to change the research direction is either raising or reducing for different meditation techniques. Many research works have been conducted by systematically examine the fundamental neuro-electrical function of the cerebral cortex after mindfulness practice, although meditation's neuro-psychological impact is now an ongoing problem. Some of the research is based on the psychological and neuroscientific approach [32]. Although it has been observed that even the theta, and alpha power can be reduced in at most alpha frequency and transmitted of alpha coherence around the brain has been increased by EEG pattern of mediation, its recognition requires more study [31].

23.7 Conclusion

Today, alternative treatments such as mindfulness are beginning to be used during diagnostic procedures. An important contribution of this study is clarifying the neuro-electrical impacts of mindfulness, particularly various applications of mindfulness meditation in medical care and treatments. Research into EEG brain activity has shown that health, meditation, and relaxation methods can be adapted to increase mental ability, depending on the region of the brain. A wide range of cognitive concepts under different styles of mindfulness has been discussed in this study. The detailed summary of various meditation styles such as Vipassana, Yoga, Gayatri Mantra, Qigong, Samatha, and Mindfulness has been described as that means of EEG signal and coherence. Particularly in comparison to task-free relaxing during meditation practices, it has been reported that enhanced autonomous brain functions are observed. Concepts of signal processing in the area of cognitive neurology were explored. The categorization of EEG samples with respect to various procedures needs a qualitative study. As discussed in the previous section, there are systematic limitations in design, research methods, and other key factors that contribute to the quality of research.

References

1. Tan, Y.-Q., Tan, L.F., Mok, S.-Y., Goh, S.-Y. (2015). Effect of short term meditation on brain-computer interface performance. J. Med. Bioeng. **4**, 135–138. https://doi.org/10.12720/jomb.4.2.135-138
2. Lo, P.C, Wu, S.D., Wu, Y.C.: Meditation training enhance the efficacy of BCI system control. In: Proceedings of the International Conference on Networking, Sensing and Control, Taipei, pp. 825–828 (2004)
3. Kasala, E.R., Bodduluru, L.N., Maneti, Y., Thipparaboina, R.: Effect of meditation on neurophysiological changes in stress mediated depression. Complement. Ther. Clin. Pract. **20**(1), 74–80 (2014)
4. Leung, N.T.Y., Lo, M.M., Lee, T.M.C.: Potential therapeutic effects of meditation for treating affective DYS regulation. Evid. Based Complement. Altern. Med. **2014**, 7 (2014). Article ID 402718
5. Tan, L.F., Dienes, Z., Jansari, A., Goh, S.Y.: Effect of mindfulness meditation on brain-computer interface performance. Conscious. Cogn. **23**, 12–21 (2014). https://doi.org/10.1016/j.concog.2013.10.010. Epub 2013 Nov 22
6. Kozasa, E.H., Sato, J.R., Lacerda, S.S., Barreiros, M.A.M., Radvany, J., Russell, T.A., Sanches, L.G., Mello, L.E.A.M., Amaro, E. Jr.: Meditation training increases brain efficiency in an attention task. NeuroImage **59**, 745–749 (2012)
7. Chow, T., Javan, T., Ros, T., Frewen, P.: EEG dynamics of mindfulness meditation versus alpha neurofeedback: a Sham-controlled study. Mindfulness **8** (2016). https://doi.org/10.1007/s12671-016-0631-8
8. Xue, S.-W., Tang, Y.-Y., Tang, R., Posner, M.I.: Short-term meditation induces changes in brain resting EEG theta networks. Brain Cogn. **87**(1), 1–6 (2014)
9. Chan, A.S., Han, Y.M.Y., Cheung, M.-C.: Electroencephalographic (EEG) measurements of mindfulness-based triarchic body-pathway relaxation technique: a pilot study. Appl. Psychophysiol. Biofeedback **33**(1), 39–47 (2008)
10. Dennison, P.: Samatha meditation and neurofeedback (2015). https://doi.org/10.13140/RG.2.1.1250.5761
11. Davidson, R.J., Lutz, A.: Buddha's brain: neuroplasticity and meditation. IEEE Sig. Process. Mag. **25**(1), 174–176 (2007)
12. Cahn, B.R., Delorme, A., Polich, J.: Occipital gamma activation during Vipassana meditation. Cogn. Process. **11**(1), 39–56 (2010). https://doi.org/10.1007/s10339-009-0352-1

13. Lynch, J., Prihodova, L., Dunne, P.J., Carroll, Á., Walsh, C., McMahon, G., White, B.: Mantra meditation for mental health in the general population: a systematic review. Eur. J. Integr. Med. **23**, 101–108 (2018)
14. Thomas, S., Rao, S.L., Effect of Gayatri mantra meditation on meditation Naive subjects: an EEG and fMRI pilot study. Int. J. Indian Psychol. **3**(2(7)) (2016)
15. Henz, D., Schöllhorn, W.I.: EEG brain activity in dynamic health qigong training: same effects for mental practice and physical training? Front. Psychol. **8**, 154 (2017). https://doi.org/10.3389/fpsyg.2017.00154
16. Pan, Z., Su, X., Fang, Q., Hou, L., Lee, Y., Chen, C.-C., Lamberth, J., Kim, M.-L.: The effects of Tai Chi intervention on healthy elderly by means of neuroimaging and EEG: a systematic review. Front. Aging Neurosci. **10** (2018). https://doi.org/10.3389/fnagi.2018.00110
17. Field, T., Diego, M., Hernandez-Reif, M.: Tai chi/yoga effects on anxiety, heart rate, EEG and math computations. Complement. Ther. Clin. Pract. **16**(4), 235–238 (2010). https://doi.org/10.1016/j.ctcp.2010.05.014
18. Tiwari, A., Tiwari, R.: Monitoring and detection of EEG signals before and after yoga during depression in human brain using MATLAB. In: 2017 International Conference on Computing Methodologies and Communication (ICCMC), pp. 329–334. Erode (2017). https://doi.org/10.1109/ICCMC.2017.8282702
19. Ingle, R., Awale, R.N.: Impact analysis of meditation on physiological signals. Int. J. Inf. Visual. **2**(1) (2018)
20. Jirayucharoensak, S., Pan-Ngum, S., Israsena, P: EEG-based emotion recognition using deep learning network with principal component based covariate shift adaptation. Sci. World J. **2014**, 10 (2014). Article ID 627892. https://doi.org/10.1155/2014/627892
21. Hassouneh, A., Mutawa, A.M., Murugappan, M.: Development of a real-time emotion recognition system using facial expressions and EEG based on machine learning and deep neural network methods. Inf. Med. Unlock. **20** (2020). https://doi.org/10.1016/j.imu.2020.100372. http://www.sciencedirect.com/science/article/pii/S235291482030201X
22. Kam, T., Suk, H., Lee, S.: Non-homogeneous spatial filter optimization for EEG-based brain-computer interfaces. In: 2013 International Winter Workshop on Brain-Computer Interface (BCI). Gangwo, 2013, pp. 26–28. https://doi.org/10.1109/IWW-BCI.2013.6506618
23. Asghar, M.A., Khan, M.J., Rizwan, M., Mehmood, R.M., Kim, S.-H.: An innovative multi-model neural network approach for feature selection in emotion recognition using deep feature clustering. Sensors **20**, 3765 (2020). https://doi.org/10.3390/s20133765
24. Zngeneh Soroush, M., Maghooli, K., Setarehdan, K., Motie Nasrabadi, A.: Emotion classification through nonlinear EEG analysis using machine learning methods. Int. Clin. Neurosci. J. **5**, 135–149 (2018). https://doi.org/10.15171/icnj.2018.26
25. Al-Nafjan, A., Al-Wabil, A., Hosny, M., Al-Ohali, Y.: Classification of human emotions from electroencephalogram (EEG) signal using deep neural network. Int. J. Adv. Comput. Sci. Appl. **8**(9) (2017)
26. Guo, K., Candra, H., Yu, H., Li, H., Nguyen, H. Su, S.: EEG-based emotion classification using innovative features and combined SVM and HMM classifier. In: Conference Proceedings: Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference, pp. 489–492. (2017). https://doi.org/10.1109/EMBC.2017.8036868
27. Gevins, A.: Electrophysiological imaging of brain function. In: Toga, A.W., Mazzoitta, J.C. (Eds.) Brain Mapping. The Methods, 2nd edn, pp. 175–188. Elsevier, New York, NY, USA (2002)
28. Sanei, S., Chambers, J.A.: EEG Signal Processing. Wiley (2007).
29. Fingelkurts, A.A., Fingelkurts, A.A.: Operational architectonics methodology for EEG analysis: theory and results. In: Modern Electroencephalographic Assessment Techniques. Neuromethods, vol. 91, pp. 1–59. Springer, New York, NY, USA (2015)
30. Mamunur, R. Norizam, S., Abdul Majeed, A.P.P., Muazu, M.R.. Fakhri, A.N.A., Sama, B.B., Sabira, K.: Current status, challenges, and possible solutions of EEG-based brain-computer interface: a comprehensive review. Front. Neurorobot. **14**, 25 (2020). https://doi.org/10.3389/fnbot.2020.00025. https://www.frontiersin.org/article/10.3389/fnbot.2020.00025

31. Travis, F., Shear, J.: Focused attention, open monitoring and automatic self-transcending: categories to organize meditations from Vedic, Buddhist and Chinese traditions. Conscious. Cogn. **19**(4), 1110–1118 (2010)
32. Khobragade, R.N.: Issues and methods in meditation research. Int. J. Recent Innov. Trends Comput. Commun. **5**(8), 110–112 (2017)

Chapter 24
Design and FPGA Implementation of Vedic Notch and Peak Filters

Meenakshi Agarwal and Madhur Garg

Abstract These days, in digital filters, high speed and low complexity are the key requirements. In this paper, notch and peak filters are proposed to design using high speed Vedic multipliers. Vedic multipliers are designed using high speed carry save adders and carry select adders. A design example is included to analyze the effectiveness of the proposed filters using direct form II structure and has been tested and verified on Xilinx ZYNC field programmable gate array (FPGA) device. The performance of the proposed Vedic notch and peak filters (VNPF) is compared in terms of utilized hardware and power dissipation. It is shown that the proposed VNPF is consuming less power and occupying small hardware compared to the conventional notch and peak filters.

24.1 Introduction

In the smart world elegance eon, a major role is going to be played by the wireless fifth generation (5G) technology and the Internet of Things (IoT). Signal processing is a major component of evolving technology. In considerable signal processing relevance, it is required to pass or stop a specific component of frequency of a signal. However, at the same time, broadband or narrow signals are maintained to be unmodified. Generally, peak and notch filters are a perfect choice for this task and find applications in the fields of control systems, audio processing, communication, biomedical engineering, etc. [1]. To remove powerline interference in a straightforward example of an electrocardiogram (ECG), notch filters are frequently used [1]. On the other hand, to stop the single-frequency interference, a notch filter is tuned at that appropriate frequency [2]. Peak filter finds applications in graphic equalizer for audio [3]. In case, if interfering frequencies are more than one, multiple notch filter is required to get rid of these prescribed frequencies [4]. To design these filters

M. Agarwal (✉) · M. Garg
BML Munjal University, Gurugram, India
e-mail: meenakshi.agarwal1@bmu.edu.in

M. Garg
e-mail: madhur.garg.17ece@bml.edu.in

T. Senjyu et al. (eds.), *IOT with Smart Systems*, Smart Innovation, Systems and Technologies 251, https://doi.org/10.1007/978-981-16-3945-6_25

either finite impulse response (FIR) or infinite impulse response (IIR), filters can be chosen [5, 6]. Digital filter design using non-recursive as well as recursive structures have been largely investigated by several research groups around all over the world [7, 8]. In applications where phase linearity is foremost concerned, FIR filters are employed. It is notable that the IIR filter has better statistical performance and is computationally efficient compared to the FIR filter. It has an additional advantage while realizing low cost and low power circuits. Therefore, in the applications where the phase should not be strictly linear, the IIR filters are recommended as it requires lower order to realize the filter with the same specifications [9–11].

In FIR and IIR filters, multiplication is the most essential arithmetic operation. Additionally, in multiplication, speed is always a constraint. Therefore, reduction of the computational steps results in improved speed. For quick mathematical calculations, Vedic mathematics have been practiced in India for thousand of years. It has offered different Sanskrit sutras (sutras are pearls of wisdom) [12] and has been validated to be a vigorous technique for different mathematical operations such as addition, subtraction, multiplication, and division [13]. Furthermore, Vedic Maths has significantly contributed to solve the complex problems also where a considerable number of mathematical operations are involved. In literature, Vedic multipliers have been used in numerous domains such as the design of image processing, less power hungry, small area circuits [14], and in the signal processing domain [16]. S. Sharma et al. have delve into novel design of Vedic multiplier and its implementation using 90 nm technology [17]. Sivanandam et al. have modified the basic Vedic multiplier using 3-1-1-2 compressor and implemented in FFT applications on FPGA [18]. In other applications such as encryption/decryption, Vedic multipliers are used, and IIR filters with floating-point coefficients have been investigated [19–21] and implemented on FPGA. However, notch and peak filter design using fast Vedic multipliers is still an unconquered area.

The objective of this work is to show an easy way to design and implement notch and peak filters simultaneously using Vedic multipliers. The presented work shows that Vedic multiplication reduces the complexity and enhances the speed of the multipliers. In this paper, the implementation process is divided into two steps. To begin with, in first step, first and second-order notch and peak filters are obtained using allpass filters (APF). APF are realized using direct form II structure. The adder, multiplier, and delay components of the notch and peak filters are replaced with minimum number of adders, Vedic multipliers, and registers, respectively. In the second step, these filters are implemented on Xilinx Zync XC7z007sclg400 FPGA device with speed grade -1. The coding of the proposed filters is written in Verilog HDL. For simulations and verification of the coded filters, all possible input vectors are applied. The comparison results show that the notch and peak filter designed using a modified Vedic multiplier are more efficient than their conventional IIR counterparts in regards of used hardware and dissipated power.

The rest of the paper is organized as follows. The existing notch and peak filters are described in Sect. 24.2. Section 24.3 explains the architecture of Vedic multipliers. In Sect. 24.4, a design example of FPGA implementation of the proposed VNPF is demonstrated. Result analysis is executed in Sect. 24.5. Section 24.6 concludes the paper.

24.2 Notch and Peak Filter

The notch and peak filters are easily designed using allpass filters and shown in Figs. 24.1 and 24.2 [22]. $Q(z)$ is the transfer function of notch filter and shown as

$$Q(z) = \frac{1}{2}[1 + A(z)] \tag{24.1}$$

and $W(z)$ is presenting the transfer function of a peak filter, given below.

$$W(z) = \frac{1}{2}[1 - A(z)] \tag{24.2}$$

where $A(z)$ is an allpass function. Therefore, the notch and peak filter design are turned into an allpass filter problem. For all frequencies, the magnitude response of an APF is equal to unity, that is,

$$|A(e^{j\omega})|^2 = 1, \quad \text{for all } \omega \tag{24.3}$$

and the frequency response is given in Eq. 24.6.

$$A(e^{j\omega}) = e^{j\theta_A(\omega)} \tag{24.4}$$

In the above expression, $\theta_A(\omega)$ is the phase response. It is observed that for stability, the phase response $\theta_A(\omega)$ varies from 0 to -2π when ω changes from 0 to π.

Compared to the conventional methods, this structure offers a mirror image symmetry relation in allpass filter; i.e., the numerator and the denominator polynomials

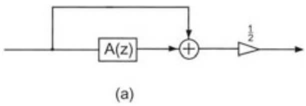


Fig. 24.1 Digital notch filter

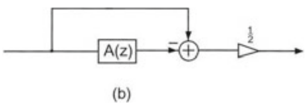


Fig. 24.2 Digital peak filter

of $A(z)$ can be utilized. As shown in Figs. 24.1 and 24.2, the only difference in the notch and peak filter pair is an output adder or subtractor; therefore, these filter pairs can be easily implemented as a single filter.

24.2.1 Notch and Peak Filter Design Using First Order APF

The notch filter can be obtained by utilizing the first order APF $A_1(z)$ in the design as shown in Fig. 24.3. The notch filter transfer function using first order allpass is given by

$$Q_0 = \frac{1}{2}[1 + z^{-1}A_1(z)] \tag{24.5}$$

where

$$A_1(z) = \frac{k_1 + z^{-1}}{1 + k_1 z^{-1}} \tag{24.6}$$

with $k_1 = -\cos\omega_0$ where ω_0 is the notch frequency. In the above equation, $A_1(z)$ is multiplied by z^{-1} in order to change the phase of the second term. For the first order notch filter, $\frac{\pi}{2}$ is the fixed value of rejection band width, and it does not depend upon the notch frequency [6]. It is observed that the first order peak and notch filters are interchangeable by simply replacing the adder with subtractor in Fig. 24.3. The transfer function of peak filter using first order APF is given by

$$W_0 = \frac{1}{2}[1 - z^{-1}A_1(z)]] \tag{24.7}$$

24.2.2 Second-Order Notch and Peak Filter

It is observed that the second-order allpass filter can also be used to implement notch and peak filter pair and provides control over frequency and bandwidth of notch or peak filter [2]. The second-order allpass function is chosen, which results into change in phase of $A(e^{j\omega})$ for ω ranges from 0 to π is -2π radians. The notch filter frequency response is as follows

$$Q(e^{j0}) = Q(e^{j\pi}) = 1$$

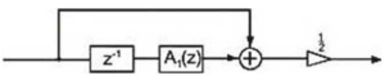


Fig. 24.3 Notch filter using first order allpass

$$Q(e^{j\omega_0}) = 0$$

where the angular frequency is ω_0. The allpass filter provides a phase shift of π radian at ω_0. The notch filter transfer function using second-order allpass is written as

$$Q(z) = \frac{1}{2}[1 + A_2(z)] \tag{24.8}$$

where

$$A_2(z) = \frac{k_2 + k_1(1 + k_2)z^{-1} + z^{-2}}{1 + k_1(1 + k_2)z^{-1} + k_2 z^{-2}} \tag{24.9}$$

The second-order allpass filter allows autonomous tuning of the notch frequency ω_0 and the 3-dB attenuation bandwidth Ω [22] and given as follows.

$$k_1 = -\cos\omega_0 \tag{24.10}$$

$$k_2 = \frac{1 - \tan\frac{\Omega}{2}}{1 + \tan\frac{\Omega}{2}} \tag{24.11}$$

The peak filter transfer function is given by

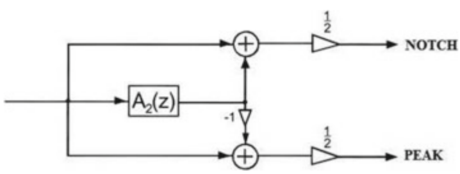
$$W(z) = \frac{1}{2}[1 - A_2(z)] \tag{24.12}$$

The implementation of the complementary notch and peak filters is shown in Fig. 24.4. Two filters are said to be complementary if the passband of one match the stopband of the other. Notch and peak filers are power complimentary as they satisfies the condition

$$|Q(e^{j\omega})|^2 + |W(e^{j\omega})|^2 = 1, \quad \text{for all } \omega \tag{24.13}$$

Complementary filters are used in discrete signal processing applications where specific frequency bands are processed separately. All the above-discussed realizations are structured using direct form II and implemented using registers, Vedic multipliers and different adders on FPGA. The proposed Vedic notch and peak filters are consuming less power and small hardware compared to conventional notch and peak filters.

Fig. 24.4 Complementary notch-peak filter as a sum and difference of APF



24.3 Vedic Multiplier Architecture

In this paper, the multipliers are based on a Vedic sutra 'Urdav Tiryagabhyam' (vertical and crosswise). In Vedic multiplication, the generation and additions of partial product are done in parallel. It results into enhanced speed of multiplier. As depicted in Fig. 24.5, the Vedic algorithm generates the partial products for 2-bit binary numbers. The lines indicate multiplication which is achieved using AND gates. The partial products are then added using adder circuit [17]. A 2-bit multiplier shown in Fig. 24.6 is designed using above-mentioned Vedic sutra. Similarly, a 4-bit Vedic multiplier is designed using 2-bit multiplier and shown in Fig. 24.7 [17].

24.3.1 High Speed Vedic Multiplie

In conventional designs of multipliers, ripple carry adders (RCA) are used. When implemented for large word-lengths, RCAs contribute a large delays. Therefore, several architectures have been proposed to where RCAs are replaced with high speed adders [23]. It results in reduced delay of overall multiplier. In presented work, 4×4 and 8×8 fast Vedic multipliers are used to design VNPF and implemented on FPGA. The generation of the partial outputs from the low order multipliers occurs concurrently. Their performance is analyzed in terms of power and area (hardware utilization) attributes. The high speed 4×4 and 8×8 Vedic multipliers are depicted in Figs. 24.8 and 24.9, respectively. The adders referred in figures can be replaced with any high speed adder of their respective lengths.

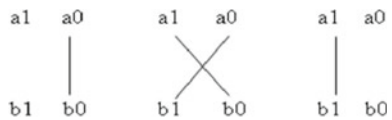


Fig. 24.5 Vdic Sutra Urdhva Tirakabhyam for 2-bit binary numbers

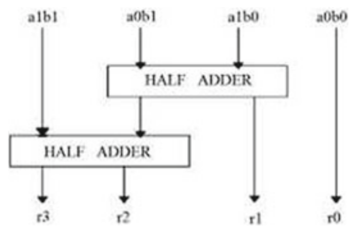


Fig. 24.6 2-bit Vedic multiplier using Urdhva Tiryakabhyam

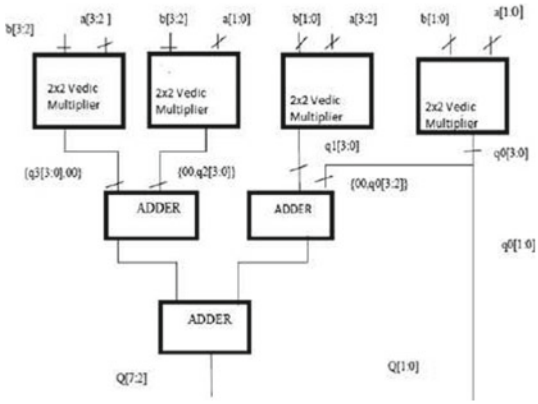


Fig. 24.7 4-bit Vedic multiplier using Urdhva Tiryakabhyam

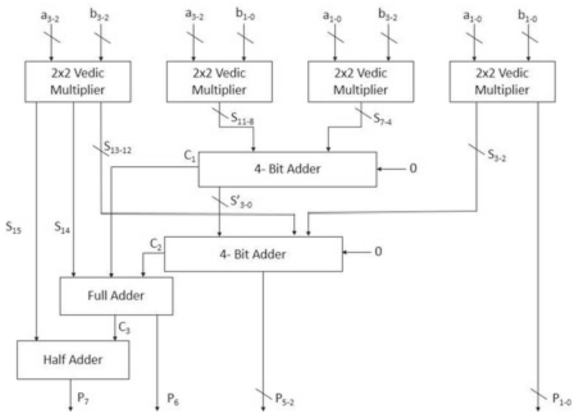


Fig. 24.8 4-bit high speed Vedic multiplier

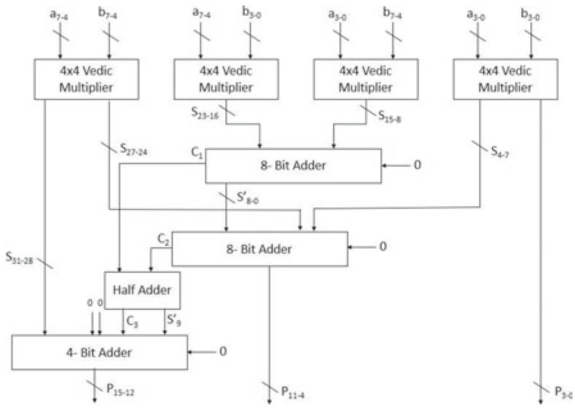


Fig. 24.9 8-bit high speed Vedic multiplier

24.4 Hardware Implementation of Proposed Notch and Peak Filters

24.4.1 Example

In this example, implementation of proposed Vedic notch and peak filter is shown by an example. Verilog HDL language is used to code the proposed VNPF, and implementation is carried on Xilinx Zync XC7z007$sclg$400 FPGA board. The input signals are a global clock, the parallel bits, and a reset. The notch and peak frequency specifications and band width of notch and peak filter are selected as:

$$\omega_0 = 50 \text{ Hz or } 0.667\pi \text{ radians/sample, } \mathrm{f_s} = 150 \text{ Hz} \tag{24.14}$$

$$Q = 20 \text{ and BW} = \frac{\omega_0}{\mathrm{Q}} = 0.0333 \tag{24.15}$$

The normalized frequency is calculated as

$$\left(\frac{\omega_0}{2 * f_s}\right) \times \pi \text{ radians} \tag{24.16}$$

The numerator and denominator coefficients obtained from the above specifications are depicted in Table 24.1.

The transfer function of the allpass shown in Eq. 24.17 is utilizing the coefficients specified in Table 24.1. The magnitude versus frequency response of the corresponding VNPF is presented in Fig. 24.10. In addition, pole-zero analysis is also made to check the stability of both the notch and peak filters and depicted in Figs. 24.11 and 24.12, respectively. As the poles are lying inside the unit circle, therefore, the above shown filter is stable.

$$A(z) = \frac{0.9567 + 0.9788z^{-1} + z^{-2}}{1 + 0.9788z^{-1} + 0.9567z^{-2}} \tag{24.17}$$

Table 24.1 Numerator and denominator coefficients of allpass filter

Numerator coefficients	Decimal value	8-bit binary
b_0	0.9576	11110101
b_1	0.9788	11111010
b_2	1	00000001
Denominator coefficients	Decimal value	8-bit binary
a_0	1	00000001
a_1	0.9788	11111010
a_2	0.9576	11110101

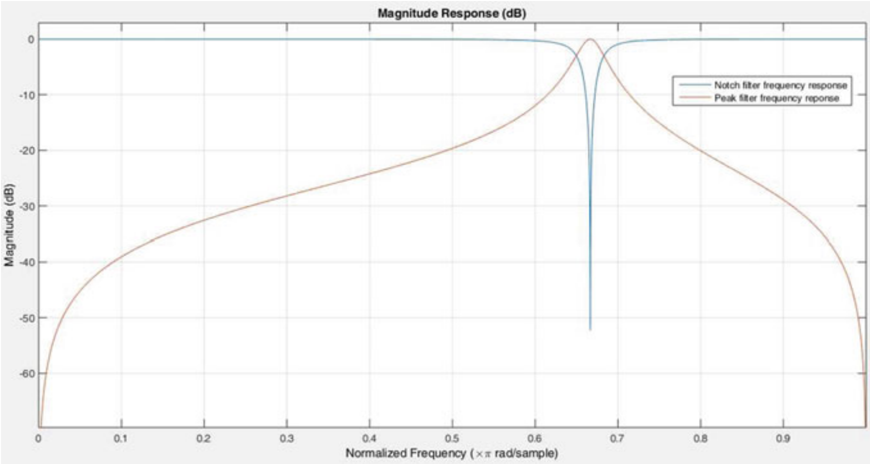
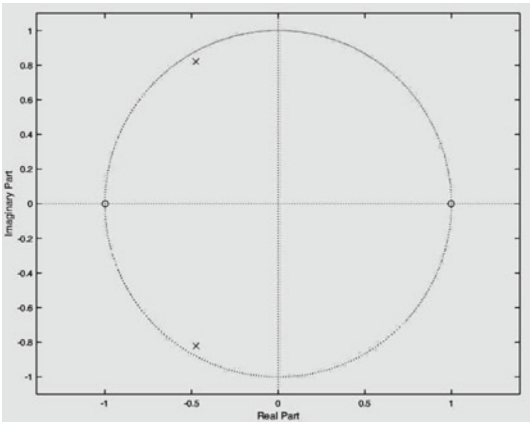


Fig. 24.10 Frequency response of the proposed VNPF

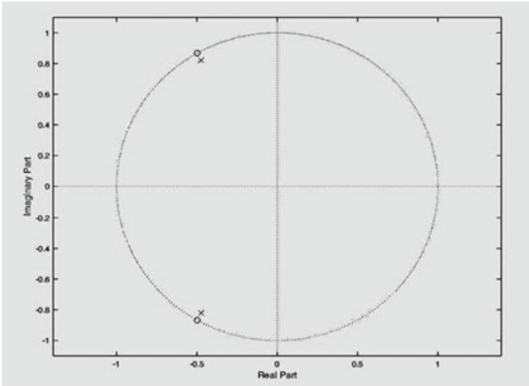
Fig. 24.11 Pole-zero plot of Notch filter



$$Q(z) = \frac{1}{2}(1 + A(z)) \quad W(z) = \frac{1}{2}(1 - A(z)) \tag{24.18}$$

where $A(z)$ is transfer function of allpass filter and $Q(z)$ and $W(z)$ are the transfer functions of notch and peak filters, respectively. In this example, the multipliers in notch and peak filters are implemented with Vedic multipliers. Further, to compare the performance of different adders in proposed filters, the adders in multipliers and overall filters are implemented with carry save adders (CSA), carry select adders (CSLA), and RCAs, respectively. Additionally, all three VNPFs are compared with conventional notch and peak filter implemented with parallel multipliers. The comparison results are summarized in Table 24.2.

Fig. 24.12 Pole-zero plot of peak filter



24.5 Result Analysis

All the proposed VNPF and conventional notch and peak filters are written in Verilog HDL and implemented on the Zync FPGA device. A comparative study of filters given in design example is also evaluated for performance and design efficiency. All the notch and peak filters are having coefficient wordlength of 8-bit and 150 Hz sampling frequency. Both the input and output word-lengths are scaled to 8-bit. The simulation waveforms for Vedic multiplier depicted in Fig. 24.13 show the multiplication results of two decimal numbers. In this figure, ip1 and ip2 are the inputs, while op is the output of the Vedic multiplier. For example, one fixed number 63 as ip1 is multiplied with different numbers in ip2, and result is shown in output op. The simulation results for notch and peak filter designed using Vedic multiplier (VNPF) are shown in Fig. 15. Additionally, their performances are compared with the existing notch and peak filters in terms of hardware utilization and power dissipation. The comparison results shown in Table 24.2 recapitulate that proposed VNPF is utilizing small hardware(LUT and slices) and less power compared to the existing notch and peak filters. Furthermore, CSA-based VNPF is consuming less power compared to CSLA based VNPF but more than RCA based VNPF. The hardware performance of CSLA-based VNPF is lying between other two VNPF. RCA-based VNPF is the best in terms of both hardware utilisation and power. However, RCA offer large delays for large word-lengths. Therefore, for high speed applications, CSA-based VNPFs are preferred (Fig. 24.14).

24.6 Conclusion

Novel Vedic notch and peak filters based on Vedic multipliers have been proposed. The performance of the proposed filters is compared with existing notch and peak filters. It is noticed that the proposed VNPF has better performance compared to

Table 24.2 Hardware utilization and total power summary of the proposed VNPF and existing notch and peak filters using Zync-7000 (xc7z007clg400) FPGA

VNPF	Slices (13300)	LUTs (53200)	Registers (106400)	Power (W)
Using CSA	40	140	30	26.05
Using CSLA	37	131	30	26.69
Using RCA	37	121	30	25.09
Conventional Notch/peak	Slices (13300)	LUTs (53200)	Registers (106400)	Power (W)
	48	146	30	29.08

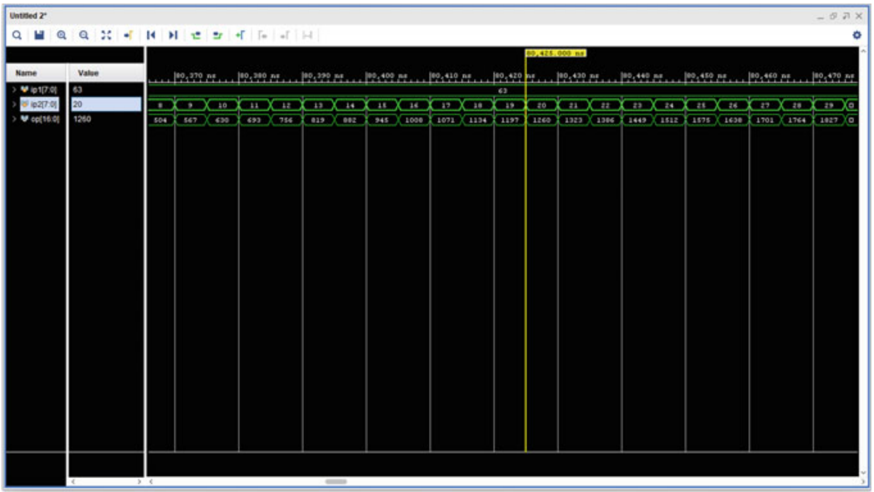


Fig. 24.13 Output of 8-bit Vedic multiplier

conventional notch and peak filters. In addition, VNPF is designed with CSA, CSLA, and RCA, respectively, and their performance comparison concludes that for power consumption, CSA-based VNPF is lying between CSLA and RCA based VNPF. However, CSLA-based VNPF is utilizing small hardware compared to CSA-based VNPF but more than RCA based VNPF. For both hardware and power performances, RCA-based VNPF is the best. However, RCAs offer large delays for large word-

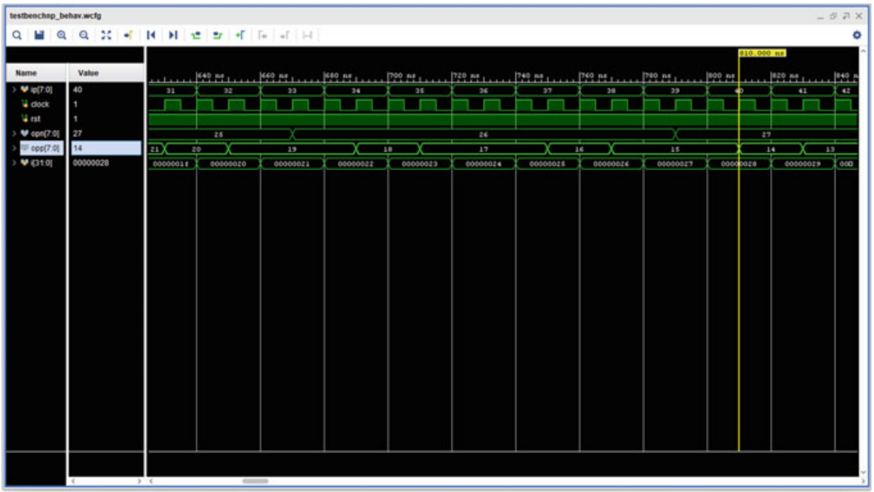


Fig. 24.14 Simulation output of VNPF

lengths. Therefore, for high speed applications, CSA-based VNPFs are preferred. On the other hand, in applications where low complexity is needed and speed is not a major factor, RCA-based VNPF can be used.

References

1. McManus, C.D., Neubert, K.-D., Cramer, E.: Characterization and elimination of AC noise in electrocardiograms: a comparison of digital filtering methods. Comput. Biomed. Res. **26**(1), 48–67 (1993). https://doi.org/10.1006/cbmr.1993.1003
2. Hirano, K., Nishimura, S., Mitra, S.: Design of digital notch filters. IEEE Trans. Commun. **22**(7), 964–970 (1974). https://doi.org/10.1109/tcom.1974.1092311
3. Abel, J., Berners, D.P.:. Filter Design Using Second-Order Peaking and Shelving Sections. ICMC (2004)
4. Pei, S.-C., Tseng, C.-C.: IIR multiple notch filter design based on allpass filter. IEEE Trans. Circuits Syst. II: Analog Digit. Signal Process. **44**(2), 133–136 (1997). https://doi.org/10.1109/82.554450
5. Dutta Roy, S.C., Kumar, B., Jain, B.S.: Fir notch filter design: a review. Facta Univ. Ser.: Electron. Energ. **14**(3), 295–327 (2001)
6. Joshi, Y.V., Dutta Roy, S.C.: Design of IIR digital notch filters. Circuits Syst. Signal Process. **16**(4), 415–427 (1997). https://doi.org/10.1007/bf01198059
7. Barsainya, R., Agarwal, M., Rawat, T.K.: Design and FPGA implementation of multiplierless comb filter. Int. J. Circuit Theory Appl. **45**(11), 1497–1513 (2017). https://doi.org/10.1002/cta.2324
8. Barsainya, R., Aggarwal, M., Rawat, T.K.: Minimum multiplier implementation of a comb filter using lattice wave digital filter. In: 2015 Annual IEEE India Conference (INDICON) (2015). https://doi.org/10.1109/indicon.2015.7443491

9. Milić, L.D., Lutovac, M.D.: Efficient algorithm for the design of high-speed elliptic IIR filters. AEU—Int. J. Electron. Commun. **57**(4), 255–262 (2003). https://doi.org/10.1078/1434-8411-54100168
10. Farooq, U., Jamal, H., Khan, S.A.: Transformation to implement computationally efficient IIR decimation filters. Digit. Signal Process. **19**(1), 33–44 (2009). https://doi.org/10.1016/j.dsp.2008.01.004
11. Poučki, V.M., Žemva, A., Lutovac, M.D., Karčnik, T.: Elliptic IIR filter sharpening implemented on FPGA. Digit. Signal Process. **20**(1), 13–22 (2010). https://doi.org/10.1016/j.dsp.2009.04.012
12. Poornima, M., Patil, S.K., Shivukumar Shridhar, K.P., et al.: Implementation of multiplier using Vedic algorithm. Int. J. Innov. Technol. Expl. Eng. **2**(6), 219–223 (2013)
13. Tirtha, S.B.K.: Vedic Mathematics or Sixteen simple Mathematical Formulae from the Vedas. Motilal Banarsidass, Delhi (1970)
14. Suma, S., Sridhar, V.: Design of multiplier for medical image compression using Urdhava Tiryakbhyam Sutra. Int. J. Electr. Comput. Eng. (IJECE) **6**, 1140 (2016). https://doi.org/10.11591/ijece.v6i3.9031
15. Gaikwad, K., Chavan, M.: Vedic mathematics for digital signal processing operations: a review. Int. J. Comput. Appl. **113**, 10–14 (2015)
16. Kunchigi, V., Kulkarni, L., Kulkarni, S.: Simulation of Vedic multiplier in DCT applications. Int. J. Comput. Appl. **63**, 27–32 (2013)
17. Sharma, S., Sharda, V.: Design and analysis of 8-bit Vedic multiplier in 90nm technology using GDI technique. Int. J. Eng. Technol. **7**(3.12), 759 (2018). https://doi.org/10.14419/ijet.v7i3.12.16496
18. Kaliannan, S., Kumar, P.: Design and performance analysis of reconfigurable modified Vedic multiplier with 3-1-1-2 compressor. Microprocessors Microsystems **65** (2019). https://doi.org/10.1016/j.micpro.2019.01.002
19. Savadi, A., Yanamshetti, R., & Biradar, S. (2016). Design and Implementation of 64 Bit IIR Filters Using Vedic Multipliers. Procedia Comput. Sci. **85** (2020). https://doi.org/10.1016/j.procs.2016.05.267
20. Anjana, S., Pradeep, C., Samuel, P.: Synthesize of high speed floating-point multipliers based on Vedic mathematics. Procedia Comput. Sci. **46**, 1294–1302 (2015). https://doi.org/10.1016/j.procs.2015.01.054
21. Bianchi, V., De Munari, I.: A modular Vedic multiplier architecture for model-based design and deployment on FPGA platforms. Microprocessors Microsystems **76** (2020). https://doi.org/10.1016/j.micpro.2020.103106
22. Regalia, P.A., Mitra, S.K., Vaidyanathan, P.P.: The digital all-pass filter: a versatile signal processing building block. Proc. IEEE **76**(1), 19–37 (1988). https://doi.org/10.1109/5.3286
23. Nehru, K.K., Linju, T.T.: Design of 16 bit Vedic multiplier using semi-custom and full custom approach. J. Eng. Sci. Technol. Rev. **10**, 220-232 (2017). https://doi.org/10.25103/jestr.102.25

Chapter 25
Nearpod: An Effective Interactive ICT Tool for Teaching and Learning Through Google Meet

Anil S. Naik, Pravin N. Kathavate, and Shivappa M. Metagar

Abstract In the digital era of education, the use of ICT tools is important for all human activities. School children, college students, graduates of various streams, and teachers are dependent on digital technology for teaching and learning. Student interaction and engagement is one of the challenging tasks in an online classroom for teachers of various streams to deliver the knowledge efficiently. An interactive Web-based platform improves the learning experiences of students and reduces the teacher effort. The transition from traditional lecture-based teaching methods to interactive learning can be accomplished by using various interactive tools such as Nearpod, Kahoot, and Socrative. Nearpod is a Web-based tool that allows the teacher to engage students during the online virtual sessions by creating contents with the addition of video, slide, Web content, Nearpod 3D, slideshow, audio, PhET Simulation School-level mathematics concepts, VR field trip, and Sway and also to create interactive lessons by including various activities such as Time to Climb, Quiz, Flipgrid, Poll, Fill in the Blanks, Memory Test, Matching Pairs, Collaborate Board, Open-Ended Question, and Draw It. This paper describes the various features of Nearpod and content creation, using Nearpod for effective teaching and learning.

25.1 Introduction

Digital technology facilitates teaching and learning methods in education using Google Meet, Zoom platform, etc. Students participate in virtual lecture sessions and finds a lack of interaction, lack of communication skills, and lack of face-to-face communication. There are various approaches presented by authors to improve the interaction between students and teachers.

The use of interactive technology such as Google Meet and the Zoom platform for teaching, meeting, and interaction in visual and audio appearance is more efficient, effective, and safe. The challenge of handling the teaching process in engineering and technology subjects to cover the syllabus in a stipulated time [1].

A. S. Naik (✉) · P. N. Kathavate · S. M. Metagar
Walchand Institute of Technology, Solapur, Maharashtra, India

T. Senjyu et al. (eds.), *IOT with Smart Systems*, Smart Innovation, Systems and Technologies 251, https://doi.org/10.1007/978-981-16-3945-6_26

The author described that large class teaching in higher education is a challenging aspect to interact with everyone by adopting teaching–learning practices to bring quality education and finds a lack of practices and policy in teaching and learning [2].

The author discussed cloud-based technologies Nearpod to provide active and collaborative teaching–learning for large classes [3].

The students are expected to actively participate in a virtual platforms-based class, and effective learning can be ensured if students apply the gained knowledge from sessions. To create an effective teaching–learning environment, there should be meaningful interaction between teachers and students.

A study of student's participation in the virtual classroom through Google Meet is presented with feedback received from students based on survey questions conducted for a class of 60 students and 48 students contributed to the survey. Students do not actively participate in a virtual classroom session due to various reasons, but few of them are represented through the graphs.

Figure 25.1 shows participation of students on virtual classroom for 40 min. The attention of 15 (7 + 8) students is between 30 and 40 min. It means that teacher needs to engage students by questionaries' or through activity.

Figure 25.2 shows the few reasons for loss of attention in a long continuous virtual session. Thirty-six (22 + 14) students mentioned that due to continuous long sessions and it indicates teacher should allow students some kind of break or any interaction.

Figure 25.3 shows the student's behaviours during the virtual classroom are collected and most of the students are hesitant to interact with the teacher.

In the digital era, students have adopted technology in academics to improve the performance. To cope up with these, the teacher needs to use different strategies for teaching and learning. Based on the feedback survey, student's participation in the virtual session is addressed by using Nearpod applications.

The main focus is to engage students with various assessment methods, active participation strategy, and content creation. Since the inception of smartphones' use of educational applications, WhatsApp application, Web-based Kahoot and Socrative tools, etc. [4], are fast-growing in the assessment of student activities in the classroom

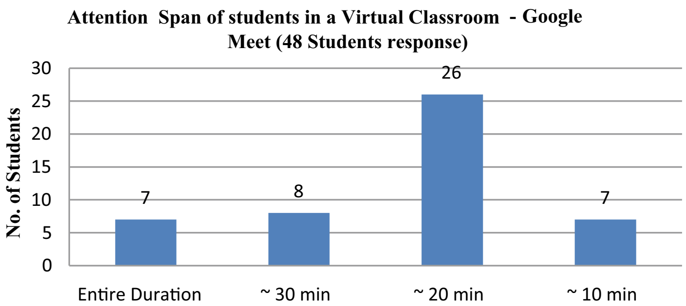


Fig. 25.1 Attention span of students in a virtual classroom—Google Meet

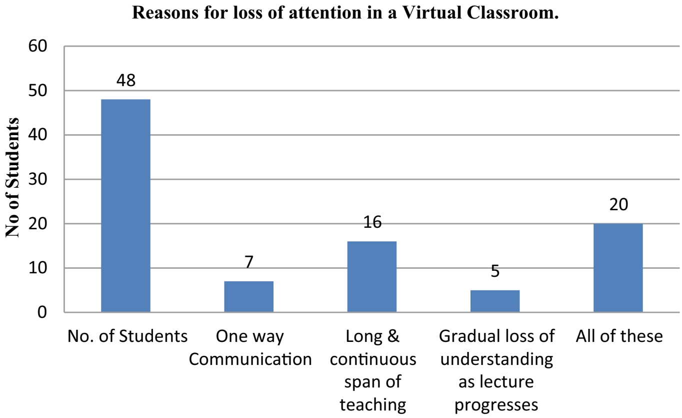


Fig. 25.2 Reasons for loss of attention in a virtual classroom

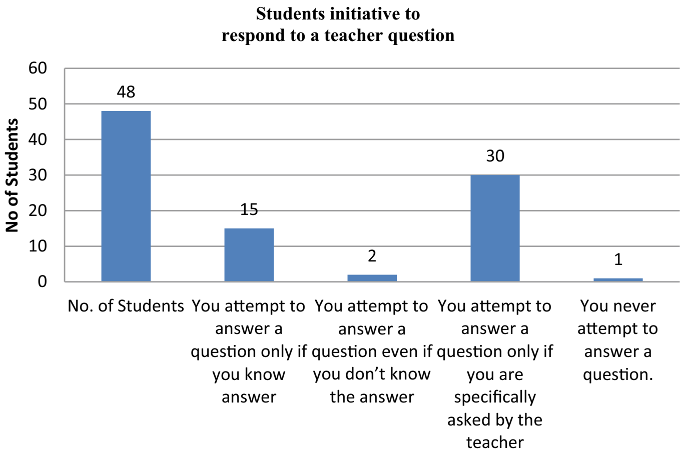


Fig. 25.3 Students initiative to respond to a teacher question

or teaching–learning procedure. All these are facilitated by wireless technology and smart devices available at a low cost (Table 25.1).

Nearpod is the most interactive, engaging, and motivation tool compared with Kahoot and Socrative as per students and teacher experience. It is also easy to use for teachers to create content by embedding various activities.

Table 25.1 Various features of **Nearpod** application

S. No.	nearpod
1	Free online assessment and interactive tool for teachers to create content by embedding features video, slide, web content, Nearpod 3D, slideshow, audio, PhET simulation, and VR field trip [7]
2	Desktop computers, laptop, mobile devices, and tablet. It supports all browser types
3	Activities of gaming type—"Time to Climb", "Collaborate Board", "Draw It" and other activities Quiz, Flipgrid, Poll, Fill in the Blanks, Memory Test, Matching Pairs
4	All features accessed by "unique code" can share through email, social, link, Google Classroom, Microsoft Teams
5	Results of the whole class and individual students are available with progress and statistics reports
6	Students can use teacher presentations by unique code shared with students

25.2 Use of Nearpod

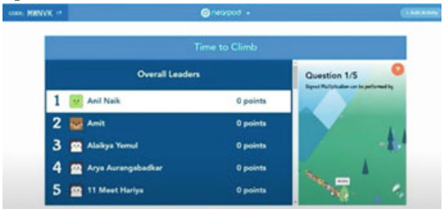
Nearpod is a Web-based application accessible in various low-cost available devices, is easy to use, and supports all browsers with an Internet connection [5–7].

Steps to Use Nearpod

- Teachers can share unique codes to access created presentations in the Nearpod application.
- Students log in by inputting unique code shared by the teacher.
- After successful login, students can see the teacher's presentation on his device screen.
- The teacher is having control of the presentation with restrictions to view the next slide by students.
- The teacher can view the number of students active during the session.
- The teacher can also monitor individual student progress and participation in all activities in real-time and also in a graded report.
- Activities such as Time to Climb, Quiz, Polls, Draw It, and Open-Ended Question can be added by the teacher to his presentations, and it allows students to actively participate in a live session.
- The teacher can monitor student progress and understanding of concepts based on student response in activities.

Table 25.2 describes the traditional classroom-based teaching and interactive teaching using Nearpod.

Table 25.2 Difference between traditional classroom teaching and Nearpod interactive teaching and learning

S. No.	Traditional classroom teaching 	Nearpod interactive teaching and learning [5–7, 9] 
1	Students in class use their study materials such as notes, textbooks, and pen. Teachers use blackboard and display contents using LCD projectors to visualize and textbooks to describe topics	Students in a live session at their place can use a smartphone, laptop, or any device using the Nearpod application. Students can view the presentation of the teacher and can participate in activities
2	During lectures, students can take notes and listen to the teacher but the teacher cannot assure student participation in the activity	During online virtual sessions, students can view teacher study content on a device with restrictions to move slides and can monitor the student participation in a quiz, poll, and team collaboration activity
3	After the lecture, the teacher asks questions or no activity	After the session, the teacher can conduct a quiz or poll to know student's understanding of the topic and activity participation of a session in the graded report

25.3 Result Analysis

The virtual session is delivered using the Nearpod application through the Google Meet platform. During the session, *Time to Climb, Matching Pairs, Open-Ended Question,* and*Draw It* activities are conducted to students to assess the topic understanding and to monitor the student progress. With Nearpod, a grading report is generated at end of all activities automatically to the teacher-registered email ID by software. As per students' understanding, teacher can discuss the answer for each question at the end as formative feedback. The teacher can realize the topic understanding based on student response through grade reports. Figure 25.4 shows a screenshot of the activity, and a video link for the activity conducted is provided [8].

Benefits of Nearpod activity with teacher perspective are to design and create content by embedding activities on a topic and to control presentation during the session. The teacher can also view how many students are actively participating and idle during the activity. The teacher also hides the student name during activity so students should not feel uncomfortable, if he/she shows poor performance in an activity. Students also expressed their positive opinion on the use of Nearpod application compared to Socrative and Kahoot application.

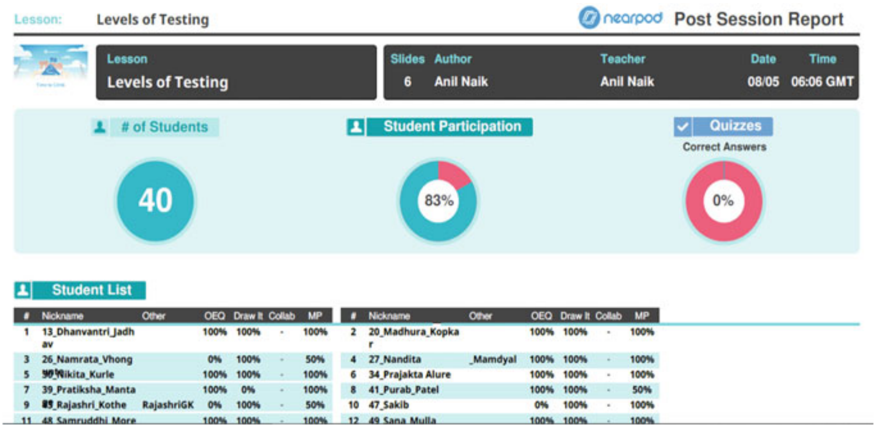


Fig. 25.4 Activity overall summary report

Six virtual sessions through Google Meet are conducted, and feedback survey of activity participation is collected. A total of 48 students are contributed and have shown interest to participate in future sessions with overall rating 4.5 out of 5 for Nearpod application (Table 25.3).

25.4 Conclusion

The use of the Nearpod application is a great learning experience for teachers and students. Interactive teaching–learning can be achieved in virtual teaching by using the Nearpod application compared to the traditional teaching method to engage students in a session for effective teaching and learning experience. Based on the survey outcomes, it is concluded that Nearpod is easy to use in teacher and student perspectives and increases student engagement. Nearpod application has effectively addressed issues such as improving the attention span of students, the participation of students in the class, the conceptual understanding of students.

The limitations of Nearpod applications are the free version limiting to 40 students for a session, and if you have weak Internet, then Nearpod will be slow and lagging.

The teacher can monitor individual student progress in real time through the generated report, can maintain the student participation report of all activity, and also can find out students who are showing poor performance during session activities.

Table 25.3 Benefits derived by using the Nearpod application

Have you participated in Nearpod activities?		Did you understand the topic conducted through Nearpod application in a Google Meet			How did you find the activity through Nearpod?	Does Nearpod application make you boring in a Google Meet		Do you want future sessions based on Nearpod application through a Google Meet		Overall rating on activity
Yes	No	Fully	Partially	Not understood	Average out of 5	Yes	No	Yes	No	Average out of 5
48	0	30	16	2	4.2	0	48	48	0	4.5

References

1. Pratama, H., Azman, M.N.A., Kassymova, G.K., Duisenbayeva, S.S.: The trend in using online meeting applications for learning during the period of pandemic COVID-19: a literature review. J. Innov. Educ. Cult. Res. **1**(2), 58–68 (2020)
2. Hornsby, D.J., Osman, R.: Massification in higher education: large classes and student learning. In: Higher Education. Springer (2014)
3. McClean, S., Crowe, W.: Making room for interactivity: using the cloud-based audience response system Nearpod to enhance engagement in lectures. J. Invset. Sci. FEMS Microbiol. Lett. **364** (2017)
4. Nirgude, M., Naik, A. (2017). Whatsapp application: an effective tool for outof-class activity. J. Eng. Educ. Transform. Spec. Issue. eISSN 2394-1707
5. Mekota, T., Marada, M. (2020). The influence of the Nearpod application on learning social geography in a grammar school in Czecha. In: Education and Information Technologies, part of Springer Nature (2020)
6. Sanmugam, M., Selvarajoo, A., Ramayah, B., Lee, K. (2020). Use of nearpod as interactive learning method. In:Conference Paper.
7. Shehata, N., Mitry, C., Shawki, M., El-Helaly, M.: Incorporating Nearpod in undergraduatefinancial accounting classes in Egypt. Accounting Education **29**(2) (2020)
8. Naik, A.S.: Video link "software testing | Nearpod | make every lesson interactive": https://youtu.be/UrOQjZ2GWK0
9. Nearpod homepage, https://nearpod.com/

Chapter 26
Enhancement in Security for Intercloud Scenario with the Help of Role-Based Access Control Model

Rashmi Dixit and K. Ravindranath

Abstract Cloud according to the latest buzz is the third wave of digitization. Cloud brought all things back at the center. Cloud redefines the way of computing—"where you computing is not important than how you compute". Sharing of public resources like platform, software and infrastructure easily depending upon requirements. The connectivity of cloud eroded the division of resource sharing. The ability to access data from the cloud—anywhere, anytime—increases the security risk. Security is an essential ingredient to reduce security risk. Trust on technology is good, but as data owner is not with data all the time, so control is better. Cryptographic scheme, identity-based proxy re-encryption, is used to transform data before transferring into public environment. Data owner is not with data, so rule-based data-centric solution is proposed to keep data secure. For preventing unauthorized access of data, role-based authorization access control is developed. The solution is developed using OpenStack.

26.1 Introduction

Cloud computing is the demand availability of resources over the Internet. Instead of physically buying, owning, and maintaining resources, we can use computing power, storage, and databases, on an as-needed basis from a service provider.

When any individual first consider about storage on cloud, i.e., storing of any music or any multimedia file or software on a remote machine which is accessed through many devices any time depending on need. Basically, cloud storage is a structure which gives you facility to store data on the cyberspace, same like saving on your PCs hard disk. Whether you are considering any cloud, viz. Google Drive, Dropbox, or iCloud, the basics of storing data on data remains the same.

R. Dixit (✉)
Walchand Institute of Technology, Solapur, India

K. Ravindranath
Department of CSE, Koneru Lakshmaiah Education Foundation, Vaddeswaram, India
e-mail: ravindra_ist@kluniversity.in

T. Senjyu et al. (eds.), *IOT with Smart Systems*, Smart Innovation, Systems and Technologies 251, https://doi.org/10.1007/978-981-16-3945-6_27

Cloud provides facility to upload data which can be used by anyone (to whom owner provides access), anywhere using the Web. When we are storing data on to cloud, data protection methods are required.

Different data protection methods all are applicable when like any traditional data center. Cloud totally depends on cloud service provider (CSP) for applying various security measures. Service-level agreement takes place between data owner and CSP before uploading data or use of any service, so that owner can trust on CSP. Moreover, data owner provides access control for other user who wants to access data. But scenario is different in case of intercloud when data flows from one cloud to another. At this stage, access control privileges are not applicable to another cloud which may be operated by another CSP. This condition leads to improvement in privileges defined by access control. Making data self-protected whenever it resides is good option to improve security. Whenever we talk security, encryption comes in mind to protect data in the cloud which avoids undesired access. Cloud security alliance also provides. Guideline to provide protect to data at any stage may be at relax, moving and in operate mode. But it gives new issue: access control. Access control with rule is the best way to provide expressiveness. But, how to apply this for data-centric approach? Who will compute rule since data has no computation ability? CSP can evaluate, but owner may not trust. Data owner may evaluate, but for this data owner should be online for each access request evaluation based on RBAC. RBAC with data at centric approach is used in which data is encrypted and self-protected.

This paper presents DCRBAC, a data-centric access control solution in untrusted environment, and provides extended expressiveness for encrypted data. The proposed solution provides data-centric RBAC scheme, where roles are defined for access control of the resources. Authorization model defines hierarchies to providing rule expressiveness that apply to multiple users. Along with data self-protection, novel cryptographic techniques such as identity-based proxy re-encryption (IBPRE) are used which enhance the security by re-encrypted data using owner identity.

Original message using keys is encrypted to ciphertext, and at the same time, keys are associated with authorization model. To allow data access only to privileged user, rules are defined and these rules are also encrypted, so that unauthorized access or misbehavior is prevented. User-centric policy is implemented to define authorization rules as data owner has permission to decide access control rights.

Two main proposals are presented in

- Improved solution for data which is pivotal with rule-based approach.
- Role-based access control model for authorization rule.

26.2 Literature Review

This paper analyzes trust relationships among clients using cloud. Author proposes a classification level for trust. Confidentiality and possible risk decide trust relationships [1]. The author proposed solution based on data-centric approach. They also provide approach based on access control based on role [2]. Identity based proxy re-encryption in which third party entity, called proxy re-encrypt cipher too provide guaranteed security as compared to encryption. This scheme has advantage over other methods [3]. Access control along encryption avoids unauthorized access to data. Confidentiality is achieved through encryption, but authorized access control makes data secure at rest, motion, or use [4]. The author proposed model for OpenStack which has been proved with performance results [5]. This author proposes Security Ontology for the Intercloud (SOFIC) which gives detailed explain security aspect of intercloud security [6]. In this paper, author comment threats in cloud or intercloud scenario for data security along with solution [7]. This paper gives guideline the important attacks and threats for security of cloud. As well as they provide possible antidote to fix it. Awareness of different cloud security problems makes it feasible to solve them that is helpful to allay the risk link with the use of cloud computing [8].

26.3 Implementation Details

Not only to protect data from unauthorized user in cloud, novel cryptography schemes are used, but at the same time to provide privileges authorization model is used which is also secure. To make secure authorization model, proxy re-encryption which is based on identity method is used in which an entity, i.e., trusted third party, is used to re-encrypt data without seeing it with the help of public key.

Use of cryptography to protect data when moved to the cloud is the basic requirement. To avoid access by any unauthorized user, identity-based proxy re-encryption technique is used to protect the authorization model. A proxy re-encryption (PRE) method uses an entity called *proxy* to re-encrypt data from one key to another without being able to decrypt it (Fig. 26.1).

One-way, non-interactivity, multi-use features are satisfied because of identity-based proxy re-encryption. The management of proxy with IBE bypasses the requirement of public and private key pair. Instead, straight the identities can be worn within the cryptographic activity. Fusion of IBE and PRE makes a third party (proxy) to again encrypt a cipher using identity of particular user, i.e., re-encryption with identity of another one. In this way, identities of user are used by a Master Secret Directory

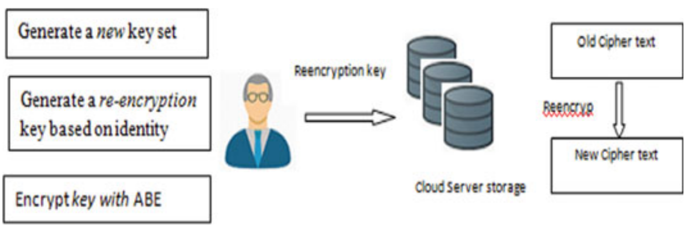


Fig. 26.1 Identity-based proxy re-encryption

(MSD) to develop secret keys. Secret keys used are identical to private keys in IBE. Public keys are not used, as identities are used instead of them in the cryptographic activity.

List of activities is mentioned below for the above-mentioned cryptography scheme to make secure data and rules:

$$\text{setup}\,(p_u,\ \text{key}) \rightarrow (p_u, msk) \tag{26.1}$$

/ initialization of cryptographic activity which takes input k and generate msk and public parameter p */*

$$\text{Gen\_key}\,(p_u, \text{msk}, id_a) \rightarrow s_a \tag{26.2}$$

/ With generated p and msk along with identity of use, secret key is generated* sk_α */*/*

$$\text{encrypt}\,(p_u,\ id_a, m) \rightarrow c_a \tag{26.3}$$

/ with public parameter p and identity, piece of data m is converted to cipher* c_α **/*

$$rkgen\,(p_u, s_a, id_a, id_b) \rightarrow rek_{a\rightarrow b} \tag{26.4}$$

/* Generates re encryption key, takes identity of both user along with secret key and public parameter */

$$\text{reencrypt}\,(p_u,\ rek_{a\rightarrow b},\ c_a) \rightarrow c_b \tag{26.5}$$

/* Reencrypt data */

$$\text{decrypt}\,(p_u,\ s_a,\ c_a) \rightarrow m \tag{26.6}$$

/* with public parameter, secret key cipher text is decrypted to original plain text */

Using trusted third party, i.e., proxy re-encryption with identity scheme is used for encryption of data. Blending of role based access control with rule based data centric solution for avoiding access from unauthorized user in Cloud.

Authorization model with high expressiveness is described which manages advanced characteristic like access control based on role. Semantic Web technology has the capability to build formal logic, and this is used by control model. Access control model uses this capability to design infrastructure beneath and authorization model. It also controls resource access in the cloud by employing rules. Moreover model with trust to define fine-grained access regarding available information for a particular. In the proposed paper, authorization model is implemented to control unauthorized data access. To implement this particular, to grant particular rights to a user (subject), concept of role is used.

One or more roles can be allocated to group of subject with set of privileges.

A data-centric security technique provides encryption of data to provide confidentiality. Access control mechanism will take care of authentication; i.e., only authorized users are able to decrypt it.

Figure 26.2 shows the relationship between role and identity. Identity may be mapped different roles via managed element which is scanned while assigning privileges. And according to that, set of privileges are permitted.

Authorization rules are defined with the base of tuples. Tuples are assigned with name which describes its role. This role plays an important role in defining grants of access. Figure 26.3 shows concept of role with particular rights.

For accessing data, user can be anyone like principal, professor, clerk, and student. Depending upon the user, roles are already defined and for each role privileges like READ ONLY ACCESS, READ WRITE ACCESS are already assigned. "Which data can be access by particular user or which user or roles have which access rights "are defined. This matrix is created in the form of this. End user already defines such rule, but modification to the existing rule or addition of new rule is also possible.

In our authorization model, authorization rules with the 3-tuple can be represented by either an identity or a role. An identity is a user that can be approved during the authentication stage, whereas a role is a set of responsibilities within an organization.

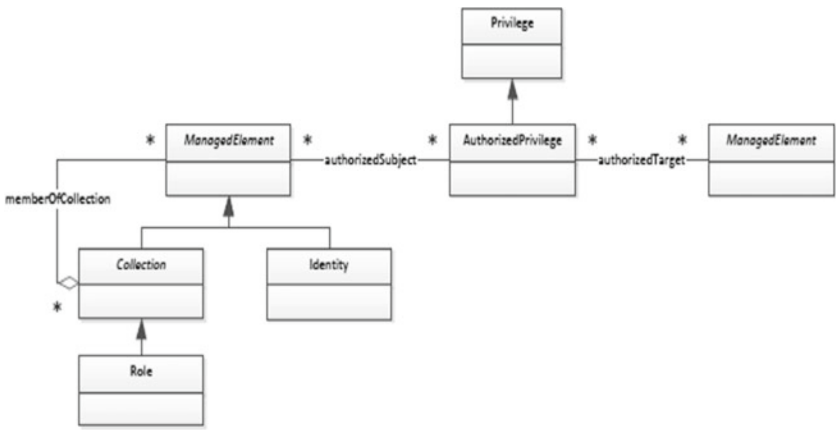


Fig. 26.2 UML diagram of authorization concepts

Fig. 26.3 Relation between user and privileges

Identities can be assigned to a role. Above all of these is a privilege which is the base for all types of activities/actions which are granted or denied. We are going to select any role having some associated privileges.

As already discussed above, collection of tuples is present in definition of particular authorization rule. For this particular proposal, 3 tuples described rules which are represented with an identity or a role. User is an identity which is validated at stage of authentication only. After that, set of responsibilities in the form of role is organized. Roles are assigned to particular identity. None other than privilege which decides mapping of allowing particular activity by particular user is above all. Identity with role has associated privileges. Identity and role selection are represented as shown below.

Role $(\pi) \wedge$ AuthorizedPrivilege (π)
Identity $(\pi) \wedge$ memberOfCollection $(\pi$, Role$(\pi))$

Any identity without role is never going to happen. All identities are always associated with role. New privileges can be added to particular identity by just switching of that identity to a new set. Because of this, two identities may share same privileges as they both have same role. Activation or deactivation of particular privileges to a particular identity is just by switching. Depending upon the available access rights, user can perform various actions like read/write/read–write. Before uploading/writing any new data in the cloud, encryption is performed which provides security to data. At the same time, same user is going to data who can operate on that data with the definition of rules to describe authorization model. And this authorization model decides privileges for a particular user.

Ultimate user may be either from same cloud or from a different cloud and is able to access data if and if only if evaluation of rules gives permit. If for that user access is allowed, then only that mapping allows either read or write. Decryption is performed on the basis of identity of user.

26.4 System Architecture

Figure 26.4 shows the proposed architecture. Data owner while uploading data to the cloud generates self-protected package which contains three things: encrypted data, access control definition in the form of authorization rule, and keys required for re-encryption. Data owner before moving data to cloud performs encryption using random symmetric key generated to provide security. Key which is used for encryption of data is also encrypted. Symmetric algorithm AES is used which is more sufficient. Next thing, i.e., authorization model, contains rules to provide access control. After the evaluation of rules, keys for re-encryption are generated. Data owner is only responsible with the help of generator and identities of elements as described above. Identity-based scheme is used which generates keys for re-encryption with avoiding extraneous attributes and storage requirements. After getting green signal in the form of re-encryption keys, ultimate user is able to perform any operation. CSP

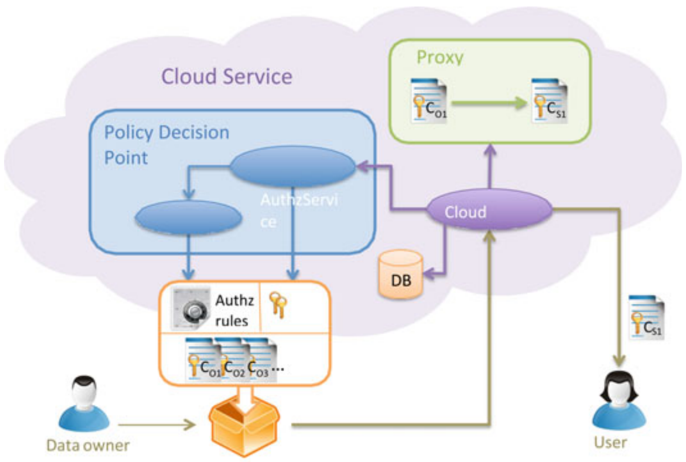


Fig. 26.4 System architecture

is not able to perform any misbehavior to access data. This results in data-centric solution as data carries all required things with itself only in either same service provider or intercloud scenario.

As shown in the above diagram, two components need to be installing, decision point to control authorization model and PRE to perform all cryptography-related actions. Authorized cloud service provider communicates with these components to run smoothly all cryptographic operations. Each and every package deployed by data owner is placed at database system of service provider with all things related to that object.

Ultimate user wants to use data and send a request to access that particular part of data. This request in the form of query sent to authorization service component acts as first step in access, which decides whether that particular user has rights or privileges to access that data. After evaluating rules, this decision returns statement about either grant or denied permission. In case of granting permission, re-encryption keys are also supplied because of this re-encryption keys. This information is sent to trusted third party, i.e., proxy which then performs proxy re-encryption. After that, wants to use data performs decryption with its set of keys. At any cost, service provider cannot misbehave with data as information for re-encryption is only with legitimate user.

26.5 Result and Discussion

Different tests measure the time to take decisions about access rights. Logic based on sematic Web ontology is used as described in above section. Subject, object decides

Table 26.1 Number of individuals by population

Population	Individuals
1	500
2	900
3	1400
4	2300
5	3700
6	6100

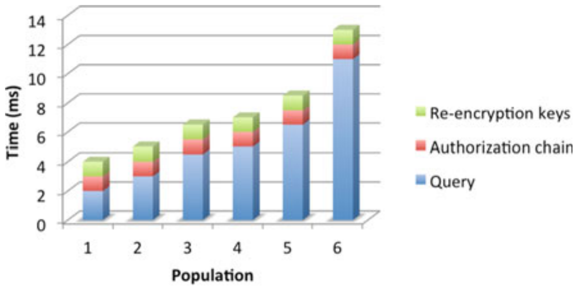


Fig. 26.5 Authorization decision times

decision time taken by authorization model. Knowledge base helped in deciding access rights based on formal logic prepared using inference stored in that. To check time taken by authorization model, 6 executions have been implemented as shown in Table 26.1.

Hundred sets of requests are executed for each population. Time taken can be divided into different operations; first part is to evaluate rule to decide privileges which executes query, and second to retrieve re-encryption keys. Figure 26.5 shows the decision time taken by each population.

It is observed that as population grows, time increases. But time for re-encryption keys and chain is same independent of individuals.

26.6 Conclusion

Rule based data centric Role based access control authorization model has been presented. It provides expressiveness including role and object hierarchy. Data-centric solution is provided for intercloud scenario whether data in use, in motion, or at rest. Access control model grants privileges to avoid unauthorized access to data. Authorization rules are also applied to access control for its security. While uploading data in the cloud, identity-based proxy re-encryption is used to double security.

Future lines of research include the use of obfuscation of the authorization model for privacy reasons.

References

1. Bernabe, J.B., Alcaraz Calero, J.M., Clemente, F.J.G., Perez, G.M.: Taxonomy of trust relationships in authorization domains for cloud computing (2014)
2. Marin Perez, J.M., Perez, G.M., Skarmeta Gomez, A.F.: SecRBAC: secure data in the clouds (2015)
3. Dixit, R.K., Ravindranath, K.: Encryption techniques and access control models for data security: a survey (2018)
4. Dixit, R.K., Ravindranath, K.: Identity based proxy re-encryption with null deduplication for secure communication in cloud environment (2019)
5. Bernabe, J.B., Calero, J.M.A. Clemente, F.J.G., Perez, G.M.: Semantic-aware multi-tenancy authorization system for cloud architectures (2012)
6. Bernabe, J.B., Perez, G.M., Skarmeta, A.: Intercloud trust and security decision support system: an ontology-based approach (2015)
7. Velumadhava Rao, R., Selvamanib, K.: Data security challenges and its solutions in cloud computing (2015)
8. Alhenaki, L., Alwatban, A., Alahmri, B., Alarif, N.: Security in cloud computing: a survey system design (2019)

Chapter 27
Professional Learning Community (PLC)—A Model Incorporating ET Practices for Continuous Improvement in Blended Teaching-Learning Process

Manisha Nirgude, Pratibha Yalagi, and Shashikant Halkude

Abstract Innovation in teaching-learning (T-L) is always necessary in education to help students to explore their full potential and to enhance creativity and critical thinking. Teachers' urge for innovation is one of the key factors for its success. It is possible to strengthen conventional T-L process through blending education technology (ET) practices for active learning. To facilitate disruptive innovations in T-L process, we have established professional learning community (PLC) in 2015. PLC focuses on e-content creation, continuous use of instructional strategies and ICT tools during course delivery, use of learning management system (LMS): Moodle for publishing courses, assessment and feedback. For creating contents in the form of videos, we have established recording studio titled 'E-Learning Centre.' As on date, we have created @3000 videos of 12-min average duration, which are published on institute's YouTube channel. Overall at institute level, on an average 15 instructional strategies and 10 ICT tools are employed in T-L process, leading to enhancement in student's learning while making T-L process a joyous experience. These blended T-L process is benefitting both students and faculty members to a great extent heralding the dawn of disruptive innovation in T-L process. This inculcation of blended T-L under PLC resulted into smoothly shifting overnight from physical classroom to an online mode of teaching during COVID-19 pandemic.

27.1 The Context/Background

Sudden expansion in engineering education has led to paucity of faculty members. Delivery of engineering education through large number of untrained faculty members has suffered significantly due to lower-quality course delivery and imparting poor skill set, leading to many fresh graduates remaining unemployed. However, for institution, it is possible to overcome such weaknesses and to ensure quality of T-L process, by empowering faculty members for quality course delivery

M. Nirgude (✉) · P. Yalagi · S. Halkude
Walchand Institute of Technology, Solapur, Maharashtra, India
e-mail: manirgude@witsolapur.org

T. Senjyu et al. (eds.), *IOT with Smart Systems*, Smart Innovation, Systems and Technologies 251, https://doi.org/10.1007/978-981-16-3945-6_28

through blended T-L process, while simultaneously bridging the industry–academia gap. With this objective, faculty members now have become proficient in ET practices and blending T-L process. Such training and practices in education technology (ET) are further being strengthened through an idea of localizing 'professional learning community' (PLC) platform. Various learning communities are functioning at school level for T-L process; however, there is no reporting of formation of learning community in engineering colleges [1].

27.1.1 Objectives of the Practice

- To provide a platform for faculty members to share and exchange views/suggestions regarding ET research and experiments, a way to improve T-L process continuously.
- To blend various learner-centric strategies and ICT tools in T-L process to enhance learning outcomes, while addressing diverse student groups.
- To review ET activities periodically and provide expertise/guidance through regular meetings and follow-ups.
- To engage teachers in peer discussions and assessments of the resources created, to make available course contents in the form of videos, OER, etc., and to create Learner-Centric MOOC (LCM).
- To facilitate students to learn anywhere, any times and many times.
- To facilitate faculty members to save some time due to creating e-contents and use of ICT for utilizing it efficiently for other academic activities.

Section 27.2 discusses formation and structure of PLC and disruptive innovations in T-L. In Sect. 27.3, we discussed assessment and feedback. Section 27.4 is about journey of PLC. The awards and recognition are presented in Sect. 27.5 followed by conclusion in Sect. 27.6.

27.2 Professional Learning Community (PLC)

PLC at our institute is constituted to continuously enhance T-L process, which is a blended formulation of administrator, ET experts and ET researchers, ET practitioners, departmental coordinators and novice learners. PLC organizational structure is shown in Fig. 27.1.

This organizational structure and leadership practices served to create a foundation for intra- and inter-department collaboration. Regular meetings facilitate interaction through group dialogue to a significant extent, peer review and individual counseling, if required and connected through departmental coordinators for implementation. Through local study groups, members actively get engaged in peer discussion

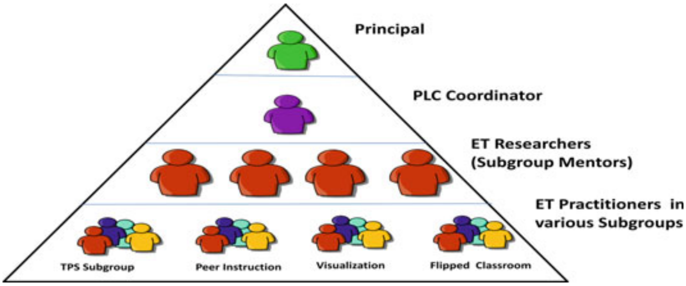


Fig. 27.1 PLC organizational structure

and review of the study resources created. Rubrics are developed for assessment employing collective wisdom of PLC.

This PLC served as the platform for Coordinated Faculty Professional Development Activity (CFPDA) model [2] which found to be effective in making MOOC successful in ET.

PLC focuses on disruptive innovations in T-L through:

1. Creation of e-contents and Learner-Centric MOOCs.
2. Continuous use of different instructional strategies and ICT tools during course delivery.
3. Use of learning management system (LMS): Moodle.
4. Continuous improvisation of the system through students' feedback on T-L process and assessment of students' learning.

27.2.1 Creation of E-contents and Learner-Centric MOOCs

For generating quality course content, we have established recording studio titled '**E-Learning Centre**' at our institute. The e-learning center has facility of lecture capturing and live lecture streaming. Faculty members use various ways of content presentation in video creation such as PowerPoint presentation, demonstrations of various tools/simulators/coding in various programming languages, animations and simulations. Faculty members also use document camera for demonstrating derivations, problem solving similar to conventional teaching methodology. Through continuous monitoring, peer review, experimentation and counseling, the results have been achieved to empower teachers to create course content in the form of videos for the benefit of learners. Standards and guidelines are set for video creation by PLC for uniformity and ensuring minimum quality standards. Some of the guidelines are: every video shall contain Learning outcome(s) aligned to Bloom's Taxonomy, reflection spots to trigger students to think and reflect, citations, plagiarism check and common creative license and follow presentation format.

Every semester, faculty members keep on adding videos in a phased manner. On completion of MOOC course ET702x: Designing of Learner-Centric MOOCs [3] by IIT Bombay, every faculty member is creating Learner-Centric MOOC (LCM) partially using these created videos on Institute Moodle server since 2018. In addition to watching the videos, students also carry out activities based on the content of the videos and accordingly they are assessed. Institute has set a well-defined process for video creation and LCM implementation as mentioned below.

27.2.1.1 Planning

- Design of course contents by subject teacher.
- Identification of topics/units for creating videos (from single unit/ muddiest points/problem solving, etc.).
- Content verification and authentication by departmental experts based on rubrics (learning outcome, content quality, reflection spot, duration, citations, presentation, etc.) set by PLC.
- Schedule for creation of video and posting on YouTube after department screening and approval.
- Planning of LCM.
- Periodic review of activities and accordingly seeking improvement.

27.2.1.2 Implementation

Implementation of LCM includes design of course, schedule for release of course and feedback. These LCMs are deployed on Institute Moodle spread over the semester.

- Design of LCM components: Learning dialogues (LeD)/videos, learning by doing (LbD) activity and assignment.
- Schedule for release of LCM.
- Collection of feedback from students on completion of LCM.
- Review on completion of activities and further planning.
- Students' learning assessment through direct and indirect tools.
- Reforms/addition/alteration **in LCM** implementation for improvement.

Thus, 360° feedback by stakeholders assists in overall improvement of the PLC activities. We have created more than @3000 videos since 2017–18. We have been publishing these videos on 'WIT Solapur—Professional Learning Community' YouTube Channel [4]. Due to E-Learning Centre, the quality of e-content generation has improved to a remarkable level enabling disruptive innovation in higher education, while enabling students to learn anywhere, anytime and many times at their pace and convenience. Students' learning has improved considerably which is also obvious from the utilization of e-contents by the stakeholders which is demonstrated through 17,452 subscribers and 2,496,913 views.

During COVID-19 lockdown period also, e-content generation, LCM creation and implementation and flipped classroom are effectively implemented. @100 videos are created during this lockdown. Thus, in the pandemic situation of COVID-19, learning happened uninterruptedly due to these videos.

As an outreach program, we have also trained faculty members from arts, science, commerce, management institutes on how to create videos and courses on Moodle. @100 faculty members are creating videos and facilitating not only students from their institute but also other learners.

27.2.2 Continuous Use of Different Instructional Strategies and ICT Tools During Course Delivery

Nearly, all faculty members took formal training in ET through various FDPs/MOOCs. Now, they are employing instructional strategies and ICT tools during content delivery. Thus, the conventional T-L process is strengthened by blending following ET practices for active learning.

Planning of Instructional Strategies/Tools/Innovations for every course in addition to LCM: At the commencement of semester, every faculty member submits course plan along with ET activity planning. Activity planning template includes name of topics, name of instructional strategies and/or ICT tools to be employed, schedule for activity, etc. Some of instructional strategies and ICT tools used are:

Instructional Strategies: Gamification, project-based learning, problem-based learning, role play, jigsaws, visualization (animation, simulation, live coding, etc.), flipped classroom, think-pair-share, team-pair-individual-share, team-pair-solo, etc., pair programming, peer instruction, teaching by example, collaborative competitive learning, competitive learning, collaborative learning.

ICT Tools: Virtual Lab, Virtual Programming Lab (VPL), BodhiTree, Socrative, Kahoot, Bubblino, ModelSim, Selenium, LogicSim, Weka, Parsing Emulator, JFLAP, Visualization Tutor, RapidMiner, Pentaho, JasperReports, Tableau, Cassendra, Wireshark, NS2, Simulation Tools (8085, Keil, MPLab, Proteus, VLSI Design, etc.)

On an average, 15 instructional strategies and 10 ICT tools are employed in T-L process in an academic year leading to enhancement in student's learning and making T-L process joyous experience.

27.2.3 Use of Learning Management System (LMS): Moodle

Two Moodle servers are installed at our institute, which are accessible remotely. All students (@2400) and 160 faculty members are enrolled on these servers. This platform is effectively used for publishing of courses and their assessments, feedback on T-L process, placement, employability enhancement activities, etc.

Virtual Programming Lab (VPL) on Moodle is effectively used for all programming languages benefitting both students and faculty. Students can run programs interactively. Also, faculty members evaluate programming assignments by setting various test cases. VPL is also used for programming contests effectively.

Apart from Institute Moodle server, Google Classroom, Gnomio and Google Forms are used for communication, assessment and deployment for material.

27.3 Assessment and Feedback

27.3.1 Assessment

Formative and summative assessments are carried out to assess students' engagement and learning using various proven forms such as MCQ tests, crossword puzzle, quizzes, survey questionnaire and rubrics. In addition to conventional ways of assessment, different ICT tools are employed for effectiveness and immediate feedback on content delivery. Tools such as Socrative and Kahoot are used for this purpose. As Moodle was accessible from outside campus, assessment of students' learning through internal semester examination (ISE), programming tests, etc., were conducted on Moodle smoothly during this lockdown.

27.3.2 Feedback

It is needless to emphasize importance of feedback for continuously improving the quality of T-L process. Objectives of the feedback are as follows:

- To provide students with an opportunity to put their views on the quality of their learning experiences, as required in preparation for and as part of review processes.
- To assess the success of academics in relation to the expectations of students.

Student feedback is taken twice in every semester to assess the quality of T-L, and corrective measures are taken by the faculty based on the feedback. A first feedback of each course is taken in the first half of the semester on the basis of which immediate action, if any, is implemented for betterment. Second feedback is taken at the end of course to assess content delivery as well as the effective implementation of the corrective measures taken based on first feedback.

Initially, to see the effect of instructional strategies/tools on students' learning and engagement in classroom, various instruments like pretests, posttests, perception surveys, interviews, etc., are used. Now, all instructional strategies are already proven; hence, there is no need to collect comprehensive feedback on the use of instructional strategies/tools separately. Therefore, this has been included in regular feedback.

Regular Feedback on T-L Process. It comprises four parts: academic, general, infrastructure and newly added online teaching. This includes course delivery, support and motivation to students, use of ICT, fulfillment of expectations, environment, etc.

Feedback on Videos and LCM Activity. To know effectiveness of videos and LCM activity, feedback is collected and corrective actions are taken, if any at the end of every unit. This feedback includes parameters such as quality of video, alignment of LbD and assignments to video contents, difficulty level of activities and recommendation of topics for video creation in the future and its usefulness in learning. The suggestions are incorporated to the extent possible while creating new videos. Also, rerecording is done, if required.

27.4 Journey of PLC

Initially in 2015, majority of faculty members in our institute were novice and did not receive any formal training in ET. Now, almost all faculty received training in ET and practicing during content delivery. Progressive journey of PLC is shown in Fig. 27.2.

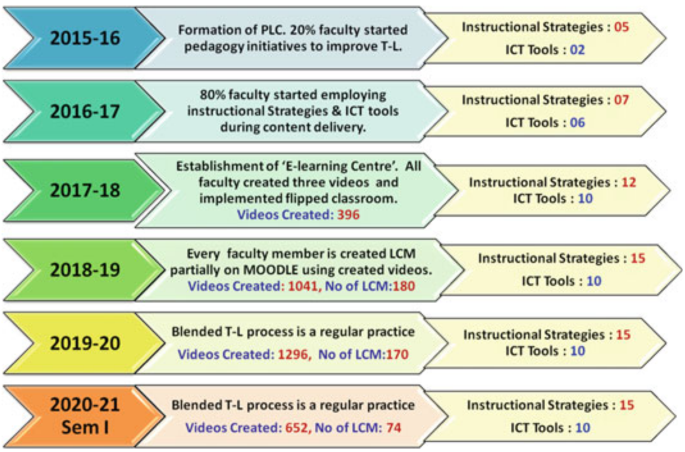


Fig. 27.2 Progressive journey of PLC

27.5 Awards and Recognition

27.5.1 Awards

- One faculty member is recipient of IUCEE outstanding Engineering Educator Award among 5 from all over India, and 4 faculty members are recipient of IUCEE Faculty Fellow Award all over India for year 2016 and 2017.
- 29 faculty members of our institute are among top 253 nationwide recipients of SAP Fellowship Award for Top Performers in a four-week FDP on 'Use of ICT in Education for Online and Blended Learning' conducted by IIT Bombay and received cash prize of Rs. 5000/– each.
- Winners of Inspire—Infosys Campus—connect Award for Content Guru and Distinguished Facilitator Tracks during three years.
- Appreciation and monitory benefit to faculty members on creation of quality e-contents. Screening is done on various parameters such as content delivery, quality of contents, alignment of contents, presentation, and audio and video quality.

27.5.2 Recognition

- Activities were highly appreciated by gathering of the Fourth International Conference on Learning and Teaching in Computing and Engineering (LaTiCE 2016) at IIT Bombay, and organizers have disseminated this success story to stakeholders for implementations as a role model.
- As on date (08-01-2020), subscribers of this channel are 17,542 and views are 2,496,913 for our WIT Solapur—Professional Learning Community YouTube Channel.
- 23 faculty members have worked as Associate Faculty of IIT Bombay for FDP on Foundation Program in ICT for Education and Pedagogy for Online and Blended T-L Process and received Rs. 12,000/– each.
- Substantial ET practices are emerging as some of the best practices after due iterations.
- Total 52 number of research publications on ET.

27.5.3 Grants Received

- Received grant of Rs. 2,505,000/– for establishing e-learning center from AICTE, New Delhi.
- Faculty members are recipient of various scholarships for presenting their ET research at international ET conferences.

27.6 Conclusion

PLC provides a platform for faculty members for sharing and exchanging views/suggestions for enhancing ET Research and T-L process. Leveraging technology for education in traditional classroom will not only scale up and improve the T-L process but also help in effective course content generation. Students' learning has improved considerably which is also obvious from the utilization of e-contents by the stakeholders. Thus, PLC is benefitting both students and faculty members to a great extent heralding the dawn of disruptive innovation in T-L process.

Because of all these efforts of blended T-L, during COVID-19 pandemic, we could smoothly shift overnight from physical classroom mode of teaching to an online mode of teaching.

Acknowledgements We would like to thank the Management, Principal and Heads of all departments of Walchand Institute of Technology, Solapur, Maharashtra, for helping in forming PLC. Also, we are thankful to all departmental ET Coordinators and Faculty Members for being part of PLC.

We would like to thank AICTE for sanctioning grant of Rs. 2,505,000/– for 'E-Learning Centre' Project for 3 years.

References

1. Halkude, S.A., et al.: Sustaining faculty professional development: evidences from a localized professional learning community. In: 2016 International Conference on Learning and Teaching in Computing and Engineering (LaTICE), pp. 89–93. Mumbai (2016). https://doi.org/10.1109/LaTiCE.2016.11
2. Mistry, R., et al.: Coordinated faculty professional development activity: key for high retention in MOOCs. In: Proceedings of the 24th International Conference on Computers in Education. Asia-Pacific Society for Computers in Education, India
3. https://www.iitbombayx.in/courses/designing-learner-centric-mooc
4. https://www.youtube.com/channel/UCb9okJF6NGPDUGgAQxu3TcA

Chapter 28
Video Shot Retrieval Using Multi-feature Approach

H. M. Nandini, H. K. Chethan, and B. S. Rashmi

Abstract The evolution of Internet and mobile technologies has raised the growth of video data that spur demand for efficient browsing and retrieval technologies. Content-Based Video Retrieval (CBVR) is evolved progressively to retrieve desired videos proficiently from large repositories based on the video content. In this aspect, an efficient CBVR method is presented using multiple features from video shot keyframes. The approach constructs histogram for Sobel magnitude of *V*-channel from HSV image and also local binary cumulative sum variance pattern for grayscale image. Histograms, thus, constructed are concatenated to build multiple feature vector database. Further, shot matching process is established by applying Euclidean distance between shot keyframe and query keyframe features. Experimentation was carried out on UCF YouTube action benchmark dataset to analyze the efficiency of the proposed algorithm. Video shot retrieved results depict significant improvement of the presented CBVR approach in comparison with baseline algorithm in terms of evaluation metric.

28.1 Introduction

In recent years, advancement in technologies and usage of Internet have led to extensive growth of video data on storage platforms. Consequently, such huge video repositories must be accompanied by efficient video retrieval and analysis technologies. Traditional search schemes are based on textual information of videos and show limited retrieval capability when they lack to understand the video content [1]. Thus, CBVR approaches have attained vital consideration in research field which exploits video content for the purpose of retrieval. CBVR systems have an extensive scope of implementation like digital museums, news event exploration [2], computer-aided retinal surgery [3], and CCTV surveillance systems [4].

H. M. Nandini (✉) · H. K. Chethan
Maharaja Research Foundation, Mysuru, India

H. M. Nandini · B. S. Rashmi
Karnataka State Open University, Mysuru, India

T. Senjyu et al. (eds.), *IOT with Smart Systems*, Smart Innovation,
Systems and Technologies 251, https://doi.org/10.1007/978-981-16-3945-6_29

The content in videos is broadly categorized into visual, textual, and audio contents. Out of these, visual content provides abundant information, and it is most frequently chosen for video retrieval [5]. There are numerous ways to extract visual content/feature of images such as color, texture, shape, and edges. The image/frame content is efficiently represented using edge feature and is utilized in various research works [6, 7]. Contents of image can be obtained by extracting features either globally or locally [8]. Global feature represents the visual information of the whole image in a single vector and is sensitive to noise, illumination, variation, etc. Their drawbacks are fixed by the use of local features that describe visual content in patches and encode the local information to get the finest details of the image [9]. Hence, several existing approaches have amalgamated both local and global features to obtain better results in various domains like retrieval of images [10] and detection of objects [11]. Use of multiple features enhances the accuracy of CBVR in terms of performance evaluation metric compared to classical method that utilizes color information only [12]. Motivated from these advantages, authors have carried out the proposed CBVR approach.

Generally, features are extracted from keyframes, as it contains most of the salient information of a video shot. This concept decreases computational complexity and preserves memory as well [13]. Thus, in this proposed approach, keyframes are selected from shot-based method proposed by Nandini et al. [7] to extract features for retrieval purpose. The proposed retrieval method extracts multiple features by incorporating both global and local features of shot keyframes. Two different approaches are explored for feature extraction: In the first method, initially, RGB images are transformed into HSV color images. Further, Sobel gradient magnitude for V-channel of HSV color images is extracted to construct the histograms. In the second method, RGB images are transformed to gray images, and Local Binary Cumulative Sum Variance (LBCSV) patterns are extracted to construct the histogram. These two histogram features are concatenated to construct the feature vector database. Subsequently, multi-feature vectors are used to match the query keyframe and shot keyframes by measuring Euclidean distance. The efficacy of the presented algorithm is verified using evaluation metrics by conducting experiments on UCF YouTube action benchmark datasets.

The article is organized as follows: Sect. 28.2 reports overview of some existing works. Section 28.3 provides proposed methodology covering description of feature extraction and similarity measurement. Section 28.4 presents experimental results and discussion, followed by conclusion of the paper in Sect. 28.5.

28.2 Related Works

An overview of existing works on CBVR approaches by various researchers is presented in this section. The importance and popularity of CBVR systems have led to several survey papers [14, 15]. An efficient CBVR approach has been presented [16] by extracting SURF features and adding color information using C-SURF of every keyframes. Distance measure is applied to match the features between stored video data and test video to retrieve the most similar videos based on the rank. Rashmi et al. [17] proposed video retrieval approach by utilizing visual features of keyframes. Here, feature matrix is constructed by extracting color and edge features of all keyframes of the shots to retrieve similar videos. FALKON [18] is an approach for efficient CBVR system that uses deep features and distributed in-memory computing along with the power of big data technologies. Further, video query maps are introduced to enhance the efficiency of the retrieval system. An approach for effective video retrieval has been introduced [19] by utilizing spatiotemporal features. In this method, combination of motion feature vector and HSV color histogram is used to generate hybrid feature vector that plays a significant role in CBVR. Padmakala [20] has used multiple features of every frame and twelve distance measures that are optimized by gravitational search algorithm to extract relevant videos from the database based on the query. An efficient CBVR method is proposed [21] by extracting multiple features like motion, color, and texture features and employing similarity measure using Euclidean distance between query and video database. It can be observed from the above overview that utilization of multiple feature plays significant role for effective CBVR during the last few years.

28.3 Proposed Methodology

In this section, an efficient CBVR approach using multiple features of video shot keyframes has been presented. The fundamental stages of proposed CBVR method are feature extraction and similarity measurement. A schematic representation of the presented approach is illustrated in Fig. 28.1.

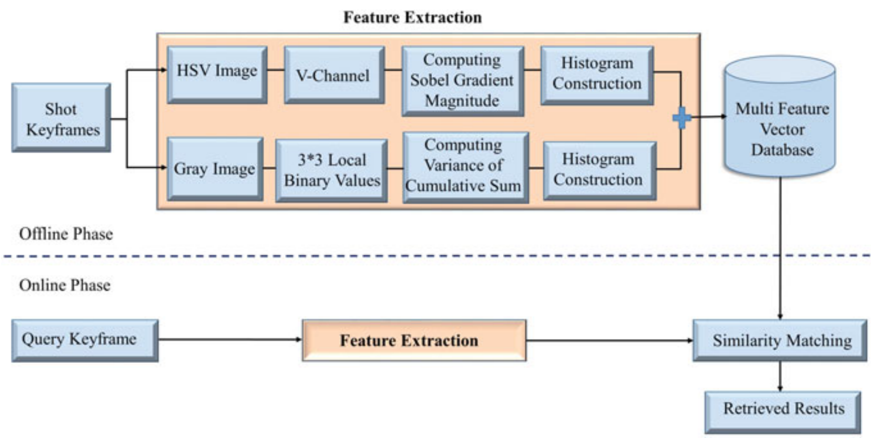


Fig. 28.1 Structure of the presented CBVR approach

28.3.1 Multi-feature Extraction

The visual features of images play a vital role in CBVR. The following subsections describe the process of multiple feature extraction from keyframes to achieve shot retrieval task.

Histogram Construction Using HSV Color Edge Approach

Edge detection helps to remove the irrelevant information and retains essential structural properties of the image [22]. Gray scale frame contains less information than color frame; thus, color edge detection is used to get more precise information [23]. There exists various types of color spaces in literature with respect to different applications. However, HSV color space has huge influence as it arranges any color image in the similar way that human eyes can perceive [8]. In the proposed method, V-channel of HSV color space is considered for edge detection, as it represents luminance of an image and stores the brightness intensity values. Sobel edge detector is employed on the V-channel of every keyframe to get Sobel magnitude as it performs better than other edge descriptors with reference to computational complexity and accuracy [24]. It also has an advantage of smoothing effect to the random noises present in the image [25]. The Sobel detector includes a two 3 * 3 kernels which are convolved with V-channel of the frame to measure derivatives approximations. Kernels produce two different measurements in vertical and horizontal directions for gradients, specifically G_x and G_y and are computed as:

$$G_x = \begin{bmatrix} -1 & 0 & +1 \\ -2 & 0 & +2 \\ -1 & 0 & +1 \end{bmatrix} \quad G_y = \begin{bmatrix} +1 & 2 & +1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} \tag{28.1}$$

The gradient magnitude provides rate of change for gradient in frame intensity at every pixel position as illustrated in Fig. 28.2 and is computed using following equation:

$$\mathrm{GM} = \sqrt{G_x^2 + G_y^2} \tag{28.2}$$

Histogram for every shot keyframe is built using gradient magnitude values.

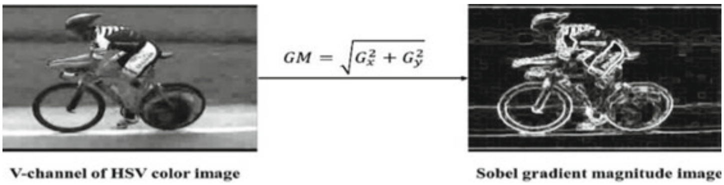


Fig. 28.2 Illustration of Sobel gradient magnitude for frame # 21_7_008 of biking video sequence

Histogram Construction Using LBCSV Pattern
The initial step of this feature extraction process starts by transforming shot keyframes from RGB to gray scale. Further, algorithm as applied in [26] is used to compute local binary values of frame by thresholding the 3 * 3 block of every pixel with the center pixel, and it is formulated as follows:

$$\mathrm{LBV}_{(P\ R)} = S(G_C - G_P) \tag{28.3}$$

$$s(a) = \begin{Bmatrix} 1 & a \geq 0 \\ 0 & a < 0 \end{Bmatrix}$$

where P and R represent number of pixels and radius, respectively. G_P is neighboring pixel values, and G_C represents center pixel value of 3 * 3 neighborhood. In this work, instead of computing LBP code on 3 * 3 local binary values, cumulative sum is evaluated using following equation:

$$\mathrm{CS}(n) = \sum_{p=1}^{n} (\mathrm{LBV}) \tag{28.4}$$

where LBV is binary values of each block and CS(n) represents cumulative sum. Furthermore, block variance of CS(n) is formulated as follows:

$$V = \frac{\sum_{k=1}^{n} (\mathrm{CS} - \mu)^2}{n} \tag{28.5}$$

where $\mu = \frac{1}{n} \sum_{k=1}^{n} \mathrm{CS}(i)$ and $n = 9$. This process is repeated at every pixel position in an overlapping fashion for the entire image. Finally, histogram of Local Binary Cumulative Sum Variance (LBCSV) values for each keyframe is constructed and is shown in Fig. 28.3.

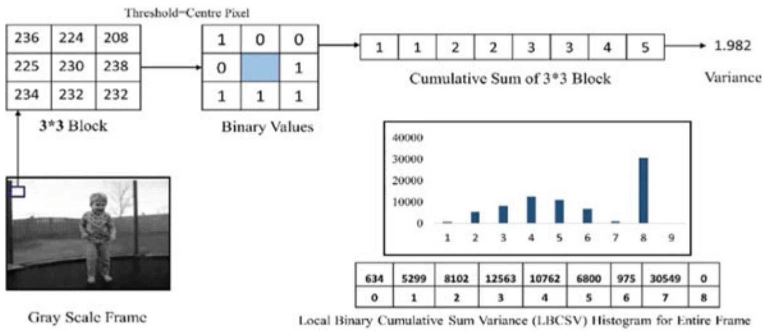


Fig. 28.3 Construction of LBCSV histogram for frame # 7_1_75 of trampoline jumping video

Histograms obtained from the above two feature representation methods are concatenated to build a multi-feature vector and are stored in database.

28.3.2 Similarity Measurement

Video shot retrieval result mainly confides in similarity measures. The similarity can be attained by matching objects, texts, features, etc., and combinations of them. However, the most direct and convenient method for measuring similarity is by matching features [15]. Extensive analysis of video retrieval similarity measures has revealed that Euclidean and Manhattan distance measures are the best in terms of their retrieval ability [27]. In the proposed approach, similarity between shot keyframe features in database and query keyframe feature is measured by Euclidean distance using following equation:

$$\mathrm{ED} = \sqrt{\sum_{i=1}^{n} \left(V_d - V_q\right)^2} \tag{28.6}$$

where, ED is the distance measure and V_d and V_q represent the multi-feature vectors. If the Euclidean distance is smaller, then database shot keyframe which corresponds to video is more similar to query keyframe and is higher in rank [28]. This concept is applied, and ranking of retrieved results is considered.

28.4 Results and Discussion

Dataset: The performance evaluation of the proposed retrieval approach is carried out considering video sequences from UCF YouTube action dataset and is described in Table 28.1. It contains 11 action categories and 1597 videos in total with 320 × 240 frame dimension.

28.4.1 Performance Evaluation

In order to check the efficiency of the presented algorithm, the query shot keyframe is randomly selected from the standard dataset. The multi-feature vector of query keyframe is matched with features in repository using similarity measurement as discussed in Sect. 28.2. Based on matching, the most similar video shots are retrieved and are ranked for analysis purpose. Performance evaluation of the presented method is determined by average precision score and compared with the existing baseline approach [29] considering top ten ranking results and is recorded in Table 28.2. The performance is graphically represented in Fig. 28.4. The result analysis based on the number of keyframes between query and repository is considered as in [29].

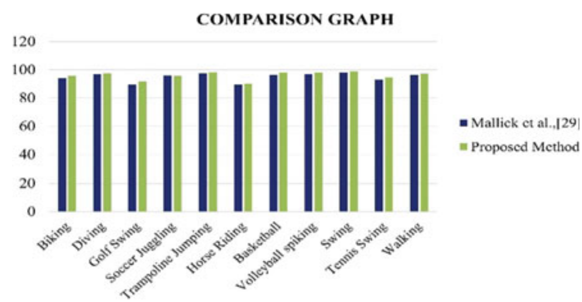
Table 28.1 Description of UCF YouTube action video dataset

Categories of video	No. of videos
"Biking"	137
"Diving"	145
"Golf_ swing"	156
"Soccer_ juggling"	142
"Trampoline_ jumping"	198
"Horse _riding"	156
"Basketball"	138
"Volleyball_ spiking"	167
"Swing"	119
"Tennis _swing"	116
"”Walking"	123

Table 28.2 Performance analysis using average precision score

Categories of video	Mallick et al. [29]	Proposed method
"Biking"	94.09	95.74
"Diving"	96.92	97.48
"Golf_ swing"	89.52	91.67
"Soccer_ juggling"	95.97	95.55
"Trampoline_ jumping"	97.74	98.34
"Horse _riding"	89.31	90.13
"Basketball"	96.43	98.11
"Volleyball_ spiking"	97.12	98.09
"Swing"	98.21	98.95
"Tennis _swing"	93.21	94.51
"”Walking"	96.32	97.23
Average	94.99	95.98

Fig. 28.4 Comparison graph of average precision score



In this approach, computation of multi-feature vector is much faster with feature dimension of 1 * 19 for each shot keyframe when compared to state-of-the-art algorithm [29] with 1 * 64 dimension. The results clearly exhibit that the proposed approach outperforms [29] with 95.98% of average precision score. The proposed approach is computationally expensive when compared with traditional system that uses single feature only.

28.5 Conclusion

In the proposed research work, video shots are retrieved utilizing visual content of shot keyframes. The method extracts multiple features by exploiting Sobel gradient magnitude for V-channel of HSV color images and extracting Local Binary Cumulative Sum Variance (LBCSV) patterns from gray images to construct the histograms. Similarity matching is employed by applying Euclidean distance between query keyframe and shot keyframe feature vectors present in the database to retrieve the similar video shots of the highest ranking. Experimental results show that the presented algorithm provides better result in comparison with baseline algorithm with 95.98% average precision score using UCF YouTube action benchmark dataset. The main drawback of this approach lies in multi-feature extraction which makes algorithm computationally complex. In the future work, the performance analysis of the presented algorithm will be investigated considering motion feature of the video sequence.

References

1. Rajendran, P.: An Enhanced Content-Based Video Retrieval System Based on Query Clip, vol. 1, pp. 236–253 (2009)
2. Peng, Y., Ngo, C.W.: Hot event detection and summarization by graph modeling and matching. Lect. Notes Comput. Sci. **3568**, 257–266 (2005)
3. Quellec, G., Lamard, M., Cazuguel, G., Droueche, Z., Roux, C., Cochener, B.: Real-time retrieval of similar videos with application to computer-aided retinal surgery. In: Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp. 4465–4468 (2011)
4. Yang, Y., Lovell, B.C., Dadgostar, F.: Content-based video retrieval (CBVR) system for CCTV surveillance videos. In: DICTA 2009—Digital Image Computing: Techniques and Applications, pp. 183–187 (2009)
5. Yasin, D., Sohail, A., Siddiqi, I.: Semantic video retrieval using deep learning techniques. In: Proceedings 2020 17th International Bhurban Conference on Applied Science and Applied Technologies (IBCAST 2020), pp. 338–343 (2020)
6. Patel, B.V., Deorankar, A.V., Meshram, B.B.: Content based video retrieval using entropy, edge detection, black and white color features. In: ICCET 2010, vol. 6, pp. 272–276 (2010)
7. Nandini, H.M., Chethan, H.K., Rashmi, B.S.: Shot based keyframe extraction using edge-LBP approach. J. King Saud Univ. Comput. Inf. Sci. (2020)

8. Thounaojam, D.M., Khelchandra, T., Singh, K.M., Roy, S.: A genetic algorithm and fuzzy logic approach for video shot boundary detection. Comput. Intell. Neurosci. **2016** (2016)
9. Kabbai, L., Abdellaoui, M., Douik, A.: Image classification by combining local and global features. Vis. Comput. **35**, 679–693 (2018)
10. Chaudhary, M.D., Upadhyay, A.B.: Fusion of local and global features using stationary wavelet transform for efficient content based image retrieval. In: 2014 IEEE Students' CEECS 2014 (2014)
11. Muralidharan, R., Chandrasekar, C.: Combining local and global feature for object recognition using SVM-KNN. Int. Conf. Pattern Recogn. Inf. Med. Eng. PRIME **2012**, 1–7 (2012)
12. Zaini, H.G., Frag, T.: Multi feature content based video retrieval using high level semantic concepts. Int. J. Comput. Sci. Issues **2**, 254–260 (2014)
13. Prathiba, T., Kumari, R.S.S.P.: Content based video retrieval system based on multimodal feature grouping by KFCM clustering algorithm to promote human–computer interaction. J. Ambient Intell. Humaniz. Comput. (2020).
14. Spolaôr, N., Diana, H., Shoity, W., Takaki, R., Augusto, L., Saddy, C., Coy, R., Chung, F.: Engineering Applications of Artificial Intelligence A Systematic Review on Content-Based Video Retrieval, vol. 90 (2020)
15. Hu, W., Xie, N., Li, L., Zeng, X., Maybank, S.: A survey on visual content-based video indexing and retrieval. IEEE Trans. Syst. Man Cybern. Part C Appl. Rev. **41**, 797–819 (2011)
16. Patel, D.H.: Content based video retrieval using enhance feature extraction. Int. J. Comput. Appl. **119**, 4–8 (2015)
17. Rashmi, M., Fernandes, R.: Video Retrieval Using Fusion of Visual Features and Latent Semantic Indexing, pp. 4272–4277 (2014)
18. Khan, M.N., Alam, A., Lee, Y.K.: FALKON: large-scale content-based video retrieval utilizing deep-features and distributed in-memory computing. In: Proceedings—2020 IEEE International Conference on Big Data and Smart Computing (BigComp), pp. 36–43 (2020)
19. Naveen Kumar, G.S., Reddy, V.S.K.: A scheme for shot detection and video retreival using spatio temporal features. Int. J. Recent Technol. Eng. **8**, 605–608 (2019)
20. Padmakala, S., Mala, G.S.A., Anandkumar, K.M.: Using diverse set of features to design a content-based video retrieval system optimised by gravitational search algorithm. Int. J. Bus. Intell. Data Min. **17**, 444–470 (2020)
21. Asha, D., Latha, Y.M., Reddy, V.S.K.: Content based video retrieval system using multiple features. Int. J. Pure Appl. Math. **118**:287–294 (2018)
22. Dhagdi, S.T., Deshmukh, P.R.: Keyframe based video summarization using automatic threshold & edge matching rate. Int. J. Sci. Res. Publ. **2**, 1–12 (2012)
23. Koschan, A.: A comparative study on color edge detection. In: Proceedings 2nd Asian Conference on Computer Vision, vol. 3, pp. 574–578 (1995)
24. Abdesselam, A.: Improving local binary patterns techniques by using edge information. Lect. Notes Softw. Eng., 360–363 (2013)
25. Rami, H., Hamri, M.: Color edge detector with Sobel-PCA. Int. J. Comput. Appl. **75**, 12–16 (2013)
26. Ojala, T., Pietikäinen, M., Harwood, D.: A comparative study of texture measures with classification based on feature distributions. Pattern. Recognit. **29**, 51–59 (1996)
27. Bekhet, S., Ahmed, A.: Evaluation of similarity measures for video retrieval. Multimed. Tools Appl. **79**, 6265–6278 (2020)
28. Patel, B.V.: Content based video retrieval systems. Int. J. UbiComp. **3**, 13–30 (2012)
29. Mallick, A.K., Mukhopadhyay, S.: Video retrieval using salient foreground region of motion vector based extracted keyframes and spatial pyramid matching. Multimed. Tools Appl. **79**, 27995–28022 (2020)

Chapter 29
Innovative Approach to Onboard Media Forensic of a Drone

Pavni Parghi, Rishi Dhamija, and Animesh Kumar Agrawal

Abstract Like the origin of aviation by enthusiasts, drone flying also started as a hobby. But unlike aviation which took a long time to mature, drone flying has proliferated very quickly across the globe. Large-scale use of drones for authorized military application and illegal use by non-state actors has blurred the distinction between them, thereby leading to an increased number of violations at vital areas/vital points. Expanded use of recreational UAVs/drones has now turned into a big security concern worldwide. Drones are being intentionally used for committing crimes and, therefore, in-depth knowledge of the ways to collect digital evidence from captured drones is a much-needed facet. In order to collect the forensic evidence, it is important to carry out forensics of onboard storage media present in the drone. This storage media which contains different artifacts like pictures, videos, log files, etc., is analyzed to gather the digital evidence to resolve a crime. This research brings out the information gathered from drone storage device.

29.1 Introduction

Today, technology is constantly evolving in every domain. The newest domain, which is highly influenced by technology, is unmanned aerial vehicles often called as drones. Military drones have been in use since many years. The technology has now matured to a level where it has reduced the cost of technology and made its easily accessible for civilian use. These two factors have resulted in posing a new problem for the Air Defense forces as well as the Law Enforcement Agencies acting against the state and non-state actors, respectively. In order to address the problem posed by illegal drones, a regulatory framework has been established which mandates registration of drone with the government. Since then, The Ministry of Civil Aviation has registered 19,553 drones comprising of 1832 nano, 13,735 micro, 2808 small, 140 medium, and 1038 large drones [1]. The proliferation of drones in civil use like hobby flying has also led to drones being used for smuggling, dropping of unauthorized weapons,

P. Parghi (✉) · R. Dhamija · A. K. Agrawal
National Forensic Sciences University, Gandhinagar, Gujarat, India

T. Senjyu et al. (eds.), *IOT with Smart Systems*, Smart Innovation, Systems and Technologies 251, https://doi.org/10.1007/978-981-16-3945-6_30

drug delivery, etc. During peacetime, the problems faced by the LEA is the most challenging and complicated because most acts by non-state actors would be from within the national boundary and the infrastructure required to fly the drone is not complex. To counter these challenges, it is very much essential that we develop a capability to first detect the drones and then capture it or destroy it. The forensic expertise would be on hardware, software as well as on all other onboard components including the air frame of the drone. Meaningful analysis should be able to answer some generic questions like how, why, when, where, etc., about the drone. Once these two capabilities of detection and forensic are developed, we would not only be able to find the source of illegal activities for punitive action, but a strong punitive action would serve as deterrence for such illegal operators. To develop effective forensic capability, an in-depth study on various facets of drone and drone flying are a must. On preliminary study, it emerged that one of the most important components which carries wealth of information is the storage media associated with the drone. Effective analysis of storage media would answer most of the questions lead to the source of any drone. This paper, therefore, focuses on the methodology of forensics on any storage devices onboard the drone.

This paper focuses on Yuneec_Typhoon_Q500_4K drone. [2] This drone has max flight time of 25 min, operates in 2.4 GHz frequency band and uses Android Operating System. Yuneec is highest stake holder in drone market after DJI making it less explored forensically as compared to commercially available DJI drones which in turns makes them suitable for ill minded people to carry out their malicious operations. The drone chosen comes with very high-end camera capabilities such as 4k recording making it ideal for spying, illegal photography, etc.

The paper is divided into four sections. Section 29.1 is introduction followed by literature review in Sect. 29.2. Section 29.3 brings out the proposed methodology to conduct forensics and experimental setup. In Sect. 29.4, results of the experiments are brought out followed by conclusion and future work.

29.2 Literature Review

In [3], the challenges associated with drone forensics were explored, and existing forensic guidelines were evaluated. As of now, there are no existing guidelines for drones which make the task of drone forensics more challenging. The aim of this paper was to evaluate existing forensic guidelines, in terms of their effectiveness for drone forensic investigations and to propose new guidelines that can be used as a guide for drone forensics investigation. Open-source tools like datcon, autopsy were used in the forensics investigation. Research in [4] was carried out using open-source tools, and some basic scripts have been developed to aid forensic analysis of two drones and for identification of operators, and extraction of data from drones. The researchers of [5] observed that drone forensics are an uncovered area, so they have presented an introductory discussion of unmanned aerial vehicle analysis and provided digital forensic investigation report of a test Parrot Bebop unmanned aerial vehicle. Investigation of drones through various forensics procedures for drone acquisition, evidence collection, forensic investigation, and reporting is given by [6]. The objective of this research [7] is to identify the drone forensics SOP to get sources of potential artifacts. In [8] authors focus on forensics of the drone flight data logs. In this paper, a methodology for collecting the necessary information, analyzing it, and constructing the corresponding timeline is proposed. The research is focused on the investigation of two specific logs, the dataflash, and telemetry logs for investigation. Motive behind this research was to gather information and conduct an analysis of digital evidence contained in drones along with its controller by using static forensic and lives forensic. As drones are being used extensively, their expanded use has become a threat to security. So, authors in [9] have focused on the estimation of the practicality of the use of UAV and AI in crime scene investigation, feasibility of UAV in forensic science and its logistical implementation. Researchers in [10] have presented the extraction and identification of artifacts from the recorded flight data as well as the associated mobile devices using open-source tools and some basic scripts developed for analysis of two popular drone systems-the DJI Phantom 3 Professional and Parrot AR. The researchers in [11] observed that there is no standard technical forensic process for drones so they have proposed a forensic process consisting of various phases for the analysis of forensic artifacts of drones using computer forensics. Researchers in [12] analyzed the essential log parameters of the autonomous drone and proposed drone forensics-related software architecture with preliminary results. A user-friendly graphical user interface is also provided that allows extracting and examining the onboard flight information. Keeping the current challenges faced in drone forensics, authors in [13] discuss the extraction and interpretation of important artifacts found in recorded flight logs on both the internal memory of the UAV and the controlling application for effective analysis of drones. The work in [14] aims to further explore drone forensics in terms of challenges, forensic investigation procedures, and experimental results through a forensic investigation study performed on a Parrot AR drone 2.0. Researchers of [15] propose a robust digital drone forensic application with log parameters of drones through a graphical user interface (GUI) developed using JavaFX 8.0 which allow users to extract and examine onboard flight information. It includes a file converter created for 3D flight trajectory visualization. The paper [16] presents a report of potential attacks against the Parrot Bebop 2 drone and the ability to collect evidence about the attacks on the drone. In [17], authors focus on drone forensics and present methodology of investigating physical and digital evidence to acquire all the possible data from drones. The challenge of extracting data from the mobile app from an unrooted phone can be overcome as brought out in [18].

29.3 Methodology

Presented methodology is based on the fact that the following forensic data can be obtained from the drone. These include (a) associated file systems and firmware, (b) digital camera image files (c) video files, (d) telemetry data, and (e) GPS coordinates (f) drone operator information, etc. This methodology enumerated in Fig. 29.1 contains following three steps:

Step 1: As a first step toward drone forensic information about various drones available and which drone is getting more used worldwide will be gathered. After gathering information about drone manufacturers, the location of data storage of each would be determined so that suitable methodologies can be applied to obtain data.

Step 2: In this step, forensic analysis of storage devices would be done in order to find various artifacts like drone's serial number, MAC address, attached payload, etc. Open-source tools would be used for the analysis maintaining data integrity at all times.

Step 3: A detailed report highlighting the evidences gathered would be part of the third step, so that they can be presented in court of law. Effort would be made to establish the ownership of the drone so that the correct crime perpetuator can be identified.

As this paper focuses on forensics of onboard storage device, the file system structure of media storage is illustrated in Fig. 29.2 where possible extraction of the media files can be achieved for through analysis. In addition, analyzing the artifacts

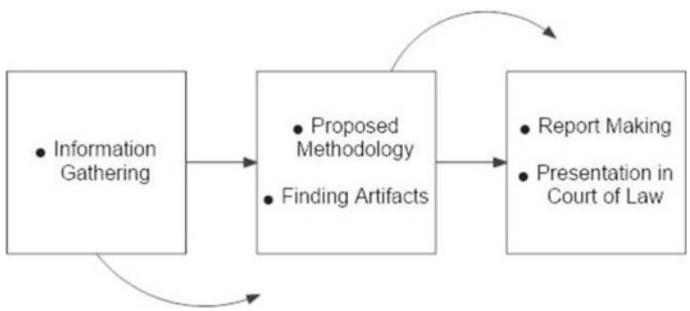


Fig. 29.1 Proposed data extraction methodology

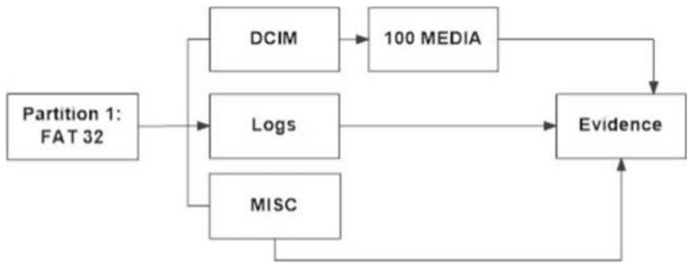


Fig. 29.2 File system of SD card used in Yunnec drone

found in media files is necessary because they contain relevant meta-data related to the flight activity.

For the forensic analysis, a Yunneec Typhoon H drone has been used because Yuneec is second highest stake holder in drone market after DJI making it less explored forensically as compared to commercially available DJI drones. COVID pandemic has resulted reduced availability of Yuneec Typhoon H drones in the market. Hence, file of Yuneec Typhoon H drone was downloaded from VTO LABS website [19] which has a huge drone database. Since the research work entails analysis of onboard storage media card, it is assumed that the internal SD Card was used, while flying the drone and data downloaded was recorded in the internal SD Card. The aim is to extract maximum evidence from the storage of the drone and analyze it so that it can be presented as useful forensic evidence.

The image of the drone is analyzed using the following open-source tools:

- Autopsy 4.17.0
- HxD 2.4.0.0
- FTK IMAGER 4.5.0.3.

29.4 Results and Discussion

The following information could be obtained from the image file of drone.

1. Controller Information

While investigating Yunnec drone, it was found out that mobile phone was used as a controller for remotely operating the drone. The file system for the same is shown in Fig. 29.3.

Figure 29.4 shows the language which was used to pass commands to the drone from the controller. For extracting the above information, FTK Imager and HxD editor were used.

2. Internal Flight Data

Figure 29.5 shows the email id which was used to log into the mobile phone used as the drone controller.

```
030 | 01 00 06 00 00 00 00 00-00 00 00 00 00 00 00 00 | ................
040 | 80 00 29 01 02 03 04 20-20 20 20 20 20 20 20 20 | ..)....
050 | 20 20 46 41 54 33 32 20-20 20 00 00 00 00 00 00 |   FAT32   ......
060 | 00 00 00 00 00 00 00 00-00 00 00 00 00 00 00 00 | ................
```

Fig. 29.3 File system of controller

```
Name: system/tts/lang_pico/en-US_lh0_sg.bin
SHA1-Digest: OmEWL+G/OLXlfcQXpILKeHbEpwQ=
```

Fig. 29.4 Language used in controller for communication

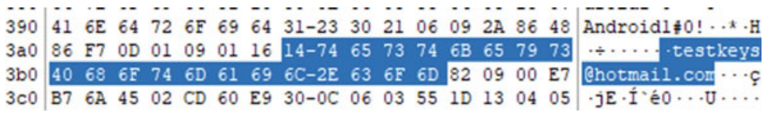


Fig. 29.5 User credentials

3. Electro Optical Sensor Data

While analyzing the drone image, it was found that the media files captured by the electro optical sensor contained critical data about the drone position at any given time in space. These included the latitude, longitude, altitude, accuracy, speed, and angle of flight. Same is brought out in Fig. 29.6. The probable location where the drone was used is California Mountain as brought out in Fig. 29.7.

Drone controller's serial numbers are unique for each drone controller that is manufactured. With the serial number, a drone controller can be linked to a given drone thereby establishing ownership of the drone (refer Fig. 29.8).

Table 29.1 shows the data found from the drone image by analyzing it with three different tools. After analyzing, it was found that out of these freely available tools maximum data was recovered from FTK Imager. HxD which is a hex editor is used to manually search information from the image hex which is a time consuming task.

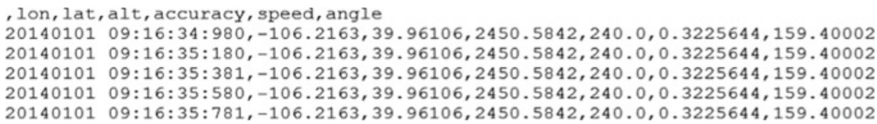


Fig. 29.6 Displays the latitude, longitude, speed, angle, etc., of the drone

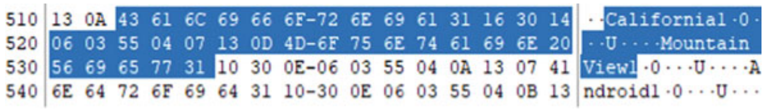


Fig. 29.7 Probable location of the drone

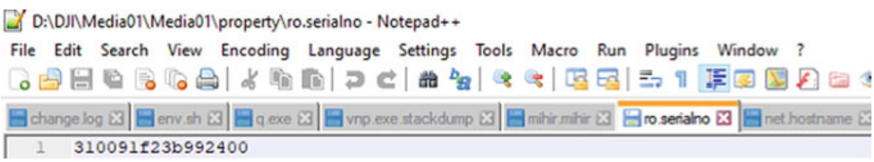


Fig. 29.8 Serial number of drone

Table 29.1 Comparative analysis of extracted artifacts

Artifact name	Autopsy	FTK imager	HxD
Application details	–	YES	YES
File system	–	YES	–
Language name	–	YES	YES
Coordinates	YES	YES	–
User credential	YES	YES	–
Location	YES	YES	YES
Camera details	YES	–	–
Telemetry	YES	–	YES

29.5 Conclusion and Future Work

The work brings out the methodology to carry out forensics of the onboard storage media of a drone. Three open-source tools have been used to compare the results of the recovered artifacts. These tools give us raw information like file system, telemetry information, time stamp, camera details, route flown, database details, etc. However, there are still some points which need to be addressed to complete the entire storage device forensics. Manual search techniques need to be adopted to gather more artifacts and then automate the process to decrypt the propriety format do a flight path reconstruction can form part of the future work.

References

1. U.S. Department of Transportation, Drone users details available at https://www.faa.gov/uas/resources/by_the_numbers/
2. Yunnec Typhon Q500 in depth details available at https://www.mydronelab.com/reviews/yuneec-typhoon-q500.html
3. Roder, A., Choo, K.R., Le-Khac, N.: Unmanned aerial vehicle forensic investigation process: DJI phantom 3 drone as a case study (2018). ArXiv, abs/1804.08649
4. Barton, T.E.A., Hannan Bin Azhar, M.A.: Forensic analysis of popular UAV systems. In: 2017 Seventh International Conference on Emerging Security Technologies (EST), pp. 91–96. Canterbury (2017)
5. Horsman, G.: Unmanned aerial vehicles: a preliminary analysis of forensic challenges. Digit. Investig. **16**, 1–11 (2016)
6. Nilay, M., Sanghvi, H.: Drone forensics: investigative guide for law enforcement agencies. Int. J. Electron. Secur. Digit. Forensics
7. Prastya, S., Riadi, I., Luthfi, A.: Forensic analysis of unmanned aerial vehicle to obtain GPS log data as digital evidence. Int. J. Comput. Sci. Inf. Secur. 280–285 (2017)
8. Mantas, E., Patsakis, C.: GRYPHON: drone forensics in Dataflash and telemetry logs. In: Attrapadung, N., Yagi, T. (eds) Advances in Information and Computer Security. IWSEC 2019. Lecture Notes in Computer Science, vol. 11689. Springer, Cham (2019)
9. Sharma, B.K., Chandra, G., Mishra, V.P.: Comparitive analysis and implication of UAV and AI in forensic investigations. In: 2019 Amity international conference on artificial intelligence (AICAI), pp. 824–827. Dubai, United Arab Emirates (2019)

10. Azhar, M.A.H.B., Barton, T., Islam, T.: Drone forensic analysis using open source tools. J. Digit. Forensics, Secur. Law **13** (2018). https://doi.org/10.15394/jdfsl.2018.1513
11. Salamh, F., Karabiyik, U., Rogers, M.: RPAS forensic validation analysis towards a technical investigation process: a case study of Yuneec Typhoon H. Sensors (2019)
12. Renduchintala, A., Albehadili, A., Javaid, A.: Drone Forensics: Digital Flight Log Examination Framework for Micro Drones, pp. 91–96 (2017)
13. Barton T.E.A., Azhar, M.A.H.B.: Open source forensics for a multi-platform drone system. In: Matoušek, P., Schmiedecker, M. (eds.) Digital Forensics and Cyber Crime. ICDF2C 2017. Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, vol. 216. Springer, Cham (2018)
14. Kamoun, F., Bouafif, H., Iqbal, F.: Towards a better understanding of drone forensics: a case study of parrot AR drone 2.0. Int. J. Digit. Crime Forensics, 1–23 (2019)
15. Renduchintala, A., Jahan, F., Khanna, R., Javaid, A.: A comprehensive micro unmanned aerial vehicle (UAV/drone) forensic framework. Digital Investigation (2019)
16. Iqbal, F., et al.: Drone forensics: examination and analysis. Int. J. Electron. Secur. Digit. Forensics 245–264 (2019)
17. Bouafif, H., Kamoun, F., Iqbal, F., Marrington, A. Drone forensics: challenges and new insights. In: 2018 9th IFIP International Conference on New Technologies, Mobility and Security (NTMS), pp. 1–6. Paris (2018)
18. Agrawal, A.K., Sharma, A., Khatri, P., Sinha, S.R.: Forensic of unrooted mobile device. Int. J. Electron. Secur. Digit. Forensics **12**(1), 118–137 (2019)
19. https://www.vtolabs.com/drone-forensics, Last accessed 31 Dec 2020

Chapter 30
Enhanced Security of Windows Executables for Intelligent Systems

S. Raja Prabhu and Animesh Kumar Agrawal

Abstract Information and Communication Technologies (ICTs) are changing our lives in unforeseen ways. The systems, however, intelligent it may be but still suffer from the flaws of misuse, abuse, unauthorized access, etc., that are common to the underlying executable. Hardening the executable files, in the context of Windows Operating System, is the focus of this research paper. The techniques to harden the Windows executables could be adopted from malicious software; but this time for a noble cause, to protect the intelligent systems. Malware reverse engineering is an exclusive domain that helps in understanding the obscure functionalities of the malicious code, obtain Indicators of Compromise (IoC) and signatures to lessen the ill effects of malware. Malware researchers and reverse engineers constantly strive to dissect malware to make it inept to create any large scale disruption. On the other hand, the malware authors are constantly on a pursuit to develop innovative techniques to prevent the reverse engineering. Numerous such techniques have, so far, been developed to protect the malware; ranging from simple anti-debugging to encoding to advanced code virtualization packers. In this research, code cave and anti-debugging techniques, usually employed by malware, are used to create password authentication mechanism that could be embedded inside any Windows PE file. It is expected that this method would prevent the unauthorized execution of critical Windows executable programs from malicious insiders.

S. R. Prabhu (✉) · A. K. Agrawal
National Forensic Sciences University, Gandhinagar, Gujarat, India

T. Senjyu et al. (eds.), *IOT with Smart Systems*, Smart Innovation,
Systems and Technologies 251, https://doi.org/10.1007/978-981-16-3945-6_31

30.1 Introduction

Windows is predominantly the most used operating system be it intelligent systems, Human Machine Interface (HMI), research labs, banks, or government/private organizations. The simple GUI and ease of use makes it user friendly. However, the Windows application programs that are installed in any Windows PC are susceptible to cyber-attacks. Hence, it is imperative that the proprietary software applications are hardened adequately so that they can be accessible only to authorized users. While it is easy to add an authentication mechanism to an application whose complete source code is available, the same becomes almost impossible if only the executable file is available.

This paper discusses how a password authentication mechanism could be added to the executable file of the target Windows application by applying the portable executable (PE) infection techniques [1] used by malicious authors. Also, various advanced techniques learnt from malware reverse engineering could be implemented to harden the authentication mechanism. While the effort involved in securing an application is phenomenal, the advantages are far too many to be ignored. Some of them are:

- It allows only the authorized user to access and use the application.
- The same PC could be used by other users without the anxiety of the secured application being misused.
- The proposed security feature can be easily implemented thereby obviating huge expenses in procuring multiple authentication software.
- The application of advanced techniques being employed by malware to attack a legitimate application is used to protect the same thereby making it more robust.

30.2 Literature Review

Malware is a term coined by merging two words—malicious and software, which is used to define a broad range of software that disrupt computer services, steal data, or compromise user safety. Malware is software designed for malicious purposes and to deliberately cause harm to its target [2]. Malware can take the form of an executable, script, code, or any other software [3]. Malware defines a range of malicious applications and includes computer virus, worms, Trojan horses, spyware, adware, rootkit, etc. [4].

Malware is known to employ sophisticated techniques to achieve its intended aim. There are dozens of anti-debugging techniques used by malware to find if it is being analyzed using any debugger [5–7]. It may use anti-VM techniques to detect the presence of virtual machines [7]. Inline hooking or inline patching is one technique where the API function itself is modified to redirect the API to the malicious code. Malware authors often use various obfuscation techniques viz., XOR, Base64, Caesar cipher, etc., to encode the information and to modify the malicious content to make detection and analysis difficult. Certain advanced functions include use of custom encryptions to make the reverse engineering harder.

This technique of altering the PE file by reverse engineering the executable is predominantly done by hackers and cyber criminals. This is, probably, assumed to be unethical and not much information is available in the research domain. The information, whatever, is available is only in various blogs and shady sites.

30.3 Methodology

This research is aimed at securing any executable file of our choice with a password authentication mechanism integrated into its native environment, i.e., code section and not by encrypting the entire binary with a password. To inject the authentication mechanism and additionally ensure that this mechanism is adequately robust, the techniques used by the malware are planned to be reverse engineered and built into the executable.

To secure its binaries, Microsoft has implemented AppLocker in Windows 10 and Windows Server versions. AppLocker helps to control the files and apps the user can run. By creating an allowed list of files and apps, AppLocker prevents non-authorized user apps from running. It also helps in app usage only by a specific user [8]. The drawback, however, is that the AppLocker is only available in the enterprise-level version of Windows limiting its usage in small and medium enterprises (SME), academic institutions, and research labs with heterogeneous IT environment. Win utilities is another software that allows one to lock any executable programs with password to prevent unauthorized usage of the software [9]. PC Tune Up tool has got similar functionality to lock executables [10]. Though both are free utilities, these software seem to be encrypting the exe in-to with a password. Also, since they are closed-source, any harm caused by undisclosed vulnerability cannot be ascertained. To be able to better appreciate the concepts discussed in this paper, certain prerequisite knowledge viz., PE file format, code cave, Windows flow control, and code injection are considered essential. Windows application programs are widely known as portable executable (PE) files. The portable executable is a file format for various executable programs in Windows operating system viz., executables, object codes, and DLLs. The PE file format is a data structure that contains the necessary information for the Windows loader to execute the application [11–14].

Reverse Engineering. Reverse engineering is an advanced skill that deals with assembly language and works closely with the CPU and the other hardware. Assembly language is architecture dependent, and in this paper, it is discussed in the context of Intel 32 bit hardware architecture. Reverse engineering Windows binaries need familiarity in Windows API programming. Windows API provides an interface between the various Kernel functions and the assembly instructions. Understanding the Windows API is a key requirement, and it comes with wide experience of analyzing many programs. Working knowledge in any Windows debugger viz., Ollydbg, WinDbg, etc., is essential. The experiments discussed in this paper are carried out with OllyDbg debugger and RadASM IDE for assembling. Certain other tools that may aid are Hex Editor, Resource Hacker, ExeInfoPE, CFF explorer, etc.

Portable Executable. When the user double clicks a program icon to launch it, the PE loader will begin execution from what is known as "AddressOfEntryPoint," one of the important fields of the PE file format. Malware and backdoored programs alter this control flow and redirect the flow to the code cave where the malicious instructions are written. Once these malicious instructions are executed, the flow will again be redirected to the AddressOfEntryPoint, which will henceforth be referred as "Original Entry Point (OEP)," to resume the normal operation.

Code Cave. The concept of code cave is often employed by hackers and reverse engineers to execute arbitrary code in a compiled program. It helps in making modifications to the compiled executable including adding additional dialog boxes, removal of key validation checks, etc. A code cave can best be defined as "a redirection of program execution to another location and then returning back to the area where program execution had previously left" [15, 16]. To understand code cave, we need to first understand page size. When an executable is loaded into the RAM, a mapping is maintained between the virtual addresses and the physical addresses for the various sections of the PE file. This virtual address space is divided into pages, a contiguous span of addresses of a particular size [17]. The "SectionAlignment" field of the PE file dictates the size of the sections. This section alignment can be no less than the page size (currently, 4096 bytes on the Windows $\times$ 86) [18]. When a program has instructions that occupies more than 4096 bytes, a new page is created with a size of 4096 bytes. Understandably, if the instructions are less than 4096 bytes, the remaining bytes are filled up with zeros to maintain the page size. This part of the section that is filled up with zeros is called as code cave. Arbitrary instructions could be written to the code cave in the native assembly language, and the control flow may be redirected to execute this segment of code.

Code Injection. The program execution in a Windows environment starts from the instruction available at the AddressOfEntryPoint and then continues linearly. If the program encounters any jump or call statements, the control is redirected to the particular segment of code. Once that finishes, the control returns back to the previous point from where it continues execution again linearly. This is visually represented in Fig. 30.1.

However, in case of PE infection or code injection, the AddressOfEntryPoint is altered to point to the code cave, which in turn executes the instructions there. After it finishes, the execution flow is redirected to the OriginalEntryPoint, where the original program instructions are located, by adding suitable assembly instructions to the code cave. This makes the program to appear to behave normally (Refer Fig. 30.2). The flowchart (Refer Fig. 30.10) given in Appendix depicts the steps involved in the process.

Fig. 30.1 Regular control flow

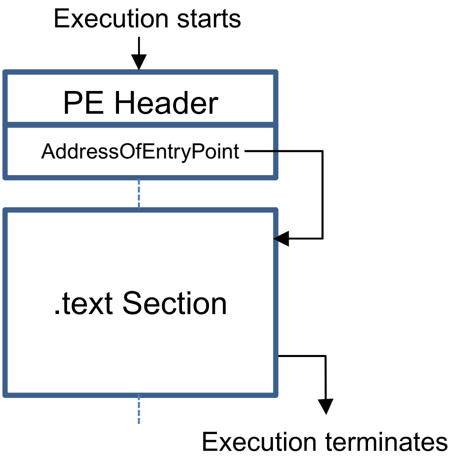
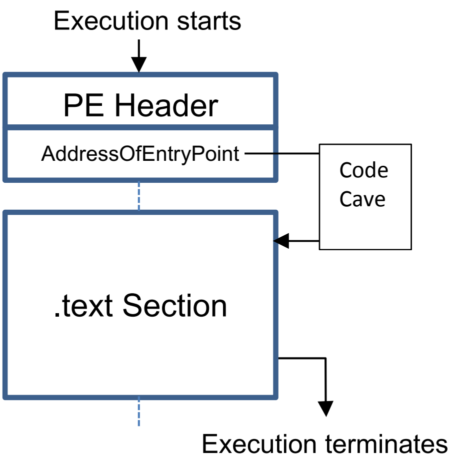


Fig. 30.2 Control redirection to code cave



30.4 Results

Windows programs are usually graphical in nature and are called as graphical user interface (GUI) programs. For the experiment, a GUI program was selected and embedded with custom codes to pop up a console which expects password for authentication [19–21]. The code injection was attempted on IIDKing, a GUI tool usually found in the arsenal of any reverse engineers. This tool is compiled using MS Visual C++ and is not packed. The code cave size is found adequate to embed our password authentication instructions. The successful implementation of the password authentication requires kernel32.dll and user32.dll which have already been imported by IIDKing. The result is a successful password authentication mechanism, which executes the program only if the user input password is correct (Refer Fig. 30.3);

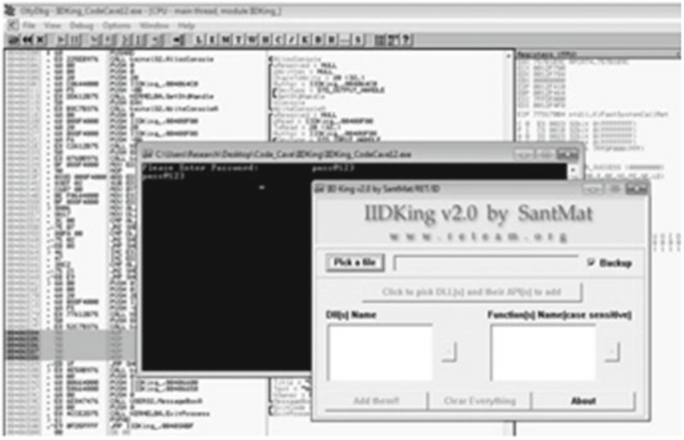


Fig. 30.3 Successful authentication

else, it pops up a message saying that the password was wrong and terminates the program (Refer Fig. 30.4).

However, if someone could open the binary in a hexeditor or any other equivalent tool, the hardcoded password would be noticeable which threatens the security of the program (Refer Fig. 30.5). To mitigate this risk, the password on-the-disk needs to be encoded, and the program should be able to decode it in-the-memory on the fly during execution. Also, the program may be made anti-debugging to prevent analysis by any novice malicious insider.

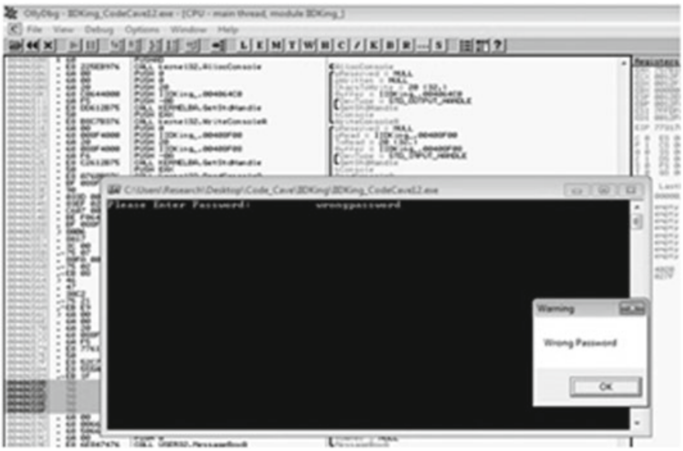


Fig. 30.4 Wrong password

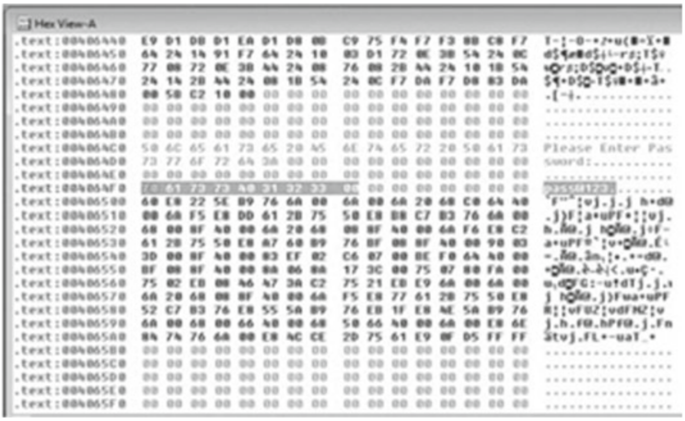


Fig. 30.5 Clear text password

A malware executable with MD5 Hash: 36527d5954bf3b2af60e6efa6398ccff uses IsDebuggerPresent() API to detect if it is being analyzed using a debugger [22]. Malware authors often use XOR function to encode the payload or the password in order to avoid detection or hide the contents from normal users [23]. Using these two techniques, the injected code has been modified to evade analysis inside a debugger and also encodes the password to make it harder to decode the same without analyzing the code, thereby providing a robust security measure. The assembly instructions to detect the presence of debugger and decode the obfuscated password are shown in Figs. 30.6 and 30.7, respectively. When this modified binary is analyzed using a debugger, it displays a warning and terminates the execution (Refer Fig. 30.8). The binary, when opened in a hex editor, now only displays an encoded password (Refer

Fig. 30.6 Assembly code to detect debugger

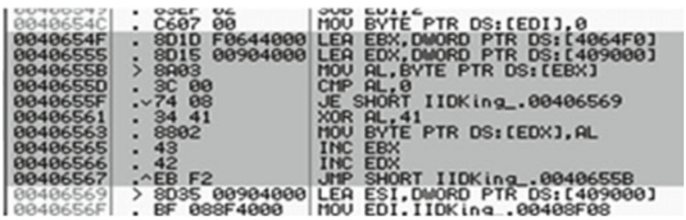


Fig. 30.7 Assembly code to decode password

Fig. 30.8 Debugging detected

Fig. 30.9 Encoded password

Fig. 30.9). These two techniques, though simple, are found effective to secure the target binary.

30.5 Conclusion and Future Work

In this paper, a method has been discussed to learn sophisticated techniques from the malware forensics and reverse engineering the malware and employ the same for hardening Windows applications. There are still many more advanced techniques used by malware authors which when implemented effectively would protect the Windows applications even from skilled hackers. This also prevents the usage of the target application from unauthorized users in the restricted workspace environment by obviating commercial software and, thus, saving huge expenses. The methodology discussed in this paper could be upscaled in the future work, and the following are few tips for achieving them:-

- The technique discussed in this paper is applicable only for 32 bit OS. The code needs to be modified for 64 bit OS in order to protect 64 bit applications.
- The PE may be added with a new section. This technique gives ample space to code a huge number of required features without worrying about the size.
- DLL injection could be another method to achieve the desired result with little overhead. This technique facilitates writing the instructions in any high level language of our choice.

A better approach would be to accept the password from the user and send it to a remote server for authentication. This kind of functionality, available only in web-based applications, could help in ensuring security of the program and monitoring and controlling who is executing the program. This hybrid application would provide the standalone functionalities of the desktop application and the accountability and auditing benefits of the web applications.

Appendix

See Fig. 30.10.

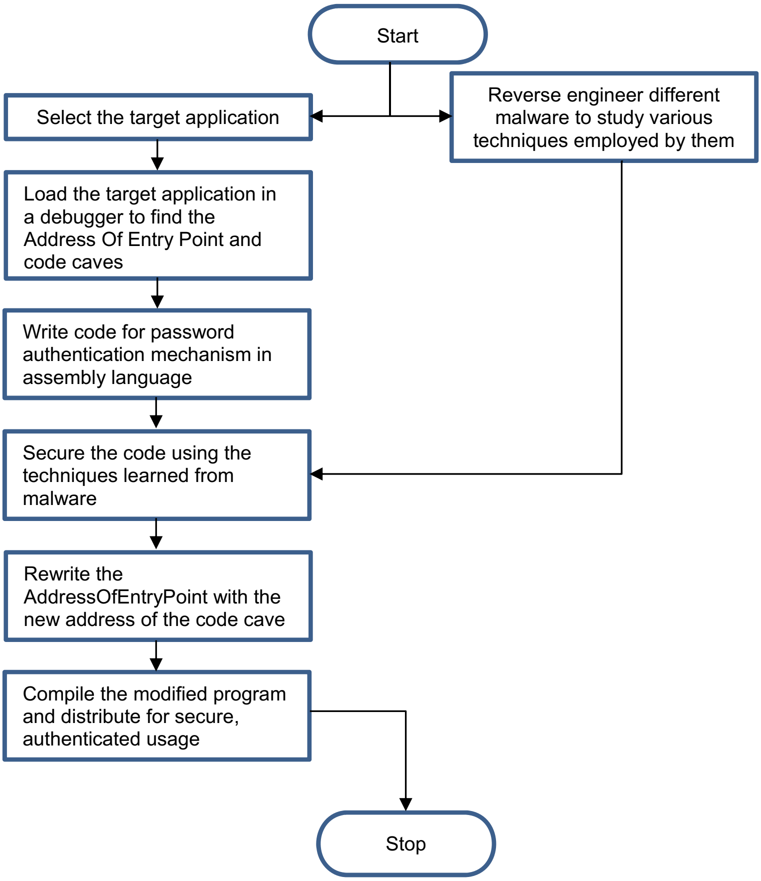


Fig. 30.10 Flowchart for hardening Windows application

References

1. Chen, C.M., Lai, G.H., Chang, T.C., Lee, B.: Detecting PE-infection based malware. In: Arai K., Kapoor S., Bhatia R. (eds) Advances in Information and Communication. FICC 2020. Advances in Intelligent Systems and Computing, vol. 1129. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-39445-5_56
2. Reddy, N.: Malware forensics. In: Practical Cyber Forensics. Apress, Berkeley, CA (2019). https://doi.org/10.1007/978-1-4842-4460-9_9
3. Oghenevo, E.E., Japhet, R.B.: Combating software piracy using code encryption technique. Int. J. Eng. Sci. (IJES) **5**(8), 18–24 (2016)
4. Goode, S., Cruise, S.: What motivates software crackers? J Bus Ethics **65**, 173–201 (2006). https://doi.org/10.1007/s10551-005-4709-9
5. Junod, P., Rinaldini, J., Wehrli, J., Michielin, J.: Obfuscator-LLVM—software protection for the masses. In: 2015 IEEE/ACM 1st International Workshop on Software Protection, Florence, pp. 3–9 (2015). https://doi.org/10.1109/SPRO.2015.10

6. Gagnon, M.N., Taylor, S., Ghosh, A.K.: Software protection through anti-debugging. In: IEEE Security & Privacy, vol. 5, no. 3, pp. 82–84 (May–June 2007). https://doi.org/10.1109/MSP.2007.71
7. Branco, R., Barbosa, G., Neto, P.: Scientific but not academical overview of malware anti-debugging, anti-disassembly and anti-VM technologies. Blackhat USA (2012)
8. https://docs.microsoft.com/en-us/windows/security/threat-protection/windows-defender-application-control/applocker/applocker-overview
9. Lock Exe Programs through Winutilities. https://www.win-utilities.net
10. https://www.pctuneupsuite.com/freeexelock/howto/how_to_lock_software_lock_exe_files_lock_applications_lock_programs.html
11. Ajit Kumar, Kuppusamy, K.S., Aghilab, G.: A learning model to detect maliciousness of portable executable using integrated feature set. J. King Saud Univ. Comput. Inf. Sci. **31**(2), 252–265 (Apr 2019). https://doi.org/10.1016/j.jksuci.2017.01.003
12. Jawwad, M.H.: A proposed system for hiding information in portable executable files based on analyzing import section. IOSR J. Eng. **04**(01), 21–30 (2014)
13. Xu, J., Feng, L.J., Ye, Y.L., Wu, Y.: An information hiding algorithm based on bitmap resource of portable executable file. J. Electron. Sci. Technol. **10**, 181–184 (2012). https://doi.org/10.3969/J.ISSN.1674-862X.2012.02.016
14. An Improved Method for Packed Malware Detection Using PE Header and Section Table Information
15. Complete Guide to Code Caves-Drew Benton
16. https://en.wikipedia.org/wiki/Code_cave
17. https://web.archive.org/web/20160611231342/, http://blog.cs.miami.edu/burt/2012/10/31/virtual-memory-pages-and-page-frames/
18. https://resources.infosecinstitute.com/2-malware-researchers-handbook-demystifying-pe-file/#gref
19. https://docs.microsoft.com/en-us/windows/console/allocconsole
20. https://docs.microsoft.com/en-us/windows/console/creation-of-a-console
21. https://www.codeproject.com/Articles/15836/Writing-to-and-read-from-the-console-From-a-GUI-ap
22. https://cybersecurity.att.com/blogs/labs-research/your-malware-shall-not-fool-us-with-those-anti-analysis-tricks
23. https://isc.sans.edu/forums/diary/Malware+and+XOR+Part+1/22486/

Chapter 31
Speed and Distance Alerting Device

Tirth Vyas, Ishan Thakkar, Yash Shah, Darpan Vasayani, Devanshi Tandel, and Radha Teredesai

Abstract This paper discusses how to tackle road accidents in school, college, hospital, and accident zones using IoT technology, and how this device will also help to maintain a safe distance between vehicles. And it will also examine the current scenario of rules and regulation proposed by the Indian Ministry of Road Transportation and Highways. Finally, this paper will suggest a device that can be applied to reduce road accidents most effectively.

31.1 Introduction

In India, around 97 thousand people died because of the road accidents prompted by over speeding in the year 2018. The number estimates for 64.4% of the total mortality in India, according to the report by the ministry of road transport and highways, under the title "Road Accidents in India, 2018" [1].

The rules provided by the government for school and accident zones are as follows: In April 2018, the Ministry of Road Transportation and Highways fixed the highest speed limit of a vehicle on controlled-access highway at 120 km/h, and the speed limit for domestic highways to 100 km/h, and for metropolitan roadways, it is 70 km/h for the heavy-duty category of vehicles.

T. Vyas (✉) · I. Thakkar · Y. Shah · D. Vasayani · D. Tandel · R. Teredesai
CSE Department, BITS Edu Campus, Vadodara, India

T. Senjyu et al. (eds.), *IOT with Smart Systems*, Smart Innovation, Systems and Technologies 251, https://doi.org/10.1007/978-981-16-3945-6_32

Institution zones are normally seen on roads next to schools and colleges where there is vital academy-related activity on and beside the road. The vehicle speed limit in these zones is 40 km/h on streets where the present limit is 50–70 km/h, and 60 km/h on paths where the limit is 80 km/h [2]. Road accidents have emerged as a major public health issue which has to be tackled in some way. Hence to implement the rules and to prevent road accidents in most effective manner, here are two devices. The first device is speed alerting device, second device is distance alerting device, and the following sections will discuss the implementation and the source code for calculating the speed of vehicle, calculating distance between vehicle and school using Android Things and also the calculation of the distance between two vehicles using Arduino.

31.1.1 Speed Alerting Device

Speed alerting device cautions the driver with a beep sound if the vehicle reaches school, hospital, or accident zones to reduce the vehicle speed if it is exceeding predefined speed limits of these zones, and if the driver reduces the speed of the vehicle to the predefined speed limit, the beep sound stops, and if the driver seems not to decrease speed, then the beep continues until the vehicle exits the zone.

31.1.2 Distance Alerting Device

Distance alerting device helps to maintain a safe distance with another vehicle in the front and alerts the driver with a buzz if the distance is less than 10 m and initiates a siren if the distance is less than 5 m and the device can also help in detection of other objects in front of a vehicle at a distance equal to or less than 10 m and can also be helpful in parking.

31.2 Related Work

The section manifests the literature review, which summarizes and explains the related work in the field of speed control of the vehicle and obstacle detection in accident zones and geofencing technology. Arslan et al. [3] has proposed an IoT Android Things application which can be implemented for controlling the lights, as this system is implemented using Raspberry Pi and Ubidots Cloud Services. Shabaz Khan et al. [4] has proposed RF communication and geofencing technology to alert the driver and control the speed accordingly. Shrivastava et al. [5] has proposed an obstacle distance measurement system using ultrasound sensors. Jyotirmoy [6] has developed a java application using geofencing which changes the device sound profiles as per the location.

31.3 Methodology and Implementation

31.3.1 Speed Alerting Device Methodology with Implementation

The methodology in the article comprises the following stages. For speed alerting device, Raspberry Pi 3 Model B can be used as a hardware component and Android Things as the software component which will help to provide sync between software and hardware tools. The first task of the device is to provide the geolocation of the vehicle for that geolocation API and a GPS module can be used for getting the precise location of the vehicle. The second task is to calculate the speed of the vehicle and for that location class can be used and integrating it with GPS. The source code and the calculation of the geofence and gelocation are as follows.

Java source code for calculating the distance between geofence and geolocation of the vehicle,

```
// Calculating distance

public static double get_both_dis(double lati_1,double
longi_1, double lati_2, double longi_2)
{final int Radi_earth = 6371; // Earth Radius

double lati_measure = Math.toRadians(Math.abs(lati_2 -
lati_1));
double longi_measure = Math.toRadians(Math.abs(longi_2 -
longi_1));
double abc = Math.sin(lati_measure/2)*
Math.sin(lati_measure/2)+
Math.cos(Math.toRadians(lati_1))*
Math.cos(Math.toRadians(lati_2))
* Math.sin(longi_measure/2)
* Math.sin(longi_measure/2);
double cad = 2 * Math.atan2(Math.sqrt(abc),
Math.sqrt(1 - abc));
double disti = Radi_earth * cad * 1000;
disti_between = Math.pow(disti, 2);
return Math.sqrt(disti_between);}
private boolean is_In_School(double a,double b)
{// the lat and long of : Crystal School Waghodia Road
   final double X = 22.2910942;
   final double Y = 73.2365132;
   // radi of 250 m checking
if(getDistance(X,Y,a,b) <= GEOFENCE_RADIUS)
{return true;}
else{ return false; }
```

The final task of the device is to compare the speed limit in the school or accident zones for that Roads API, for extracting the virtual location and creating a geofence in a 250 m of radius area, and alerting the driver accordingly with led and buzzer [7]. The visual implementation can be seen in Fig. 31.1.

Java source code for calculating the current speed of vehicle and alerting the vehicle in geofence and if speed limit is exceeding the predefined.

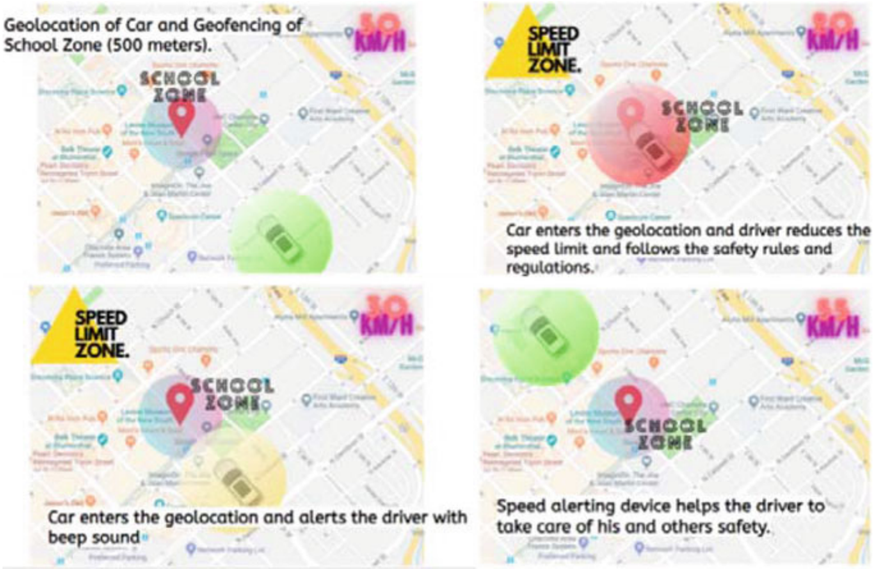


Fig. 31.1 Visual implementation of speed alerting device

```
public void on_change_of_location(Location location) {
if (location == null)
{tv_speed.setText("00 km/h");// setting default speed}
else{
int t_speedy = (int) ((location.getSpeed() * 3600) /
1000); // converting meters to kilometers
tv_speed.setText(t_speedy + " km/h"); //speed
Double lati_s = location.getLatitude();
Double longi_s = location.getLongitude();
if (t_speedy > 30 && is_In_School(lati_s,longi_s))
{Toast.makeText(this, " Please Slow Down...",
Toast.LENGTH_LONG).show();player =
MediaPlayer.create(MapsActivity.this, R.raw.slowdown);
player.start();}}}
```

31.3.2 *Distance Alerting Device Methodology with Implementation*

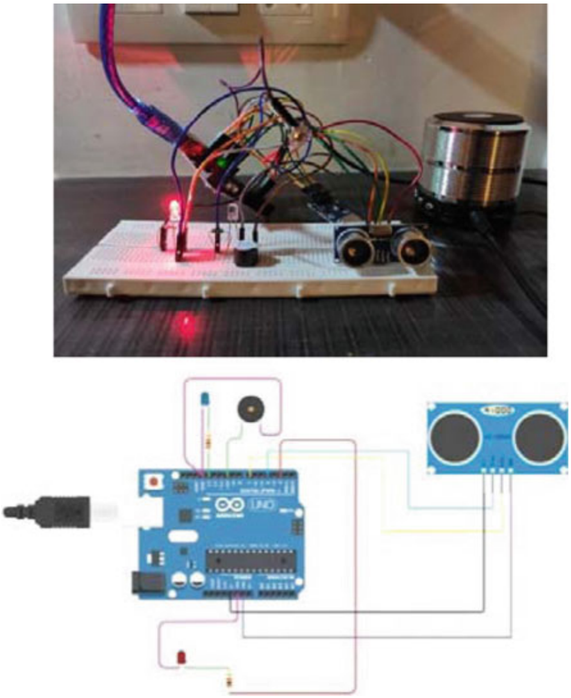
The tools for the distance alerting device will have Arduino UNO, ultrasonic sensor, led, and buzzer. The ultrasonic sensor will detect the distance between the vehicles and alert the driver in audio-visual form [8, 9]. The visual implementation can be seen in Fig. 31.2. Arduino controller to computer connection can be seen in Fig. 31.3.

The class diagram illustrated in Fig. 31.4 shows the flow of the microcontroller (Arduino UNO) with the ultrasonic sensor which works to calculate the distance between the vehicles, and the Arduino is also connected with shared led and buzzer with the Raspberry Pi, and this single board computer receives and processes the information received from the Android Things App which coordinates with the GPS module and calculate the speed, distance, and location of vehicle in different zones.

Calculating distance in distance alerting device.

Fig. 31.2 Visual implementation of distance alerting device

Fig. 31.3 Arduino controller to computer connection



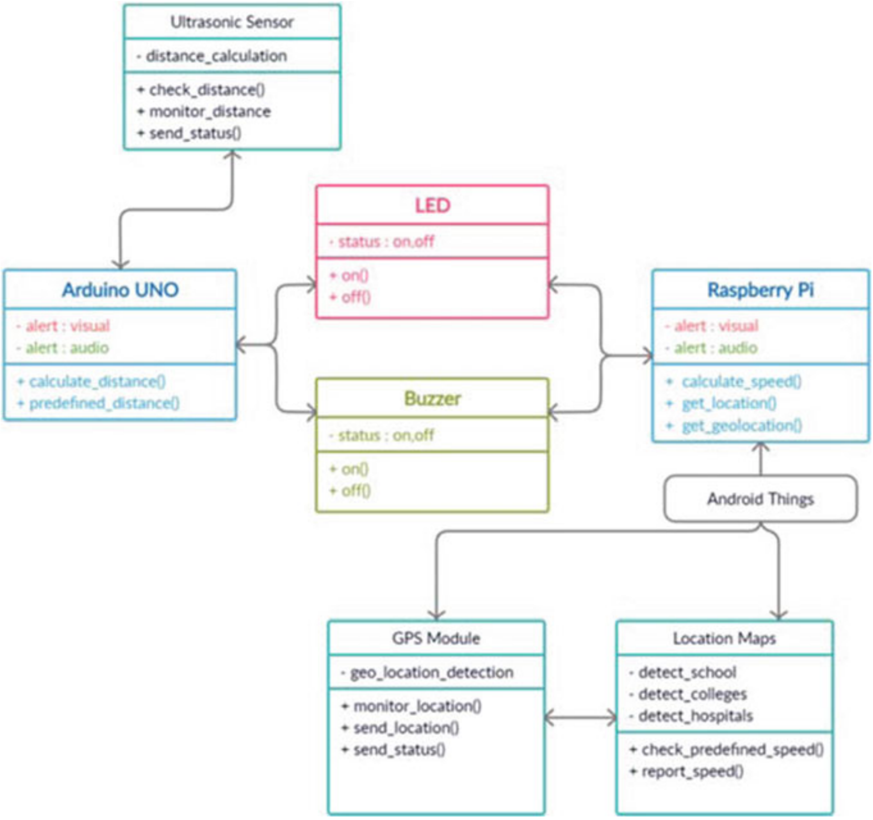


Fig. 31.4 Class diagram of speed and distance alerting device

```
// calculating the distance
distance= duration*0.0133/2;
safetyDistance = 40*distance;
timer = distance 10;
if (safetyDistance>= 300 && safetyDistance<=800)
{ digitalWrite (ledPinBlue, HIGH);
else{ digitalWrite (ledPinBlue, LOW);
if (safetyDistance <= 300)
{digitalWrite (buzzer, HIGH);digitalWrite (ledPin, HIGH);
delay (timer); digitalWrite (buzzer, LOW) ; digitalWrite
(ledPin, LOW); delay(timer);
else{digitalWrite (buzzer,LOW); digitalWrite(ledPin,
LOW);}}}
```

31.4 Conclusion and Future Scope

31.4.1 Conclusion

Speed and distance alerting device will help to reduce the road accidents and maintain a safe distance between two cars and detect the objects in front of the car, and as it also detects the vehicle speed and alert the driver if the vehicle speed is exceeding the predefined speed limit in school, college, and hospital zone.

Distance alerting device uses the ultrasonic sensor and calculates the distance between the cars and cautions the driver with LED and buzzer with a beep sound if the distance is less than 10 m.

Speed alerting device uses the GPS module for vehicle location, location class calculates the speed of the vehicle, using MAPS API, the physical location of the schools and colleges is extracted, and a geofence is created of 500 m which alerts the driver if the speed limit is more than the predefined speed limit in the zones.

31.4.2 Future Scope

Further, this device can be linked to an Android and iOS application, and data ingestion can be performed using Cloud IoT core. The important data can be abstracted, and can be used to derive driving performance, over speeding analytics, and many other useful data like frequent visit to school, college, or accident zone and alert the driver in advance to drive safe and slow. Hence, this data can be used in a variety of domains like vehicle insurance, testing driving skills, and several others.

References

1. Article on accidents caused due to over speeding by livemint.com. https://www.livemint.com/news/india/road-accident-data-more-than-97-000-people-killed-due-to-over-speeding-in-2018-11574249715762.html
2. Wikipedia page on the Ministry of Road Transport and Highways. https://en.wikipedia.org/wiki/Ministry_of_Road_Transport_and_Highways.
3. Arslan, S., Dagdeviren, O., Kardas, G.: An IoT LDR bulb application with android things operating system for smart cities. In: 2019 Innovations in Intelligent Systems and Applications Conference (ASYU) by IEEE. ISBN: 978-1-7281-2869-6. https://doi.org/10.1109/ASYU48272.2019.8946400.
4. Shabaz Khan, A., Chethan, M.N., Akash, O., Rakesh Singh, Ch., Manjunatha, D.V.: Speed control of vehicle in accident zone. Int. J. Innov. Res. Sci. Eng. Technol. **7**(4), 3274–3281. ISSN (Online): 2319-8753. ISSN (Print): 2347-6710 (Apr 2018)
5. Shrivastava, K., Verma, A., Singh, S.P.: Distance measurement of an object or obstacle by ultrasound sensors using P89C51RD2. Int. J. Comput. Theor. Eng. 1793–8201 **2**(1), 64–68 (Feb 2010)

6. Geofence Detection by jyotirmoy on GitHub. https://github.com/jyotirmoy-paul/android-geofencing
7. Android Things. https://developer.android.com/things
8. Arduino Forum. https://forum.arduino.cc/
9. Built-In Examples. https://www.arduino.cc/en/Tutorial/BuiltInExamples

Chapter 32
Design and Implementation of an Efficient Scalable Forwarding in Named Data Networking (NDN) Using Huffman Coding

Devi Kala, Anurudh Kumar, N. B. Arunekumar, and K. Suresh Joseph

Abstract This paper presents software-based name encoding technique which is used for NDN by enhancing Huffman coding. Previous findings ascertained that there is need to reduce memory access time during insertion, deletion and update in NDN. A scalable and memory efficient scheme that improves memory access of NDN during insertion, deletion and update in the PIT using Huffman coding is proposed. Using the proposed approach, components in the named URL are assigned as tokens wherein each component will be provided with a token. The encoded names are stored in the EHNRT which tends to facilitate lookup in a cost effective manner. The findings of the present research revealed that the proposed method EHNRT will reduce memory consummation 47.30% compared to the most effective scheme ENPT for up to 19 million prefixes per seconds.

32.1 Introduction

32.1.1 Named Data Networking—A Brief

In the recent years, the interest towards information-centric network (ICN) has phenomenally increased wherein there is a significant growth in the evolution of the design and development of ICN-based prototype architectures. Two of the projects that have been utilising ICN-based framework are the content-centric networking (CCN) [1] and its open source implementation named data networking (NDN) [2]. Among the two projects, NDN has gained momentum and is a newly emerging internet architecture that was developed from years of empirical research that examined the ways to manipulate network usage. The producer of the data then replies to the consumer with the data packets, wherein the reverse path of the interests is taken [3]. NDN has acquired immense attention owing to the fact that it has better advantages such as low dissemination latency, reduction of network load and better energy efficiency than the traditional TCP/IP architecture [4]. One of the major issues

D. Kala · A. Kumar · N. B. Arunekumar · K. Suresh Joseph (✉)
Pondicherry University, Kalapet, India

T. Senjyu et al. (eds.), *IOT with Smart Systems*, Smart Innovation, Systems and Technologies 251, https://doi.org/10.1007/978-981-16-3945-6_33

with NDN is that the architecture is memory intensive and the use of router's buffer will affect the cache search time negatively. When the cache hit rate is not sufficient, then the cache search time increases. Furthermore, the use of huge fast memories to serve the cause is also not affordable [5]. If the forwarding information base (FIB) is made up of slow and large memories, then it will burden the routers to complete the forwarding faster [6]. An efficient data structure needs to be designed which should reduce memory consumption to some extent, such that it lowers the complexities of computation and enables fast name lookup. Some of the data structures used to serve the cause were the Component Character Trie (CCT) [7], hash table [8] and the Name Prefix Trie (NPT) [9] used as the algorithm to compare the name present in data packet against the interest packet. Moreover, the use of hash table tends to map the component or the whole of the content name into an array using a hash function which supports fast lookup of name and incurs relatively less memory; however, the hash table has limitations such as hash collisions which increase the chance of occurrence of false positives. However, while considering the CCT, each single node is considered as a character, and hence, a complete task of name lookup might require numerous numbers of nodes which is equivalent to the number of characters except the delimiter '/' [10]. The difference between NPT and CCT is based on the fact that the granularity of the NPT is based on the component lever, whereas in contrast, the CCT is based on the character lever. Dai et al. [11] have utilised the modified version of the ENPT and simplified the STA data structure. Furthermore, authors used the radient scheme which is used to encode name in the PIT of NDN router and radient trie is used to reduce lookup cost.

32.2 Related Work

In this section, the finding presents the various lookup schemes developed. On forwarding to the next hop, a prefix entry is marked on the forwarding table. However, an important factor to consider for designing is towards minimisation of the length that could be obtained at the worst case of the linked lists for balancing the bucket population by bloom filter. For IP lookup, various hardware schemes are available of which the ternary content addressable memory (TCAM) tends to improve throughput of lookup. However, high consumption of power and poor scalability is found to be the limitations of the TCAM approach. Furthermore, most TCAMs are found to run at low speeds when compared to memory that is SRAM based. Software-based algorithms include the simple binary trie structure which was used initially to achieve lookup of IP, wherein other algorithms emerged such as the SAIL, tree bitmap and so on. For fast lookup using hardware-based solutions, GPU and TCA were used with implications of high power consumption.

32.3 Proposed Solution for Name Encoding

Huffman coding and data structures the proposed in the work is an extension of the work proposed by Saxena and Raychoudhury [5], wherein our work in the present paper has utilised Huffman coding as a greedy algorithm technique for the reduction of memory needs during name lookup in NDN. The Huffman coding is incorporated with proposed model for the reduction of lookup costs and the memory consumption in the PIT.

The proposed framework comprises of two modules which are shown in Fig. 32.1. The two modules are as follows, Enhanced Huffman-based component radix trie (EH-CRT) and enhanced huffman name radix trie (EHNRT). EH-CRT comprises of two components namely the token stack and the token frequency map. The token stack will make a record of the tokens that have been newly created or deleted, whereas the token frequency map will keep a track of the count of components on the PIT table. However, prior entering the EH-CRT, the names will be encoded using Huffman coding. In this paper, variable length encoding has been utilised.

However, the creation of the Huffman tree was based on the assignment of the smallest code for the most frequently used characters and largest code for frequently used characters. When a request arrives for content, the name will be encoded and will be lookup in the EH-CRT. However, the encoding occurs at the name encoding module. When the component searched is present in the EH-CRT, then the token corresponding to the particular component will be returned. However, in the absence of the matching component, a new token will be generated for the respective component using the token stack. Furthermore, with the new component generated, the frequency of the component will be incremented by 1. Such a process is performed until tokens are generated for all the components with the delimiters and then the encoded name will be generated. Then the encoded name will be sent to the EH-CRT

Fig. 32.1 Enhanced Huffman component radix trie (EH-CRT)

for the update or insert operations in the PIT. For the insertion of a name in the PIT table, a lookup needs to be performed to examine the presence of the name in these data structures.

32.3.1 Enhanced Huffman Component Radix Trie (EH-CRT)

This section explains the processes involved in the generation of tokens for the names using the EH-CRT data structure. Whilst the conventional CRT structure is optimised to act as a compact prefix trie with common characters sharing similar memory, the EH-CRT is a bit different. When an interest arrives at the NDN router, the name will be decomposed using the delimiter '/' wherein each component will be assigned a unique token based on the frequency of occurrence. Three of the major operations that take place in EH-CRT are insertion, deletion and lookup. The operations of EH-CRT are provided in Fig. 32.1. For the lookup of a component in the EH-CRT, the component will be encoded using a variable length encoding technique. This is an attempt to reduce the number of bits stored after encoding. For instance, the name components that occur frequently will acquire few bits, whereas the name components that occur rarely will have larger bit codes. For instance, Table 32.1 depicts the encoded codes after the generation of the Huffman tree. For the process of lookup in the EH-CRT, the component will be encoded based on its position in the Huffman tree. For example, the component 'divest' arrives more frequently whereas the word 'deviate' occurs least number of times. In that case, based on the frequency of the component, it will be added as the last leaf node. In a Huffman tree, each code depicts the path from the tree's root to a leaf node, wherein '0' bit is assigned to the branch in the right and '1' to the branch in the left. Whenever a leaf node is arrived, the entire length of traversing will provide the encoded code [6]. Suppose when a new component arrives at the EH-CRT, then the component will be inserted into the Huffman tree with a new token. When a new component is added, a new token is

Table 32.1 Encoded code table

Name components	Frequency	Encoded codes	Tokens assigned
Deviate	5	1111010	1
Deviation	7	1111011	2
Device driver	10	111100	3
Device	15	111110	4
Diverge	20	111111	5
Diverse	45	1110	6
Diver	65	110	7
Divest	110	10	8
Diversity	175	0	9

Fig. 32.2 Arrival of the interest packet for *N*

```
// Algorithm- Context: During the arrival of the
interest packet for N //
  Begin
  execute function decompose (N)
  execute lookupEHCRT (C array)
  For all Ci ∈ C array
  [if C]_i∈EHCRT
  TC =TC + 1
  Else
  Execute insertEHCRT (Ci)
  EN= EN U TCi
  TC =TC + 1
Insert Ci in EHCRT
  If DStack. isEmpty()
  TCi = pop(DStack)
  else
  TCi = TCi +1
  Return TCi
  Lookup EHNRT (EN)
  If EN ∈ EHNRT
  If nonce.new=nonce.prev
  Drop interest
  Else
  Add interface
  Else
  Execute insertEHNRT(EN)
  Insert EN in the EHNRT
Return PtrPIT
```

assigned; however, an increase in the existing component will lead to increment in the frequency of the component (Figs. 32.2 and 32.3).

Whenever a component needs to be deleted, the total number should be decremented by 1. When the availability of any component completely becomes zero, then the component is deleted. However, with the deletion of a component, the token becomes free which can be assigned with a new component. Throughout the process of name encoding, the EH-CRT generates the name that is encoded for the EHNRT. Overall, when the name arrives as 'deviate/diverge', then it will return the encoded value as '1/5'. Such a feature enables quick name lookup. However, after the generation of the encoded names from EH-CRT, the names are then passed to EHNRT, wherein these names are looked up. In case of non-existence of the names in the EHNRT, an insert operation will be performed. EN-NRT will then compare the nonce of the entry that exists with the incoming nonce of incoming interest packet and then stops processing when both the names are found to be equal. Furthermore, such a process assesses the list of interfaces that are incoming and performs no action if the interface with interest incoming already exists in the matched entry list of PIT. If both nonce not matching and incoming interface exists previously, then the interest incoming interface will be appended with the PIT to the entry list of interfaces that exist. EHNRT which is an enhancement of NRT uses tokens as well.

Fig. 32.3 During the arrival of the interest packet for *N*

```
// Algorithm- Context: During the arrival of the
data packet for N //
  Begin
  Execute function decompose(N)
  Execute lookup DeleteEHCRT(C array)
  lookupDeleteEHCRT(C array)
  For all Ci ∈ C array
  If Ci ∈ EHCRT
  Tc = Tc -1
  If Tc.isEquals(0)
  Execute deleteEHCRT(Ci)
  EN = EN U TCi
  Execute DeleteEHNRT(EN)
  DeleteEHCRT(EN)
  Delete Ci from the EHCRT
  Push DStack (TCi)
  Return
  deleteEHNRT(EN)
  For all TCi ∈ E_N
  Delete EN TCi path
  if timer expired
  Drop the data packet
  Else
Send data packet to the incoming interface
De-allocate the memory
```

Furthermore, the request for insert, update and delete is performed on the NRT which returns entries for PIT accordingly. This section discusses the following algorithms that were used for insertion, deletion and update using the proposed scheme with Huffman coding which is lossless in nature. Following are the assumptions and the variables considered in the algorithm as shown in Table 32.2. EHNRT basically works on two phases, where first phase is to build an EHNRT tree from input characters followed by the second phase to traverse the EHNRT and assign codes to characters. Table 32.2 represents some assumptions over the variables used in the algorithms. Table 32.2: Variables used in the enhanced scalable forwarding algorithm in NDN using Huffman coding.

32.4 Analysis of Complexity

In the present section, the complexities of space for the EHNRT and EH-CRT structures used for the storage of name sets and its encoded forms are examined. Analyses of the complexities in time of the various operations that are performed on PIT using the proposed approach are done. It is assumed that the requests in the form of N name arrive at the NDN in a sequential manner, wherein each name comprises of ‘C’ components which also contain ‘CChar’ characters. The encoded name will be

Table 32.2 Variable used in the enhanced scalable forwarding algorithm

Variable	Significance
N	Stored content name in data and interest packet
Nc	Total number of components in the content name *N*
Ci	Component that is present in the content name (every *i*th component) where ‘C1 + C2 + C3….. + Cn = N’
C array	Array for the storage of Ci
EN	Encoded name for the *N* for each Ci
TCi	Token value assigned for a name component
TC	Total components in the name
Darray	Array for the storage of components deleted
PtrPIT	Pointer to the PIT
DStack	De-allocated token management stack
nonce	A 4-octet long byte string present within the interest packet

denoted as ‘Ne’, wherein the same has encoded components designated as ‘EK’. The time and memory complexity of the proposed algorithms are calculated as follows:

32.4.1 Memory Complexity

Memory complexity is defined as the measure of the total space that is required for the storage of all *N* names, wherein the space consumed by a single name in the NDN is computed as $\sum Ci$. The Huffman coding is used to create the tree, wherein the allocation of memory is based on each and every name component. In that case, there is no sharing of memory for the storage of components wherein each component will have a unique memory space. In the EHNRT, the memory that is required for the storage of encoded name is depicted as ‘Ne’ which is equal to the total number of components in a name. Hence, the memory required for the storage of names would be $N * C$; however, the memory complexities are the summation of both EH-CRT and EHNRT complexities. Hence, the memory complexity would be $O\ (N * C)$.

32.4.2 Time Complexity

In the EH-CRT, the insertion of each name component takes ‘*C*’ time, and hence, the insertion of a single name will take $\sum Ci$ time. In this context, the time complexity would once again become $O\ (N * C)$. For all update, deletion and insertion operations, the time complexity remains the same.

Table 32.3 Dataset statistics

Dataset and dataset size	Number of names	Total component number
Backlist-4 M	3,736,394	16,252,070
Shallalist-4.5 M	4,535,775	13,398,702
Heavy workload-10 M	10,000,000	61,213,285
Total-18.5 M	18,272,169	90,864,057

32.5 Experimental Setup

32.5.1 Dataset and Experimental Setup

The proposed approach was developed and implemented using C# programming and Visual Studio 12. The program is executed on an i5 2.40 GHz processor with 4 virtual cores and an 8 GB DDR3 RAM storage memory over the datasets such as Shallalist, Heavy Workload and Blacklist. All these datasets were combined together which formed a dataset of 18 M. For all the URLs in the datasets, all URLs transformed into named data format. For instance: 'http://www.pondiuni.edu.in/' was converted into 'in/edu/pondiuni/'. Each dataset was identified based on their size (Table 32.3). To compare the performance of the proposed method with that of previous methods, the results of techniques that used hash table, CCT, ENPT and NPT were acquired and examined.

32.5.2 Performance Metrics

The total memory usage and the frequency of operations that are used for the insertion, deletion and update of names in the PIT are measured. The performance metrics used are as follows. Total usage or consumption of memory is the amount of EH-CRT and EHNRT memory used for the storage of encoded names in the PIT. Frequency denotes the number of names that are inserted/deleted or updated per second. The metric is measured in million/second. Memory consumption by the encoding of the named URLs is performed, wherein the results are revealed in Fig. 32.4. Using the Huffman coding scheme used for encoding, the proposed approach has reduced the memory consumption of all the datasets significantly, wherein the results are revealed in Fig. 32.4. For the better analysis of results, the insert, delete and update frequencies of the datasets are depicted in Fig. 32.5.

It is deemed that the memory consumption frequency of the proposed approach is less when compared with other schemes. This results claim that the proposed approach is way better than the other schemes compared as shown in Fig. 32.5.

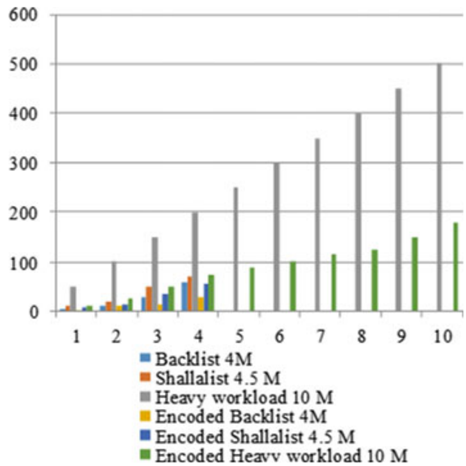


Fig. 32.4 Memory consumption of original and encoded names for the datasets

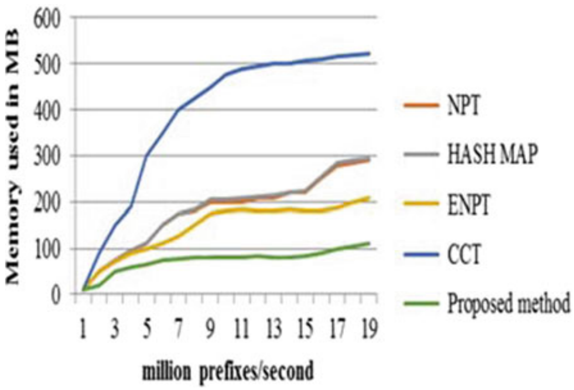


Fig. 32.5 Comparison of the proposed scheme and other schemes

32.6 Conclusion and Future Works

The present paper has revealed the utilisation of the Huffman coding as an algorithm that could improve name lookup in NDN. Using the proposed approach, components in the named URL are assigned tokens, wherein each component will be provided with a token. The encoded names are stored in the EHNRT which tends to facilitate lookup in a cost effective manner. The findings of the present research revealed that the proposed method could be used to reduce memory consumption whilst storing named data. However, future researches should look into the venues of improving the usage of Huffman coding in NDN which is least used in the named networking domain.

References

1. Ahmed, S.H., Bouk, S.H., Kim, D.: Content-centric networks: An overview, applications and research challenges. Springer, Singapore (2016)
2. Zhang, L., Afanasyev, A., Burke, J., Jacobson, V., Claffy, K., Crowley, P. et al.: Named data networking. ACM SIGCOMM Comput. Commun. Rev. [Internet] 28. **44**(3), 66–73 (2014)
3. Soniya, M.M.S., Kumar, K.: A survey on named data networking. In: 2015 2nd International Conference on Electronics and Communication Systems (ICECS) (2015)
4. Wu, H., Li, J., Pan, T., Liu, B.: A novel caching scheme for the backbone of Named data networking. In: 2013 IEEE International Conference on Communications (ICC) (2013)
5. Rezazad, M., Tay, Y.C.: NDN||MEM: An architecture to alleviate the memory bottleneck for named data networking. In: Proceedings of the 2013 Workshop on Student Workshop—CoNEXT Student Workhop '13 [Internet]. New York, New York, USA: ACM Press, pp. 1–4 (2013)
6. Li, Z., Liu, K., Liu, D., Shi, H., Chen, Y.: Hybrid wireless networks with FIB-based named data networking. EURASIP J. Wirel. Commun. Netw. (2017)
7. Wang, Y., He, K., Dai, H., Meng, W., Jiang, J., Liu, B. et al.: Scalable name lookup in NDN using effective name component encoding. In: 2012 IEEE 32nd International Conference on Distributed Computing Systems (2012)
8. Takemasa, J., Koizumi, Y., Hasegawa, T.: Toward an ideal NDN router on a commercial off-the-shelf computer. In: Proceedings of the 4th ACM Conference on Information-Centric Networking ICN '17 (2017)
9. Li, D., Li, J., Du, Z.: An improved trie-based name lookup scheme for named data networking. In: 2016 IEEE Symposium on Computers and Communication (ISCC) (2016)
10. Wang, Y., Zu, Y., Zhang, T., Peng, K., Dong, Q., Liu, B. et al.: Wire speed name lookup: A GPU-based approach. In: 10th USENIX Symposium on Networked Systems Design and Implementation (2013)
11. Dai, H., Liu, B., Chen, Y., Wang, Y.: On pending interest table in named data networking. In: Proceedings of the Eighth ACM/IEEE Symposium on Architectures for Networking and Communications Systems (2012)

Chapter 33
Analysis of Multiple Antenna Techniques for Unmanned Aerial Vehicle (UAV) Communication

Rajesh Kapoor, Aasheesh Shukla, and Vishal Goyal

Abstract Military and civil applications of unmanned aerial vehicles (UAVs) are increasing at rapid pace. These applications range from battlefield surveillance to goods transportation. Despite their usefulness, UAVs typically require appropriate communication connectivity for their successful operation, which is a very challenging task. Therefore, to ensure enhanced utilization of UAVs in future, it is important to have reliable and high data rate wireless communication links between UAVs and their ground stations. For achieving this objective, UAVs can be integrated into existing cellular networks as aerial users. Cellular communication support to UAVs is generating new research challenges and opportunities. One way of enhancing the cellular communication support to UAVs is to use multiple antennas. In this article, we present an overview of this promising technology, by initially covering UAV cellular communications. We then carry out analysis of various multiple antenna techniques to bring out the aspect of improvement of performance of the UAV communication links, particularly by increasing the number of antennas.

33.1 Introduction

An unmanned aircraft which is piloted by a remote control unit or by an on-board computer is called an unmanned aerial vehicle (UAV). Other terms used for UAV are drone, unmanned aircraft system (UAS) or remotely piloted aircraft system (RPAS). Traditionally, UAVs have been used for military applications of battlefield and airspace surveillance and patrolling. However, with the advancements in technology particularly in field of miniaturization of electronic instruments and improvements in

R. Kapoor (✉) · A. Shukla · V. Goyal
GLA University, Uttar Pradesh, Mathura 281406, India
e-mail: rajesh.kapoor_phd17@gla.ac.in

A. Shukla
e-mail: aasheesh.shukla@gla.ac.in

V. Goyal
e-mail: vishal.goyal@gla.ac.in

T. Senjyu et al. (eds.), *IOT with Smart Systems*, Smart Innovation, Systems and Technologies 251, https://doi.org/10.1007/978-981-16-3945-6_34

control systems, civil applications of UAVs are growing. Civil applications include search operations during natural disasters, crowd or traffic surveillance, wildlife conservation, goods transportation, etc. [1–3]. Easily available low-cost quadcopters drones are being regularly used for drone racing and aerial photography. Many such civil applications will emerge in future [4–6]. UAVs can also be put into use as aerial platforms for communication such as airborne base stations (BSs) or flying relays. Such arrangement is commonly called UAV-assisted communication, which is realized by mounting the communication transceivers to UAVs, for provisioning or enhancing communication services to the users on ground [7, 8]. Similarly, cellular-connected UAVs are the ones which are used as aerial users/nodes [9, 10]. A swarm of UAVs creating flying ad hoc networks (FANETs) is being used to support high rate wireless communications in a large geographical area [11, 12]. In this article, review of cellular communication support to UAVs is covered in Sect. 33.2, and analysis of various multiple antenna techniques for UAV communication is Sect. 33.3.

33.2 Review of Cellular Communication Support to UAVs

Wireless connectivity is required by aerial UAVs to support their communication needs. Wireless fidelity (Wi-Fi), ZigBee, etc., are the existing wireless technologies that are being used by UAVs to communicate to their respective remote control station situated on ground. Such technologies have few drawbacks, such as low mobility, extremely short range and low throughput, which restrict their usage primarily for wireless access for indoor scenarios. Communication link between UAV and its ground station is generally established by use of unlicensed spectrum. The range between a ground station and UAV is primarily limited by unlicensed frequency spectrum used and transmits power. Enhanced ranges can be achieved by already deployed cellular networks, which are no more limited by unlicensed frequency spectrum. If the cellular network delivers the data rate required by a UAV, then UAV will have same range as the coverage of cellular network. Figure 33.1 shows the connectivity of UAV with ground station using one to one link. Figure 33.2 shows the connectivity of UAV with a cellular network.

The existing cellular communication networks are found to be non-appropriate for few typical requirements of UAVs communications. The inadequacies of existing cellular networks can be removed by utilization of multiple antennas in base stations.

33.3 Analysis of Multiple Antenna Techniques for UAV Communication

Goal of digital communication is to transfer information bits, as these bits cannot be sent directly through communication channels. Thus, these bits are required to

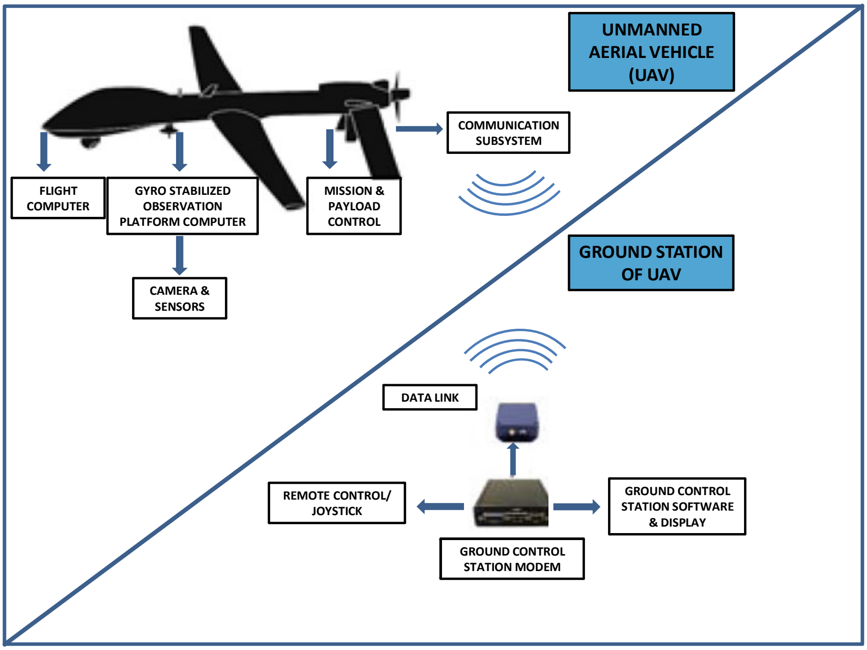


Fig. 33.1 UAV connectivity with ground station

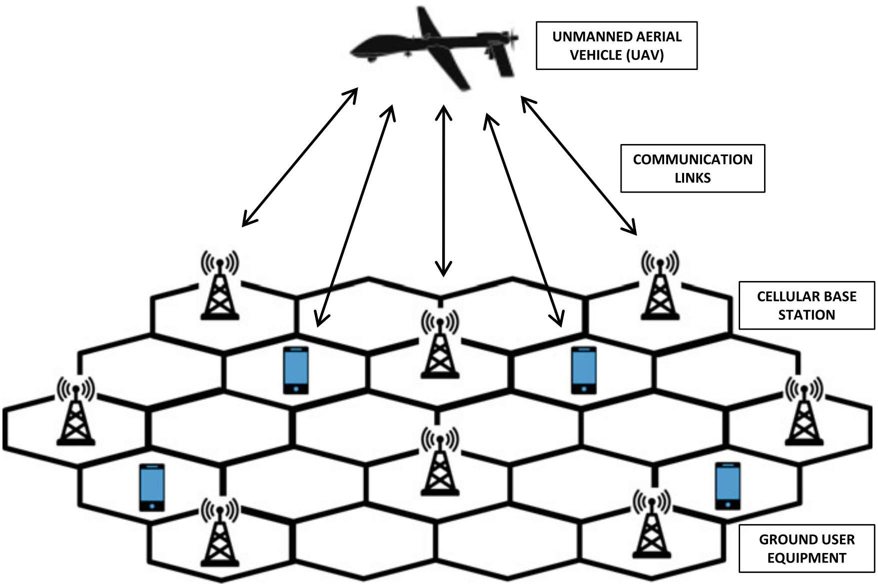
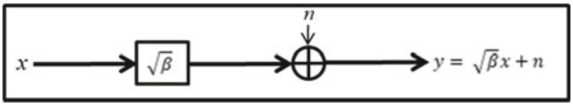


Fig. 33.2 UAV connectivity with cellular network

be mapped to modulation symbols. For example, two bits are transmitted with one symbol in QPSK modulation. Basic example of communication channel is shown below.



Here, x is transmitted symbol, β is channel gain, n is noise, and y is received signal. All values are complex valued numbers. The receiver guesses the value of x based on received signal y. This guess is called $\hat{x}$, and the probability Prob$\{x \neq \hat{x}\}$ is called symbol error rate or symbol error probability, which is ought to be zero. This is the probability of one symbol, whereas generally one symbol is not sent, but packets of symbols consisting of header and information symbols are sent. The header consists of channel code that protects against errors. For particular choice of modulation scheme and channel code, block error probability is computed. Block error probability is not a performance metric, but a constraint. 0.1 or 1% block error probability is acceptable. It is desired to squeeze in as much as possible into one block. Real systems have long blocks. Theory developed for measuring how much good performance can be achieved in a long block is called channel capacity. Thus, channel capacity is the highest number of bits per symbol transmitted without error for infinitely long blocks. Channel capacity takes different form depending on the type of channel. For basic communication channel, $y = \sqrt{\beta}x + n$, where x is energy per symbol, i.e. q power, the channel capacity C is given as, $C = \log_2\left(1 + \frac{q\beta}{N_o}\right)$, where, q is energy per symbol, β is channel gain, and N_o is noise variance. Channel capacity is number of bits per symbol that can be transmitted without error. Therefore, channel capacity is the performance metric for analysing communication systems. For a pass band channel, having carrier frequency f and bandwidth B, with $B \ll f$, continuous time signal is represented as $2B$ samples per second. Real signals will be real valued, i.e. 2B real samples per second. However, communication theory is not dependent on which frequency is being used. Therefore, if bandwidth is moved to the centre around zero, then it is baseband channel. The bandwidth decreases as bandwidth is now counted between 0 and largest frequency, i.e. $B/2$. However, the signal is now complex valued. As the real-valued signals have to be symmetric around zero, in frequency domain, which is not the case here. So, we have signal with complex samples per second. Traditionally, sampling is a way of representing continuous time signal in discrete time. However, in digital communication, it is opposite, as samples are called modulation symbols where information is put and based on that continuous time signals are generated, with the allowed usable bandwidth. Signal is sent over channel, and samples are taken at receiver and that is where we get back discrete time and use all such models in digital communication. For basic communication channel, $y = \sqrt{\beta}x + n$, channel capacity is given by $C = \log_2\left(1 + \frac{q\beta}{N_o}\right)$ bits per symbols. Here, the number of symbols is B symbols per second. Therefore, channel capacity

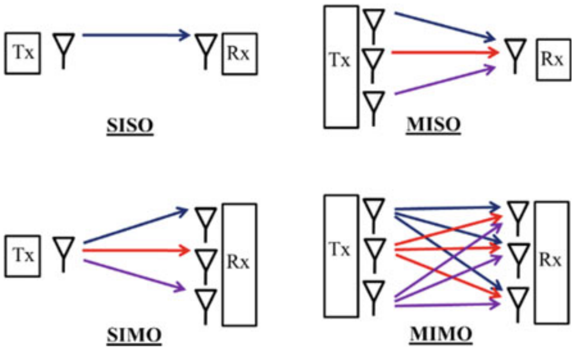


Fig. 33.3 Diagrammatical representation of the techniques

expression becomes $C = B\log_2\left(1 + \frac{q\beta}{N_o}\right)$ bits per second. Here, $q = \frac{P}{B}$, P is power, and B is number of symbols per second. Therefore, channel capacity expression becomes $C = B\log_2\left(1 + \frac{P\beta}{BN_o}\right)$ bits per second, where $\frac{P\beta}{BN_o}$ is signal-to-noise ratio (SNR). Single input single output (SISO), single input multiple output (SIMO), multiple input single output (MISO) and multiple input multiple output (MIMO) are basic forms of multiple antenna techniques. Diagrammatical representation of these techniques is shown in Fig. 33.3.

33.3.1 Single Input Single Output (SISO) and Single Input Multiple Output (SIMO)

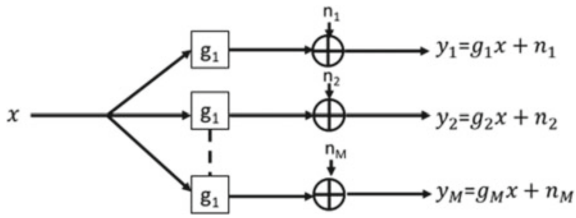
When x is energy per symbol, g is channel response and β is channel gain, i.e. $\sqrt{\beta} = g$ or $\beta = g^2$. For SISO technique, the communication set-up has single input antenna to transmit and single output antenna to receive. Channel capacity is given as $C = \log_2\left(1 + \frac{q|g|^2}{N_o}\right)$ bits per symbol.

33.3.2 Single Input Multiple Output (SIMO)

For SIMO technique, the communication set-up has single antenna to transmit and M antennas to receive. The set-up is shown in Fig. 33.4.

In vector form, $y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_M \end{bmatrix}$ $g = \begin{bmatrix} g_1 \\ g_2 \\ \vdots \\ g_M \end{bmatrix}$ $n = \begin{bmatrix} n_1 \\ n_2 \\ \vdots \\ n_M \end{bmatrix}$. Therefore, received signal vector can be written as $y = gx + n$, where x is scalar information signal. Geometrically, it can be shown in Fig. 34.5a and b. All information is in same direction

Fig. 33.4 SIMO



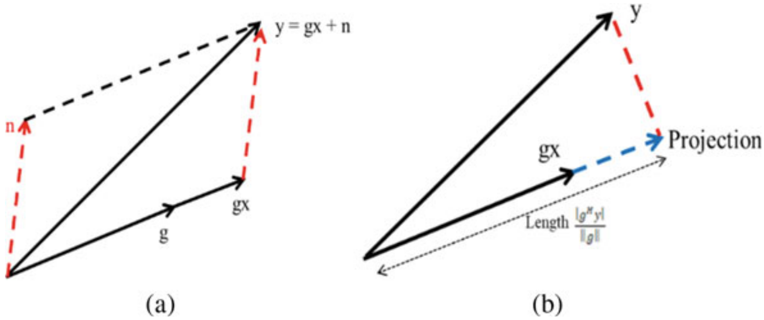


Fig. 33.5 Received signal vector

as g in vector space, and anything pointing in other directions will be noise. Therefore, at the receiver side, y is projected back to the direction of g that means the received signal vector is projected on to the channel vector. In order to do that mathematically, length one version of channel vector g called $v = \frac{g}{g}$ is created, where v is combining vector and g is norm of g. Inner product $v^H y$ is taken, in order to know how long the projected version y be when projected on g. This projection is receive combining. This type of receive combining when projecting it back/down to channel vector is maximum ratio combining. Because the SNR is maximized in this way, we get all the received signal power left but get only some of the noise power (if noise is in same direction). As rest is pointing in different directions, these disappear in projection. In the product $v^H y$, the combining vector is v^H. The received signal is effectively taken and projected to the scalar. Therefore, one value is received, as if there is one antenna. But it is the best combination of different antennas in reality. For scalar received signal $v^H y$, we know how to deal with system with one receive antenna $v^H y = x + v^H n$, where g is a constant (earlier there was channel response and now its combination of all of them) and n is noise with variance. Channel capacity of SIMO system is given by $C = \log_2\left(1 + \frac{qg^2}{N_o}\right)\frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$ bits per symbol. As compared to SISO channel capacity, it is found that SISO channel capacity is $C = \log_2\left(1 + \frac{q|g|^2}{N_o}\right)$ bits per symbol, where $|g|^2$ is channel response of one receive antenna, whereas in SIMO channel capacity expression of $C = \log_2\left(1 + \frac{qg^2}{N_o}\right)$ bits per symbol, g^2 is squared norm of channel vector, i.e. sum of absolute values of square of channel responses for each of M antennas $g^2 = \sum_{m=1}^{M} |g_m|^2$, if the channel responses are same, we get M times strong signal (beamforming gain) $\frac{|g^H y|}{g}$.

33.3.3 Multiple Input Single Output (MISO)

For MISO, the communication set-up has M transmit antenna and one receive antenna, as shown in Fig. 33.6.

It is convenient to describe channels using vectors transmit signal vector *x*, channel vector *g* and noise vector *n*.

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_M \end{bmatrix} \quad g = \begin{bmatrix} g_1 \\ g_2 \\ \vdots \\ g_M \end{bmatrix} \quad n = \begin{bmatrix} n_1 \\ n_2 \\ \vdots \\ n_M \end{bmatrix}.$$

$$y = \sum_{m=1}^{M} g_m x_m + n.$$

Therefore, received signal scalar $y = \begin{bmatrix} g_1 & \ldots & g_M \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_M \end{bmatrix} + n$ or $y = g^T x + n$.

We are finding which part of pointed vector *x* points in the direction of channel vector *g* with a conjugate, i.e. $g*$. Everything in x that is pointing in other direction, i.e. orthogonal to $g*$ will disappear, will be projected away. Therefore, there is no point in projecting anything other than the direction of channel vector, i.e. transmit in the direction of channel. Direction of channel vector has nothing to do with physical direction. Thus, transmit with precoding vector. Transmit vector *x* and multiply with precoding vector *w*. This is called maximum ratio transmission (MRT), which maximizes the SNR. Here, $x = w\tilde{x}$, where w is precoding vector and $\tilde{x}$ is information signal. The precoding vector points to the direction of channel (normalized version of channel) $= \frac{g^*}{g}$, length of one vector pointing in same direction as g conjugate. With MRT, $g^T w = g$. Therefore, $y = g^T w\tilde{x} + n$ or $y = g\tilde{x} + n$. Channel capacity

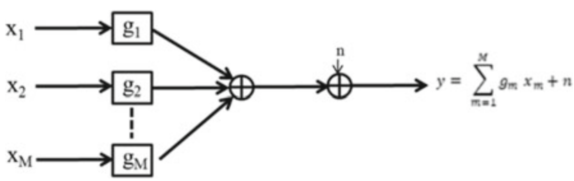


Fig. 33.6 MISO

of MISO is given by $C = \log_2\left(1 + \frac{qg^2}{N_o}\right)$ bits per symbol. For SIMO and MISO channel capacity, M times larger SNR is achieved, when we are transmitting with M antennas and receiving with M antennas. When we transmit with M antennas, we transmit in directive way using beamforming towards user, i.e. M copies of signals are constructively added at the receiver side without using more power. When we receive with M antennas, we are actively transmitting with one antenna (isotropic), we are observing different copies of signal, with different channel responses, and all copies are added constructively using MRC. M noise terms are not constructively combined. In both transmission and reception using multiple antennas, beamforming gain proportional to M is achieved.

33.3.4 Point-to-Point Multiple Input Multiple Output (MIMO)

Consider K transmit antennas and M receive antennas. Between transmit and receive antennas, we have scalar channel responses from transmit antenna k to receive antenna m, $g_{m,k}$. There are total of MK channel responses to be described. Therefore, it is convenient to put them into matrix, where rows describe receive antenna and columns describe transmit antenna. $G = \begin{bmatrix} g_{1,1} & \cdots & g_{1,K} \\ \vdots & \ddots & \vdots \\ g_{M,1} & \cdots & g_{M,K} \end{bmatrix}$. The received signal at mth antenna is given by $y_m = \sum_{k=1}^{K} g_{m,k}x_k + n_m$, where $g_{m,k}$ is channel response from kth transmit antenna to mth receive antenna, x_k is signal from kth transmit antenna, and n_m is noise at mth receive antenna. Here, $y = Gx + n$ and $y_m = \sum_{k=1}^{K} g_{m,k}x_k + n_m$. In the vector form, $\begin{bmatrix} y_1 \\ \vdots \\ y_M \end{bmatrix} = \begin{bmatrix} g_{1,1} & \cdots & g_{1,K} \\ \vdots & \ddots & \vdots \\ g_{M,1} & \cdots & g_{M,K} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_M \end{bmatrix} + \begin{bmatrix} n_1 \\ \vdots \\ n_M \end{bmatrix}$.

To find the capacity of this MIMO channel, let us consider model given in Fig. 33.7.

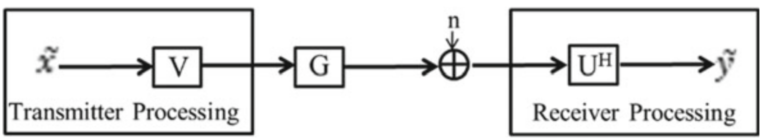


Fig. 33.7 Channel model

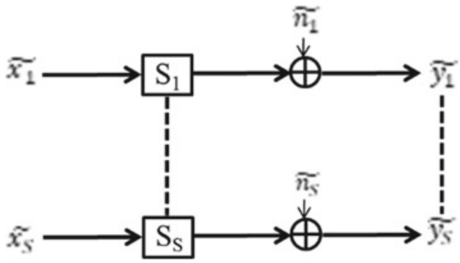


Fig. 33.8 Point to point MIMO as parallel SISO channels

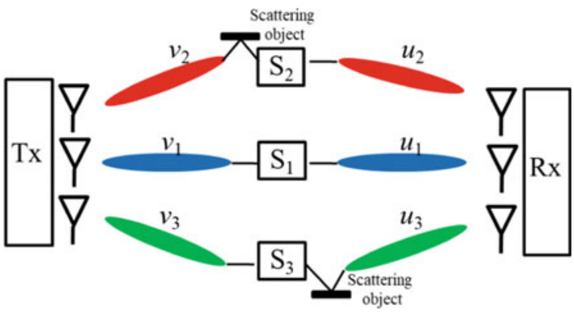


Fig. 33.9 Representation of singular vector decomposition

The transmitted signal is $x = V\tilde{x}$, and receiver processing is $\tilde{y} = U^H y$. Ideally, $\tilde{y}$ should be equal to $\tilde{x}$. We can use V and U by singular value decomposition $G = U\sum V^H$, where U and V are left and right singular vectors for matrix G. If S is the rank of the channel matrix G such that $S \leq \min(M,K)$, the singular values $S_1 \geq \ldots \geq S_{\min(M,K)} \geq 0$. Then, we can parallelize the channel, i.e. by this the point-to-point MIMO is converted into S parallel SISO channels having no interference in between (Fig. 33.8).

The representation of singular vector decomposition for $S = 3$, meaning three transmit antenna and three receive antenna, is shown in Fig. 33.9. The channel is divided into three parallel sub-channels. For each one of them, there is a precoding vector $V = \begin{bmatrix} v_1 & v_2 & v_3 \end{bmatrix}$, which is selected from their matrix V (one of the columns). Combining vector $U = \begin{bmatrix} u_1 & u_2 & u_3 \end{bmatrix}$, is at the receiver side, and the strength of sub-channel is given by singular value S_1, S_2, S_3. Here, S_1 is stronger than S_2 and S_2 is stronger than S_3. Blue is direct path, whereas green and red are following scattered paths. Singular value decomposition creates independent parallel sub-channels. The capacity is known for each of the independent channels. Rate of sub-channel k is $C = \log_2\left(1 + \frac{q_k S_k^2}{N_o}\right)$ bits per symbol, where q_k is transmit power, S_k is singular value, and N_o is noise power spectral density. Sum rate of all sub-channels, i.e. S parallel channels are given by $\sum_{k=1}^{S} \log_2\left(1 + \frac{q_k S_k^2}{N_o}\right)$. We need to maximize it by selecting $q_1, \ldots, q_S$. Total transmit power q and divide it into S parallel channels. Therefore, the capacity of point-to-point MIMO channel is given by $C = \max\limits_{q_1 \geq 0 \ldots q_s \geq 0} \sum_{k=1}^{S} \log_2\left(1 + \frac{q_k S_k^2}{N_o}\right)$ bits per symbol, $\sum_{k=1}^{S} q_k = q$. As per water

Table 33.1 Comparative analysis of antenna techniques

Antenna technique	Channel capacity (bits per symbol)
SISO	$C = \log_2\left(1 + \frac{q\lvert g\rvert^2}{N_o}\right)$
SIMO	$C = \log_2\left(1 + \frac{qg^2}{N_o}\right)$
MISO	$C = \log_2\left(1 + \frac{qg^2}{N_o}\right)$
Point-to-point MIMO	$C = \max_{q_1 \geq 0 \ldots q_s \geq 0} \sum_{k=1}^{S} \log_2\left(1 + \frac{q_k S_k^2}{N_o}\right)$

filling method of power allocation, in case of low SNR, use only one sub-channel, and in case of high SNR, give equal power to all sub-channels.

For SIMO and MISO channel capacity $C = \log_2\left(1 + \frac{qg^2}{N_o}\right)$ bits per symbol, the main benefit is beamforming gain, where SNR grows proportionally with number of antennas. This means a lot for low SNR, but not for high SNR. For MIMO capacity $C = \max_{q_1 \geq 0 \ldots q_s \geq 0} \sum_{k=1}^{S} \log_2\left(1 + \frac{q_k S_k^2}{N_o}\right)$ bits per symbol, the main benefit is multiplexing gain, which is a much larger thing. It is summation over number of channels. If a number of transmit and receive antennas are increased, the capacity is increased linearly.

33.3.5 Comparative Analysis of Multiple Antenna Techniques

After carrying out the analysis of various multiple antenna techniques for provision of cellular communication support for UAVs, it is ascertained that there is an opportunity to improve the communication performance of the UAV communication links, particularly by increasing the number of antennas. In future, simultaneous utilization of hundreds of antennas in 5G technologies such as massive MIMO would be able to provide enhanced communication support to UAVs (Table 33.1).

33.4 Conclusion

This article has surveyed various aspects of provisioning of cellular communication support to UAVs. It has brought out the opportunities, challenges and inadequacies of integrating UAVs with existing cellular networks. By utilizing multiple antennas, these inadequacies of existing cellular networks can be mitigated. After carrying out the performance analysis of basic forms of multiple antenna techniques, it is perceived that 5G technology, particularly massive MIMO, is better suited to deal with the challenges of provisioning of UAV communications. The development of technologies for communication support to UAVs through cellular networks is at a nascent stage, and future research in this field would enable speedy implementation of same.

References

1. Hayat, Yanmaz, Muzaffar: Survey on unmanned aerial vehicle networks for civil applications: a communications viewpoint. IEEE Commun Surv Tutorials **18**(4), 2624–2661 (2016)
2. Gupta, Jain, Vaszkun, Survey of important issues in UAV communication networks. IEEE Commun Surv Tutorials **18**(2), 1123–1152 (2016)
3. Zeng, Y., Zhang, R., Lim, T.: Wireless communications with UAVs: Opportunities and challenges. IEEE Commun. Mag. **54**(5), 36–42 (2016)
4. Akram, Markantonakis, Mayes, Habachi, Sauveron, Steyven, Chaumette, Security, privacy and safety evaluation of dynamic and static fleets of UAVs. IEEE Conf. 1–12 (Sept 2017)
5. Mozaffari, M., Saad, W., Bennis, M., Nam, Y., Debbah, M.: A tutorial on UAVs for wireless networks: Applications, challenges, and open problems. CoRR (2018). [Online]. Available: http://arxiv.org/abs/1803.00680
6. Asadpour, Van den Bergh, Giustiniano, Hummel, Pollin, Plattner, Micro aerial vehicle networks: An experimental analysis of challenges and opportunities. IEEE Commun. Mag. **52**(7), 141–149 (July 2014)
7. Mozaffari, Saad, Bennis, Debbah: Mobile unmanned aerial vehicles (UAVs) for energy-efficient internet of things communications. IEEE Commun. **16**, 7574–7589 (2017)
8. Li, Fei, Dai, Zhang: Secrecy-optimized resource allocation for UAV-assisted relaying networks. IEEE GLOBECOM, UAE (Dec 2018), pp. 1–6
9. Zeng, Lyu, Zhang: Cellular-connected UAV: Potentials, challenges and promising technologies. arXiv preprint arXiv:1804.02217 (2018)
10. Mozaffari, Kasgari, Saad, Bennis, Debbah, Beyond 5G with UAVs: Foundations of a 3D wireless cellular network. IEEE Commun. Early Access. https://ieeexplore.ieee.org
11. Hayat, Yanmaz, Muzaffar, Survey on unmanned aerial vehicle networks for civil applications: A communications viewpoint. IEEE Commun. 18, 2630–2661 (2016)
12. Shi, Zhou, Li, Xu, Zhang, Shen: UAV assisted vehicular networks: Architecture, challenges and opportunities. IEEE Netw. **32**, 130–137 (2018)

Chapter 34
Feature Selection for Chili Leaf Disease Identification Using GLCM Algorithm

Asha Patil and Kalpesh Lad

Abstract In agriculture, detecting and diagnosing leaf diseases are a major concern. Tracking crop fields and identification of symptoms of the disease is important for farmers. Image processing is an aid to the identification and classification of leaf diseases. For leaf disease identification, there are three image features, i.e., texture, color, and shape. Texture features are more important elements of them. The feature selection process is critical to get the best accuracy and minimum time measurement. There are 2500 samples of chili leaves with five diseases in this analysis are train, and 1000 samples are gathered in the research dataset are a test. In this job, using the GLCM algorithm, two classifiers analyze texture features. The extracted texture features and target value are given during training as an input to the SVM and KNN classifier. The texture features. Contrast, energy, correlation, entropy, cluster_shade, cluster_provience, kurtosis, skewness are used for disease identification, respectively. We are labeled as Cercospora leaf spot, chili mosaic, powdery mildew, leaf curl, and healthy leaf in five groups. SVM provides 87.04% accuracy, and KNN provides 94.04% accuracy using the k-fold method

34.1 Introduction

India is a cultivated region, relying on agriculture for about 70% of the population [1]. In choosing different suitable crops and finding the best herbicides and pesticides for plants, farmers have a wide range of variations. Disease on plants contributes to the convincing reduction in both the quality and productivity of agricultural products. In the main horticultural products, chili is included. The demand for chilies is often very high because the supply is limited. Indeed, business chili is part of the high-risk plants. Strategies, technological expertise, and the field have therefore become an important

A. Patil (✉)
S.T.Co.Op., Edu Society Science Sr. College, Shahada, Maharashtra, India

K. Lad
Shrimad Rajchandra Institutes of Management and Computer Application, UTU, Bardoli, India
e-mail: kalpesh.lad@utu.ac.in

T. Senjyu et al. (eds.), *IOT with Smart Systems*, Smart Innovation, Systems and Technologies 251, https://doi.org/10.1007/978-981-16-3945-6_35

issue to master. Systematic and systematic growth should be conducted in such a way that operators can use it to improve overall efficiency. Owing to the rise in chili disease, many farmers refused to grow chili in the rainy season to become a high risk for quality control and productivity. Pathogens caused by fungi, viruses, and bacteria are the cause of many plant diseases. Exact disease finding is a difficult task for farmers which results in loss of income to the farmers and the state [2]. The diseases in chili plants reduce the yield and also go down its extract from the cultivation. Disease detection can make easy the control of disease using proper management way and disease-specific chemical applications can increase production. Experts in the classical method have detected naked eyes of leaf diseases that are very costly for farmers. The research aims to enhance the economic life of the farmer by increasing the production of chili. It is important to recognize the disease at each stage [3]. The disease is primarily responsible for reducing the production of any crop type. The effort and cost of farmers to automatically make the system run and identify the disease are reduced. The goal of the study is to identify early stage remedial solutions for chili leaf disease with following objective:

1. To identify the name of a disease.
2. To select crucial features from the image for the identification of chili leaf disease.
3. To choose the correct classifier with the help of result comparison of accuracy.

Any research work based on the identification of leaf disease is reviewed with some conclusions in the literature. These findings assist in discovering new results for this work. A significant and successful part of finding accuracy is the selection of features from the image [4]. The previous work helps us to find the solution with appropriate features set for a given problem and guides us. The paper is organized into four sections such as two sections give literature review and three sections give methodology. Section 34.4 gives the result and discussion. Section 34.5 gives the conclusion and future scope.

34.2 Literature Review

Thirteen features for texture features are usually described by Harlick. Classifiers are given these measured values and graded as leaf disease [3]. The RGB color image is converted into HIS. Specially detailed information about an affected portion of the leaf is given to the only hue. The function is extracted and entered into the GLCM algorithm from the Hue variable [1]. There are distinct sizes, forms, and disease symptoms of the various crop leaves. The selective features are time-consuming and provide an important outcome for disease detection [4]. The leaf disease identification is more appropriate with texture features of energy, entropy, and contrast, correlation, and homogeneity features of GLCM. The values of the extracted features are entered into classifiers for disease classification [2] Accuracy findings directly affected the sample size and extracted features values in the classification. KNN and

SVM are primarily appropriate classifiers for the identification and classification of leaf diseases [5]. The paper is organized into four sections. Three sections give methodology. Section 34.4 gives the result and discussion. Section 34.5 gives the conclusion and future scope.

34.3 Proposed Method

In this study, feature extraction technique is performed. The texture and color features of the identification and classification method for chili leaf disease (CLDDCS) are introduced here. The first steps are the acquisition of pictures, and then the image is resized to 255 × 255 [6]. Then the image is filtered to reducing the noise or unnecessary pixel from the image. Image is converted into RGB to HSI format. Then focused on a region of interest, i.e., ROI is used for extracting the features [7].

The image has three features: shape, color, and texture. Image texture features obtained by the gray level co-occurrence matrix (GLCM) algorithm calculation. Extracted features are entered in the image classification and disease recognition classifier in classifiers to the specified performance, and the best result classifier is identified. A total of 2500 samples were tested in this analysis, and 1000 samples were tested [8]. All these samples of disease symptoms have been confirmed by agricultural sector experts. Five datasets of chili leaf disease present in the Cercospora leaf spot, chili mosaic, leaf curl, powdery mildew, and healthy leaf data collection were added to the method proposed here. Samples of chili leaf disease are collected from three types of chili varieties, such as the Angara, VNR, and Roshani, from open farms in the Nandurbar district of Maharashtra with different diseases stages, such as early, middle, and final disease stages. These affected leaf disease samples were collected between July and February of last year. A digital camera (Canon EOS REBEL T2i, 20 Megapixels) was used for image acquisition. All of the images are stored in the standard format of JPEG. In general, leaf disease identification consists of image processing and classification using the algorithm of machine learning. An important function in image processing is the extraction and selection of features. The small size of the feature set will help reduce time and help increase the accuracy of machine learning. SVM and KNN classifiers are used for two machine learning algorithm classifiers. Sixteen features are tested using classifiers for accuracy. We have found that eight textures contribute to better classification as a sharp feature, as this is essential for disease identification (Fig. 34.1).

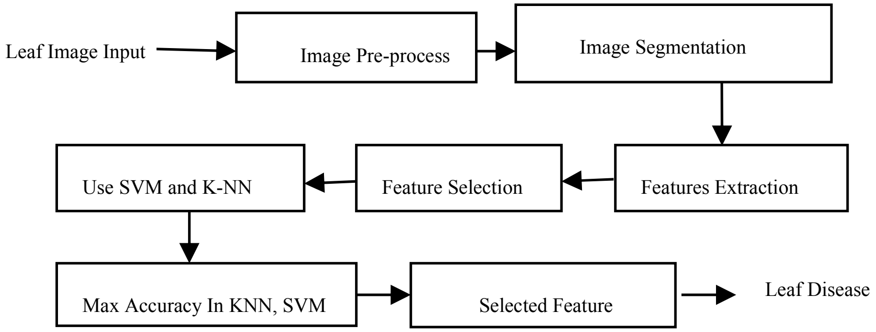


Fig. 34.1 Methodology of proposed system

34.4 Results and Discussion

Detect and classify the chili leaf diseases after extracting color, shape, and texture features from an image. The proposed system was tested on the chili leaf with different viral, bacterial, and fungal disease symptoms. After preprocessing these images, the cluster containing the diseased part of the leaf is chosen using the K-mean clustering algorithm [2, 5]. Features are derived from using GLCM, and this is used for disease classification. One way to strengthen the holdout process is the k-fold cross-validation method. We are now going to divide the dataset into a training dataset and a test dataset. We are going to use 70% of our data for training and the remaining 30% for testing.

Support Vector Machine (SVM)

To identify chili plant disease affecting chili production, a support vector machine is used. We use nonlinear SVM when we cannot isolate data from a straight line. We have kernel functions here. This converts information into low to high dimensions to classify the data. The original optimal hyperplane algorithm in nonlinear SVM proposed a way to construct nonlinear classifiers by applying the maximum margin hyperplane. For multiple classifications, two methods are followed: one-versus-one and one-versus-all. One-versus-all includes splitting N classes with labeled; SVM uses the training dataset to predict the multisvm() function testing dataset. SVM provides more accuracy when texture features are used. To classify chili leaf disease, SVM uses the RBF kernel. In an appropriate feature space, the kernel functions return the inner product between two points [2]. Linear, polynomial, radial basis function (RBF), etc., are kernel functions. In this outcome, we conclude that the accuracy of SVM using ‘rbf’ (radial basis function) kernel, other output measurement parameters such as precision, recall, and F1 measure are measured at 120 iterations with twofold as shown in Table 34.1.

For this analysis, the parameter tuning parameter kernel function and iteration will be checked with the *k*-fold process. We got this SVM result with accuracy, recall, and F1 values after some experiments.

Table 34.1 Performance measurement of SVM

Iteration	Accuracy	Precision	Recall	F1_Measure
30	79.85	0.82	0.59	0.58
50	87.12	0.86	0.74	0.73
70	90.41	0.88	0.8	0.8
90	93.91	0.91	0.87	0.87
100	92.12	0.89	0.84	0.84
120	93.04	0.94	0.92	0.91

K-Nearest Neighbor (KNN)

A semi-supervised learning algorithm is a k-nearest neighbor so that training data and a predefined k-value are needed to find the k-nearest data based on distance computation [5]. If k data has different classes, the class of unknown data is predicted to be the same as the majority class by the algorithm. Using the Euclidean distance, the closest class will be defined. This algorithm uses information from the neighbor point to predict the target class. For this algorithm, the training process involves storing feature vectors and marking the training images [9]. The unlabeled point is simply assigned to the label of its k-nearest neighbors in the classification process. Usually, chili leaf disease is categorized by the maximum number of votes based on the labels of its k-nearest neighbors. By contrasting features from the image and the dataset, classification is carried out (Table 34.2).

The table shows that with precision, recall, and F1 measurements, accuracy is also improved as the values of k increase. If the value of $k = 1$, then the prediction is less stable, but if we increase the value of k, then prediction becomes more accurate due to the majority of voting [10]. Then we get an accuracy of 87.04% of the $k = 12$, which is higher values than others. The graph shows the result comparison of SVM and KNN (Fig. 34.2).

In machine learning algorithms, SVM gives 93.04% accuracy and KNN gives 87.04% accuracy as shown in the above graph. If we take the dataset from high-dimensional online platform images, then the result will be more suitable. The distance between the camera and the leaf has a greater effect on image quality with less image noise.

Table 34.2 Performance measurement of KNN

KNN	Accuracy	Precision	Recall	F1_Measure
3	79.85	0.76	0.74	0.68
5	80.12	0.80	0.74	0.73
7	80.41	0.88	0.78	0.76
9	83.91	0.81	0.80	0.82
10	82.12	0.81	0.79	0.76
12	87.04	0.8	0.79	0.78

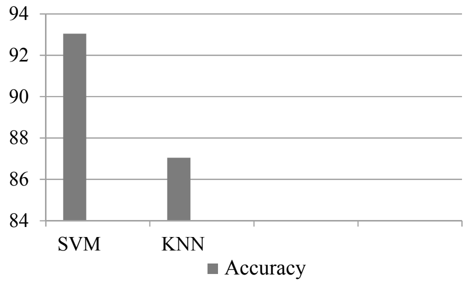


Fig. 34.2 SVM and KNN result comparison

34.5 Conclusion and Future Scope

The proposed paper compares the two machine learning algorithm KNN and SVM classifiers to define, identify and recognize five chili leaf diseases, namely Cercospora leaf spot, chili mosaic, leaf curl, healthy leaf, and powdery mildew images. The suggested solution was compared based on both classifiers' performance analysis. Eight GLCM features extracted from the chili leaf of the 'H' portion of the affected leaf region. The texture features. contrast, energy, correlation, entropy, cluster_shade, cluster_provience, kurtosis, skewness are used for disease identification, respectively SVM offers 93.04% accuracy and KNN gives 87.04% accuracy. In the proposed system, the size of the image is changed then the effect on the result. In a natural disease scenario, the chili leaf may be single or multiple leaves, also affecting the extraction and selection of texture features. The limitation of the proposed model is moisture variation, distance variation from camera to leaf, heavy rain affect the brightness of the image and contrast will directly affect the accuracy result of finding the name of the disease. The image noise makes reason for lower image classification accuracy. In the future, early stage disease identification provides farmers with a remedial solution regarding pesticide. The model consistency can also be implemented by the large data collection with advanced image classification techniques such as CNN.

References

1. Pujari, J.D., Yakkundimath, R., Byadgi, A.S.: Automatic fungal disease detection based on wavelet feature extraction and PCA analysis in commercial crops. Byadgi I.J. Image, Graph. Sig. Process. (2014)
2. Badugu1, S.K., Kontham, R.K., Vakulabharanam, V.K., Prajna, B.: Calculation of texture features for polluted leaves. Int. J. Sci. Res. Res. Paper. Comput. Sci. Eng. **6**(1), 11–21 (Feb, 2018)
3. Harlick, R.M., Shanmugam, K., Dinstein, I.: Texture feature for image classifications. Trans. Syst. Man Cybern. IEEE 610–621 (Nov, 1973)
4. Pujari, J.D., Yakkundimath, R., Byadgi, A.S.: Recognition and classification of produce affected by identically looking powdery mildew disease. Acta Technol. Agric. (Nitra, Slovaca Universitas Agriculturae Nitriae) 29–34 (2014)

5. Thakur, H., Jindal, S.K., Sharma, A., Dhaliwal, M.S.: Chilli leaf curl virus disease: a serious threat for chilli cultivation. J. Plant Dis. Prot. (2018)
6. Baghel, J., Jain, P.: Disease detection in soya bean using K-means clustering segmentation technique. Int. J. Comput. Appl. (0975–8887) **145**(9) (July, 2016)
7. Anand, R., Veni, S., Aravinth, J.: An application of image processing techniques for detection of diseases on brinjal leaves using K-means clustering method. In: Fifth International Conference on Recent Trends in Information Technology 978-1-4673-9802-2/16/$31.00c IEEE-(2016)
8. Dixit, S., Nema, S.: Wheat leaf disease detection using machine learning method-a review. Int. J. Comput. Sci. Mobile Comput. **7**(5), 124–129. ISSN 2320-088X (2018)
9. Pathanjali, C., Salis, V.E., Jalaja, G., Latha, A.: A comparative study of Indian food image classification using K-nearest-neighbor and support-vector-machines. Int. J. Eng. Technol. **7**(3.12) 521–525 (2018)
10. Saadatfar, H., Khosravi, S., Joloudari, J.H., Mosavi, A., Shamshirband, S.: A new K-nearest neighbors classifier for big database on efficient data pruning. www.mdpi.com/journal/mathematics (2020)

Chapter 35
"Faculty eCourseBook": A Digitized Faculty Course File as a Green Campus Initiative

Pratibha S. Yalagi, P. S. R. Patnaik, and Shasikant A. Halkude

Abstract Record keeping is an essential part of an academic activity for a teacher for the various quality initiatives. The course file is a document for keeping the record of academic activities of a faculty and student's performance continuously. To ascertain the overall standard of academic process, the efficacy of course delivery and timely maintenance of assessment records is very important. These records contain a teaching plan and its execution, assessment, and other allied co-curricular and extracurricular activities. Earlier faculties were maintaining the records manually. A lot of time was required to be spent for maintaining records, which in turn used to affect their performance and productivity as a teacher. As a first step towards the betterment, faculty members used to maintain records in the form of hard copy which was written manually on a printed form. This form of record keeping had a variety of disadvantages: involving physical storage, inefficient information retrieval, and repeated paper and printing costs. To save time and as a green initiative, Institute has developed a digitized product namely "Faculty eCourseBook" for compiling, organizing, and monitoring academic activities, which is implemented through G Suite: Google Sheets. Due to which record keeping, accessing and managing has become easier, while ensuring uniformity in the record keeping. G Suite's sharing and the relevant access permission have made performance monitoring instant, simple and easy. This practice of using "Faculty eCourseBook" at our institute is recognized as the Best Practice by AICTE New Delhi.

35.1 Introduction

For an academic setup, a course file is essentially a document that includes all the necessary details regarding the delivery of the course and students continuous assessment. Course file generally includes information like the student details, course information, assignments, assessment metric and course outcomes, etc. A course file becomes an integral part for academic and administrative audit. Outcome-Based

P. S. Yalagi (✉) · P. S. R. Patnaik · S. A. Halkude
Walchand Institute of Technology, Solapur, Maharashtra, India

T. Senjyu et al. (eds.), *IOT with Smart Systems*, Smart Innovation, Systems and Technologies 251, https://doi.org/10.1007/978-981-16-3945-6_36

Education focuses on assessment and its process containing identifying, collecting, and preparing data to evaluate the attainment of course outcomes (Cos) and program outcomes (Pos). This requirement is being facilitated by the course file keeping record of all student activities and faculty academic activities as documented evidence of teaching and learning (T-L). It supports faculty to review and reiterate the T-L process. A course file is not only a collection of curriculum delivery, assignments, and evidence of students' performance, but also includes an agenda for pedagogy, assessment methods and techniques in order to improve the T-L process.

As a green campus initiative, the institute has come up with a prototype "Faculty eCourseBook" management framework for compiling and organizing the course files utilizing Google Sheets [1, 2]. It is an in-house customized digital product leveraging Google Docs and Apps, to streamline the process for better governance through standardization for Outcome-Based Education (OBE). It is a time saving and efficient tool for faculties. Semester wise "Faculty eCourseBook" for each faculty is generated in PDF form. It is a web-based Google entry sheet which is shared through individual faculty mail id.

35.2 Related Work

During a study in Nigeria [3] that surveyed the problems associated with records keeping at universities, it was found that ineffective means of retrieving records, inadequate computer terminals, improper security of records, lack of record keeping policy, inadequate resources, and a lack of record retention and disposition cycle were some of the many problems encountered for maintaining records. To solve a few of these issues, digitization of all records and training the staff in this regard was suggested as a course of action.

Program assessment has gradually become a predominant part of the accreditation procedure for every institution. All organizations need to establish targets and objectives before determining the measures based on outcome to facilitate fulfilling the set goals of the organization [4]. Assessment methodologies have also evolved from manual work on paper toward automation of the process with the help of computational tools, thus making the process faster, efficient, and convenient.

But even with the best computers and tools, it is not guaranteed that it will be beneficial. To make the most of the technology and resources, it is important to follow best practices to keep track of records and information in an efficient manner. Toward the end of twentieth century, technological and social advancements have challenged traditional record keeping methods, either making them insufficient or obsolete [5]. Considering the importance of well-based records, best practices are essential for efficiency.

The current shortage of computer science faculty at institutions has reduced the time available for creating and maintaining an effective tool for assessment. For an assessment method to be successful, a variety of carefully chosen and well-timed assessment instruments need to be developed so as to be effective while also not burdening the faculty. A good assessment also requires a process model that provides faculty support for the process [6].

35.3 About "Faculty eCourseBook"

35.3.1 Timeline

The need for a solution like the "Faculty eCourseBook" is well-founded with an objective of consistent record keeping. Initially, the course material and records were very subjective, and every faculty had their own way of keeping records. This had several disadvantages like the lack of standards for maintaining records and difficulty aggregating and evaluating the records and compiling the results. To address these issues and to improve assessment, the idea of a generalized, standard printed course book in hardcopy was put forth in the year 2015–16. While this solved some issues like the lack of standard for record keeping and found a way to generalize the results, it came with other challenges such as entry of data into the book by hand, manual calculations of results, and preserving its condition among other things. This was very tedious and time-consuming, taking up a significant portion of the faculty's time and patience. Accessing the data and borrowing it from faculty by making another copy of the record has made it difficult.

Many of the problems arising due to the printed version of hardcopy of the course book could be solved by just digitizing the records and using a soft version of the course book. This would mean files could be shared, and additions were easier to make. Hence in the year 2016–17, our next course of action was to make a Microsoft Word version of the same template and enter data into the file. This came with many advantages as it saved printing costs, prevented loss of records, and accidental damage made it easier to share and consolidate records, and it was easier to make corrections and implement changes in the assessment process quickly and effortlessly. All in all, this was a more versatile solution and definitely a step up from the printed version of the course book.

Later in the year 2017–18, the Google Docs version has made it accessible through mails for increasing privacy, collaborative accessibility, and multi-device synchronization.

While the Google docs version of the course book had made record keeping much more efficient than before, it was still not an ideal solution for record keeping as it lacked many data handling features such as calculating sums and averages, conditional formatting, data validation, and so on. All of these could be solved with the use of spreadsheets. Hence, in the year 2018–19, the Google Spreadsheet version of "Faculty eCourseBook" is implemented.

In the year 2019–20, the "Faculty eCourseBook" is enhanced by adding more features in the attendance module for easy attendance monitoring and time saving in attendance tanking in the class and its recording.

35.3.2 *Objectives and Design*

Our digital product “Faculty eCourseBook” for maintaining record of faculty course file has the following objectives:

- Bringing uniformity in maintaining academic records
- Digitizing the faculty academic records to make documentation easier and simpler for continuous improvement
- Developing a sustainable solution for paper and time saving
- Managing, monitoring, and utilizing the data effectively for continuous improvement.

The “Faculty eCourseBook” is in modular form containing: faculty performance, academic schedule, assessment, continuous improvement, personal details of a faculty member, and other academic activity details. These modules allow faculty members to keep track of the attainment of program outcomes (POs) and course outcomes (COs).

Records pertaining to continuous evaluation of course delivery, records of corrective measures taken to improve attainment of COs based on student’s feedback, and student’s performance are also maintained in the “Faculty eCourseBook.” It comprises of five modules, and the details are spread over with sub-modules as shown in Fig. 35.1.

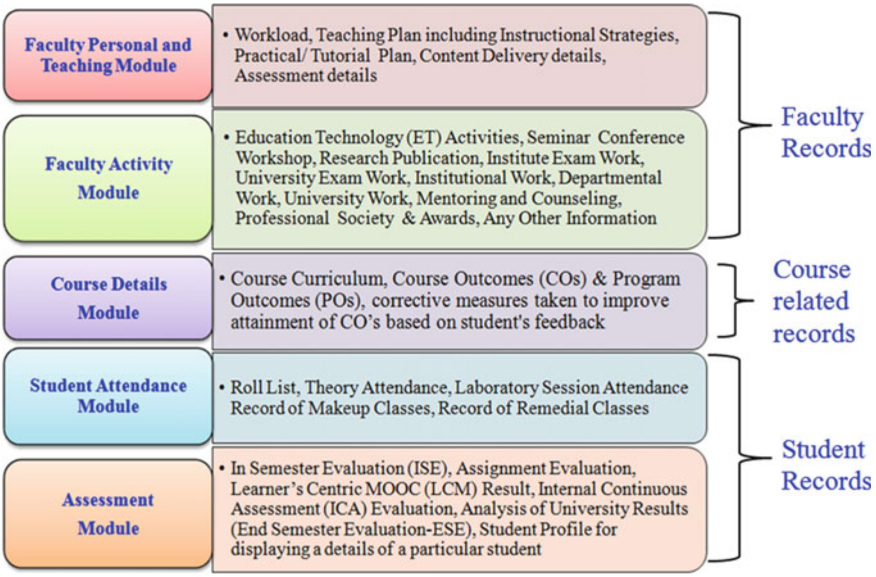


Fig. 35.1 Implementation modules of “Faculty eCourseBook”

35.4 Implementation

35.4.1 Product Development

The "Faculty eCourseBook" has been implemented using the services offered by Google's G Suite for Education. The implementation relies on collaborative sharing and access controls inherent to the G Suite's organizational structure. The "Faculty eCourseBook" uses Google spreadsheets and Google Drive with custom written Google Apps Script code for custom functionalities not available in Google spreadsheet. G Suite ensures multi-device access with real-time synchronization of the data across multi-devices, and this makes it good choice for academic record keeping purposes [7, 8].

35.4.2 Distribution of "Faculty eCourseBook"

The distribution structure of "Faculty eCourseBook" is shown in Fig. 35.2. At commencement of every semester, coordinator creates a copy of the "Faculty eCourseBook" for every faculty member for editing assigns "right to view" to Principal, Heads of the Departments, Departmental coordinator's. This "Faculty eCourseBook" is distributed to the faculties individual mail ids, and they can access it from anywhere through their Gmail ids. It is visible to only concerned faculty members, Head of Department, Departmental Coordinator, Principal and "Faculty eCourseBook" Coordinator. Modifications are done before distributing to faculty by considering all suggestions and grievances from faculty users and Departmental coordinators.

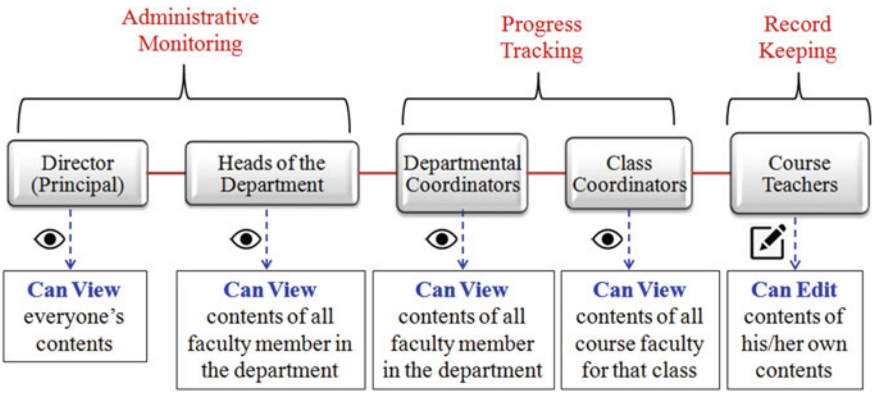


Fig. 35.2 Distribution of "Faculty eCourseBook"

35.4.3 Usage of "Faculty eCourseBook"

Faculty is maintaining all the records of their academic activities in the "Faculty eCourseBook" during the semester as a record keeping of academic activities. As it is accessible through their Gmail id, they can access it through all the compatible devices for editing. Daily record of student attendance, content delivery and assessment records are maintained. Other academic activity details are also recorded as and when required by the faculty. They can also record their personal records like research publication publications, university work, training received, etc., through this.

35.4.4 Activity Monitoring

Various academic activities take place throughout every semester. These activities need to monitor for ensuring uniformity and adhering to the schedule without compromising with the quality. Director (Principal) of the Institute is monitoring all faculties by viewing "Faculty eCourseBook." Heads of the departments are monitoring the respective faculty activity, and Class Coordinators and Departmental coordinators are tracking the progress by taking follow-ups of timely activity by viewing their respective class and department "Faculty eCourseBook."

As the administrators can view the academic activity of faculty through "Faculty eCourseBook," they can monitor faculty activity by accessing it from time to time. This timely monitoring makes the faculty to work timely, effectively, and honestly. This facilitates a proper monitoring of various activities and academic work. Student's continuous improvement is also monitored by the faculty as and when required. It includes student's attendance, laboratory attendance, in semester evaluation (ISE) and internal continuous assessment (ICA), and other assignment submission details. A separate student profile sheet is prepared for viewing it. Attendance module shows the monthly record in percentage.

It serves the objective of maintaining the various records which enables compiling the information and archival of the same for deriving the various matrices by the institution. This assists to submit information to various authorities as and when required, leading to green engineering and paperless models. The "Faculty eCourseBook" is being incrementally extended to improve usability experience and ease of use [9].

35.5 Feedback Analysis

A feedback on usage of "Faculty eCourseBook" since its inception has been provided by nearly all faculty members of the institute. Primary objective of feedback is to arrive at the various matrices from user's perspective and accordingly to improve the

“Faculty eCourseBook” after a stipulated time duration and to know the perception of the faculty members toward the various aspects of “Faculty eCourseBook” and its usability. Also, the feedback of the administration from the monitoring perspective is obtained.

For the feedback question on ease of using the “Faculty eCourseBook” in comparison with hardcopy, more than 50% faculty members found that using “Faculty eCourseBook” is much simpler to maintain and easy to use w.r.t. hardcopy for record keeping. Regarding time saving achieved 50% respondents reported more than 45% of time savings, whereas others found time saving up to the extent of 30%.

Regarding achievement toward digitization, more than 58% faculty reported that digitization achieved is more than 70%. Also, more than 90% of faculties agree to the fact that “Faculty eCourseBook” helped them to keep track of student’s continuous progress. Figures 35.3 and 35.4 show these responses.

Additionally, feedback by administrators of the institute and department showed improved ease of timely monitoring and increase in productivity of teachers by the virtue of reduced efforts and time spent in academic record up keeping. Responses given by the administrators are shown in Fig. 35.5.

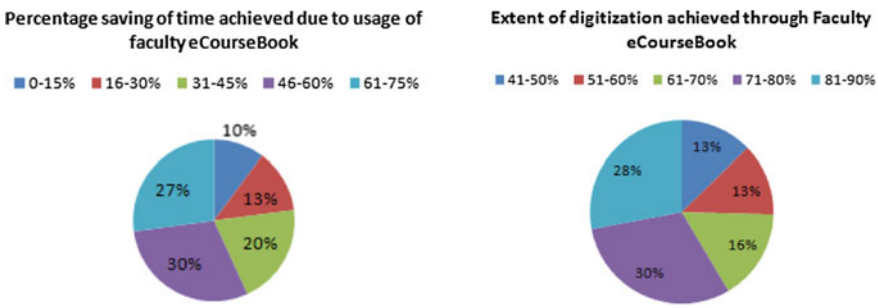
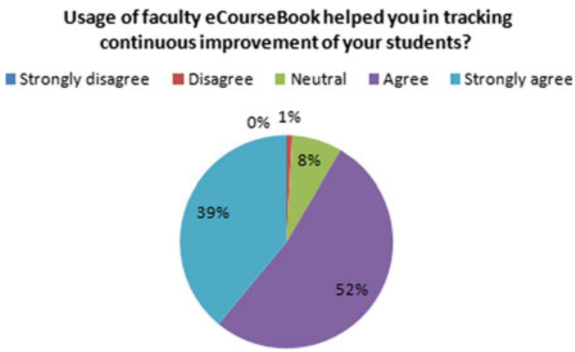


Fig. 35.3 Feedback on time savings and digitization aspect of “Faculty eCourseBook”

Fig. 35.4 Feedback on tracking student’s progress aspect of “Faculty eCourseBook”



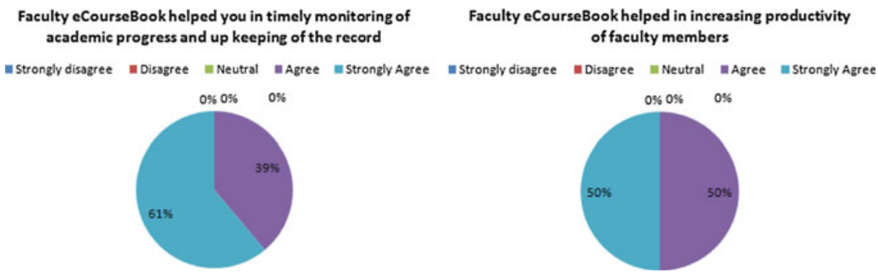


Fig. 35.5 Administrator's feedback on "Faculty eCourseBook"

35.6 Recognition

The "Faculty eCourseBook" is recognized by All India Council for Technical Education (AICTE) as a Best Practice and is published by AICTE (https://www.aicte-india.org/flipbook/Best_Practices/#p=6).

35.7 Conclusion

The "Faculty eCourseBook" initiative has been well received by the faculty of the institute and assisted administrators in timely monitoring and data compilation for quality initiatives, the survey conducted also reflects the same. Some of the major achievements of the "Faculty eCourseBook" are as follows:

- Both faculty and administrator found that it was possible to bring uniformity in record keeping due to "Faculty eCourseBook."

From the Faculty perspective

- Time saving achieved in record keeping due to "Faculty eCourseBook" is @ 50%.
- Maintenance and retrieval of records have become much simpler.
- Monitoring of students continuous improvement, attendance, and assessment is instantaneously available.

From the Administration Perspective

- Data is available at all the time and easily accessible.
- Monitoring faculty performance has become easier, resulting in to the improvement of performance of faculty member.

This digital product, namely "Faculty eCourseBook," developed using G Suite which is powered by green computing [10] has proved to be a green and sustainable product as it is saving paper worth 14 trees and earns carbon credit of value 0.386 annually.

In the future version of this product, it is planned to incorporate OBE-based course outcome attainment and program outcome attainment calculations.

References

1. Hooi, K.K., Hassan, P., Jami, N.A.: Sustainable education: an assessment of carbon footprint at UCSI university and proposed green campus initiative framework. In: 3rd International Conference on Information and Financial Engineering, Singapore, pp. 342–347 (2011)
2. Oracle Content Management: Go Green With Oracle Content Management [Online]. Available: http://www.oracle.com/us/products/middleware/contentmanagement/059500.pdf (20 Nov 2012) (2009)
3. Egwunyenga, E.J.: Record keeping in universities: associated problems and management options in South West geo-political zone of Nigeria. Int. J. Educ. Sci. **1**(2), 109–113 (2009)
4. Maxim, B.R.: Closing the loop: assessment and accreditation. J. Comput. Sci. Coll. **20**(1), 7–18 (Oct 2004)
5. Schreier, A.A. et al.: Academic research record-keeping: best practices for individuals, group leaders, and institutions. Acad. Med. J. Assoc. Am. Med. Coll. **81**(1), 42–7 (2006)
6. Hogan, T.R., Harrison, P.R., Schulze, K.G.: Developing and maintaining an effective assessment program. SIGCSE Bull. **34**(4), 52–56 (Dec 2002)
7. Kakoulli-Constantinou, E.: Teaching in clouds: Using the G Suite for Education for the delivery of two EAP courses. J. Teach. English Specif. Acad. Purp. (2018).
8. Cohen, S.G., Bailey, D.E.: What makes teams work: Group effectiveness research from the shop floor to the executive suite. J. Manag. **23**(3), 239–290 (1997)
9. Sharp, L.: Green campuses: The road from little victories to systematic transformation. Int. J. Sustain. High. Educ. **3**(2), 128–145 (2002)
10. Google Cloud Environment Go Green https://cloud.google.com/sustainability

Chapter 36
Performance Evaluation of Novel Moth Flame Optimization (MFO) Technique for AGC of Hydro System

Shamik Chatterjee and Ahmed Nura Mohammed

Abstract To enhance the infrastructures of the future, the complexities associated with the power system need to be addressed continuously to keep up with the trajectory of the ever emergent challenges associated with power system. Since the crucial role played by AGC in this regard is tremendous, then its study is indispensable. Therefore, in this study, an equal three-area, PID-controlled hydropower system optimized by the novel moth flame optimization (MFO) that is known to give propitious and competitive outcomes is considered for AGC investigation. The control mode used is adaptive, comparing FOD objective function to three other performance indices to ensure consistency of operation. The dynamic performance realized from the first two case scenarios was compared with the most popular stochastic optimization algorithms that is known for providing high-quality solution to optimization problems in the AGC literature, and the transformative results obtained clearly point out the dominance of MFO algorithm over genetic algorithm (GA). However, the slight shift recorded in the third case scenario demonstrates that the proposed control approach is flexible to parameter uncertainty. Thus, the proposed algorithm can be considered as an efficient optimizing tool over a wide range of AGC application.

36.1 Introduction

36.1.1 Background

Electric power system is a network of electrical components employed primarily to generate and consume electrical power. Hence, the remaining components that manifest between these two are mainly for the purpose of conveyance, monitoring or control. The power generated must be in sufficient quantity, quality and at minimum cost possible. Generally, appropriate operation of power system demands that the power generated must be the same to the power consumed or load demand; this is from the point of view of stability, reliability, efficiency and cost effectiveness of

S. Chatterjee · A. N. Mohammed (✉)
School of Electronics and Electrical Engineering, Lovely Professional University, Punjab, India

T. Senjyu et al. (eds.), *IOT with Smart Systems*, Smart Innovation, Systems and Technologies 251, https://doi.org/10.1007/978-981-16-3945-6_37

the system. However, load is stochastic in nature, it keeps on changing all the time, and whenever load changes, generation is designed to make an effort to compensate such changes in order to avoid or minimize the consequences of the changes to ensure quality, reliability and cost effectiveness. Quality here indicates the ability of power systems particularly generating units to constantly maintain both frequency and voltage within specified value [1]. The frequency in the power system network is largely associated with real power while voltage is mainly associated with reactive power; hence frequency control influences the real power and voltage control influences the active power in the network. These controls are carried out separately as real and reactive powers in the network are largely independent of each other [2].

According to [3], the frequency control is carried out primarily by load frequency control (LFC) which depends on governor speed control, and this generally makes the power system interconnection possible. However, LFC operation is slow and often associated with steady-state errors; hence a secondary loop insensitive to rapid load and frequency changes, mostly with controller that automatically detects and fine-tunes the steady-state error quickly recognized as automatic generation control (AGC) is added to the system. Even though, all through the literature, AGC and LFC are used interchangeably but one may see that the aim of AGC is to constantly allocate the load to the various areas of the system and correctly control the area error (ACE) to steady-state value.

However, the ever-growing demand of power, the stringent requirement of reliability and complexities of power system together with many other changes that are taking place overtime such as advancement from regulated to deregulated environment and hybrid scheme to mention are few that make AGC study necessary to explore and accommodate the changes for innovation of new ideas, schemes, methods and/or strategies of dealing with AGC problems so as to make the infrastructure of future more flexible and smarter. Thus, since time in memorial, a wide range of AGC studies with different assumptions and considerations are well documented in the literature.

36.1.2 Literature Review

AGC functions had progressed swiftly over the years from the time when the task was performed manually, and through the era of simple analog control to the years of advanced digital control scheme, the pioneering work on AGC proposed by Edgerd and Fosha in 1970 sets the ground for modeling, design and optimization of interconnected power system. Subsequently, AGC studies that discussed the techniques and idea of modeling the elements of power systems, highlighting their essential features, assessing the dynamic behavior of the models and extensively discussing the outstanding role played by AGC, the doubt of its aptness in highly interconnected areas and the vital considerations that are needed for achieving the finest control were carryout. Also the main differences between thermal and hydro systems were highlighted, for instance it was identified that for a step load change in the hydro area, the maximum transient response deviation in tie-line power and the settling time are considerably greater than for a similar change in the thermal area [4]. However, the main ideas behind AGC of hydropower system such as generation control modes and its operations, hydro allocator operation and its similarities to thermal systems were highlighted in [5]. Furthermore, the characteristics and properties of cascaded hydro plants together with its principles, approaches, processes and operational experiences were fully discussed in [6], whereas multi-area hydro–hydro system considering integral control single-frequency approach was proposed in [7].

In essence, quite a large number of AGC studies with different strategies and schemes were carried out, thus Pandey et.al. carried out critical literature survey on AGC, highlighting different power system types, controls techniques, strategies, collaboration of distributed generation and the inclusion of energy storage systems and facts devices as a whole part of power systems [8]. Also, Kumar et al. [9] make effort to critically discuss these different strategies and schemes from the time of simple classical control schemes to the adaptive and self-turning schemes and then the most recent intelligent control and optimization schemes, stating their key features. The changes associated with power system recently such as in-cooperation of BES, SMES, wind turbine and AGC in the deregulated environment are also highlighted. Furthermore, similar discussion was carried out by Singh et al. [10] but with greater emphasis on the intelligent control scheme, inferring they are more suitable for nonlinear systems modeling. That is why intelligent techniques particularly nature-inspired optimization techniques mostly in combination with classical control are widely used in AGC literature. Nature-inspired optimization techniques are well discussed in [10] and [11]. However, nature-inspired optimized classical control superiority has been established by the excellent result obtained in [12], where decentralized three-area GA optimized thermal power system considering three level of constraints was investigated, [13] where three unequal hydrothermal systems using intelligent nature-inspired firefly optimized PID controller with generation rate constraints and BES was investigated, [14] which examined interconnected two unequal area hydrothermal power system using craziness-based particle swarm (CRPSO)-optimized integral controller incorporating TCPS and CES [15] that investigate multi-area power system using simulated annealing-based PI controller considering system parametric uncertainties and numerous loads conditions. Other works that further stressed the performance of nature-inspired optimized classical control can be seen in [16] for ACO, [17] for CRPSO, [18] for SA, [19] for ATS, [20] for HSA and [21] for QOHS to mention a few. Therefore, for proper tuning of the controller in this work, an efficient optimization technique, the novel moth flame optimization (MFO) algorithm that is known to give propitious and competitive outcomes is employed [22] and compared with the most popular stochastic optimization algorithm (genetic algorithm, GA).

36.1.3 The Contribution of the Present Study

The major contribution of this study is as follows.

(a) To design three-area interconnected hydropower system model,
(b) To consider MFO algorithm and employ it as an optimization tool of the designed system for enhancing the performance of AGC,
(c) To consider the system transient responses obtained with the MFO algorithm and compare it to that obtained by GA algorithms,
(d) Also to perform sensitivity analysis for valuation of the flexibility and sophistication of the proposed approach.

36.2 System Investigated

36.2.1 Structure of PID Controller

Proportional integral and derivative controller commonly referred to as PID controller is an industrial control instrument that uses feedback control loop to control process variable; it is the most accurate and stable controller with a very simple structure, that is why it is convenient to use in AGC domain. In AGC of power system particularly multi-area system, both frequency and tie-line power deviation should be restored to their set limits, thus the fusion of these two normally refers to area control error (ACE) given by (37.1a) which is used as a driver to the PID controller, while a transfer function representation for the PID controller ($K_{\text{PID}}(s)$) may feasibly be presented in (37.1b).

$$\text{ACE}_n = B_n \Delta f_n + \Delta P_{\text{tie}} \quad n = 1, 2, 3 \tag{37.1a}$$

$$K_{\text{PID}n} = K_{Pn} + \frac{K_{\text{in}}}{s} + \frac{sK_{dn}}{\xi s + 1} \quad n = 1, 2, 3 \tag{37.1b}$$

where K_{pn}, K_{in} and K_{dn} are, respectively, the nth area proportional, integral and derivative gain of the PID controller. It should be noted here that a suitable first-order time constant filter ξ in the order of 0.01 is applied to the derivative path to lessen the shift of the output due to integral action of the controller.

36.2.2 Designed AGC Model

Characteristics of Hydro-generating Units

As the characteristics of hydro turbine defer from that of steam turbine, different dynamic models are required for hydraulic turbine; the characteristics of hydraulic turbine that determine its transfer function are the dynamics of water flow in the penstock; the physical arrangement of hydraulic turbine can be found in [2]. Technically, just like any other power system model, hydropower system models are nonlinear systems, but considering small perturbation about a steady-state operating condition AGC study of hydraulic power system can be carried out by linearly relating the various parameters, as given in (36.2) and (36.3), where q, h, n, g and m are per unit deviation in flow, head, speed, gate position and torque, respectively.

$$q = a_{11}h + a_{12}n + a_{13}g \tag{36.2}$$

$$m = a_{21}h + a_{22}n + a_{23}g \tag{36.3}$$

Design of Three-Area Hydropower System Model

The system under investigation is three equal hydropower-generating systems represented by closed loop model shown in Fig. 36.2, and the system contains the controller (K), speed governor control R_{h}, hydro-generating unit with governor G_{h} that contains hydraulic amplifier and turbine and power system G_{p} components, represented by the transfer function portrayed in the blocks of Fig. 36.1 and the system model parameters reported in Appendix [23].

36.3 Problem Formulation

36.3.1 Objective Function

The objective function employed to obtain the optimal gain of the controller in this study is the figure of demerit (FOD) which uses ITAE benchmarks, and its requirements are settling time, should be small, and oscillations should die out quickly. Hence, to lessen the oscillations, the product of absolute error value and time are obtained, thus FOD can be expressed as (36.4).

$$\mathrm{FOD} = \mathrm{ITAE} = \int_0^{t_{\mathrm{sim}l}} \{|\Delta f_n| + |\Delta P_{\mathrm{tie}nm}|\} t\,\mathrm{d}t \tag{36.4}$$

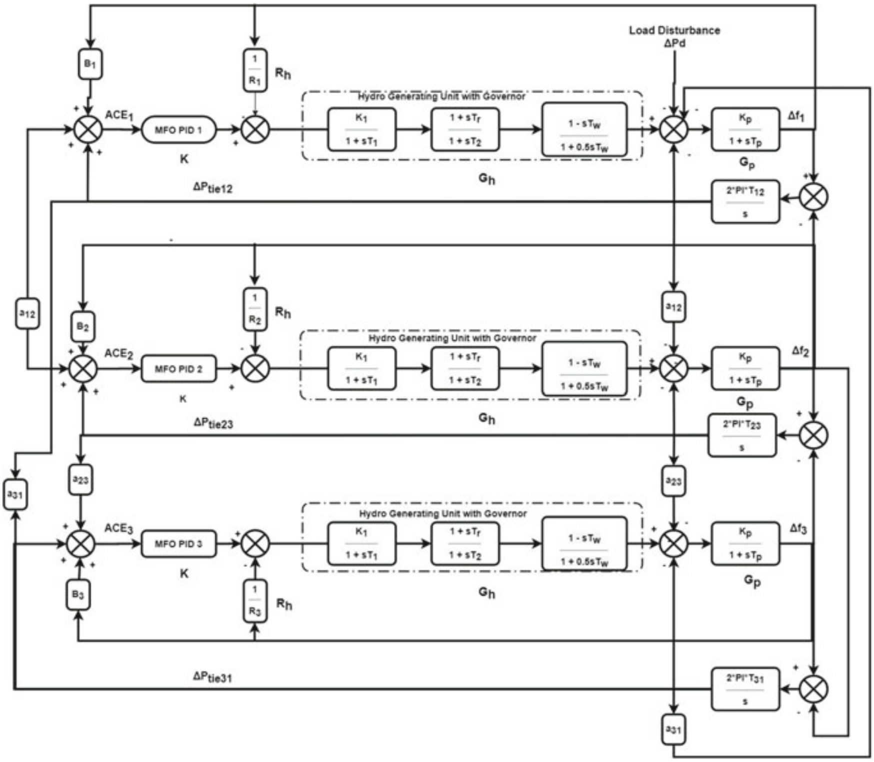


Fig. 36.1 Three-area-based hydropower system model

where Δf_n is the frequency deviation of the nth area in question, $\Delta P_{\text{tie}nm}$ is the net power deviation of tie-line among any two interacting nth and mth areas, and Δt_{siml} is the time taken for simulation.

36.3.2 Problem Constraints

The constraints of this design problem are the gain of the controller parameter bounds, which are formulated by (36.5) and optimized by the proposed MFO techniques as the set of minimum and maximum design variables.

$$\left.\begin{array}{r} K_{Pn}^{\min} \leq K_{Pn} \leq K_{Pn}^{\max};\ n = 1, 2, 3 \\ K_{\text{in}}^{\min} \leq K_{\text{in}} \leq K_{\text{in}}^{\max};\ n = 1, 2, 3 \\ K_{dn}^{\min} \leq K_{dn} \leq K_{dn}^{\max};\ n = 1, 2, 3 \end{array}\right\} \tag{36.5}$$

Step-1: Initialization and parameter specification
Population size (P), maximum number of iteration ($iter_{max}$) and Controller Variables $(d=3)$

Step-2: Calculation of flames number using $Flame\ number = round\left(N - l \times \frac{N-1}{T}\right)$

Step 3: Specify lower bound lb and upper bound ub of ith variable
Hence, for any random distribution, $M(i,j)$ is calculated

for $i = 1 : n$
  for $j = 1 : d$
    $M(i,\ j) = (ub\ (i) - lb\ (i)) * rand\ (\ ) + lb\ (i);$
  end
end

$$OM = FitnessFunction(M)$$

Step-4: sort the best flame position (K_p, K_i, K_d) and the fitness function

Step-5: Calculated $S(M_i, F_j) = D_i \cdot e^{bt} \cdot \cos(2\pi t) + F_j$ and $D_i = |F_j - M_i|$
by considering t as a random number [-1, 1]

Step-6: update the flame position using $Flame\ number = round\left(N - l \times \frac{N-1}{T}\right)$

Step-7: examine the flame boundary condition

Step-8: sort the flame position (K_p, K_i, K_d) and fitness function

Step-9: stop when a termination criterion is reached

Fig. 36.2 Generalized form of MFO algorithm as used in AGC

36.3.3 Measure of Performance

Eigenvalues calculation

The stability of the system is scrutinized by eigenvalue techniques; the concept is that for complete stability the whole of the eigenvalues lies on the position to the left of the imaginary axis. Therefore, judgment will be carried out based on the position of the eigenvalues; hence, oscillatory mode of the eigenvalues that give the precise transient response profile is calculated.

Performance Indices' Calculation

In order to ensure best possible mode of operation is obtained despite parameter uncertainties, robustness of the controller is important and is given emphasis on this work; hence, adaptive control modeling is employed. Therefore, to ensure the consistency of performance of the designed controller, extra performance indicators, namely integral of square error (ISE), integral of time square error (ITSE) and integral of absolute error (IAE) are also given due considerations in this study. These indices can be given, respectively, by (36.6–36.8).

$$\mathrm{ISE} = \int_{0}^{t_{\mathrm{sim}l}} \{(\Delta f_n)^2 + (\Delta P_{\mathrm{tie}nm})^2\}\mathrm{d}t \tag{36.6}$$

$$\mathrm{ITSE} = \int_{0}^{t_{\mathrm{sim}l}} \left\{(\Delta f_n)^2 + (\Delta P_{\mathrm{tie}nm})^2\right\} t\,\mathrm{d}t \tag{36.7}$$

$$\mathrm{IAE} = \int_{0}^{t_{\mathrm{sim}l}} \{|\Delta f_n| + |\Delta P_{\mathrm{tie}nm}|\}\mathrm{d}t \tag{36.8}$$

36.4 MFO Algorithm

36.4.1 Basic Concept of MFO

Moth flame optimization commonly referred to as MFO was inspired from navigation characteristics of moths at night, moths bear much resemblance to the family of butterflies but mostly fly at night, these insects just like butterflies are decorative in nature with larvae and adult as the major life time stages, and the larvae are transformed into moth by cocoons which is an important process in the life of a moth. The exceptional navigation technique of moth at night that allows them to travel long straight distance by maintaining constant angle in reference to the moon called transverse orientation has proven to be very effective. However, in the presence of the artificial lights, moths stuck in a spiral path and consequently converge toward the light source; hence, Mirjalili in [22] take advantage of this undesirable behavior and mathematically modeled it as MFO algorithm. MFO algorithm is population based and presumes that the candidate solutions are moths given by (36.9), and the problem's variables are their location in space; hence, the moths can hover in different dimensional space with varying location vectors.

$$M = \left\{ \begin{array}{cccc} m_{i,j} & \ldots & \ldots & m_{i,d} \\ : & : & : & : \\ : & : & : & : \\ m_{n,j} & \ldots & \ldots & m_{n,d} \end{array} \right\} \begin{array}{l} \text{for } i = 1, 2, 3, \ldots n \\ \text{and } j = 1, 2, 3, \ldots d \end{array} \tag{36.9}$$

Generally, n and d are the number of moths and variables (dimension), respectively, and the array for storing the subsequent value of fitness function for the moths is given in (36.10).

$$\mathrm{OM} = \left[\, \mathrm{OM}_i \; \ldots \; \ldots \; \ldots \; \mathrm{OM}_n \,\right]' \text{ for } i = 1, 2, 3, \ldots n \tag{36.10}$$

However, the essential components of the algorithm and flames specified in (36.11) are the best location the moths attained, thus moths upgrade their location in reference to the location of the flame as established in (36.14).

$$F = \begin{Bmatrix} F_{i,j} & \dots & \dots & F_{i,d} \\ : & : & : & : \\ : & : & : & : \\ F_{n,j} & \dots & \dots & F_{n,d} \end{Bmatrix} \begin{matrix} \text{for } i = 1, 2, 3, \dots n \\ \text{and } j = 1, 2, 3, \dots \text{d} \end{matrix} \tag{36.11}$$

Since the dimension of (36.9) and (36.11) is the same, then the array for storing the corresponding value of fitness function for all the flames may be given by (36.12). Thus, the global optimal of MFO algorithm may be assumed by a sequence of three elements expressed by (36.13).

$$\text{OF} = \left[\text{OF}_i \dots \dots \dots \text{OF}_n \right]' \text{ for } i = 1, 2, 3, \dots n \tag{36.12}$$

$$\text{MFO} = (I, P, T) \tag{36.13}$$

$$M_{i,j} = S\left(M_i, F_j\right) \tag{36.14}$$

Hence, exploitation and exploration of the quest are balanced by gradual decrement of number of flames, and consequently utilization of the best auspicious solutions is guaranteed by adaptive mechanism. The set of relevant equations and algorithm necessary for MFO algorithm implementation are well documented in [22], and the generalized form of MFO algorithm as used for AGC is given in Fig. 36.2 [24].

36.5 Simulation Results and Analysis

MFOA-based simulation results are considered for three-area hydropower system and compared with GA-based designed controller which is known in literature for providing high-quality solution to optimization problems. The variables of interest are the convergence swiftness of the proposed approach, frequency and tie-line power deviation; the load perturbation profiles used for studying the response of the examine power system are shown in Fig. 36.3. However, three case scenarios are considered in this simulation study, the first and second case scenarios are performance evaluations, while third case is robustness analysis in confined range.

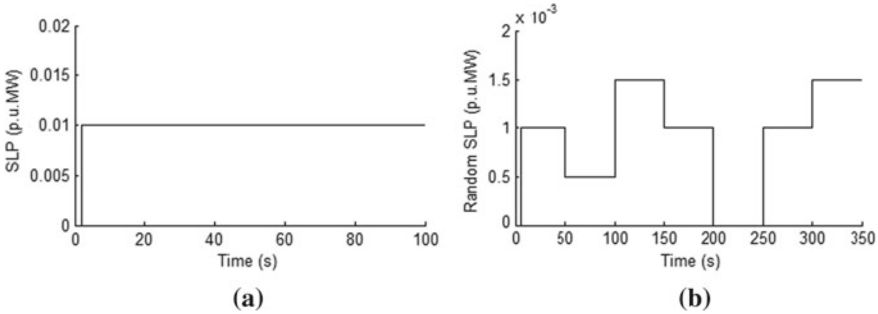


Fig. 36.3 Load perturbation: **a** step load and **b** random step load

36.5.1 Case Scenario 1: Application of Per Unit Load Perturbation in Area 1

In this case, the dynamic performance evaluation is carried out using 0.01 per unit step load in area 1, and the optimized controller gain and oscillatory modes obtained are documented in Tables 36.1 and 36.2, respectively. Table 36.2 clearly indicates how the proposed approach eigenvalues are far spaced out from the imaginary axis in contrast to its GA-PID counterpart, and also Fig. 36.4 delineates further how the proposed approach transients decay faster than that of the GA-PID controller. Table 36.4 makes available the transient parameter values of the proposed approach against that of GA-PID approach. Also, all performance indices values FOD inclusively reported in Table 36.3 clearly indicates lower values for the proposed approach as against GA-PID approach, which is sign of superior designed controller. However, the outstanding convergence characteristic displayed by the proposed approach in Fig. 36.4d against that of GA-PID is exemplary. Thus, it can be deduced from Tables 36.2, 36.3, 36.4 and Fig. 36.4 that the proposed approach increases the stability of the system.

36.5.2 Case Scenario 2: Application of Random Load Perturbation in Area 1

A random load shown in Fig. 37.3b is perturbed to area 1 of the model with same parameter as the previous case, and this is carried out to evaluate how flexible the system could be in case of any uncertainty especially stochastic loading. However, the simulation study carried out revealed that the proposed control technique is capable of suppressing the oscillations more effectively after each load perturbation application as compared to GA-PID, and this can readily be comprehended from Fig. 36.5 for the frequency deviations, tie-line power deviation and area control error response. Thus, the proposed approach is capable of driving the system into steady-state condition efficiently despite unpredictable loading.

Table 36.1 Optimized controller gains for the investigated system

Controller type	Area 1			Area 2			Area 3		
	$K_{P1}(-\text{ve})$	$K_{i1}(-\text{ve})$	$K_{d1}(-\text{ve})$	$K_{P2}(-\text{ve})$	$K_{i2}(-\text{ve})$	$K_{d2}(-\text{ve})$	$K_{P3}(-\text{ve})$	$K_{i3}(-\text{ve})$	$K_{d3}(-\text{ve})$
GA-PID [Studied]	0.0593	0.0728	0.9705	0.0549	0.0773	0.9435	0.0549	0.0100	0.8313
MFO-PID [Proposed]	0.6314	0.8960	0.0010	1.0651	0.0010	1.8681	0.0121	0.0010	2.7456

Table 36.2 Oscillatory modes for the investigated system

GA-PID (1×10^2)	Proposed MFO-PID (1×10^2)
$0.0014 \pm i0.0116$ $0.0010 \pm i0.0113$ $-0.0002 \pm i0.0045$ $-0.0011 \pm i0.0001$	$0.0005 \pm i0.0114$ $0.0004 \pm i0.0114$ $-0.0011 \pm i0.0026$ $-0.0001 \pm i0.0004$ $-0.0002 \pm i0.0001$

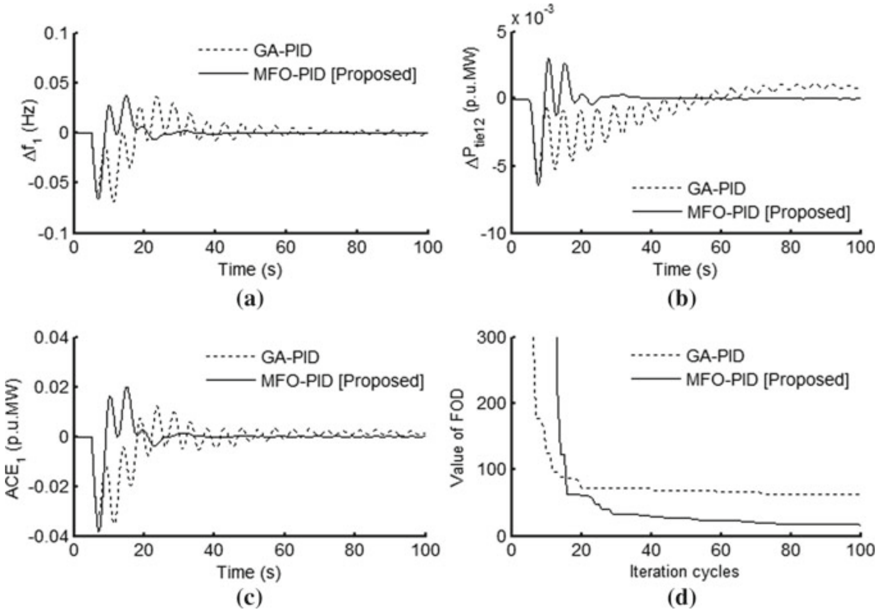


Fig. 36.4 Transient response for the studied system at 0.01 per unit step load in area 1; **a** Δf_1, **b** ΔP_{tie12}, **c** ACE_1 and **d** comparative convergence profile of FOD

Table 36.3 Performance indicators for the investigated system

Control approach	FOD	ISE	ITSE	IAE
GA-PID	61.9563	0.05605	0.5366	2.391
Proposed MFO-PID	**17.4426**	**0.03948**	**0.2164**	**1.228**

Bold shows result of interest

36.5.3 Case Scenario 3: The Sensitivity Analysis of the Proposed Approach

In this section, the dynamic performance of the system to change in some of the important characteristics such as model parameters condition and loading of the

Table 36.4 Transient response specification for the investigated system

AGC response profiles	GA-PID				Proposed MFO-PID			
	$T_r(s)$	$T_s(s)$	$M_p(\%)$	$T_p(s)$	$T_r(s)$	$T_s(s)$	$M_p(\%)$	$T_p(s)$
Δf_1	3.8194	99.0270	0.0691	11.1494	**7.80 × 10^{-5}**	**32.9017**	**0.0659**	**6.8669**
ΔP_{tie12}	18.9243	97.2758	0.0057	7.5675	**9.49 × 10^{-5}**	**33.8943**	**0.0064**	**7.7309**
ACE_1	0.1273	99.1589	0.0353	11.4246	**6.16 × 10^{-9}**	**33.3622**	**0.0384**	**7.1289**

Bold shows result of interest

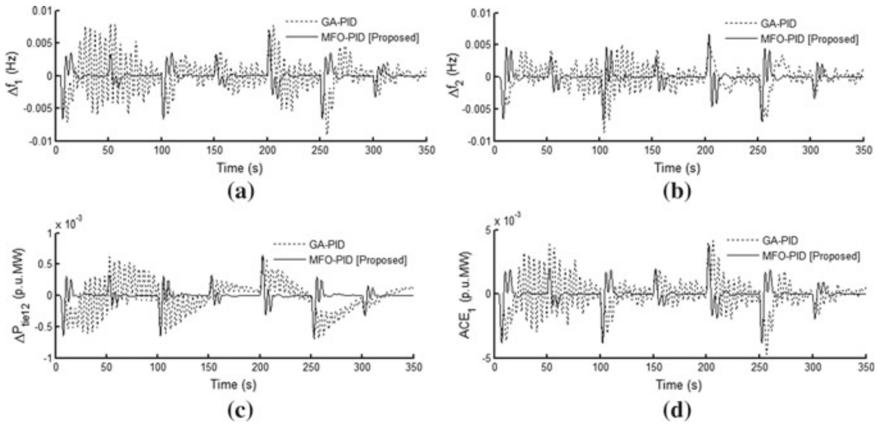


Fig. 36.5 Comparative transient for the studied system at random perturbed load in area 1; **a** Δf_1, **b** Δf_2, **c** $\Delta P_{tie12}$and **d** ACE_1

system called robustness analysis is implemented. This is accomplished by considering ± 25% variability in loading condition, tie-line and governor constant from their nominal value and then subjected to 0.01 per unit step load in area 1. The dynamic characteristics obtained specifically, first area frequency deviation, respectively, with ± 25% variation in nominal loading, T_{12}, T_1 and R_{h1} constants are depicted in Fig. 36.6. These results show slight difference to that of nominal loading despite substantial variation in some of the important characteristics of the system which is an indication of how robust the proposed controller is to parameter uncertainty. Thus, the study carried out here also confirms the effectiveness of the proposed controller.

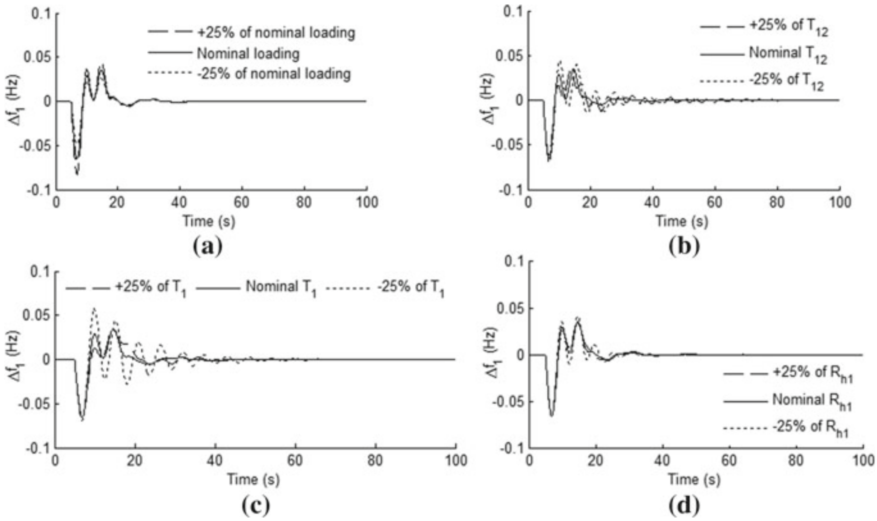


Fig. 36.6 Frequency deviation at 0.01 per unit step load obtained with the proposed approach for $\pm 25\%$ variation in; **a** nominal loading **b** T_{12}**c** T_1 and **d** R_{h1}

36.6 Conclusion

The performance evaluation of moth flame optimization algorithm considering three-area interconnected hydropower system model is offered for AGC investigation. At the onset, the convergence swiftness of the proposed algorithm was studied, and then the dynamic transient characteristics of the designed model were examined with SLP and random SLP. However, comparative examination revealed better convergence and faster dynamic transient characteristics decay of the proposed MFO algorithm over GA, and this is the strong indication of high performing controller. Finally, robustness analysis in confined range was carried out to the design model and the small shift registered in the dynamic characteristics despite variation in some of the important parameters of the system demonstrate the flexibility of the proposed approach, and this is also another indication of high performance. Hence, in this study the proposed MFO technique stands out to show high and competitive performance by quickly and adaptively converging toward the optima and, thus, pronounced its dominance over the well-known genetic algorithm technique.

Appendix

System Model Parameter

$\mathrm{s}f = 50\,\mathrm{Hz},\ P_{r_1} = P_{r_2} = P_{r_3} = 2000\,\mathrm{MW},\ \mathrm{Total\,area\,load} = 1000\,\mathrm{MW}$ $\mathrm{Base\,.\,rating} = 2000\,\mathrm{MW},\ \mathrm{Initial\,loading} = 50\%,\ a_{12} = -1,\ B_1 = B_2 =$

$B_3 = 0.425\,\text{p.u MW/Hz}$, $R_1 = R_2 = R_3 = 2.4\,\text{Hz/p.u. MW}$, $T_1 = 48.7\,\text{s}$, $T_2 = 0.513$, $T_w = 1.0\text{s}$, $T_r = 5.0\,\text{s}$, $K_1 = 1.0$, $T_{12} = T_{23} = T_{31} = 0.0707\,\text{p.u}$, $K_{P_1} = K_{P_2} = K_{P_3} = 100\,\text{Hz/p.u. MW}$, $T_{p1} = T_{p2} = T_{P3} = 20\,\text{s}$.

References

1. Sivanagaraju, S., Sreenivasan, G.: Power System Operation and Control. Pearson India (2003)
2. Kundur, P.: Book Power System Stability and Control. McGraw Hill, Inc. (1994)
3. Saadat, H.: Power System Analysis. McGraw Hill Companies, pp. 527–585 (1999)
4. Nanda, J., Kothari, M.L., Satsangi, P.S.: Automatic generation control of an interconnected hydrothermal system in continuous and discrete modes considering generation rate constraints. IEE Proc. D Control Theor. Appl. **130**(1), 17–27 (1983). https://doi.org/10.1049/ip-d.1983.0004
5. Kusic, G.L., Sutterfield, J.A., Caprez, A.R., Haneline, J.L., Bergman, B.R.: Automatic generation control for hydro systems. IEEE Trans. Energy Convers. **3**(1), 33–39 (1988)
6. Zadeh, K.N., Villada, Z.J.H., Weibull, A.: Operation planning and control of cascaded Hydro Plants. IEEE Trans. Power Syst. **6**(2), 835–841 (1991)
7. Divya, K.C., Rao, P.S.N.: A Simulation Model for AGC Studies of Hydro—Hydro Systems, vol. 27, pp. 335–342 (2005) https://doi.org/10.1016/j.ijepes.2004.12.004
8. Pandey, S.K., Mohanty, S.R., Kishor, N.: A literature survey on load—Frequency control for conventional and distribution generation power systems. Renew. Sustain. Energy Rev. **25**, 318–334 (2013). https://doi.org/10.1016/j.rser.2013.04.029
9. Ibraheem, P.K., Kothari, D.P.: Recent Philosophies of Automatic Generation Control Strategies in Power Systems. IEEE Trans. Power Syst. **20**(1), 346–357 (2005)
10. Singh, O., Tiwari, P., Ibraheem, Singh, A.K.: A survey of recent automatic generation control strategies in power systems. Int. J. Emerg. Trends Electr. Electron. **7**, 2 (Jan 2013) (2018)
11. Yang, X.: Optimization and metaheuristic algorithms in engineering. In: Yang, X.S., Gandomi, A.H., Talatahari, S., Alavi, A.H. (eds.) Metaheursitics in Water, Geotechnical and Transport Engineering. Elsevier, pp. 1–23 (2013)
12. Golpîra, H., Bevrani, H., Golpîra, H.: Application of GA optimization for automatic generation control design in an interconnected power system. Energy Convers. Manag. **52**(5), 2247–2255 (2011). https://doi.org/10.1016/j.enconman.2011.01.010
13. Dash, P., Saikia, L.C., Das, S.K.: AGC of a multi-area hydro-thermal system with BES and firefly optimized PID controller. 2014 18th Natl Power Syst. Conf. NPSC **2014**, 5–10 (2015). https://doi.org/10.1109/NPSC.2014.7103819
14. Bhatt, P., Roy, R., Ghoshal, S.P.: GA/particle swarm intelligence based optimization of two specific varieties of controller devices applied to two-area multi-units automatic generation control. Int. J. Electr. Power Energy Syst. **32**(4), 299–310 (2010). https://doi.org/10.1016/j.ijepes.2009.09.004
15. Boroujeni, S.M.S., Delafkar, H., Behzadipour, E., Boroujeni, A.S., Hemmati, R.: Multi area load frequency control using simulated annealing. Aust. J. Basic Appl. Sci. **5**(7), 644–651 (2011)
16. Jagatheesan, K., Anand, B., Baskaran, K., Dey, N., Ashour, A.S., Balas, V.E.: Effect of nonlinearity and boiler dynamics in automatic generation control of multi-area thermal power system with proportional-integral-derivative and ant colony optimization technique. Stud. Syst. Decis. Control **109**, 89–110 (2018). https://doi.org/10.1007/978-3-319-58996-1_5
17. Gozde, H., Taplamacioglu, M.C.: Automatic generation control application with craziness based particle swarm optimization in a thermal power system. Int. J. Electr. Power Energy Syst. **33**(1), 8–16 (2011). https://doi.org/10.1016/j.ijepes.2010.08.010

18. Vijaya Chandrakala, K.R.M., Balamurugan, S.: Simulated annealing based optimal frequency and terminal voltage control of multi source multi area system. Int. J. Electr. Power Energy Syst. **78**, 823–829 (2016). https://doi.org/10.1016/j.ijepes.2015.12.026.
19. Oonsivilai, A., Marungsri, B.: Optimal PID tuning for AGC system using adaptive Tabu search. In: 7th WSEAS International Conference on Power Systems, Beijing, China, pp. 42–47 (2007)
20. Omar, M., Ebrahim, M.A., Ghany, A.M.A., Bendary, F.: Tuning of PID Controller for Load Frequency Control Problem via Harmony Control Problem via Harmony Search Algorithm, pp. 255–263 (2016). https://doi.org/10.11591/ijeecs.v1i2.pp255-263.
21. Shiva, C.K., Shankar, G., Mukherjee, V.: Automatic generation control of power system using a novel quasi-oppositional harmony search algorithm. Int. J. Electr. Power Energy Syst. **73**, 787–804 (2015). https://doi.org/10.1016/j.ijepes.2015.05.048
22. Mirjalili, S.: Moth-flame optimization algorithm: A novel nature-inspired heuristic paradigm. Knowledge-Based Syst. (July, 2015). https://doi.org/10.1016/j.knosys.2015.07.006
23. Chandrakala, K.R.M.V., Balamurugan, S., Sankaranarayanan, K.: Electrical power and energy systems variable structure fuzzy gain scheduling based load frequency controller for multi-source multi area hydro thermal system. Int. J. Electr. Power Energy Syst. **53**, 375–381 (2013). https://doi.org/10.1016/j.ijepes.2013.05.009
24. Mohanty, B.: Performance analysis of moth flame optimization algorithm for AGC system. Int. J. Model. Simul. **6203**, 0 (May, 2018). https://doi.org/10.1080/02286203.2018.1476799

Chapter 37
Identification of Breast Abnormality from Thermograms Based on Fractal Geometry Features

Aayesha Hakim and R. N. Awale

Abstract Breast cancer is a common cause of mortality among women globally. Thermography is a prognostic modality currently under research that has shown potential of providing early information on developing breast cancer based on temperature changes. In this work, we have extracted fractal texture features, namely Hurst coefficient, fractal dimension and lacunarity, from the segmented breast region for abnormality identification. These features are fed to classification algorithms like support vector machine (SVM), logistic regression, k-nearest neighbours (KNN) and Naïve Bayes to classify thermograms into two classes—healthy and sick. The best accuracy obtained for identification of breast abnormality using fractal features is 94.53% (Naïve Bayes classifier) as compared to accuracy of 92.74% obtained using texture and statistical features (PCA-SVM classifier) in our previous study. This further opens the scope for thermography to be used clinically for breast examination on a large scale.

37.1 Introduction

Over the last few years, breast cancer has become a huge problem in global health sector [1]. In India, this disease has a high mortality rate, especially in rural regions. Approximately, 60% deaths are due to delay in diagnosis. Therefore, early detection is the only way to improve patient survival rate by providing treatment in time. Techniques like physical examination, ultrasonography, mammography, magnetic resonance imaging (MRI) and thermography [2, 3] are used to diagnose breast cancer. Mammography [2] is the most used detection method. However, it misses cancer for dense breasts which are common in women aged younger than 40. Further, it exposes the patient to harmful radiation and causes uneasiness to the patient due to breast compression. Thermography is an upcoming screening method with zero risk and

A. Hakim (✉) · R. N. Awale
Veermata Jijabai Technological Institute, Mumbai, India

R. N. Awale
e-mail: rnawale@el.vjti.ac.in

T. Senjyu et al. (eds.), *IOT with Smart Systems*, Smart Innovation,
Systems and Technologies 251, https://doi.org/10.1007/978-981-16-3945-6_38

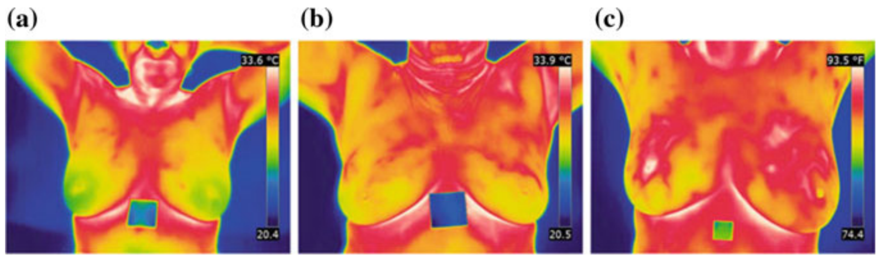


Fig. 37.1 Sample breast thermograms **a** normal, **b** benign, **c** malignant [6]

improved efficiency of detection [4]. Lawson [5] was the first to use thermography for detecting breast cancer by studying heat distribution of the breasts. Malignancy changes the local temperature of the breast by 2–3 °C as compared to normal breast tissue. Some sample thermograms of normal, benign and malignant cases are shown in Fig. 37.1.

37.1.1 Related Works

Fractal geometry concepts are used to analyse complex patterns in images. In thermal imaging, these concepts can be exploited to differentiate between healthy and sick cases as abnormal thermograms have irregular boundaries. Authors in [7] suggested the use of statistical features to predict breast abnormality. Authors in [8] used fractal analysis for the detection of breast cancer using mammograms. Etehad-Tavakol et al. [9] computed the fractal dimension of first hottest regions of eight benign and seven malignant thermograms to detect benign and malignant tumours. Their method showed significant differences among the two patterns. Borchartt et al. [10] extracted features like fractal dimension (FD), the local lacunarity (LL) and the succolarity (LS) from 4 healthy and 24 abnormal thermograms and fed them to different machine learning techniques for classification. Highest ROC area of 96% was achieved. Results obtained by Serrano et al. [11] showed that the abnormal breast thermograms had higher values of lacunarity as compared to the normal one. The authors in [12] extracted asymmetry features to classify breast thermograms into malignant and benign classes. In [13], statistical significance of these features was tested using SPSS tool, and only significant features were fed as input to various classifiers. Highest accuracy of 94.4% was obtained using random forest classifier [14].

The goal of this paper is to investigate the use of fractal features for prediction of breast abnormality by comparing features from normal and abnormal thermograms using machine learning algorithms. The main contributions of this work are as follows: (a) extraction of the fractal features from thermograms of the Database of Mastology Research (DMR), (b) classification of thermograms as healthy or sick based on fractal indices and (c) comparison of our classification results with the results obtained using statistical and texture features from our previous study [12]. The paper is organized as follows: Sect. 37.2 describes the methods and materials used to conduct the investigation. Classification using machine learning models is presented in Sect. 37.3. Section 37.4 covers the results, and Sect. 37.5 concludes the paper with discussion on future work.

37.2 Methods and Materials

We used thermal images from the Database of Mastology Research (DMR) [15] which contains anonymous breast images of varying size and shape of women from age 29–85 years. In thermograms, the abnormalities in one or both the breasts are represented by the presence of asymmetric vascular patterns due to increased metabolism [4]. This results in surface texture differences between the two breasts. To quantify these textural differences, we use the image fractal geometry properties [16].

37.2.1 ROI Extraction

The RGB images from DMR are converted into grayscale images, and only the breast area is extracted from each image as shown in Fig. 37.2. We have employed the procedure described in our previous study [12] for ROI extraction. Further, we resized the final images to 256×256 pixels for faster computation.

37.2.2 Extraction of Fractal Features

Fractal geometry concepts link each texture to numerical values that can further help to classify the textures. In this study, three measures are extracted from the breast ROIs—Hurst coefficient [11], fractal dimension [16] and lacunarity [17–19].

37.2.2.1 Hurst Coefficient (H)

For Hurst coefficient calculation, we used the approach employed in [11]. From the segmented left and right breasts, H values were extracted. We choose square windows of size 5, 7, 9, 11, 13 and 15 pixels and calculate different H values depending on the window size w. Then, the average of the Hurst coefficient for each w, for each breast, is calculated. So, each image of $N \times N$ size has H values calculated Σ $(N -$

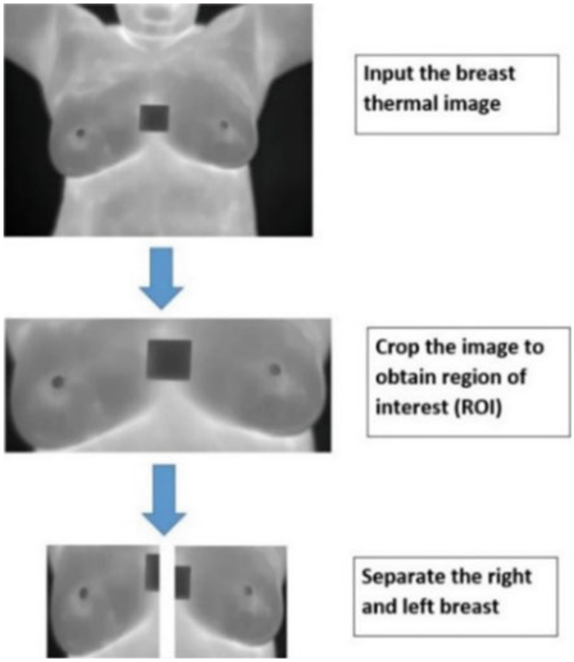


Fig. 37.2 Preprocessing and ROI extraction of the breast thermograms

$(2i + 3))^2$ times. *i* ranges from 1 to 6 [11]. In abnormal cases, the *H* values are less than 0.5. For healthy cases, *H* is greater than 0.5 [20].

37.2.2.2 Fractal Dimension (F)

Fractals are defined as the deviation in repetitive patterns in an image as the observing window's scale changes. We calculate the fractal dimension of a thermogram using the box-counting method by observing how the number of boxes deviates as the grid becomes finer [16]. The values of fractal dimensions for healthy cases were found to be near 1, whereas that of abnormal cases containing irregular patterns were considerably higher than 1. In the equation below, $N\delta$ (F) is the minimum number of diameters sets δ which can cover *F* [16].

$$\dim_B(F) = \lim_{\delta \to 0} \frac{\log N^{\delta}(F)}{-\log(\delta)}$$

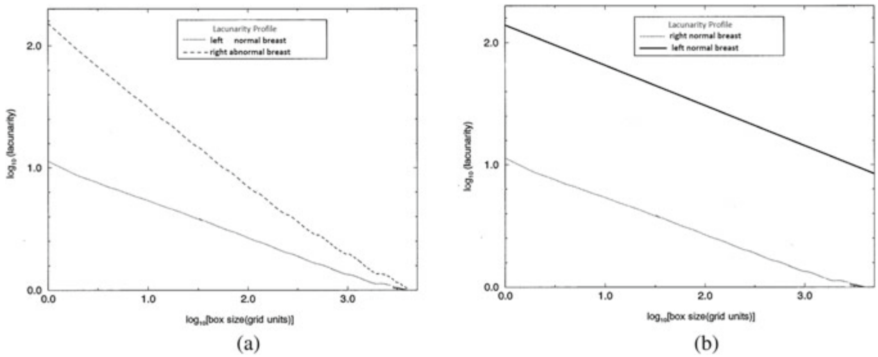


Fig. 37.3 Lacunarity profile for **a** abnormal case, **b** normal case

37.2.2.3 Lacunarity

Lacunarity refers to the way the space or gap is taken by a pattern in the image. This feature helps in detection of the presence of homogeneous gap distributions or random gaps depicting correlation between colours to distinguish between healthy and sick thermograms. Large heterogeneous gaps mean abnormal texture leading to higher lacunarity values, and homogeneous gaps (normal case) give lower lacunarity values [17–19]. It is a ratio of variation from the mean to the square of the probability of gap occurrence for each box size as shown below [17]

$$\Lambda(r) = \frac{m_2(r)}{(m_1(r))^2}$$

Figure 37.3 shows lacunarity profiles of abnormal and normal cases, respectively. If the behaviour of the log–log curve is two translated parallel lines, then both the breasts have similar vascular patterns, i.e. normal case. As the box sizes vary, the lacunarity values change and are plotted in logarithmic scale. When the plots intercept, it depicts different thermic distribution in the two breasts, i.e. abnormal case.

37.3 Classification Using Machine Learning Models

After the extraction of fractal features, a.csv file is created. This file is fed as an input to machine learning algorithms [14] like support vector machine (SVM), logistic regression, k-nearest neighbours (KNN) and Naïve Bayes. We partitioned the data set into 80:20 ratio for training and testing, respectively, for all algorithms.

37.4 Results and Discussion

The results obtained in Fig. 37.4 and Table 37.1 depict that there is a noticeable difference in the fractal dimension and Hurst coefficient values, whereas lacunarity does not show significant difference in values for healthy and abnormal cases. Accuracy and F1-score plot for various machine learning algorithms is shown in Fig. 37.5. Receiver operating characteristic (ROC) plots for various machine learning algorithms are shown in Fig. 37.6, and performance metrics are tabulated in Table 37.2. As seen from Table 37.2, Naïve Bayes classifier performs better than rest of the classifiers. The comparative results of this investigation with our previous work [12] confirm that fractal features are more efficient than texture and statistical features in discriminating thermograms. The reasons for misclassification can be hormone or age related, or due to the presence of a cold tumour.

Table 37.3 gives the confusion matrices for all machine learning algorithms used. The diagonal entries for each algorithm show the correctly identified number of patients for malignant and benign classes. Our findings confirm that thermograms can provide insights about the breast abnormalities with high efficacy. However, this should be further confirmed with a real time, large database that encompasses a

Fig. 37.4 Plot of fractal features for an abnormal case

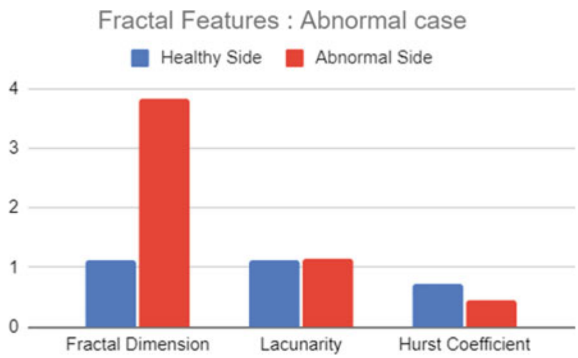


Table 37.1 Sample values of fractal dimension, lacunarity and Hurst coefficient for **a** normal and **b** abnormal case

Fractal features	Case type: normal	
	Healthy side	Healthy side
Fractal dimension	1.312	1.2261
Lacunarity	1.1021	1.1142
Hurst coefficient	0.9154	0.8581
Fractal features	Case type: abnormal	
	Healthy side	Abnormal side
Fractal dimension	1.1254	3.8261
Lacunarity	1.1121	1.1542
Hurst coefficient	0.7154	0.4581

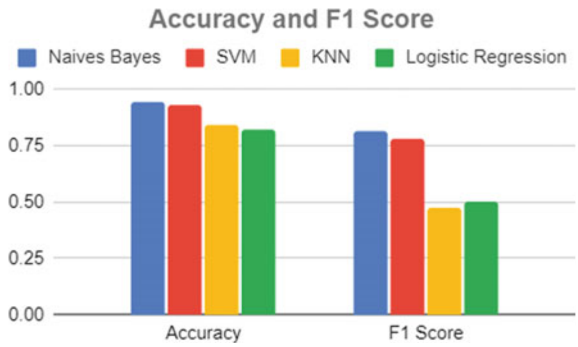


Fig. 37.5 Accuracy and F1-score plots for various machine learning algorithms

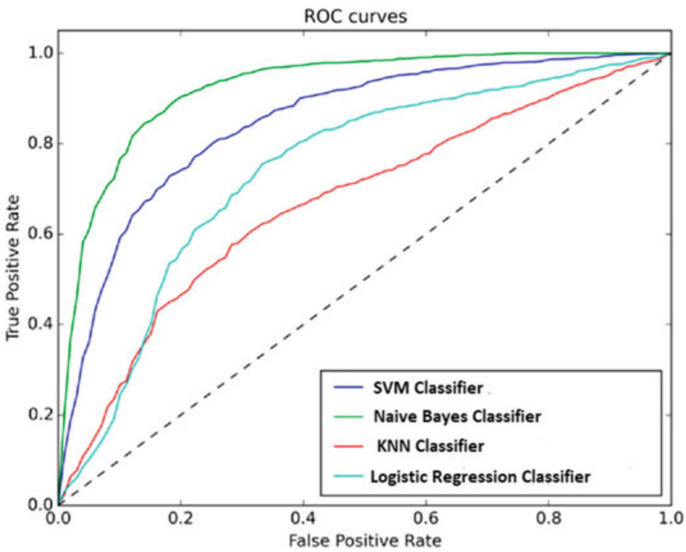


Fig. 37.6 Receiver operating characteristic (ROC) graph for various classifiers

Table 37.2 Performance evaluation parameters of various classifiers

Method/parameters	SVM	Logistic regression	KNN	Naive Bayes
Accuracy %	92.74	82.2	84.35	94.53
Sensitivity %	77.77	55.55	44.44	86.25
Specificity %	95.83	87.5	91.67	97.75
F1-score	0.7777	0.5	0.4706	0.812
AUC	0.85	0.74	0.67	0.92

Table 37.3 Confusion matrices of various classifiers

		Predicted class	
Algorithm used	Actual class	Malignant	Benign
Naive Bayes	Malignant	8	1
	Benign	2	46
Logistic regression	Malignant	5	4
	Benign	6	42
KNN	Malignant	3	6
	Benign	3	45
SVM	Malignant	7	2
	Benign	2	46

variety of cases. In our experiment, we used all the images from DMR for classification using fractal geometry, whereas previous studies [9–11] used limited samples.

37.5 Conclusion

Despite significant advances in cancer treatment, detection of breast cancer remains a top biomedical investigation urgency. Thermography is a low-cost emerging screening tool for the early prognosis of breast cancer for all age groups, especially younger women. The abnormality information is only in the breast region; hence, we performed texture analysis based on the extraction of fractal features from the breast ROI. We used various machine learning algorithms for abnormality identification and compared the results with the actual diagnosis. The results indicate the possibility of employing fractal features for a reliable detection of breast cancer using infrared imaging. The comparative results of this investigation with our previous work confirmed that fractal features are more effective than texture and statistical features in discriminating thermograms. Naïve Bayes classifier outperforms other decision-making models with highest accuracy of 94.53% and AUC of 0.92. The role of numerical simulation with thermograms can also be explored further. As a future improvisation to this study, neural networks and CNN can also be employed for classification.

Acknowledgements The authors of this paper would like to thank the assistance of TEQIP-III project and V.J.T.I, Mumbai, for financial assistance in this research work.

References

1. Breast Cancer India: Pink Indian Statistics. Available at: http://www.breastcancerindia.net/statistics/stat_global.html
2. Prasad, S., Houserkova, D.: The role of various modalities in breast imaging. Biomed. Papers Med. Fac. Univ. Palacky, Olomouc, Czechoslovakia **151**(2), 209–218 (2007)
3. Head, J., Wang, F., Lipari, C., Elliott, R.: The important role of infrared imaging in breast cancer. In: IEEE Engineering in Medicine and Biology Magazine, pp. 52–57 (2000).
4. Hakim, A., Awale, R.N.: Thermal imaging—An emerging modality for breast cancer detection: A comprehensive review. J. Med. Syst. **44**, 136 (2020)
5. Lawson, R.N.: Implications of surface temperatures in the diagnosis of breast cancer. Can. Med. Assoc. J. **75**, 309–310 (1956)
6. Visual Lab. A Methodology for Breast Disease Computer-Aided Diagnosis using dynamic thermography. Available online: http://visual.ic.uff.br/dmi [Accessed on 15 Dec 2020]
7. Hakim, A., Awale, R.N.: Detection of breast pathology using thermography as a screening tool. In: 15th Quantitative InfraRed Thermography Conference, Portugal (2020)
8. Rangayyan, R.M., Nguyen, T.M.: Fractal analysis of contours of breast masses in mammograms. J. Digital Imaging 223–237 (2007) (Springer)
9. Etehad Tavakol, M., Lucas, C., Sadri, S., Ng, E.Y.K.: Analysis of breast thermography using fractal dimension to establish possible difference between malignant and benign patterns. J. Healthc. Eng. 27–44 (2010) (Multi Science Publishing)
10. Borchartt, T.B., et al.: Combining approaches for early diagnosis of breast diseases using thermal imaging. Int. J. Innov. Comput. Appl. **4**(3/4), 163 (2012)
11. Serrano, R.C., Ulysses, J., Ribeiro, S. et al.: Using Hurst coefficient and lacunarity to diagnosis early breast diseases. In: Proceedings of 17th International Conference on Systems, Signals and Image Processing. Osijek, Croatia (2010), pp. 550–553
12. Hakim, A., Awale, R.N.: Harnessing the power of machine learning for breast anomaly prediction using thermograms. Int. J. Med. Eng. Inf. (IJMEI) [in press] (2021)
13. Hakim, A., Awale, R.N.: Statistical analysis of thermal image features to discriminate breast abnormalities. In: Fifth International Conference on Information and Communication Technology for Competitive Strategies (ICTCS-2020), 11–12 Dec 2020 (virtual), India
14. Hakim, A., Awale, R.N.: Predictive analysis of breast cancer using infrared images with machine learning algorithms. In: Analysis of Medical Modalities for Improved Diagnosis in Modern Healthcare. Taylor & Francis, CRC press book (2020) [in press]
15. Silva, L.F., Saade, D.C.M., Sequeiros, G.O., Silva, A.C., Paiva, A.C., Bravo, R.S., Conci, A.: A new database for breast research with infrared image. J. Med. Imaging Health Inf. **4**(1), 92–100 (2014)
16. Falconer, K.: Fractal Geometry: Mathematical Foundations and Applications. Wiley (2004)
17. Melo, R.H.C., Vieira, E.A., Conci, A.: Characterizing the lacunarity of objects and image sets and its use as a technique for the analysis of textural patterns, Advanced Concepts for Intelligent Vision Systems. Lecture Notes in Computer Science, ACIVS—IEEE Benelux Signal Processing Chapter, vol. 4179, pp. 208–219, Springer Berlin/Heidelberg, Antwerp, Belgium (2006)
18. Allain, C., Cloitre, M.: Characterizing the lacunarity of random and deterministic fractal sets. Phys. Rev. A. **44**, 3552–3558 (1991)
19. Melo, R.H.C., Vieira, E.A., Conci, A.: Comparing two approaches to compute lacunarity of mammograms. In: Proceedings of 13th International Conference on Systems, Signals and Image IWSSIP, Budapest, Hungary, pp. 299–302 (2006)
20. Marusina, M., Mochalina, A., Frolova, E., Satikov, V., Barchuk, A., Kuznetsov, V., Gaidukov, V., Tarakanov, S.: MRI image processing based on fractal analysis. Asian Pac. J. Cancer Prev. **18**(1), 51–55 (2017)

Chapter 38
A Concise Study on IoT-Based Health Care

M. B. Lakshmi, Rehna Baby Joseph, Salini Suresh, V. Suneetha, and R. Sunder

Abstract Internet of Things (IoT) simply refers to taking and connecting all the surrounding smart devices (things) to the Internet. It is one of the most hopeful future technologies with a wide range of applications. The implementation of the IoT in the healthcare sector plays a significant role in handling various chronic diseases, and also in the prevention and control of diseases. IoT devices provide remote health monitoring, emergency notification, and human well-being systems. These health tracking devices extend from monitoring health parameters to advanced wearable devices that upgrade the quality of human life. Smart health care can be accomplished at all stages, starting from observing pediatric patients to tracing of chronic conditions in the aged. Different technologies and constituents are used in the Internet of Things for providing better health. This paper gives a glance at various applications of IoT that serves in the healthcare systems for providing a better quality of life.

38.1 Introduction

Internet of Things (IoT) refers to a concept that connects all devices to the Internet that also allows their communication with each other over the same. It extends the power of the Internet beyond computers and smartphones to an entire range of other things, processes, and environs. IoT is a colossal network of connected devices that gather and share data about their usage and the environments in which they are employed. Kevin Ashton is known as the Father of Internet of Things. He is a founder director of MIT's Auto-ID Center which is an alumni to Procter and Gamble. IoT has undergone a transformation from interrelated embedded computing devices to that of smart sensor devices. IoT networks use a wide variety of sensors and actuators to make the system smart and intelligent. IoT gives users the ability to organize their

M. B. Lakshmi (✉) · R. B. Joseph · R. Sunder
Sahrdaya College of Engineering and Technology, Kodakara, Kerala, India

S. Suresh · V. Suneetha
Dayananda Sagar College of Arts, Science Commerce, Bengaluru, Karnataka, India

T. Senjyu et al. (eds.), *IOT with Smart Systems*, Smart Innovation, Systems and Technologies 251, https://doi.org/10.1007/978-981-16-3945-6_39

daily life and integrate real-world entities such as smart gadgets that provides both physical and virtual communications. IoT has an extensive range of applications in areas, such as health care, water management, cities, waste management, traffic control, parking activities, agriculture, surveillance and security, and industry sector. Low storage and restricted processing capacity are some of the issues faced by IoT when applied to the smart city environs.

Fog computing is one of the most hopeful technologies that reinforce IoT applications. IoT systems that are based on fog computing consists of three layers, namely device layer, fog layer, and cloud layer. It acts as an intermediate between IoT devices and the cloud computing framework forming a cloud–fog–thing network. Cloud computing is the delivery of different services including resources, tools, and applications like data storage, servers, databases, networking, and software through the Internet. IoT–cloud technologies require a conceptual infrastructure that provides patient-centric, and high-quality smart health care economically.

IoT applications are suffering from some major challenges: (i) Fail in reliability, data integrity, data provenance, and security for IoT applications through the Internet secures the communication channels, (ii) Lack of support of the broadcast nature of wireless media for IoT communications, (iii) Efficient handling of a large amount of data generated by IoT devices, (iv) How to assure sensor data, and (v) How to issue both local and global communication efficiently among numerous heterogeneous devices. However, IoT is in its initial stage in the healthcare sector which requires a thorough awareness of current research that could provide useful guidance for various researchers. This paper provides a brief description of the trends in IoT-based healthcare research and the available systems for this purpose.

38.2 Smart Health Care

Traditional health care fails in many cases due to a terrible increase in population. Here, comes smart health care as a solution for it. Smart health care refers to the technology that helps in upgrading the quality of life [1]. It helps users by making them aware of their medical status. The transformation of health care to smart health care is a slow process because of the lack of updation of the digital era within healthcare professionals. Smart health care enables users to manage certain emergency situations on their own. The emphasis given on personal health care has switched from traditional hospital care to smart home care. This provides both improved quality and user experiences. Smart health care with the help of IoT devices provides remote health monitoring, emergency notification, cost-effective treatment, and availability of medical services without any geographical barriers. These health monitoring devices range from monitoring health parameters such as fitness trackers to advanced wearable devices that improve the quality of life. Several technologies and constituents are used in Internet of Things for the betterment of health. Smart health care can be accomplished at all stages, starting from monitoring the temperature of pediatric patients to tracing of chronic conditions in the aged population. Just now, "Personalized and Connected Health" improved the healthcare sector, claiming that people can examine and track their health by using smart devices.

Connected health care refers to remote healthcare solutions with continuous monitoring and emergency notifications. It focuses on improving healthcare quality and efficiency through self-care remotely. The main difference between smart health care and connected health care is that smart health care refers to autonomously operating solutions, whereas the latter provides users with feedback from clinicians. Connectivity technologies play a key role in smart healthcare systems. There are certain security requirements for smart health care. Some of them are data privacy, data confidentiality, integrity, authentication, service availability, access control, location privacy, etc. Monitoring of environmental factors such as humidity and temperature provides improved quality of smart health care.

38.3 IoT in Smart Health Care

IoT-based healthcare systems can be employed for taking care of children and elderly patients, supervision of chronic diseases, and for self-health care and fitness. Monitoring the health of people has now become one of the important applications of IoT devices. IoT is designed so as to provide an extensive range of healthcare facilities in which all of them delivers a set of solutions for healthcare issues. We require IoT–cloud amalgamation to manage high demands in smart healthcare facilities. IoT plays a key role in providing health care, both globally and locally. It is been necessary to have a system at home that helps people to keep track of their health status without affecting their daily life. They help in managing chronic diseases to monitoring daily physical activities helping in maintaining an individual's fitness goals. It comprises of a system for remotely communicating between systems, applications, and gadgets that help both the patients and doctors to track, examine, and record the important medical data and history of a patient.

IoT can be employed as a tool that supports examining and control within a healthy ecosystem, and data analysis technologies can be used to assist decision-making. The central idea of smart health care is real-time communication and patient monitoring. So, it is essential to have IoT-enabled health monitoring systems for this purpose which also generates alerts to healthcare service providers in case of any emergency. Human-like intelligence is introduced to smart health care with the help of artificial intelligence and deep learning technologies. IoT can also be used for tracking the production and delivery of medical equipments.

Internet of Health Things (IoHT) is basically considered as an IoT-based solution that incorporates a network architecture enabling a connection between a patient and various healthcare services [2]; i.e., IoHT provides interaction between the doctor and the patient through remote access. This network is recognized as Intelligent Healthcare Network (IoThNet). It is mainly utilized for sending and receiving information. Specialized doctors are capable of accessing medical reports and records, and providing suggestions. IoHT includes interconnected devices for remote health monitoring, X-ray, and imaging devices. It also provides smart ambulances or smart clinics like services in case of emergency situations. Moreover, IoT acts as a link between doctor and patient by offering continuous monitoring and consultations through remote access.

38.4 Technologies Used in IoT-Based Health Care

There are numerous technologies available for IoT-based healthcare solutions. Some of the core technologies are listed below.

38.4.1 Wireless Technologies

Wireless technologies are said to be the keystone of any smart healthcare networks. Different wireless technologies are used to gather and send data through specific communication protocols such as Wi-Fi, Bluetooth, Global Positioning System (GPS), Wireless Local Area Network (WLAN), Wireless Personal Area Network (WPAN), 6LoWPAN, Radio Frequency Identification (RFID), ZigBee, and Near-Field Communication (NFC) plays a key role in information exchange among different physical entities within the healthcare network.

38.4.2 Wearable Technology and Medical Sensors

Wearables are smart devices with their supporting environment as either the human body or a piece of clothing. Wearable medical devices have the capability to collect, store, process, and analyze data to provide necessary feedback and generate alerts in case of any emergency situations. Developments in the field of various communication technologies and wearables have led the way to the evolution of numerous instantaneous healthcare monitoring systems. Wearable devices are designed in such a way that they can monitor the vital health parameters of the users and to provide the necessary support and medical aids. These are mainly used for patients with temporary or permanent disabilities, elderly people, or infants. A common use of these wearable devices is monitoring the day-to-day physical activities of the users. IoT wearable devices have reached such a level that it has merged itself making everything possible enabling patients to record their health information on mobile applications.

Sensors on the patient's clothes collect data and provide a digital electrical output that can be used for monitoring their health parameters. A sensor is a small analytical device that combines with a biological component that creates event recognition [1]. Based on the monitoring systems, sensors or actuators vary. It plays an important role in IoT-based health care as they are accountable for gathering data from any embedded gadgets or machines. These sensors can discover physical, chemical, and biological signals and provide a method for their measurement and recording. They can gather and send data, such as ECG, location, and humidity, to the gateway (i.e., a device that converts one protocol or format to another) via a specific communication protocol such as Wi-Fi, Bluetooth, or 6LoWPAN. Sensors attached to the human body can be used for monitoring physiological or biomechanical conditions that include heart rate and muscle activities, respiration, body temperature, blood pressure, and body posture, motion, and acceleration. The output of smart sensors and IoT devices is usually complicated, and hence, they need advanced technologies like deep learning, cloud computing, and big data analytics for its processing. Temperature sensors, ECG (electrocardiogram), blood pressure, heart rate, etc., are commonly used medical sensors.

38.4.3 Body Sensor Network

Body sensor network (BSN) is a wireless sensor network (WSN) that is used for examining the human body. It is an interconnected assembly of wearable (programmable) sensor nodes that can interact with each other, and also with other smart devices and other ambient sensors. These sensor nodes have the capability of computation, storage, wireless transmission, and sensing. BSNs use cloud computing-based frameworks for adaptable storage and scalable processing facilities. Even though BSN-based systems have a vast range of applications, these can be used for continuous and non-invasive monitoring of essential signs, as tiny wireless sensors are positioned on the skin and sometimes implanted with the clothes. This helps in the early detection and diagnosis of diseases. Data related to human body movement, body temperature, heart rate, skin conductivity, brain and muscle activities, and biomarkers are commonly sensed using these sensors. BSNs are also used in various application areas such as e-Sport, e-Fitness, e-Wellness, and e-Social.

38.5 Edge–Fog–Cloud Network Computing

The architecture of edge–fog–cloud network comprises edge computing, fog computing, and cloud computing.

38.5.1 Edge Computing

Edge computing is a computing framework that exists at the edges of network or data sources. It is a disseminated information technology (IT) infrastructure where the client data is processed at the network edge (boundary). It performs this as adjacent as possible to the originating source. The concept of edge computing came from the ideas of mobile computing, the dropping cost of computer parts, and the number of networked gadgets in IoT, as per studies. The time-sensitive data in this architecture may get processed either at the origin by smart device or forwarded to an intermediate server which is situated in adjacent geographical proximity to the client depending on its implementation. Later, the data which is less time-sensitive will be forwarded for further processing to the cloud.

38.5.2 Fog Computing

CISCO invented the concept of fog computing which permits software applications to run on the edge of the network devices rather than on data centers of cloud computing. Fog computing (FC) is a computing structure that allows the data and applications to be located anywhere between the data source and the cloud. It is also known as fogging. It is an extension of cloud computing in which the data, storage, and applications are on a distant server. FC acts as an intermediate between IoT devices and cloud computing forming a cloud–fog–thing network. Fog node is the main component in this network. It is located in a smart router or a gateway device allowing the data to be processed on these smart devices. Hence, only the necessary data are transmitted to the cloud. FC is anticipated to assist an extensive range of IoT applications, incorporating inter-device data sharing, wearable cognitive assistance, editing and sharing of videos, vehicular systems, etc.

38.5.3 Cloud Computing

Cloud computing is liable for systematically handling, stowing, and studying the entire data gathered by the system that is used for health monitoring purposes (i.e., cloud computing integrated with IoT-based healthcare technology) as Internet plays a key role in today's technology. It utilizes the remote server networks provided on the Internet for storing, managing, and processing data instead of a local server. Cloud storage consists of a large amount of repository to stow all the intermediate and ultimate outcomes, the medical information of each user, emergency notifications, etc. Memory plays a significant role in smart health care as it requires stowing user-related particulars. The incorporation of cloud computing with IoT technologies in health care provides access to allocated resources, offers facilities based on requests in the network, and executes operations consistent with the varied requirements.

38.6 Big Data Analytics

Big data analytics is a complex process of collecting, organizing, and analyzing large sets of data (called big data) to discover information—such as hidden patterns, unknown relations, market trends, and customer selections that provide decision-making possible. Data mining (also known as Knowledge Discovery or Knowledge Extraction) is simply referred to as the process of extracting functional data from a huge amount of raw data. It is a procedure that uncovers patterns in huge data sets which comprises of techniques that intersects machine learning, statistics, and database systems. This can be applied to any kind of data. For example, data in Data Warehouses, Transactional Databases, Relational Databases, Multimedia Databases, Spatial Databases, Time-series Databases, or World Wide Web. Temporal mining is a data mining technique used for taking out data sets in time series pattern (TSP). Here, the extraction of useful data from fog data services takes place instantaneously. Time-sensitive health parameters of patients used for remote health monitoring are extracted using the temporal mining technique.

38.7 Medical Applications of IoT and Related Work

There are several applications and works available that are related to IoT-based health care. We can divide the applications of IoT into single-condition and clustered-condition applications [3]. A single-condition application refers to a particular disease or any bad state of health like monitoring of blood sugar level, ECG, blood pressure, body temperature, and oxygen saturation. A clustered-condition application refers to numerous diseases or entire conditions. Some of the IoT-based healthcare applications are as follows:

- Blood pressure monitoring
- Pain monitoring
- ECG monitoring
- Sleep monitoring
- Heart rate variability monitoring
- Nutrition monitoring
- Pathology monitoring
- Cardiac auscultation monitoring, etc.

Now, we can see each of these applications and their related works.

The first application we are discussing here is blood pressure monitoring. Blood pressure (BP) problems are found in people if there is a variation in the normal flow rate of blood pumping out from the heart. Hypertension, also five known as high blood pressure, is a worldwide health problem that is produced on account of exalted BP in the arteries. Several circumstances leading to this elevation of BP involve an unhealthy diet, inadequacy of exercises or physical activities, human emotions, immediate surrounding conditions, and geographical locations. Long-term high BP leads to heart failure, chronic kidney diseases, vision loss, etc. IoT healthcare systems build on the cloud computing platform are trending over the years, which helps patients to monitor and control their blood pressure using IoT devices. The data from IoT-based wearable medical sensors are stowed at the cloud for further computation and examination.

A feasible solution for continuous cuffless blood pressure measurement and calibration that has broad application prospects in the fields of athlete scientific training and medical care was developed in 2020 [4]. In their work, they have developed a blood pressure measuring and calibrating system that was based on a mobile crowdsensing platform named CrowdOS. The system was capable of continuously collecting the blood pressure from the patient, and then they have used the regression model for calibrating the collected blood pressure.

Next application is pain monitoring, which is an important application in treating patients. The detection of human emotions is important for providing emotional care that improves the quality of life. Direct communication with patients or traditional methods used may not provide proper interaction. Mainly children, elderly, and mentally ill people need this type of interaction. Facial expressions are a behavioral indication of pain. Since the experience of pain causes changes in facial expressions, they can be used for assessing human pain as an automatic tool. Thus, it can be used as a substitution method to the traditional self-report methods and for those who are ineffectual to do so, for instance, patients in the intensive care unit (ICU) and infants. Parents often monitor their infants' facial expressions, as they provide information about their health state.

In 2020, an automated acute nociceptive pain recognition system was proposed to objectively measure nociceptive pain using physiological signals and a hybrid deep learning network [5]. This hybrid deep learning network that was used in the proposed system consists of a convolutional neural network that is capable of extracting the essential information related to the pain that is obtained from the physiological signals. The featured matrix that is extracted is then given to the LSTM network for further concatenation of the features. BioVid heat pain database was utilized for this system. Another new postoperative pain assessment model based on pulse contour analysis was developed and also evaluated its effectiveness in postoperative pain assessment [6] in 2019. In their study, they have extracted several candidate features from PPG waveform and then they have developed a model that assesses the pain that is based on multiple logistic regression with a combination of these features.

The next application in our discussion is ECG monitoring. Electrocardiogram (ECG) is an examination that measures the electrical activity of our heart that shows whether it is working normally or not. Sensors attached to the skin are used to detect the electrical signals produced when each time our heart beats. These electrical signals are utilized to monitor the patient's health state by the doctors. Traditionally, ECG is extensively used in clinics by positioning electrodes (up to 12) to the chest. This arrangement is not wearable. Here smart devices came as a solution to this problem.

There are numerous devices and solutions in practice for remote ECG examining purposes. In 2018, an ultra-low-power wrist-worn wearable was proposed for examining the user's heart rate via ECG [7]. It uses some extracted heart rate variability readings and the noticeable morphological components of the ECG waveform. In this, two electrodes are used. One electrode is used for providing ECG monitoring by asking the users to touch the wearable face with a finger. The other one is located at the back of the wearable coming in contact with the wrist which allows it to have connection points on either side of the heart for recording traces of ECG with large amplitude. Another ECG monitoring system based on IoT embedded platform was proposed in 2020. In this method, the features from ECG signals of patients are taken out using a modified version of discrete wavelet transform (DWT) that is implemented using fast Fourier transform (FFT) [8]. These features are then classified with the help of twin support vector machine (TSVM) classifiers that are based on particle swarm optimization (PSO). This method can be implemented as a warning system that can be employed in both hospitals and remote monitoring systems to provide users with a healthier life free from cardiac diseases.

Sleep monitoring is another important application of IoT since nowadays many people are suffering from sleep disorders. Sleep is a natural and periodic state of rest of both mind and body. It is a state of muscle relaxation and reduced interactions with the surroundings. Good sleep is required for the proper functioning of our body. Any interruption of the normal sleep pattern is often referred to as sleep disorders. There are several varieties of sleep disorders, namely insomnia, sleep apnea, obstructive sleep apnea, etc. Obstructive sleep apnea (OSA) is a menacing respiratory disorder that happens in the course of sleep. It affects the quality of life causing behavioral and personality disorders, etc.

There are innumerable systems available for detecting OSA. A wearable in-ear electroencephalography (ear-EEG) for overnight sleep monitoring as a 24/7 continuous and unobtrusive technology for sleep quality assessment in the community was proposed in 2020 [9]. In their study, 22 healthy participants were selected for taking part in overnight sleep monitoring study. The overnight sleep monitoring was done with the help of simultaneous ear-EEG and conventional full PSG recordings. The features that were extracted were used for automatic prediction of sleep stage that was done using supervised machine learning.

Heart rate variability monitoring is yet another important application of IoT. Heart rate can be simply interpreted as the rate at which heart beats. It is measured in terms of the number of heart beats per minute (bpm). The heart rate varies throughout a person's lifetime, according to their age, fitness, and even whether they are frightened or not. This may also vary according to the basic physical needs of body, including oxygen intake and carbon dioxide expulsion. Heart rate variability (HRV) deals with the disparity in the beat-to-beat interval, i.e., the difference in the time interlude between heart beats. It indicates the current heart-related abnormalities and a warning of threatening cardiac diseases. HRV can be used for six diagnosing many intricate cardiac diseases, for instance, arrhythmia, myocardial ischemia, and long QT syndrome.

An automated remote cloud-based heart rate variability monitoring system [10] was proposed in 2018 for this purpose. Web applications were developed by them to extract medical data and some personal particulars like gender and age from the user's wearable sensors. These systems then alert the doctors in case of emergency after detecting any abnormalities in the heart. Homomorphic encryption (HE) technique was used for data encryption to assess the security of remote monitoring. The proposed system was efficacious, reliable, and flexible. It could also extract heart abnormalities from multiple patients simultaneously.

Nutrition monitoring, another important application of IoT plays a key role in today's busy life. Nutrition is the science that deals with the proper intake of food required for better health in all living organisms. It also includes the absorption, metabolism, and excretion of the consumed food. In other words, nutrition is the science of consuming and utilizing foods. A properly balanced diet is very important since nutrient deficiency leads to serious health issues in both children and adults. Imbalanced nutrition in newborns and youngsters can lead to health issues such as weak immunity, cognitive disorders, enfeebled skeletal structure, thinning hairline, and bleeding gums in adulthood. Nutrition imbalance may be either due to undernourishment (i.e., the sufficiency of nutrients in the consumed food) or due to overeating (i.e., excessive consumption of non-nutrient-rich food). So, diet monitoring is an essential part of the healthcare system for maintaining a proper diet and a healthy lifestyle.

An automated nutrition monitoring system, named SmartLog, was proposed in 2018 [11]. It was a consumer electronic system that was used for diet monitoring in smart homes for improving the quality of life. SmartLog consisted of Wi-Fi-enabled sensors for the qualification of food nourishment, and a smartphone application for collecting the nutritional contents of the food ingredients. An open IoT platform was used for performing data analytics and repository operations. It also made suggestions to reduce the risk of an imbalanced diet in addition to the nutritional content in meals. It was mainly designed for infants but can also be utilized for adults by enlarging its food database stored in the cloud storage.

The next application under discussion is pathology monitoring. Pathology is the study of the causes and consequences of disease or injury. It refers to the study of disease in general. Pathology acts as a bridge between science and medicine. Electroencephalogram (EEG) is an examination that is used to perceive issues that are associated with the electrical activities of the brain. This is done with the help of small metal disks with narrow wires attached to the scalp, and subsequently, signals are sent to a computer to store the outcomes. EEG is a commonly used technique for this purpose as it provides economical and non-interfering nature. Certain diseases like epilepsy and stroke which are brain-related diseases are diagnosed using EEG. Patients with such diseases require an instant response in case of emergency because any detain can be life-threatening. In such cases, a smart healthcare infrastructure that helps in examining the patient's status is necessary.

A cognitive smart healthcare model was proposed in 2019 for pathology detection and monitoring [12]. In this framework, communications within smart cities were made possible by using smart sensors, and proper decisions were made with the help of deep learning. The state of the patients was determined based on the sensor readings such as facial expressions, speech, gestures, and EEG. The obtained EEG signals were categorized as pathological or normal with the help of deep learning-based real-time decisions made by the cognitive module. The results are shared with caregivers so they can take necessary actions.

Cardiac auscultation monitoring is the last application of our discussion under this section. Heart sounds are the noises that are generated when heart beats and resultant blood flows through it. It delivers physiological and pathological evidence about health. Cardiovascular disease is considered to be one of the major health problems as any delay in the proper diagnosis and treatment leads to death. ECG is the most admired method used for examining the proper working of the heart by detecting the electrical signals. In earlier days, physicians practiced heart auscultation by directly placing their ears to the patient's chest. It helps in the early detection of heart diseases since the heart sound contains a lot of information that helps in diagnosis. Electronic stethoscope uses ambient noise reduction (ANR) technology that makes heart auscultation more convenient with reduced noise. The most extensively used substance of the heart sensor is piezoelectric.

A wireless cardiac auscultation monitoring system was proposed in 2018 for continuous cardiac monitoring and analysis of people without any need for manual healthcare services [13]. It involves the integration of Hilbert–Huang transform that was used to study coronary artery disease, and double-threshold method for preprocessing, and extraction of heart sound signal features and physiological parameters. Hidden Markov model was utilized for the categorization purpose. In other words, preprocessing, segmentation, and clustering technique was performed for the proper elucidation of available health information.

38.8 Conclusion

It is transparent that the future of IoT is enormous, and surely, it will reach a stage where people will not be able to imagine life without IoT devices. Progress in the field of various wireless technologies and wearables has led the way to the evolution of numerous instantaneous healthcare monitoring systems. IoT has also enabled the improved performance of the traditional healthcare system and making it adjustable with smart devices. This paper focuses on smart health care and the role of IoT in smart health care. The various technologies that are utilized in IoT-based health care and certain ways of using IoT devices to enhance the health of people were also discussed. Certain examples of such devices were discussed that helps every individual, mainly children, elderly, and physically disabled people in maintaining their physical wellness. The objective of this paper was to give a comprehensive outlook rather than a deep understanding of IoT usage in the healthcare sector and the various healthcare monitoring systems available. Both research works and devices that are commercially and currently available for the purpose of examining and exploring health conditions were mentioned. The discussion on this paper is intended to impart a basis for each and everyone who is working in the field of IoT and for further advancements of various healthcare technologies.

References

1. Sundaravadivel, P., Kougianos, E., Mohanty, S.P., Ganapathiraju, M.K.: Everything you wanted to know about smart healthcare: evaluating the difference technologies and components of the internet of things for better health. IEEE Consum. Electron. Mag. **7**(1), 18–28 (2018)
2. Rodrigues, J.J.P.C., Segundo, D.B.R., Junqueira, H.A., Sabino, M.H., Prince, R.M., Al-Muhtadi, J., de Albuquerque, V.H.C.: Enabling technologies for the internet of health things. IEEE Access **6**, 13129–13141 (2018)
3. Riazul Islam, S.M., Kwak, D., Humaun Kabir, Md., Hossain, M., Kwak, K.-S.: The internet of things for health care: a comprehensive survey. IEEE Access **3**, 678–708 (2015)
4. Dong, Z., Yian, Z., Lanqing, W., Junhua, D., Jiaxuan, H.: Continuous blood pressure measurement platform: a wearable system based on multidimensional perception data. IEEE Access **8**, 10147–10158 (2020)
5. Subramaniam, S.D., Dass, B.: Automated nociceptive pain assessment using physiological signals and a hybrid deep learning network. IEEE Sens. J. **21**(3), 3335–3343 (2020)
6. Seok, H.S., Choi, B.-M., Noh, G.-J., Shin, H.: Postoperative pain assessment model based on pulse contour characteristics analysis. IEEE J. Biomed. Health Inform. **23**(6), 2317–2324 (2019)
7. Beach, C., Krachunov, S., Pope, J., Fafoutis, X., Piechocki, R.J., Craddock, I., Casson, A.J.: An ultra-low power personalizable wrist-worn ECG monitor integrated with IoT infrastructure. IEEE Access **6**, 44010–44021 (2018)
8. Raj, S.: An efficient IoT-based platform for remote real-time cardiac activity monitoring. IEEE Trans. Consum. Electron. **66**, 106–114 (2020)
9. Nakamura, T., Alqurashi, Y.D., Morrell, M.J., Mandic, D.P.: Hearables: automatic overnight sleep monitoring with standardised in-ear EEG sensor. IEEE Trans. Biomed. Eng. **67**(1), 203–212 (2020)
10. Hussein, A.F., Arun Kumar, N., Burbano-Fernandez, M., Ramrez-Gonzalez, G., Abdulhay, E., De Albuquerque, V.H.C.: An automated remote cloud-based heart rate variability monitoring system. IEEE Access **6**, 77055–77064 (2018)
11. Sundaravadivel, P., Kesavan, K., Kesavan, L., Mohanty, S.P., Kougianos, E.: Smart-log: a deep-learning based automated nutrition monitoring system in the IoT. IEEE Trans. Consum. Electron. **64**(3), 390–398 (2018)
12. Amin, S.U., Shamim Hossain, M., Muhammad, G., Alhussein, M., Abdur Rahman, Md.: Cognitive smart healthcare for pathology detection and monitoring. IEEE Access **7**, 10745–10753 (2019)
13. Ren, H., Jin, H., Chen, C., Ghayvat, H., Chen, W.: A novel cardiac auscultation monitoring system based on wireless sensing for healthcare. IEEE J. Transl. Eng. Health Med. **6** (2018)

Chapter 39
Substrate Integrated Waveguide Antenna for High-Frequency Application

Inderpreetkaur, Banani Basu, Anil Kumar Singh, and Suchita Saxena

Abstract In the given research paper, a low-cost antenna is designed which can be used for high-frequency applications. The proposed antenna operates at the frequency of 18.7 GHz (resonant frequency) with bandwidth ranges from 18 to 22 GHz. The proposed Substrate Integrated Waveguide antenna of rectangular shape is considered to achieve the wider bandwidth with microstrip feedline. The overall size is 16 * 16 mm^2 with thickness 1.6 mm and for designing this patch antenna, substrate material FR4 epoxy is used which is commercially available. Using high-frequency structure simulator (HFSS) tool, the value of return loss, gain, front-to-back ratio, radiation efficiency and current distribution has been analyzed. The designed SIW antenna achieves the return loss of 29.8 dB at the resonant frequency.

39.1 Introduction

In the present time, communication is mostly human-driven but in the future, and communication will be with the help of large numbers of machines [1–3]. Today Internet of things (IoT) is playing a vital role in the field of communication and in future also IoT and machine-type communication (MTC) will continue to play an important role in controlling the communication traffic. Machine-type communication plays an important role in 5G systems. MTC denotes the broad area of

Inderpreetkaur
Department of Electronics and Communication Engineering, FET, MJP Rohilkhand University, Bareilly, Uttar Pradesh, India

B. Basu
Department of Electronics and Communication Engineering, NIT Silchar, Silchar, Assam, India

A. K. Singh
Department of Electronics and Instrumentation Engineering, FET, MJP Rohilkhand University, Bareilly, Uttar Pradesh, India

S. Saxena (✉)
Department of Electronics and Communication Engineering, MIT Moradabad, Moradabad, Uttar Pradesh, India

T. Senjyu et al. (eds.), *IOT with Smart Systems*, Smart Innovation, Systems and Technologies 251, https://doi.org/10.1007/978-981-16-3945-6_40

wireless communication with sensors, actuators, physical object, and other devices not directly operated by humans. It is seen that need of wireless communication is increasing rapidly, at a very fast rate, in various applications. Moreover, a number of users using this technology are increasing resulting in the spectrum shortage problem. Thus, study of millimeter wave is becoming important for applications like wireless mobile cellular networks, wireless local and personal area networks (LAN and PAN), wearable networks, and communication in vehicles and radar.

As we know that without an antenna system any communication system is incomplete, so antenna designing is gaining much attention and also becoming very challenging task. For transferring data at high rate, the need of wideband directional microstrip patch antennas having constant gain and high bandwidth is increasing. For this, designing a low-profile directional patch antenna becomes very challenging. Large research has been done on increasing the bandwidth of the antenna. It is even more difficult to achieve high gain in an antenna which works on multiple frequencies. To overcome this problem, substrate integrated waveguide (SIW) is nowadays replacing the conventional waveguide for designing of antennas. Also, many other important advantages include integrating all the components on the same substrate with low insertion loss and low profile.

39.2 Antenna Design

The proposed design of the antenna uses FR4 substrate. The geometry of the antenna has size of $16 \times 16\ \text{mm}^2$ with a substrate having thickness of 1.6 mm, and dielectric constant, ε_r of 4.4. The proposed structure consists of a slot of rectangular shape of length S_l and width S_w (Fig. 39.1).

The radius of the vias of SIW is optimized rigorously. A remarkable problem which arises in the microstrip antenna is the generation of surface waves which results in the decrease in the antenna efficiency. The antenna having SIW structure effectively controls these surface waves and provide self-contained electromagnetic

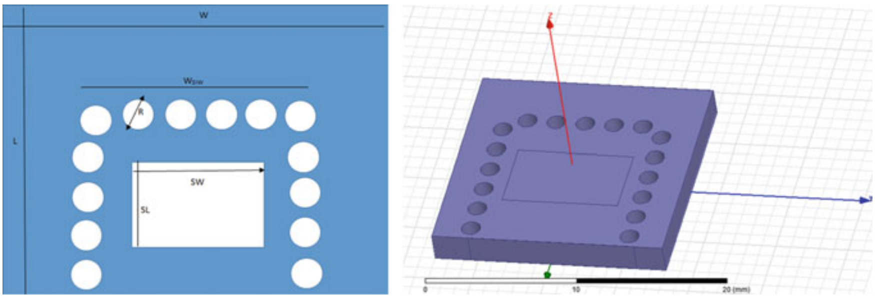


Fig. 39.1 Design geometry of the proposed structure

Table 39.1 Antenna dimensions

Parameters	Value (nm)
L_{SIW}—Length of SIW	14
W_{SIW}—Width of SIW	14.5
L—Length of patch	16
W—Width of patch	16
S_L—Length of slot	8
S_W—Width of slot	5
p—Spacing between centers of vias	1.8
r—Radius of via	0.65

shielding. The dimensions of slot are optimized to enhance the performance of the SIW antenna. The dimensions of the antenna are given in Table 39.1.

39.3 Parametric Analysis

39.3.1 Effect of Radius of Via

Antennas bandwidth can be strengthened by two techniques, i.e., thickness of the antenna is increased, which further enhance the effect of surface wave, and by lowering down the permittivity but the value of dielectric constant is limited to 1. Here, in this proposed work, bandwidth is enhanced because of SIW structure in which the holes resonate with each other to merge the modes and merging of modes enhance the bandwidth. In Fig. 39.2 the variation of return loss S_{11} with frequency

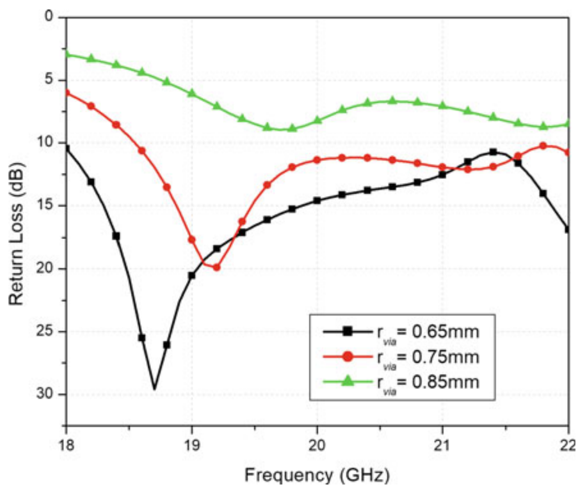


Fig. 39.2 Return loss varies with frequency with change in the value of r_{via}

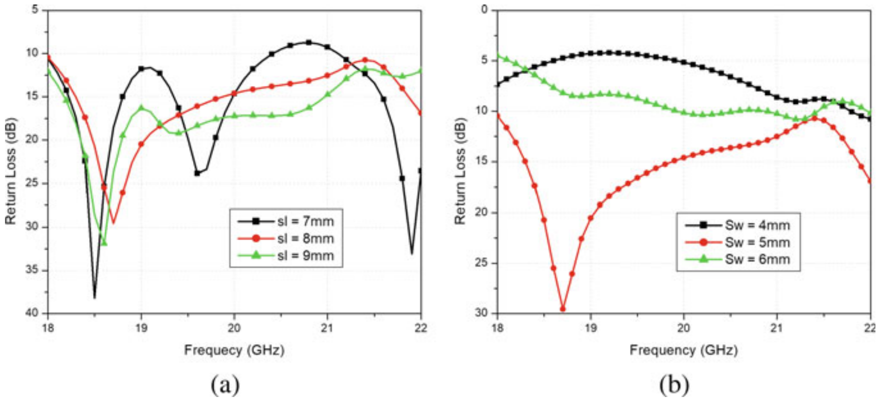


Fig. 39.3 **a** Variation of return loss, S_{11} with frequency on changing the value of S_l and **b** variation of return loss, S_{11} with frequency on changing the value of S_w

for different r_{via} values is shown. Radius of designed via is optimized from 0.65 to 0.85 mm, and it is clear from the figure that at $r_{via} = 0.65$ mm, there is maximum impedance matching. Also, at $r_{via} = 0.65$ mm, the bandwidth is maximum which ranges from 18.2 to 22 GHz.

39.3.2 Effect of Dimension of Slot

Figure 39.3a, b shows the change in the value of return loss with frequency on changing the values of S_l and S_w, respectively. Based on the observations derived from Fig. 39.3b it is clear that the width of the slot defines the impedance matching of the structure. The passing hole in the substrate generates the parasitic capacitance of value C and parasitic inductance of value L.

$$C = 1.41\varepsilon r H \frac{d_1}{d_2 - d_1} \tag{39.1}$$

$$L = 5.08H\left[\ln\left(\frac{4H}{d}\right) + 1\right] \tag{39.2}$$

The resonance which is generated between C and L of a single hole or that produced between holes greatly affects the performance of the excitation antennas. The expression for series resonance is

$$f_0 = 1/2\pi\sqrt{LC} \tag{39.3}$$

$$Q = w_0 \frac{L}{R} = \frac{1}{w_0 C R} = \frac{1}{R}\sqrt{L/C} \tag{39.4}$$

The expression for parallel resonance is

$$f_0 = 1/2\pi\sqrt{LC} \tag{39.5}$$

$$Q = w_0\, C R = R/w_0\, L = R\sqrt{C/L} \tag{39.6}$$

39.4 Result and Discussion

Antenna performance is measured by return loss. The ratio of the reflected signal to the launched signal is called return loss. More the signal is reflected, less it is delivered to the load. Signal reflection is due to the discontinuity of the transmission line. The return loss of the proposed SIW antenna is illustrated in Fig. 39.4a and is observed that return loss obtained at resonance frequency of 18.70 GHz is 30 dB.

In studying electromagnetics, antenna's power gain is a key factor that defines the directivity and electrical efficiency of an antenna. In the case of antenna transmitter system, the calculated value of gain shows how well the antenna is converting the input power into radio waves moving in a particular direction. Figure 39.4b shows that the graph which is clearly showing that the gain of the proposed antenna is 6.84 dBi at resonance frequency.

Efficiency is defined as the ratio of the total power that is radiated by an antenna to the net power that is accepted by the antenna from the connected transmitter. It is sometimes represented as a percentage (less than 100) and this is frequency

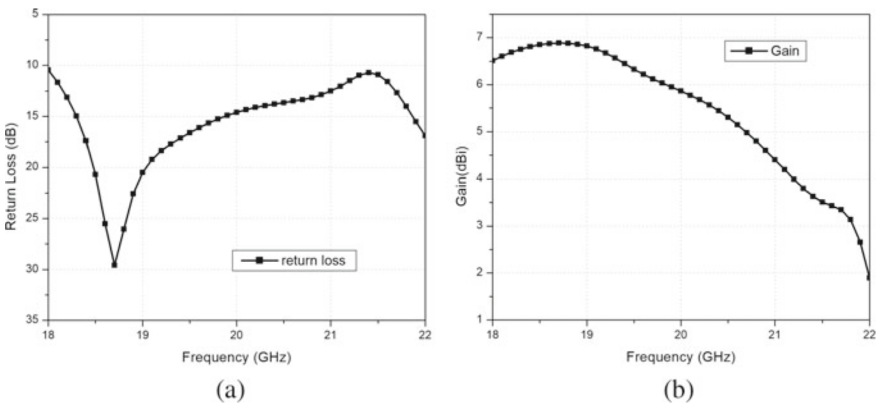


Fig. 39.4 **a** Return loss of the proposed antenna and **b** gain of the proposed antenna

dependent. In Fig. 39.5a, it is shown that the efficiency of the designed antenna is 80.3%. The term front-to-back ratio is defined as the ratio of power gain between the front and back of a directional antenna. It can also be defined as the ratio of signal strength that is transmitted in a forward direction of the antenna to that transmitted in a backward direction of the antenna. The FBR of the proposed work is above 25 dB at resonance and above 15 dBi along the bandwidth. Figure 39.6 shows the measured surface current density at the resonance frequency of 18.70 GHz.

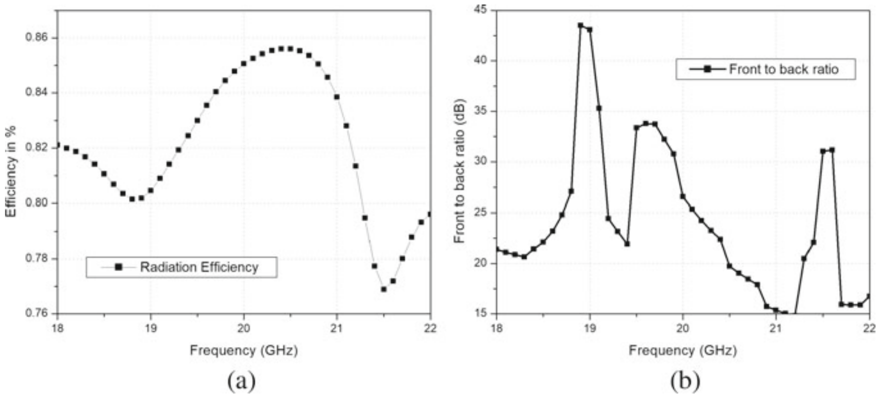


Fig. 39.5 **a** Radiation efficiency of the designed antenna and **b** FBR of the designed antenna

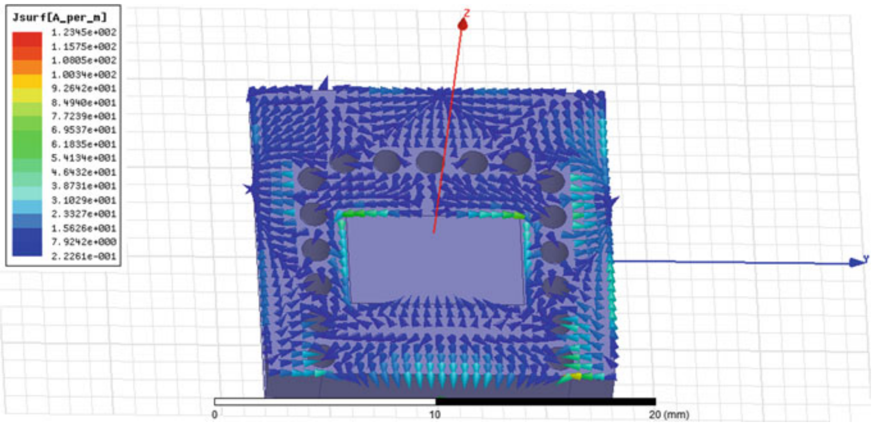


Fig. 39.6 Surface current density at the resonance frequency at 18.70 GHz

39.5 Conclusion

The proposed antenna has been analyzed with various characteristics and has been successfully simulated. SIW is introduced to increase the gain and improve the efficiency. Due to use of dielectric into the waveguide structure, dielectric loss will occur compared to air used in normal rectangular waveguide; moreover, it depends on frequency and hence mm wave application of SIW needs to be carefully considered. The closed structure of the SIW also prevents undesired backward radiation. The total radiation size of 16 * 16 mm^2 has been achieved having return loss value of -29.8 dB at the resonance frequency of 18.7 GHz. The gain of 6.84 dBi is achieved at resonance frequency. By optimizing the radius of via and slot dimensions, maximum radiation efficiency achieved is 82% with front-to-back ratio (FBR) of 25 dB at resonance and more than 15 dB for the complete bandwidth.

References

1. Shi, Y., Liu, J., Long, Y.: Wideband triple-and quad-resonance substrate integrated waveguide cavity-backed slot antennas with shorting vias. IEEE Trans. Antennas Propag. **65**(11), 5768–5775 (2017)
2. Wi, S.H., Lee, Y.S., Yook, J.G.: Wideband microstrip patch antenna with U-shaped parasitic elements. IEEE Trans. Antennas Propag. **55**(4), 1196–1199 (2007)
3. Zhou, S.G., Huang, G.L., Chio, T.H., Yang, J.J., Wei, G.: Design of a wideband dual-polarization fullcorporate waveguide feed antenna array. IEEE Trans. Antennas Propag. **63**(11), 4775–4782 (2015)
4. Sun, L., Feng, H., Li, Y., Zhang, Z.: Compact 5G MIMO mobile phone antennas with tightly-arranged orthogonal mode pairs. IEEE Trans. Antennas Propag. **66**(11), 6364–6369 (2018)
5. Mighani, M., Akbari, M., Felegari, N.: A CPW dual band notched UWB antenna. Appl. Comput. Electromagn. Soc. **27**(4), 352–359 (2012)
6. Abdollahi, M.M., DaliliOskouei, H.R., Akbari, M., Mighani, M.: A novel compact UWB antenna with triple band-notches for WiMAX/WLAN/ITU bands. Appl. Comput. Electromagn. Soc. **27**(12), 1014–1021 (2012)
7. Lin, C.C., Jin, P., Ziolkowski, W.: Single, dual and tri-band-notched ultrawideband (UWB) antennas using capacitively loaded loop (CLL) resonators. IEEE Trans. Antennas Propag. **60**, 102–109 (2012)
8. Ansoft High Frequency Structure Simulation (HFSSTM), Ver. 13, Ansoft Corporation, USA (2010)
9. Babaeian, F., Karmakar, N.C.: A high gain dual polarized ultra-wideband array of antenna for chipless RFID applications. IEEE Access **6**, 73702–73712 (2018)

Chapter 40
A Comparative Study of Recent Feature Selection Techniques Used in Text Classification

Gunjan Singh and Rashmi Priya

Abstract As we all know, handling large amounts of data is a problem these days. Despite having so many resources to store, train and process the data, still it is required to reduce these datasets in order to reduce computational complexity, save time, cost and retrieve valuable information from large text documents. The presentation of a machine learning algorithm relies upon the dataset utilized. When the dataset is large, the learning algorithm tries to accommodate all the features which increases the dimensionality of the data. This high-dimensional data is not useful as it might contain irrelevant and redundant features. It becomes important to remove these features. Thus, pre-processing of data is required to compress and analyse the dataset for the purpose of text classification (TC). This can be achieved by using feature selection (FS) techniques. The fundamental goal of FS techniques is to acknowledge pertinent features and to get rid of repetitive attributes w.r.t. high-dimensional data (Shroff and Maheta in 2015 International Conference on Computer Communication and Informatics (ICCCI), Coimbatore, India, 2015 [1]). Nowadays, major FS methods use optimization algorithms (Brownlee in https://machinelearningmastery.com/. 23 Dec 20 [Online]. Available: https://machinelearningmastery.com/tour-of-optimization-algorithms/. Accessed 3 Feb 2021 [2]) to get an ideal component subset from high-dimensional information from feature space which decreases computational expense and builds classifier precision. Some of the recent feature selection techniques have been discussed in this paper which can prove to be useful for text classification (TC).

40.1 Introduction

The development of Internet innovations has prompted a gigantic expansion in the quantity of electronic records utilized overall [3]. Text reports accessible on the Internet contain the high business esteems. The mining of valuable data covered up in content advantages from text classification which intends to dole out content reports

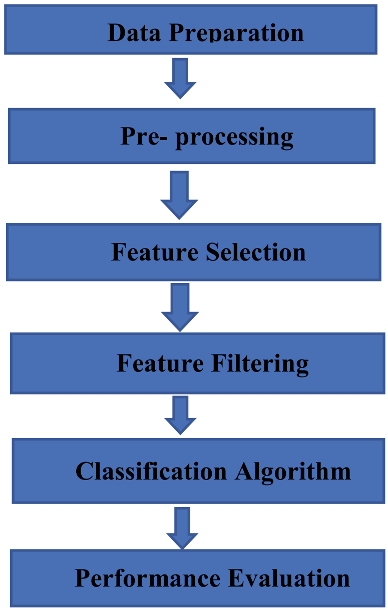
G. Singh (✉) · R. Priya
G.D. Goenka University, Gurugram, India

T. Senjyu et al. (eds.), *IOT with Smart Systems*, Smart Innovation, Systems and Technologies 251, https://doi.org/10.1007/978-981-16-3945-6_41

with obscure categories to a set number of categories by specific categorizations [4]. The supervised learning [5] approach, TC allocates the content archive to fitting classes [5]. TC is a traditional issue in NLP to allot tags to text-based elements, for example sentences, questions, passages and archives [6]. There may be two ways for TC, i.e. manual interpretation or via automated labelling [6]. As per the developing size of text information of contemporary operations, automated TC is turning out to be progressively significant [6]. Text characterization has been engaged with different applications, for example, distinguishing proof of specific points and grievance criticism of specific administrations in informal organization, TV projects and sound music sorts order in amusement content choice, mishap discovery for illness flare-ups and industry oil mishaps, assessment of exposition answers and arrangement of materials of various perusing levels for understudies in instruction, acknowledgment of extreme paediatric asthma from experience notes in biomedicine [4], email sifting [7, 8], email grouping [9, 10], page grouping [11, 12], creator sex ID and theme ID [5]. The size of the component vector of the content collection is lofty [5]. The execution of text classifiers is impacted by high-dimensional feature size. [5]. It is important to manage these text documents effectively and efficiently. This can be helped by feature selection. The process of text classification can be shown in Fig. 40.1.

Feature selection involves selecting relevant features from the dataset or it can also be taken as a subset of the original data. Its purpose is to remove redundant and non-informative variables from the dataset. It is divided into four categories,

Fig. 40.1 Workflow of text classification [13]



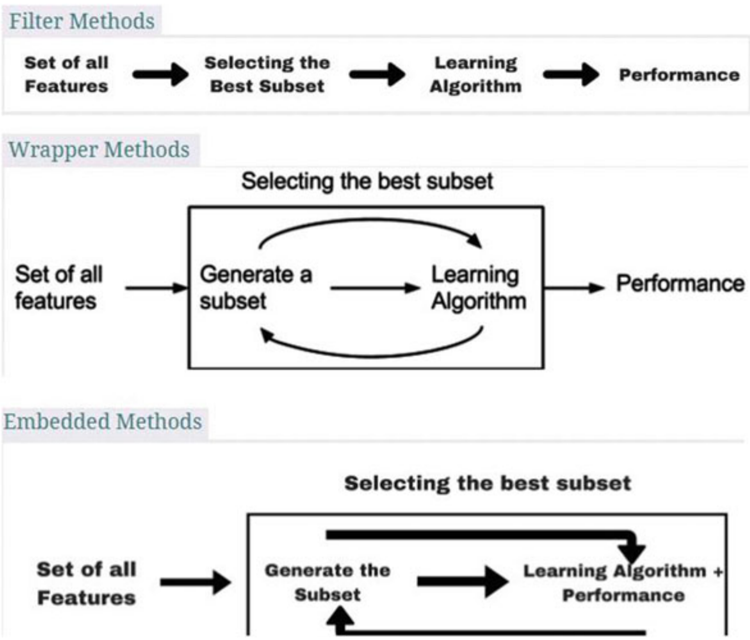


Fig. 40.2 Feature selection model [15]

namely the filter model [14], wrapper model [14], embedded model [14] and the hybrid model [14] (Fig. 40.2).

The filter method quickly wipes out unimportant features. In any case, it neglects to dispose of excess between the features. The various filter methods are information gain, chi square test, fisher score, correlation coefficient and variance threshold [16]. The wrapper technique gives better performance; however, it is time-consuming. The different wrapper methods are recursive feature elimination, sequential feature selection algorithms and genetic algorithms [16]. The embedded techniques makes FS and learning algorithms interleaved. The different embedded methods are L1 (LASSO) regularization and decision tree [16]. The hybrid model exploits the good qualities of two or more algorithms that can solve the same problem, efficiently to make the overall model better than the individual algorithms.

Some of the new strategies for feature selection for text classification will be discussed in this paper which are improved sine cosine algorithm (ISCA) [14], quadratic programming-based numerical optimization approach [17], normalized difference measure and binary Jaya optimization (NDM-BJO) [5], Hebb rule-based feature selection (HRFS) [4], active learning-based support vector machine (SVM) [18], Lion fuzzy neural network (LFNN) [19], relative document-term frequency difference (RDTFD) method [3], limited-memoryBroyden–Fletcher–Goldfarb–Shanno (ML-BFGS) [20]. This paper comprises four sections. Section 40.1 comprises the introduction including, Sect. 40.2 comprises a literature review, which has been further divided into eight subsections and a table for comparative analysis of recent feature selection techniques, Sect. 40.3 comprises the results and discussions with six tables and Sect. 40.4 comprises the conclusion and future scope.

40.2 Literature Review

This section further consists of eight subsections, namely improved sine cosine algorithm (ISCA) [14], quadratic programming-based numerical optimization approach [17], normalized difference measure and binary Jaya optimization (NDM-BJO) [5], Hebb rule-based feature selection (HRFS) [4], active learning-based support vector machine (SVM) [18], Lion fuzzy neural network (LFNN) [19], relative document-term frequency difference (RDTFD) method [3], limited-memory Broyden–Fletcher–Goldfarb–Shanno (ML-BFGS) [20] and Table 40.1 to show the relative study of the recent FS techniques, as mentioned.

40.2.1 Improved Sine Cosine Algorithm (ISCA) [14]

ISCA is a feature selection technique which is an improvisation of sine cosine algorithm (SCA). Earlier bag of words (BOW) [14] model for feature selection was in use because of its simplicity; however, it contained redundant and irrelevant features that increase the dimensionality [14]. The limit of this strategy is that it expands the computing time of categorization algorithms, and it corrupts an exactness of outcomes [14]. Accordingly, feature selection strategies become a significant advance in pre-processing of information [14]. The sine cosine is a new continual speculative calculation wherein arrangements are refreshed next to every emphasis dependent on one or the other sine or cosine work function. SCA has been fruitful in many optimization applications [14]. In any case, it isn't great and has a few downsides. It will in general stall out in imperfect regions to get the best presentation. It restricts the investigation of the hunt space [14]. To get the arrangement from the further locales of search space, there should be a calculation [14]. Thus, ISCA was acquainted with improving the broadening of SCA to find new regions for search space [14]. It takes two positions: firstly, position of best arrangement and besides, given irregular situation from search space though, SCA zeros in just on best answer for new arrangement [14]. ISCA permits to evade untimely assembly to improve the exhibition [14].

Table 40.1 Comparison table for recent feature selection techniques

Authors/year published	Techniques	Evaluation metrics	Techniques compared to	Databases Considered	Advantages	Disadvantages
Belazzoug, M. et al. [14] (2020)	Improved sine cosine algorithm (ISCA) [14]	Precision (P) [14] Recall (R) [14] F-measure (F1) [14]	Original SCA, OBL_SCA, Levy_SCA, Weighted_SCA, MFO (Moth–Flam optimization algorithm), GA (Genetic Algorithms) and ACO (Ant Colony Optimization) [14]	Nine Text Collection from Reuters21578, TREC and OHUSMED [14]	This algorithm conceivably be an effective option for tackling feature choice issues [14]	ISCA neglected to get applied to some other component choice issues or complex issues in different areas [14]
Lim, H. and Kim, D.W. [17] (2020)	Quadratic programming-based numerical optimization approach [17]	F1 Measure, Precision (P) and Recall to evaluate classification performance Micro-F1 and Macro-F1 for Multicategory Text Classification [17]	Compared with conventional FS Metrics Chi Square (Avg.), Chi Square (Max), IG, MI (avg) and recent FS methods like IGFSS, t-test and DFS (Distinguishing Feature Selector) [17]	Three datasets: 20NG, Reuters and TDT [17]	This strategy finds the ideal harmony between two conditions for feature selection [17]. The label similarity lessens the quantity of repetitive labels chosen to enhance classification correctness [17]	It experiences high processing time necessities [17]

(continued)

Table 40.1 (continued)

Authors/year published	Techniques	Evaluation metrics	Techniques compared to	Databases Considered	Advantages	Disadvantages
Thirumoorthy, K. and Muneeswaran, K. [5] (2020)	Normalized difference measure and binary Jaya optimization [5] (NDM-BJO) [5]	Precision [5], Macro-averaging recall [5] and Macro- averaging F1 measures [5]	Compared with filter-based NDM, IG, DF, MI and DFS, also Wrapper-based GA and FFY [5]	Four sets of data, namely Web-KB, SMS, BBC and 10Newsgroups [5]	It lessens the high-D attribute expanse to enhance the classifier exactness [5]	The computational complexity of this method was high [5]
Wang, H. and Hong, M. [4] (2019)	Hebb's rule-based feature selection (HRFS) [4]	F-measure [4], ROC area [4] and classification accuracy [4]	Compared to Seven FS based IG, ECE, MI, GI, WOR, DF and NDM; Classification algorithms such as NB, SVM, MNB used in TC.[4]	Six datasets: CarF & CarR from OpinRank Review dataset, CNAE, IMDB, KDC & TTC [4]	It recognized discriminating labels with the perspective of brain cells be-twixt nerve cells [4]	For the convergence of the optimization process, more time is required [4]
Goudjil, M. et al. [18] (2018)	Active Learning Using Support vector machine (ALSVM) [18]	Training Set, Testing Set, Threshold [18]	ALMSVM [18]	Three Datasets: R8, 20 ng, WebKB [18]	It expands the classification accuracy [18] and holds awesome steadiness with all datasets [18]	In order to characterize reports with increased labels, additional labelled tests are required [18]

(continued)

Table 40.1 (continued)

Authors/year published	Techniques	Evaluation metrics	Techniques compared to	Databases Considered	Advantages	Disadvantages
Ranjan, N.M. and Prasad, R.S. [19] (2018)	Lion fuzzy neural network approach (LFNN) [19]	Accuracy & Error [19]	Compared with text classification techniques: I-BP, FI-BP [19] and I-BPLion [19]	Four datasets: 20NG, Reuters, Web-KB and RCV1 [19]	This approach has better accuracy [19]	It is further required for processing to minimize the error rate [19]
Liu, Y. et al. [3] (2020)	Relative document-term frequency difference (RDTFD) [3]	F-1 measure [3]	Six FS methods, namely CHI, IG, TFIG, TTFS, IMTFIDF and GININTF [3]; Three classifiers used: MNB, SVM and SMO [3]	Six spam datasets: PU1, PU2, PU3, PUA, CS & ES [3] from PU123A, CSDMC2010(CS) and Enron-spam3 (ES) datasets [3]	This method was more robust [3]	It failed to improve the running speed [3]
Borhani, M. [20] (2020)	Limited-memory Broyden–Fletcher–Goldfarb–Shanno (ML-BFGS) [20]	Hamming Loss [20], Precision [20], Recall [20] and F1 score [20]	ITCG, BFGS, L-BFGS-B Multi-label classifiers: ML-k-NN (Multi-Label K-Nearest Neighbour) & SVM [20]	Quranic Text [20]	It expands the performance of text characterization information [20]	The computational complexity was high [20]

40.2.2 *Quadratic Programming-Based Numerical Optimization Approach [17]*

It is utilized to decrease choice of repetitive terms by utilizing the term comparability [17]. The need is given to autonomous terms to stay away from repetitive terms [17]. It very well may be utilized to think about a harmony among positioning and closeness terms for suitable term determination from a worldwide viewpoint [17].

40.2.3 *Normalized Difference Measure and Binary Jaya Optimization (NDM-BJO) [5]*

It is a mixture include choice technique comprising of BJO which is wrapper and NDM which is filter-based feature selection method to lessen computational expense and improve text classifier accuracy [5]. Naïve Bayes (NB) and support vector machines (SVM) have been utilized as the classifiers [5].

40.2.4 *Hebb's Rule-Based Feature Selection Technique (HRFS) [4]*

It follows Hebb's standard which is one of the primary neural network learning laws which expresses that in a cell or two, the development cycle happens when the synapse of one cell is close to energize other cell more than once or it participates in termination with the goal of the effectiveness of first cell when the cell that terminates another cell gets expanded [4]. The synapse within neurons can help to recognize unfair categories [20].

40.2.5 *Active Learning Method Using SVM (Support Vector Machine) [18]*

The supervised learning algorithm, SVM requires manual labelling of information tests during preparation which is tedious and mistake inclined errand [18]. Thus, the procedure of active learning is talked about which can diminish the labelling effort by brilliantly choosing the samples to be named without bargaining the accuracy of classification [18]. The cluster of useful examples utilizing backward likelihood given across a bunch of SVM categorizations [18] is chosen by this strategy, and these samples are manually labelled by a specialist [18].

40.2.6 Lion Fuzzy Neural Network (LFNN) [19]

The supervised learning approach, LFNN utilizes factor, sense-based component extraction and fuzzy neural network which is hybrid [19]. For this network model, the feature selection depends on entropy capacity and extraction measure that uses a semantic word handling alongside context-oriented methodology and theme terms to separate important words [19]. It alters backpropagation Lion neural networks (BPLNN) by the steady learning of BPLNN, named LFNN, where it coordinates fuzzy cavort to design the fuzzy neural network (FNN) that is hybrid [19]. The optimization algorithm is received to choose ideal weight and LA for refreshing the loads of fuzzy neural networks [19]. Thus, on the grounds that the unique information base is considered in this where the classifier can get familiar with the model progressively relying upon the new dataset, it is called evolutionary classification [19].

40.2.7 Relative Document-Term Frequency Difference (RDTFD) [3]

RDTFD is a FS technique for TC that depends on searching an unconstrained feature space [3]. The overall feature selection technique involves four stages [3]. Firstly, to add all features to the original features set [3]. Secondly, to calculate the weights of all features [3]. Thirdly, to rank all features in descending order as per their weights [3]. Fourthly, selecting topmost features from feature set [3]. Then again, the RDTFD techniques additionally includes four stages [3]. Firstly, all features to be attached to the native set of features [3]. Secondly, weights of features to be calculated [3] where all features/attributes are detached as a certain or negative subset of attributes according to positive or negative assessments of the weighting [3] to improve the posh significance of the attributes to decrease the relationship inside the attributes [3] and to lessen the exploration scope of attributes expanse and keep up suitable attribute repetition [3]. Thirdly, to rank all the attributes as per loads, the positive attribute subsets are situated in sliding request and negative attribute subsets are positioned in climbing request [3]. Fourthly, the competitor attribute subsets are looked over two self-governing attribute subsets with some search strategies [3].

40.2.8 Limited-Memory Broyden–Fletcher– Goldfarb–Shanno (ML-BFGS) [20]

It is ANN with semi-Newton system for multi-label, multiclass text characterization [20]. This mathematical unconstrained training method [20], the log-loss function's multi-label extension of being utilized in ML-BFGS [20], gives an important chance to text mining and prompts critical improvement in text characterization execution [20]. In order to distribute some or many of the categories to sentences related from various accessible marks, the training process is applied [20]. It has reduced computing expense, quicker intermingling amount and considerable exactness in whole examination [20]. Mentioned, below is Table 40.1 that comprises recent feature selection techniques for text classification, with their advantages, disadvantages, evaluation metrics, databases used and the techniques they have been compared with.

40.3 Results and Discussions

Table 40.2 presents the observational assessment of precision [20], recall [20], F-Score [20], hamming loss [20] and test time [20] tried on artificial neural network for the Quranic corpus. It unmistakably shows that ML-BFGS accomplished high accuracy with elevated review while utilized even as multi-label categorization as strict aggregated data study [20]. It requires less calculation in every iteration and more stockpiling when contrasted with different techniques [20]. It meets in less cycles and results indicated that the methodology is more progressive when contrasted with previous quasi-Newton strategies [20].

ML4BFGS has more exact results than SVM and ML-k-NN with more computational expense [20]. It is the quickest technique for quasi-Newton calculations with least computational multifaceted nature [20]. When contrasted and the consequences of standard semi-Newton strategies on the Quranic dataset, ML4BFGS calculation enhances the presentation in Hamming loss [20], review, precision [20] and F1 score [20]. The most reduced computational unpredictability is accomplished in SVM, ML-k-NN individually. It outflanks other quasi-Newton procedures and different rivals regarding generally speaking performance measures [20].

Subsequently, this technique ends up being better than others with sensible computational expense and ends up being exceptionally serious with some high-level multi-label strategy for text characterization [20].

The improved sine cosine algorithm (ISCA) beats original SCA [14], OBL_SCA [14], Levy_SCA [14], weighted_SCA [14], the hunt calculation, moth–flame

Table 40.2 Performance measures of ML4BFGS, SVM and ML-kNN [20]

	Precision (%) [20]	Recall (%) [20]	F-score (%) [20]	Hamming loss (%) [20]	Test time (s) [20]
ML4BFGS [20]	88	71	87	84	18.5
SVM [20]	84	59	83	78	3.1
ML-kNN [20]	74	45	71	66	4.7

advancement calculation and (MFO), ant colony optimization (ACO) and accomplished truly preferred execution over genetic algorithm (GA), according to Belazzoug et al. [14]. It gives better execution as well as it is basic in ideas. Be that as it may, there is a necessity to improve it further to take care of hard enhancement issues. Consequently, it very well may be joined with other pursuit calculations to improve execution [14].

Quadratic programming-based numerical optimization approach beats the traditional element choice strategies. It is contrasted with basic ordinary techniques, i.e. chi square (Avg.), chi square (Max), IG, MI (avg), the quadratic programming-based mathematical improvement approach, experiences high preparation time necessities. Consequently, this limit can be tended to in the future [17] (Table 40.3).

NDM-BJO [5] produces better performance in accuracy and Macro-F1 w.r.t. NB and SVM at WebKB dataset (Table 40.4).

In Table 40.5, the performance of NDM-BJO has been demonstrated where four prestigious content corpora have been thought of, which plainly shows NDM-BJO FS approach is superior to different filter and wrapper techniques [5].

Table 40.3 Datasets used [17]

Datasets [17]	Documents [17]	Topics [17]	Terms [17]	Average terms in each document [17]	Maximum terms in a document [17]
20NG [17]	18,774	11,745	20	131.6	6216
Reuters10 [17]	7285	5204	10	48.57	464
TDT10 [17]	7456	12,867	10	174.1	1392

Table 40.4 Original feature expanse after being pruned for reducing dimensions [5]

Dataset [5]	No. of features [5]	Accuracy in % (correctly classified docs over total testing docs) [5]
WebKB [5]	7830	80.5 (1124 over 1396)
SMS [5]	4533	94.3 (2169 over 2300)
BBC News [5]	5648	92.5 (925 over 1000)
10Newsgroup [5]	8957	78.9 (3102 over 3930)

Table 40.5 Dimensionality reduction (optimal feature subset selection NDM-BJO) [5]

Dataset [5]	No. of features [5]	Accuracy in % (correctly classified docs over total testing docs) [5]
WebKB	200	88.9 (1242 over 1396)
SMS	300	97.8 (2251 over 2300)
BBC News	200	96.7 (967 over 1000)
10Newsgroup	200	90.5 (3558 over 3930)

On SMS dataset, the execution of NDM-BJO in accuracy utilizing NB categorization as superior and utilizing SVM categorization as most elevated exactness when contrasted alternative to FS strategies and Macro-F1 estimate considering NB and SVM is better [5]. As for the BBC dataset, NDM-BJO creates better execution in accuracy and Macro-F1 estimate as NB and SVM categorizations [5]. For the 10Newsgroup dataset, it produces the most elevated accuracy and better Macro-F1 value for NB and SVM [5]. Be that as it may, the computational unpredictability of this strategy is high. Consequently, it tends to be dealt with in the future [5].

Hebb rule-based component choice (HRFS) ensures great execution of text classification under few discriminative terms for F-measure, ROC region and classification accuracy [4]. It can diminish the intricacy of feature selection as it tends to be done as matrix operation [4].

ALSVM shows that the quantity of tagged specimens in the preparation set was decreased to 90% with a loss of 1.5% accuracy for R8 dataset [18]. For a 20 ng dataset, the number of labelled training samples can be reduced by 41% with an accuracy loss of only 1.5% [18]. For WebKB dataset, the number of labelled training samples can be reduced by 56.54% with an accuracy loss of 5% [18].

ALMSVM shows that for R8 dataset, the classification accuracy is improved by 0.36% utilizing only 10.77% of accessible samples [18]. For Web-KB data, the exactness got improved over 1.27% for just 48.4% labelled pool tests [18]. For a 20 ng dataset, accuracy is expanded by 2.4% by utilizing just 61.5% of the total tests [18].

Hence according to Tables 40.6 and 40.7, ALSVM performs finer than ALMSVM [18].

Lion fuzzy neural network (LFNN) execution has been contrasted and those three present strategies, namely I-BP, FI-BP and I-BPLion using metrics [19], accuracy

Table 40.6 Profit ratio using ALSVM and ALMSVM in terms of accuracy [18]

Dataset [18]	SVM (%)[18]	Profit in accuracy [18]	
		ALSVM [18]	ALMSVM (%)[18]
R8 [18]	83.33	−1.344	0.36
20 ng [18]	43.79	– 1.64	2.49
WebKB [18]	53.07	– 5.08936	1.27

Table 40.7 Profit ratio using ALSVM and ALMSVM in terms of labelled samples [18]

Dataset [18]	SVM (%) [18]	Profit in number of labelled samples [18]	
		ALSVM (%) [18]	ALMSVM (%) [18]
R8 [18]	83.33	10.11	10.77
20 ng [18]	43.79	59.19	61.34
WebKB [18]	53.07	43.45	48.4

[19] and error [19]. The test outcomes reasons that the proposed gradual methodology beats the current system taken for connection with the greatest exactness of 81.49%, 7.49%, 3.02% and 4.92% for the datasets, 20 Newsgroup, Reuter, WebKB and RCV1 individually [19].

RDTFD strategy figures out that the operation time of RDTFD has been influenced with the quantity of chosen terms on particular examples that shifts straight for expansion of the specimen as well as the quantity of chosen labels when bit number and repetition labels are consistent [3].

The NB and SVM categorizations are solicited to assess RDTFD demonstration on six sets of data, namely PU1 [3], PU2 [3], PU3 [3], PUA [3], CS [3] and ES [3]. The RDTFD technique and other six component choice strategies, namely CHI [3], IG [3], TFIG [3], TTFS [3], IMTFIDF [3] and GININTF [3], apply the Wilcoxon marked positions test on the paired F1 values [3]. The outcome represents that the accomplishment of RDTFD has been superior to additional FS techniques for 66 cases; however for 6 cases, there has been no significance variance [3]. It is more powerful than other FS strategies in spam filtering [3]. Nonetheless, other feature search procedures can be utilized to improve this strategy to improve its running velocity for the parallel operation of large scale PSO regarding high-dimensional feature space [3].

40.4 Conclusion

Artificial intelligence and machine learning are yet an arising field. There has been a great deal of work going around in this area and has the tons of scope for future work. Hence, it can be concluded that improved sine cosine algorithm (ISCA) is simple in concept and is capable of proving good performance for updating the solution's position that can be helpful to adapt with a broad range of search problems as it provides flexibility with some parameters setting. Quadratic programming-based numerical optimization finds an ideal harmony between two conditions for feature selection wherein the label similarity lessens the quantity of repetitive labels chosen to enhance classification correctness [17]. NDM-BJO lessens the high-D attribute expanse to enhance the classifier exactness [5]. HRFS can diminish the unpredictability of feature selection as it tends to be done as matrix operation [4]. Active learning-based support vector machine (SVM) [18] expands the classification accuracy [18] and holds generally excellent strength with all datasets [18]. Lion fuzzy neural network (LFNN) has better accuracy [19]. Relative document-term frequency difference (RDTFD) strategy is more robust [3]. ML-BFGS helps increment the performance of text classification [20]. Consequently, we can hybridize the characteristics of these algorithms to refine their individual impediment which gives us the extension to future work.

References

1. Shroff, K.P., Maheta, H.H.: A comparative study of various feature selection techniques in high-dimensional data set to improve classification accuracy. In: 2015 International Conference on Computer Communication and Informatics (ICCCI), Coimbatore, India (2015)
2. Brownlee, J.: https://machinelearningmastery.com/, 23 Dec 2020. [Online]. Available: https://machinelearningmastery.com/tour-of-optimization-algorithms/. Accessed 3 Feb 2021.
3. Liu, Y., Ju, S., Wang, J., Su, C.: A new feature selection method for text classification based on independent feature space search. Mathematical Problems in Engineering (2020)
4. Wang, H., Hong, M.: Supervised Hebb rule-based feature selection for text classification. Inf. Process. Manage. **56**(1), 167–191 (2019)
5. Thirumoorthy, K., Muneeswaran, K.: Optimal feature subset selection using hybrid binary Jaya optimization algorithm for text classification. Sadhana **45**(1), 1–13 (2020)
6. Minaee, S., Kalchbrenner, N., Cambria, E., Nikzad, N., Chenaghlu, M., Gao, J.: Deep learning-based text classification: A comprehensive review. arXiv preprint arXiv:2004.03705 (2020)
7. Drucker, H., Wu, D., Vapnik, V.N.: Support vector machines for spam categorization. IEEE Trans. Neural Networks **10**(5), 1048–1054 (1999)
8. Guzella, T.S., Caminhas, W.M.: A review of machine learning approaches to spam filtering. Expert Syst. Appl. **36**(7), 10206–10222 (2009)
9. Günal, S., Ergin, S., Gülmezoğlu, M.B., Gerek, Ö.N.: On feature extraction for spam e-mail detection. In: International Workshop on Multimedia Content Representation, Classification and Security, Springer, Berlin, Heidelberg, pp. 635–642, September 2006
10. Yu, B., Zhu, D.H.: Combining neural networks and semantic feature space for email classification. Knowl.-Based Syst. **22**(5), 376–381 (2009)
11. Anagnostopoulos, I., Anagnostopoulos, C., Loumos, V., Kayafas, E.: Classifying Web pages employing a probabilistic neural network. IEEE Proc.-Softw. **151**(3), 139–150 (2004)
12. Chen, R.C., Hsieh, C.H.: Web page classification based on a support vector machine using a weighted vote schema. Expert Syst. Appl. **31**(2), 427–435 (2006)
13. Markov, I.: Automatic Native Language Identification (2018). https://doi.org/10.13140/RG.2.2.15566.51520
14. Belazzoug, M., Touahria, M., Nouioua, F., Brahimi, M.: An improved sine cosine algorithm to select features for text categorization. J. King Saud Univ.-Comput. Inf. Sci. **32**(4), 454–464 (2020)
15. More, V.D.: https://moredvikas.wordpress.com/, 9 Oct 2018. [Online]. Available: https://moredvikas.wordpress.com/2018/10/09/machine-learning-introduction-to-feature-selection-variable-selection-or-attribute-selection-or-dimensionality-reduction/. Accessed 3 Feb 2021
16. Raschka, S.: https://sebastianraschka.com, 21 Dec 2020. [Online]. Available: https://sebastianraschka.com/faq/docs/feature_sele_categories.html. Accessed 3 Feb 2021
17. Lim, H., Kim, D.W.: Generalized term similarity for feature selection in text classification using quadratic programming. Entropy **22**(4), 395 (2020)
18. Goudjil, M., Koudil, M., Bedda, M., Ghoggali, N.: A novel active learning method using SVM for text classification. Int. J. Autom. Comput. **15**(3), 290–298 (2018)
19. Ranjan, N.M., Prasad, R.S.: LFNN: Lion fuzzy neural network-based evolutionary model for text classification using context and sense-based features. Appl. Soft Comput. **71**, 994–1008 (2018)
20. Borhani, M.: Multi-label Log-Loss function using L-BFGS for document categorization. Eng. Appl. Artif. Intell. **91**, 103623 (2020)

Chapter 41
Adaptive Vehicle Safety and Collision Warning System Using DSRC for Heavy-Duty Vehicle

Abhishek Malhotra and Hardil Kanabar

Abstract Passenger vehicles face many problems as heavy-duty vehicles are challenging to manage on the road, and they also block the view for smaller vehicles traveling behind them, especially on two-lane highways and roads. An adaptive system is used to avoid such collisions by giving them warning with vehicle-to-vehicle communication. Global positioning system is used to get the required information about the location and velocity of the vehicle. Each vehicle contains a wireless system that helps in vehicle-to-vehicle communication. As a result, the communication between passenger and heavy-duty vehicles was managed and improved passengers and vehicle safety.

41.1 Introduction

According to a report by the transport research wing of the ministry of road transport and highways, 151,113 people were killed in 480,652 road accidents across India in 2019, an average of 414 a day or 17 an hour. The deaths in truck accidents amounted to 25,108 in the year 2018 that is a total of 16.7% of the total accident deaths. Heavy-duty vehicles amount to the third highest number of accidents and fatalities in India. These heavy-duty vehicles also block the field of view of the drivers behind them, which amounts to a lot of accidents due to overtaking cars, integrating RFID with DSRC technology for unavoidable situations where the other vehicle is not active. We have developed an adaptive system that uses dedicated short-range communication to share information between different vehicles, such as rpm velocity brake pressure distance. Privacy of passengers is ensured by not associating the messages sent from one vehicle to another. Safety from non-active vehicles on highways is also ensured by using RFID technology. We segregate this paper into three parts where at first, we analyze the situation using MATLAB/Simulink, Ansys, and Prescan [1]. We present our solution, and at the end, we have compared the results we get from our research.

A. Malhotra (✉) · H. Kanabar
SRM IST, Chennai, India

T. Senjyu et al. (eds.), *IOT with Smart Systems*, Smart Innovation, Systems and Technologies 251, https://doi.org/10.1007/978-981-16-3945-6_42

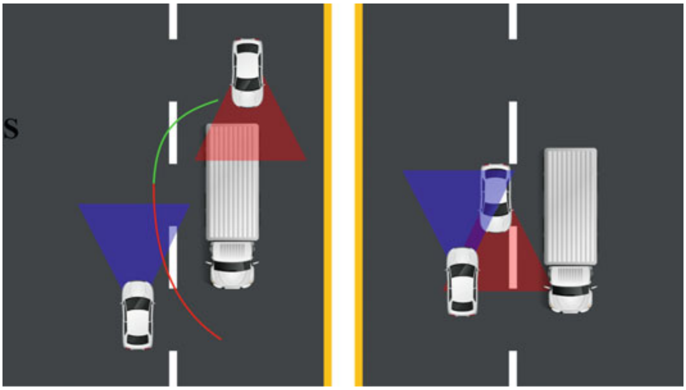


Fig. 41.1 Overtaking a heavy-duty vehicle without adaptive vehicle safety system

41.2 Collision and Blind Spot Analysis

Without the adaptive vehicle safety system, the blind spots created by heavy-duty vehicles, which are much bigger, do not allow passenger vehicles to overtake them safely and efficiently as drivers wouldn't see the road in front of the other vehicles. These types of situations are especially dangerous on roads and highways [2] with just two lanes. In India, where there are a lot of narrow highways and roads, it would be impossible for the drivers to avoid such incidents, which often lead to serious accidents. Even on highways with more lanes, drivers have to be very careful while overtaking trucks and other bigger vehicles due to the narrow field of view (Fig. 41.1).

In a situation where a vehicle is parked or stopped on a curved turn, other vehicles that are approaching the turn wouldn't be able to see the vehicle stopped. This would lead to a serious mishap on the road. Also, if a turning vehicle slows down or applies the brake suddenly, the vehicle at the back, as shown in Fig. 41.2, would get a very less amount of reaction time and thus could lead to an accident. This situation often happens when careless drivers stop/slow down on a large curved [3] portion of the roads or highways.

41.3 Adaptive Vehicle Safety System

Using global positioning system vehicle-to-vehicle communication is incorporated, and considering the case of non-active vehicle RFID is used. A warning system helps to avoid collisions caused by blind spots, as shown in Fig. 41.3.

The adaptive vehicle safety system can avoid blind spots created by heavy-duty vehicles can be avoided as drivers will get a warning whenever there is a chance of other vehicles colliding. These types of situations are especially dangerous on roads and highways with just two lanes. In India, where there are many narrow highways

Fig. 41.2 Overtaking a heavy-duty vehicle without adaptive vehicle safety system

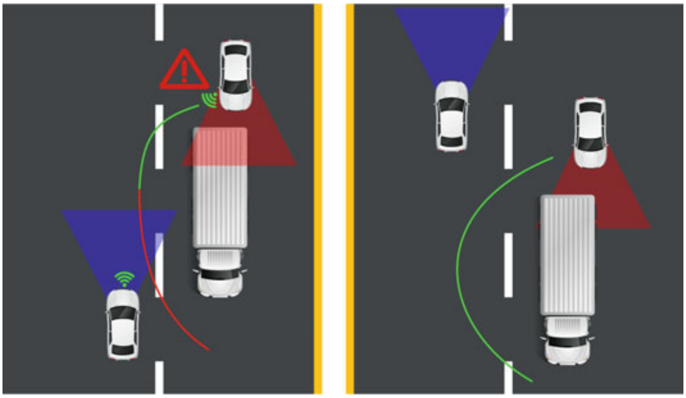


Fig. 41.3 Overtaking a heavy-duty vehicle with adaptive vehicle safety system

and roads, our adaptive system will help the drivers avoid such Incidents, which often leads to serious accidents. Even on highways with more lanes, drivers have to be very careful while overtaking trucks and other bigger vehicles due to the narrow field of view. Warning system interlinked with vehicle-to-vehicle communication using vehicle bus [4], and GPS data helps get the desired result (Fig. 41.4).

The vehicle-to-vehicle warning system contains the main central processing unit, which has five components; the central processing unit takes input from 4 and gives outputs to 2 components. The Global Positioning System data provides the processing unit with the car's location and its approximate velocity. The vehicle bus data contains all the data received from the various sensors present in the vehicle's controlling units, such as the rate per minute of the engine crankshaft, the steering angle, brake pedal pressure, and the throttle pedal pressure. All this data will be relayed to the central processing unit, which will use the communication system to transmit the data to

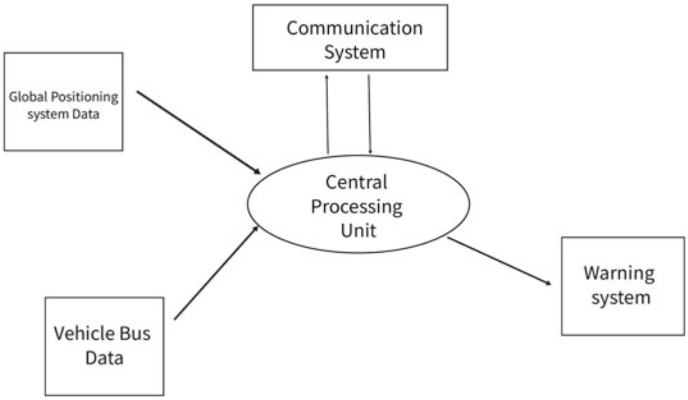


Fig. 41.4 Vehicle-to-vehicle warning system integrated with GPS

other vehicles and also receive the said data from those vehicles. This data will then be combined and processed. If any vehicles on the road are prone to an accident [5], the CPU will alert the warning system to send out warnings through visual and audio interfaces. The CPU will then monitor the vehicle bus data. It will analyze the driver's response to the above-said warning and decide whether to continue displaying the warning.

41.4 Algorithm and Calculation

Warning at a time to hit
No warning if the hit is not possible
No warning if the passenger takes precautions
No warning if Time to Hit > > Time to Prevention
When an adaptive safety system is not installed, it leads to the collision as blind spots generated by heavy-duty vehicle block the view

TTH- Time to hit
TTP- Time to prevention algorithm

1) Catch the sent signal
2) Check if a collision is possible
3) if TTH closes to TTR
send warning
else
if TTH > TTP
Don't Send
Else
goto

Pass, goto

2 vehicles at (m1,n1),(m2,n2) moving at s1,s2 & direction a1, a2 A collision is identified while overtaking expected hit at (m + ,n +)
m + = (n2-n1)-(m2 tan a2- m1 tan a1)/ tan a1 – tan a2
n + = (m2-m1)-(n2 cot a2 – n1 cot a1)/ cot a1 – cot a2
Time to Crash(TTQ) TTQ = D + - D1/S1 sign (D + -D1)*s1

TTD2 = |D + - D2| sign(D + -D2)*S2)/(S2) Dn is vector representative of coordinate of(nn, mm),Sign () sign function If time to crash of vehicle is equal to the time of the crash of other vehicles, i.e., TTQ1 = TTQ2. (Hit is possible) [5]

As shown in Fig. 41.5, we have considered the case of a collision while overtaking where a light motor vehicle is coming at a velocity of s2, and another vehicle is coming at a velocity of s1 from the opposite direction. A collision is identified while overtaking, and a hit is expected at (m+, n+). Time to hit and time to prevention is calculated, and the warning system works according to the algorithm. If time to hit is greater than time to prevention, then a warning will not be given [6], and on the other side, if both of them are close to each other warning will be sent (Fig. 41.6).

With an adaptive [7] vehicle safety system, the blind spots created by curved roads or highways can be avoided by using RFID [8] technology. Drivers who are about to enter a wide turn cannot see other vehicles that are ahead of the curve. This type of situation would be dangerous if the driver in the front suddenly slows down or applies a brake. A collision would be inevitable when vehicles break down/stop on such a curve on a highway or an expressway. Using active RFID tags and readers on all vehicles, the driver would be alerted if there's an idle vehicle in front 100 m before its location. Using RFID technology is useful in such scenarios as it would transmit data without the need for the other vehicle to be active. Drivers will also get a warning if they suddenly brake or slow down on such curves by using GPS technology present in the safety system [9].

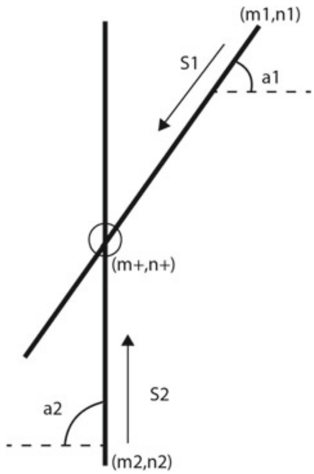
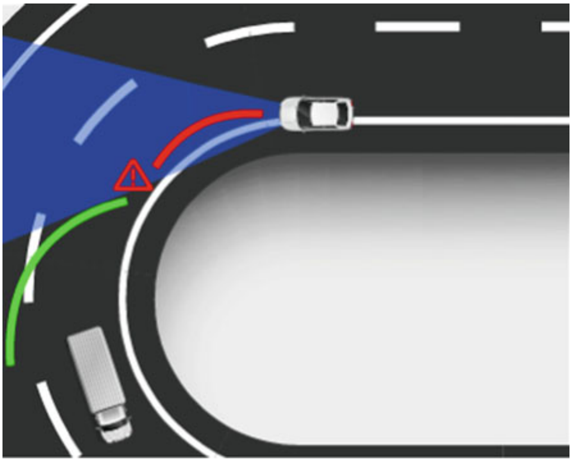


Fig. 41.5 Vehicle collision while overtaking

Fig. 41.6 Turning on blind spots with adaptive vehicle safety system



41.5 Conclusion

Adaptive vehicle safety and collision warning system with DSRC and RFID help prevent road accidents faced by passenger vehicles and heavy-duty vehicles. Covering the case of both overtaking and turning on highways through blind spots, we have analyzed the unavoidable situations and have come to this solution covering all the aspects. As a result, the communication between passenger and heavy-duty vehicles was managed and improved passenger and vehicles' safety.

References

1. Jiang, D., Taliwal, V., Meier, A., Holfelder, W., Herrtwich, R.: Design of 5.9 GHz DSRC-based vehicular safety communication. IEEE Wirel. Commun. **13**(5), 36–43 (2006). https://doi.org/10.1109/WC-M.2006.250356
2. Michahelles, F., Thiesse, F., Schmidt, A., Williams, J.R.: Pervasive RFID and near field communication technology. IEEE Pervasive Comput. **6**(3), 94–96, c3 (2007). https://doi.org/10.1109/MPRV.2007.64.
3. Bai, F., Krishnan, H.:Reliability analysis of DSRC wireless communication for vehicle safety applications. In: 2006 IEEE Intelligent Transportation Systems Conference, Toronto, Ont., pp. 355–362 (2006). https://doi.org/10.1109/ITSC.2006.1706767
4. Jansson, J., Gustafsson, F.: A framework and automotive application of collision avoidance decision making. Automatica **44**(9)
5. Jansson, J., et al.: Decision making for collision avoidance systems. SAE Trans. **111**, 197–204 (2002). JSTOR, www.jstor.org/stable/44699413. Accessed 3 Feb 2021
6. Blythe, P.: RFID for road tolling, road-use pricing and vehicle access control. In: IEE Colloquium on RFID Technology (Ref. No. 1999/123), London, UK, pp. 8/1–816 (1999). https://doi.org/10.1049/ic:19990681
7. Williams, T., Alves, P., Lachapelle, G., Basnayake, C.: Evaluation of GPS-based methods of relative positioning for automotive safety applications. Transp. Res. Part C: Emerg. Technol. **23**. https://doi.org/10.1016/j.trc.2011.08.011

8. Bevly, D.M., Gerdes, J.C., Wilson, C., Zhang, G.: The use of GPS based velocity measurements for improved vehicle state estimation. In: Proceedings of the 2000 American control conference. ACC (IEEE Cat. No.00CH36334), Chicago, IL, USA, vol. 4, pp. 2538–2542 (2000). https://doi.org/10.1109/ACC.2000.878665
9. Biswas, S., Tatchikou, R., Dion, F.: Vehicle-to-vehicle wireless communication protocols for enhancing highway traffic safety. IEEE Commun. Mag. **44**(1), 74–82 (2006). https://doi.org/10.1109/MCOM.2006.1580935

Chapter 42
Towards Enhancement of the Lexicon Approach for Hindi Sentiment Analysis

Dhanashree S. Kulkarni and Sunil F. Rodd

Abstract The amount of data from users in the form of reviews, comments and opinions in Hindi language is tremendously increasing on social media, blogs and online forums due to which sentiment analysis of Indian languages has turned out to be a predominant research area. Lexicon-based approaches are being used for analysing sentiments but they may not be as accurate as machine learning approaches. To enhance and improve the lexicon approach, a frequency score generation weighting scheme algorithm is proposed. Novel rules for handling some linguistic features such as negation, intensifiers are also implemented. Experiments performed to evaluate the performance of the proposed algorithm shows that the performance of lexicon-based sentiment analysis is increased significantly.

42.1 Introduction

Sentiment analysis has turned out to be an emerging area of research due to the increasing amount of data generated on social media, websites, forums and blogs. Opinions, emotions and sentiments of people are analysed specifically towards a particular entity. There are numerous applications of sentiment analysis such as product analytics, recommendation systems, reputation management, etc.

Most of the work related to sentiment analysis is done considering English language. But nowadays, people give reviews and comments in their native languages. So, there is a need to perform sentiment analysis in languages other than English also. Hindi being the national language of India and one amongst the most spoken languages, performing sentiment analysis in Hindi could prove to be very much useful specially for government organizations.

D. S. Kulkarni (✉)
Angadi Institute Of Technology and Managment, Belagavi, India

S. F. Rodd
Professor, Gogte Institute of Technology, Belagavi, India
e-mail: sfroddgit@git.edu

T. Senjyu et al. (eds.), *IOT with Smart Systems*, Smart Innovation, Systems and Technologies 251, https://doi.org/10.1007/978-981-16-3945-6_43

Sentiment analysis often called as opinion mining [1] can be performed at either document level, sentence level or at aspect level. Two well-known approaches that are used for analysing sentiments are lexicon approach and machine learning approach. Lexicon approach also known as the rule-based approach, depends on a dictionary in order to find out the sentiment score of a particular word. On the other hand, machine learning approach is based on the concept of feature extraction and learning from the data provided with the use of machine learning algorithms. It basically creates a system that learns from experience and has the ability to improve itself. Lexicon-based approaches are said to be time efficient but are not that accurate [2]. Machine learning in general is more accurate. Therefore, some methods need to be devised so as to increase the accuracy of lexicon-based methods. In case of Hindi sentiment analysis, lexicon-based methods are implemented using a Hindi Sentiwordnet. Hence, accuracy can be increased in such cases by making efficient use of the Hindi SentiWordNet and handling the problem of absent word and morphological variation in the SentiWordNet.

42.2 Related Work

Joshi et al. [3] proposed a strategy for performing sentiment analysis in Hindi language at document level where they developed a Hindi SentiWordNet (HSWN) for the language. Bakliwal et al. [4] developed a Hindi subjective lexicon for sentiment classification. Mittal et al. in 2013 [5] tried to improvise the HindiSentiWordNet (HSWN) and enhanced its coverage by including more opinion words into it. Sharma et al. [6] built a system for Hindi language that performed negation handling and classified reviews into positive, negative and neutral classes. Pandey et al. [7] improved the HSWN by adding new terms from the corpus, getting their polarity from English Sentiwordnet and also handled negations and discourse relations. In 2015, Jha et al. [8] developed a Hindi opinion mining system (HOMS) for data on movie reviews using machine learning algorithm, namely Naïve Bayes classifier and POS tagging.

References [9–11] addressed the issue of aspect-based sentiment analysis for Hindi language. Jha et al. [12] proposed sentiment analysis in Hindi language for movie reviews by incorporating lexicon-based classification technique and machine learning techniques. A sentiment aware dictionary was built in [13] considering multiple domain data and was compared with the labelling done by HSWN. In 2020, Garg [14] used lexicon method and performed sentiment analysis of Indian PM's "Mann Ki Baat" radio show. Shrivastava et al. [15] proposed a deep learning approach where a genetic algorithm was used to build a gated recurrent network architecture and was experimented on a Hindi movie review dataset. Verma et al. [16] mined public opinions on Indian Government policies using R where emotions such as happy, fear and anger were investigated using emotion-based lexicon technique.

Related [17–21] work show that lexicon approach has been used predominantly for performing sentiment analysis in Hindi. But very less measures have been devised to enhance and improvise the lexicon approach, especially the HSWN. Hence, this work makes effective use of the HSWN by proposing a frequency score generation weighting scheme and suggesting appropriate methods for efficient negation handling and handling of morphological variations.

42.3 Proposed Methodology

A simple lexicon approach for Hindi sentiment analysis uses the HindiSentiWordnet (HSWN) to extract the corresponding positive and negative score. But it may happen that a particular word may not be present in the HSWN which affects the lexicon performance. Also, morphological handling specific to HSWN may be needed when the HSWN has to be searched for a word. Training datasets are being used to train machine learning algorithms. In the same way, training datasets can be used for generating frequency score of a word which when combined with simple lexicon approach and can definitely increase the accuracy of the lexicon approach. The proposed approach considers Hindi reviews as input to the system which is first pre-processed. Stop words are removed. POS tagging is then performed using a Hindi POS tagger for mainly identifying verbs, adverbs, adjectives and nouns. Algorithm 1 explains the proposed methodology of using HSWN with frequency score generation term weighting scheme.

Algorithm 1 HSWN with frequency score generation term weighting scheme (HFSG)

Input: Words of a sentence.

Output: Sentiment polarity of a sentence.

Step 1: POS tagging is done and Words other than nouns, adjectives, adverbs and verbs are discarded.

Step 2: Word is searched in HSWN along with its corresponding tag. If the word is present, then its positive and negative score is obtained. If not, check for variations of word by performing HSWN specific morphological handling as below:

- From the input word, filter the root and suffix pair.
- Find out all the variation words of the root
- Search the HSWN to find out which variation word is present in it
- If one of the variations is present, then its positive and negative score is obtained.
- If none of the variations is present goto Step 3

Step 3: If the word is not present in HSWN, synonyms or synset words from Hindi wordnet are retrieved, and the corresponding positive and negative score is obtained.

Step 4: Sentiment score is calculated based on score in HSWN and conditional probability:

$$\text{PosScore} = \text{positive score of word in } HSWN * \frac{\text{No of occurrences in positive sentences}}{\text{No of all occurrences in positive and negative sentences}}$$

$$\text{NegScore} = \text{negative score of word in } HSWN * \frac{\text{No of occurrences in negative sentences}}{\text{No of all occurrences in positive and negative sentences}}$$

The maximum score out of PosScore and *NegScore* is assigned as polarity of the word.

Step 5: Finally, the polarity of the sentence is assigned depending on the polarity which majority of the words from the sentence possess.

The output of the simple lexicon approach is also affected by various other reasons such as presence of negation words, intensifiers, diminishers, etc. which need to be handled. Also, problem of word sense disambiguation pertaining to HSWN needs to be solved. There may be many senses of the same word occurring in the HSWN. The HFSG system uses the first sense occurrence and fetches its corresponding positive and negative score. To further increase the accuracy of the system, the proposed method is enhanced by including a rule-based model for handling negations and performing averaging of all senses for solving problem of word sense disambiguation (WSD).

Different linguistic rules are defined as a part of the rule-based model. Some of the rules included for negations, intensifiers and diminishers are mentioned in Table 42.1.

The problem of WSD is solved by first matching the words and the tags in the HSWN, and if the word has multiple senses of the same tag, the average of the score of all senses of the same tag is calculated.

Table 42.1 Negation handling rules

Occurrence of word	Example	Sentence Polarity
Negative before NEG	बुरा नहीं	Positive
Positive before NEG	अच्छा नहीं	Negative
Diminisher before positive	कम अच्छा	Negative
Intensifier before diminisher	काफी कम	Negative

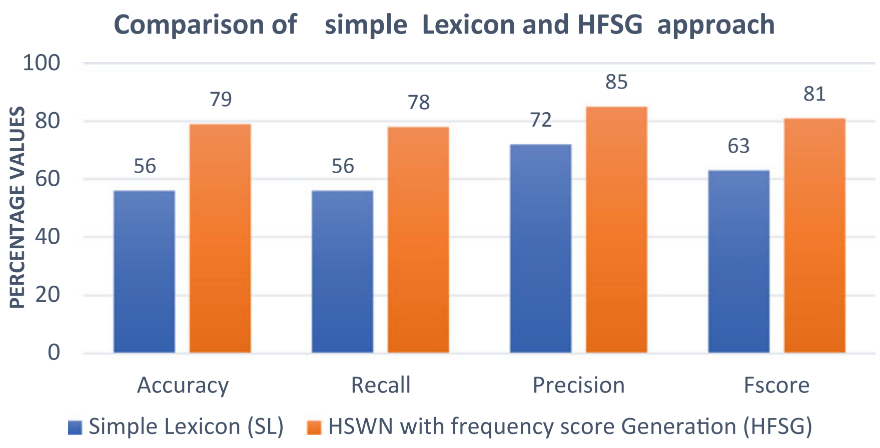


Fig. 42.1 Comparison of simple lexicon (HSWN) and HSWN with frequency score generation weighting scheme algorithm

42.4 Results Discussion

To conduct the experiments, dataset was developed which contains around 4028 sentences in Hindi that were collected from various sources such as movie reviews, product reviews and travel reviews. Output of the proposed system is evaluated to check its efficiency using four parameters, namely Accuracy, Precision, Recall and Fscore. It is compared with the simple lexicon approach as shown in Fig. 42.1. The results show a considerable increase in all the parameters when the proposed approach is used.

The implementation of rule-based negation handling and averaging multiple senses solution further has a positive impact on the performance of the HFSG proposed system as shown in Fig. 42.2. There is a small increase in Accuracy and Precision, but a considerable increase is seen in Recall and Fscore. Thus, the proposed system works much better when compared to the simple lexicon approach.

42.5 Conclusion

Sentiment analysis is an upcoming field of research with very little work done in performing sentiment analysis in Hindi language. This paper makes some empirical contributions to sentiment analysis by enhancing the lexicon approach and making efficient use of the HindiSentiWordnet. A new method called HSWN with frequency score generation weighting scheme is introduced that solves the problem of absent words in HSWN and increases the performance of the system. Also, problem of

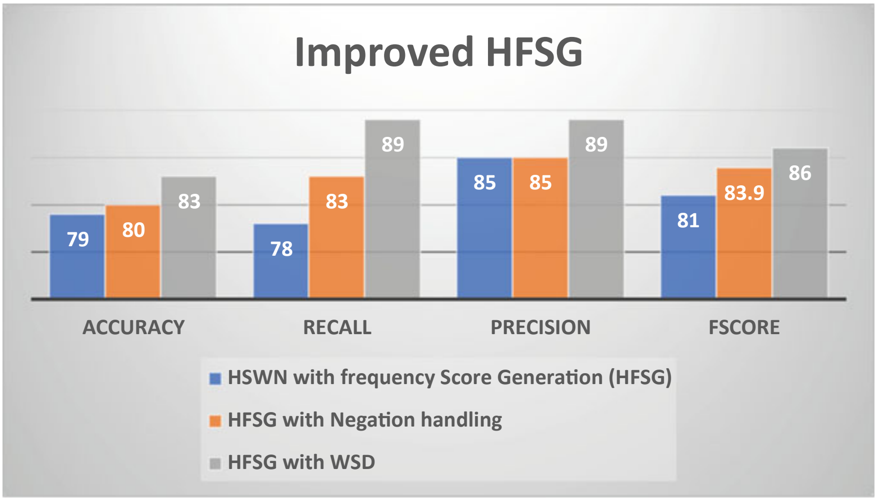


Fig. 42.2 Improvement in HFSG weighting scheme algorithm by incorporating negation handling and word sense disambiguation

word sense disambiguation is solved, and a rule-based model is presented for negation handling. Results show that the proposed approach significantly enhances the sentiment classification accuracy of simple lexicon approach.

Future work will deal with converting the rule-based approach into a fuzzy logic approach for negation handling and enhancing the HFSG scheme so as to perform multiclass classification.

References

1. Pang, B., Lee, L.: Opinion mining and sentiment analysis. Found. Trends Inf. Retrieval **2** (2008)
2. Augustyniak, Ł, Szymanski, P., Kajdanowicz, T., Tuligłowicz, W.: Comprehensive study on lexicon-based ensemble classification sentiment analysis. Entropy **18**(1), 4 (2016)
3. Joshi, A., Bhattacharyya, P., Balamurali, R.: A Fall-back Strategy for Sentiment Analysis in Hindi: A Case Study (2010)
4. Bakliwal, A., Arora, P., Varma, V.: Hindi subjective lexicon: a lexical resource for Hindi polarity classification. In: Proceedings of the Eight International Conference on Language Resources and Evaluation (LREC) (2012)
5. Mittal, N., Agarwal, B., Chouhan, G., Bania, N., Pareek, P.: Sentiment analysis of Hindi review based on negation and discourse relation. In: Proceedings of International Joint Conference on Natural Language Processing, pp. 45–50 (2013)
6. Sharma, R., Nigam, S., Jain, R.: Polarity detection movie reviews in Hindi language. arXiv preprint arXiv:1409.3942 (2014)
7. Pandey, P., Govilkar, S.: A framework for sentiment analysis in Hindi using HSWN. Int J Comput Appl **119**(19), 23–26 (2015)

8. Jha, V., Manjunath, N., Shenoy, P.D., Venugopal, K., Patnaik, L.M.: Homs: Hindi opinion mining system. In: 2nd International Conference on Recent Trends in Information Systems, pp. 366–37. IEEE (2015)
9. Akhtar, M.S., Ekbal, A., Bhattacharyya, P.: Aspect based sentiment analysis: category detection and sentiment classification for Hindi. In: 17th International Conference on Intelligent Text Processing and Computational Linguistics, pp. 1–12 (2016a)
10. Akhtar, M.S., Ekbal, A., Bhattacharyya, P.: Aspect based sentiment analysis in Hindi: resource creation and evaluation. In: Proceedings of the 10th International Conference on Language Resources and Evaluation, pp. 1–7 (2016b)
11. Akhtar, M.S., Sawant, P., Sen, S., Ekbal, A., Bhattacharyya, P.: Solving data sparsity for aspect based sentiment analysis using cross-linguality and multi-linguality. In : Association for Computational Linguistics (2018)
12. Jha, V., Manjunath, N., Deepa Shenoy, P., Venugopal, K.R.: Sentiment analysis in a resource scarce language: Hindi. Int. J. Sci. Eng. Res. **7**(9) (2016)
13. Jha, V., Savitha, R., Deepa Shenoy, P.: Venugopal, K.R., Sangaiah, A.K.: A novel sentiment aware dictionary for multi-domain sentiment classification. Comput. Electr. Eng. **69**, 585–597 (2018)
14. Garg, K.: Sentiment analysis of Indian PM's "Mann Ki Baat". Int. J. Inf. Tecnol. (March 2020) **12**(1), 37–48 (2020)
15. Shrivastava, K., Kumar, S.: A sentiment analysis system for the Hindi language by integrating gated recurrent unit with genetic algorithm. Int. Arab J. Inf. Technol. **17**(6) (2020)
16. Verma, P., Jamwal, S.: Mining public opinion on Indian Government policies using R. Int. J. Innov. Technol. Explor. Eng. (IJITEE) **9**(3). ISSN: 2278-3075 (2020)
17. Sharma, R., Bhattacharyya, P.: A sentiment analyzer for hindi using hindi senti lexicon. In: Proceedings of the 11th International Conference on Natural Language Processing (2014)
18. Sarkar, K., Chakraborty, S.: A sentiment analysis system for Indian language tweets. In: International Conference on Mining Intelligence and Knowledge Exploration, pp 694–702. Springer (2015)
19. Kumar, A., Kohail, S., Ekbal, A., Biemann, C.: Iit-tuda: system for sentiment analysis in Indian language using lexical acquisition. In: International Conference on Mining Intelligence and Knowledge Exploration, pp 684–693. Springer (2015)
20. Sharma, P., Moh, T.S.: Prediction of Indian election using sentiment analysis on Hindi twitter. In: International Conference on Big Data, pp. 1966–1971. IEEE (2016)
21. Mishra, D., Venugopalan, M., Gupta, D.: Context specific Lexicon for Hindi reviews. Procedia Comput. Sci. **93**, 554–563 (2016)

Chapter 43
Electricity Anomalies Detection and Automation in Smart Meter System

Poonam Katyare and Shubhalaxmi S. Joshi

Abstract Electricity consumption in day-to-day life has a higher impact on household expenses. This paper focuses on the various machine learning (ML) techniques for the detection of electricity anomalies which are the comparison of supplied electricity power and actual customer's electricity consumption. Electricity theft leads to economic loss and can be very harmful to grid suppliers. This non-technical loss generates a need for detection of theft or anomalies in power system industries. This paper focuses on the comparative study on detection of electricity theft by the researchers in various aspects and proposed the implementation of this study by ML techniques. This paper first provides an overview of electricity theft and the solution provided by the researchers. Paper surveys the research efforts in a field of energy consumption and analytical overview of data and ML algorithms used. It finally identifies the various problems in anomalies detection of electricity paper also focus on the usage of automation in smart meter system with wireless sensor networks (WSNs) which captures the meter readings and upload the reading data on server of grid provider which helps in generation of bill automatically. From electricity data, two aspects are applied for deep analysis of consumption of electricity. This study highlights use of clustering for data filtering and classification for the identifying the two category of consumption. The study uses various ML algorithms to measure performance of model.

43.1 Introduction

In era of digitization, every daily activity is highly depend on the presence of electricity. Electricity plays a vital role in industry, society and in household activities. It is a source of energy. The electrical power is supplied by power plants. Electrical power grid provides electricity to the consumers such as industries or residences. A

P. Katyare (✉) · S. S. Joshi
Dr. Vishwanath Karad, MIT World Peace University, Pune, India

S. S. Joshi
e-mail: shubhalaxmi.joshi@mitwpu.edu.in

T. Senjyu et al. (eds.), *IOT with Smart Systems*, Smart Innovation,
Systems and Technologies 251, https://doi.org/10.1007/978-981-16-3945-6_44

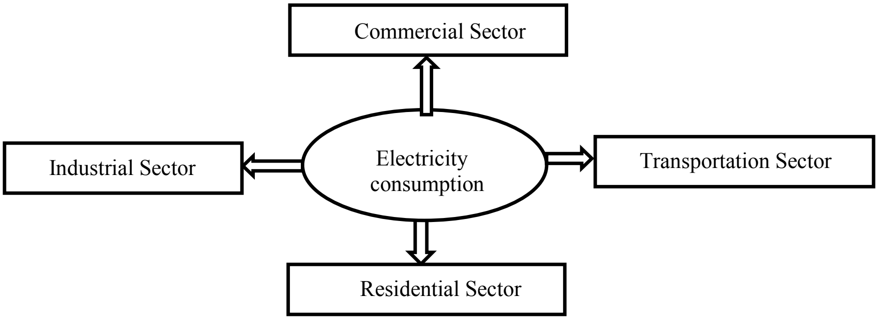


Fig. 44.1 Electricity consumption

problematic situation comes when there is no continuous supply of electricity power. Non-technical losses of electricity cause electricity theft. As the amount of electricity consumption increase or decrease per zone or thieves of electricity per zone changes reflects in the detection of anomalies in electricity consumption and which reflects to the performance of the algorithms used for detecting the anomalies and continuation in the performance of an algorithms [1]. The smart meter infrastructure also provides speedy available of electricity consumption data with high frequency. Various techniques are available to detect the electricity theft in which use of ML techniques are mostly used. Large data for consumption of electricity is provided by power grid suppliers [2]. This data is available for analysis purpose. Along with data analysis, hardware-based techniques are also available to solve the problem of electricity theft and making use of such analytical approach and hardware-based approach with smart meter would definitely lead to decrease the non-technical losses in electricity consumption and lead to satisfaction toward consumer's utilization and supply by power grid plats of electricity. Large amount of electricity is consumed in major sectors are shown in Fig. 44.1.

In traditional way of thinking, the electricity fraud is very challenging task. For detection of electricity, theft manpower is required to take readings manually, check that readings and finding the fault. It is a time consuming process with the findings [1]. Electricity anomalies may lead to economic loss to the utility suppliers. It is not possible to inspect manually such theft in large amount of data. Wireless technologies' for communication is also used for electric system applications [3]. Now a day's, demand of electricity is increases due to digitization. Electricity is very important in daily life and in industries also. Usage of electricity is more in daily applications of digital gadgets, digital learning, digital machineries, household equipment's and industrial machineries which requires continuous and stable supply of electricity. Electrical power industries provide electricity and face many challenges, while providing electricity. Loss in electricity consumption minimizes the quality of supply, increase electricity creation load and lead to pay extra electricity bills to an authenticate customers.

Use of smart grids with WSNs can reduce electricity non-technical losses and use of data analysis techniques helps to detect the fraudulent electricity consumption [4]. Previous study focuses on use of advanced metering infrastructure and some focus on the use of ML algorithms with various combinations of classification algorithms, neural network algorithms, time series analysis [5–7]. This non-technical losses issue with electricity theft detection problem is observed in all over the world with all developing and growing countries. This problem can be solved by many researchers with ML algorithms with support vector machine based fuzzy inference system. As per the area of study and dataset values of residential and non-residential consumers, electricity consumption and the performance of the model could be varies [8, 9]. Another aspect of this study provides various application based electricity meter reading systems with communication protocols and smart metering infrastructure which may be costly, while actual implementation of such systems have been applied [10–14].

Punishments for electricity theft in India [15] focus to control and preserve a form on the theft of electricity for which the Electricity Act, 2003 was enacted. To strengthen the process of detection of anomalies to the power industry, various deterrent punishment and penalties are imposed on the person committing the electricity theft. Detection of anomalies in electricity plays an important role to save electricity. For this more reliable, cheaper and useful solution would be recommended. Data analysis on electricity consumption is a major task and use of digital meter reading with automatic bill generation system. The process of classifying theft and normal consumers using ML algorithm includes dataset preparation, filtering, learning and evaluation. The dataset preparation involves the pre-processing of data of the electricity consumption and selections of features are involved. The filtering phase creates a grouping of theft consumers and normal consumers by using clustering algorithm. The next learning phase involves the classification of theft users and normal users by applying ML algorithms. The evaluation phase is used to assess applied ML algorithm. The present data pre-processing techniques as cleaning of data, replacing missing values and normalized data and feature selection technique are used [4]. The existing ML algorithms such as support vector machine (SVM), Random Forest are used to classify theft and normal consumers whose accuracy is inferior [16–18]. There are many aspects where actually theft is not present as non-residential area where only supply of electricity but no any consumption is happened, and second aspect is consideration of vacation period where no one is there at home. Most of study do not consider all aspects of theft [7, 9, 12–14, 19]. So the consideration of analysis and digital meter with sensor should be adapted for the anomalies detection in electricity consumption. This study provides use of ML techniques for the detection of theft and proposed use of digital meter readings with optical character recognition.

43.2 Related Work

Identification of anomalies in electricity consumption is an important task in energy domain [14, 19]. Empirical and deep studies with significant and substantial research have done in the area of detection of electricity theft. Some research is based on advanced smart meter infrastructure [1, 7, 12–14, 16, 17, 19–23] as well as using data analysis by ML techniques. Existing research involves use of support vector machine classifier [4, 9, 11, 16, 20, 24, 25], decision tree classifier [21], extreme gradient boosting (FA-XGBoost) technique[2], convolution neural network (CNN), logistic regression (LR) [2, 24, 25] and deep convolutional-recurrent neural network-based detector [1, 26, 27], clustering algorithms [28]. A long short-term memory (LSTM) for the detection of abnormal patterns in electricity consumption [25, 29]. Some study focus on the way of detecting outliers in smart meter dataset [30] and other study insights on the use of deep learning and random forest [18] techniques. Each study is based on dataset used for a specific area with its own constraints. Techniques in ML detect the theft as per the dataset. To automate the process of detection of electricity anomalies and generate the electricity bill our study makes a use of digital meter reading system using sensor and from the historical data, we used clustering techniques and neural networks for the fraud detection. Our study involves data preparation phase, data filtration phase and data classification phase along with use of digital meter reading with sensor based data.

43.2.1 Data Preparation Phase

This phase has two steps—Data pre-processing and feature extraction.

Pre-processing of Data—Electricity consumption dataset is needs to be pre-processed using pre-processing techniques as processing data cleaning with identification of missing data, normalization process, outlier removal, scaling of data and feature selection methods are applicable as per the dataset. These are used in [4, 8, 21, 24, 29] studies.

Extraction of features—Data mining techniques and statistical techniques are used to load electricity customer's profiles. Average of electricity data consumption on daily, monthly, quarterly and yearly is also considered as features [4, 5, 8, 16, 17]. Random forest, decision tree models are used to select features automatically [18]. Features are selected from feedback mechanisms and surveys also [20]. The model can be trained by taking input as features. Sparse auto encoder (SAE) feature extraction method is used for the price and load forecasting of electricity. It is used in [26]. Deep learning techniques are used for auto selection of features as convolution neural network (CNN), long short term memory (LSTM) and visual geometry group-16. It is used in [2, 18, 24].

43.2.2 Machine Learning and Deep Leaning Methods

This leaning phase involves various supervised algorithms, unsupervised algorithms, semi-supervised and deep learning algorithms. Most of the researchers used optimum path forest (OPF), SVM, C-SVM, one-class SVM, multiclass SVM, cost sensitive SVM(CS-SVM) and C4.5 Tree [4, 9, 16, 21] for the classification of consumers. Dataset was balanced by cross fold validation techniques which helps in improving accuracy and more than two classes used in SVM classification and consumers data from Tenaga National Berhad (TNB) in Malaysia used in [4] for study. Results of this study are inferior. As dataset was imbalance. Types of SVMs, random undersampling boosting (RUSBoost) for class imbalance [29], decision tree and SVM classifier together used in [16], deep learning convolutional neural networks (CNN) and CNN with random forest used on database provided by electric Ireland and sustainable energy authority of Ireland (SEAI) and readings from the supervisory control and data acquisition (SCADA) metering points that monitor various electrical parameters in distribution systems, respectively, used in [18, 27]. Visual geometry group (VGG-16) module with firefly algorithm based extreme gradient boosting (FA-XGBoost) [2] neural networks ensembles [5], fuzzy clustering [8, 11], statistical analysis [19], supervised approach [20], consumption pattern-based energy theft detector[21], entropy-based electricity theft detector (EBETD) [22], wide & deep convolutional neural networks (CNN) [1, 24], parameter-free incremental clustering algorithms closeness factor based algorithm (CFBA) and correlation-based incremental clustering algorithm (CBICA) [28], long short-term memory (LSTM) architecture [24, 29], ensemble bagged tree (EBT) algorithm [25], maximum information coefficient (MIC) and clustering technique by fast search and find of density peaks (CFSFDP) unsupervised approach found in [30]. Efficient sparse auto encoder nonlinear autoregressive network with exogenous inputs (ESAENARX) and differential evolution recurrent extreme learning machine (DE-RELM) are used forecasting load and price of electricity [26].

43.2.3 Filtration Methods

It gives the selection of two categories as electrical consumption in regular basis or irregular, customer load profile in typical behavior or atypical behavior content. The study uses pattern recognition on daily average consumption which helps in differing confirm Fraud Suspects and unconfirmed Fraud Suspects classes [4]. Higher usage of electricity identifies as suspect in some criteria as quarterly consumption of electricity, annually consumption of electricity, current month consumption compared with corresponding previous month consumption, current month consumption compares with previous month consumption [5], auto- regressive moving average (ARMA) [12], simple moving average (SMA) [23], low security and high-security parameter [7] which identifies normal and abnormal behavior.

Fuzzy clustering used to detect consumption of patterns [8], parameters with min, max, average consumption values selected for identification of customer selection considered as fuzzy rules [11]. Statistical evaluation of parameters with lower in consumption used in [16, 20]. K-means clustering with seasonality, change of appliances, and different usage during weekdays and weekends, singular value decomposition to find the accurate line resistance, consumption patterns used in [1, 2, 17, 21, 25, 28–30]. Entropy-based approach as the measure of the distance between two probability mass functions with min value of entropy indicate no theft and larger than threshold indicates theft [22].

43.2.4 Automation in Electricity System

Many research studies are conducted in the electricity system automation using various communication networks. Wireless automatic meter reading (WAMR) systems with wireless sensor networks (WSNs) for smart grid used for automatic measurement of consumer's consumption. It helps in reduction of cost for human readers, reduction in peak demand of energy, and the time-of-use bill generation and beneficial to consumer also as they can shift to their demand to off-peak times and feels real time pricing model [3]. WAMR has some challenges as effect of environment conditions, quality-of-service requirements, resource constraints and link quality measurement. The researcher has provided the solution to all these challenges for better utilization of WAMR system [6, 14]. It will make use of third generation (3G) communication systems, the DLMS/COSEM (Data Language Messaging Specification/Companion Specification for Energy Metering) standard and Internet Protocol-based SIP (Session Initiation Protocol) signaling [10]. Advanced metering infrastructure with efficient monitoring system along with cyber side network is used in [13]. Learning automata is self-learning model used for adjusting power requirements and power distributions at different time intervals. Learning automata for efficient power management in smart grids (LAPM) delivers the necessity of power for various allocation substations and provides abnormal usage of electricity [14]. Use of observer meter, smart meters, and smart prepaid energy metering system in distribution system [19], smart meters with data processing hardware in the cloud achieve high performance in monitoring and controlling the smart grid operations with wireless sensor network [31].

43.2.5 Limitations of Existing Techniques

Existing ML algorithms as SVM, DT and RF are inferior in case of large datasets. The problem of class imbalance and manual feature selection are found in many studies which have high-false positive rate [4, 9, 16, 21, 24, 26]. In most of the cases, used dataset has large amount missing data with outliers which causes overfitting of the classifier. Data imbalance with local and global features extraction presented by wide component and deep component from 1D data and 2D data, respectively, used for theft detection in [1]. Hybrid model with CNN and LSTM with imbalanced dataset uses synthetic minority over-sampling technique (SMOTE) leads to over fitting and differs the theft [7]. Another hybrid model CNN with RF used where CNN used for feature extraction and RF used for classification with accurate performance measures with high execution time [18].

43.3 Proposed Work

The main objective of the electricity anomalies detection is to make use of smart meters with wireless sensor networks and advanced metering infrastructure. Classification of normal users and fraudulent users is most significant task in smart grid electricity consumption. Mostly filtering methods and use of ML techniques and deep learning with neural network methods being studied in literature review with new model would help to classify the anomalies detection in electricity consumption.

We have proposed the architecture as shown in Fig. 44.2.

Step 1: Data Preparation: We recommend to use smart meters with WSNs which helps to capture the meter readings automatically through sensors and available on smart grid servers. Here, the image pre-processing techniques along with optical character recognition are used to detect and measure the readings. The public smart meter dataset is available for study. Pre-processing of data with data cleaning and feature extraction is involved.

Step 2: Data Filtration: Data filtration techniques are an important step to prepare input for the classification algorithms. For identification of theft users and normal users in data filtration uses clustering of electricity consumption values on pre-processed data to identify the labels. Normal users and theft users creates classes for the classification algorithms.

Step 3: Classification of theft and normal users—Classification methods with deep leaning and ML needs to be intended and settled to train the model which classifies the theft and normal user with electricity consumption.

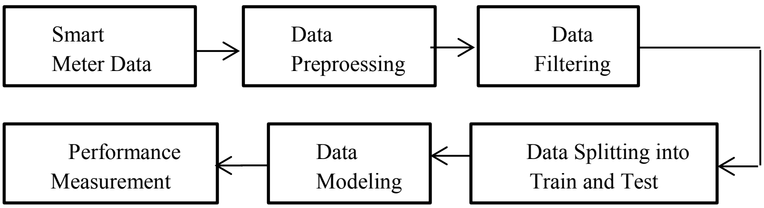


Fig. 44.2 The proposed architecture

Step 4: Evaluation of classification technique—Once the model is trained, it can be tested to check the performance of the technique.

Step 5: Generation of electricity bill amount—From available meter data, we can generate electricity bill and communicate to the consumers.

43.4 Conclusion

Most of the study in electricity anomalies detection makes use of fundamental methods to classify the normal and fraudulent users. Methods proposed in the research are found to be complex for usage in practice as these cannot be applied to any kind of datasets. Smart grids currently face the problem of anomalies detection in electricity consumption and there is a gap between actual electricity consumption by consumer and supply by smart grid. With the understanding of the different aspects of clustering and classification in ML, new algorithms need to be developed as solutions to bridge the gap between research and practice.

In our proposed work, we will attempt to read smart meter data and filter data for classification. The ML algorithm to classify the users will be designed, developed and compared with other solutions. Similarly, software to generate the electricity bill amount will be developed and assessed. Once algorithms are developed and assessed, the framework can be developed. The framework can helps in smart grid.

References

1. Zheng, Z., Yang, Y., Niu, X., Dai, H.N., Zhou, Y.: Wide and deep convolutional neural networks for electricity-theft detection to secure smart grids. IEEE Trans. Ind. Informatics **14**(4), 1606–1615 (2018). https://doi.org/10.1109/TII.2017.2785963
2. Khan, Z.A., Adil, M., Javaid, N., Saqib, M.N., Shafiq, M., Choi, J.G.: Electricity theft detection using supervised learning techniques on smart meter data. Sustainability **12**(19), 1–25 (2020). https://doi.org/10.3390/su12198023
3. Gungor, V.C., Lambert, F.C.: A survey on communication networks for electric system automation. Comput. Networks **50**(7), 877–897 (2006). https://doi.org/10.1016/j.comnet.2006.01.005
4. Nagi, J., Mohammad, A.M., Yap, K.S., Tiong, S.K., Ahmed, S.K.: Non-technical loss analysis for detection of electricity theft using support vector machines. In: PECon 2008—2008 IEEE 2nd International Power Energy Conference, no. PECon 08, pp. 907–912 (2008). https://doi.org/10.1109/PECON.2008.4762604
5. Muniz, C., Figueiredo, K., Vellasco, M., Chavez, G., Pacheco, M.: Irregularity detection on low tension electric installations by neural network ensembles. In: Proceedings of International Joint Conference Neural Networks, pp. 2176–2182 (2009). https://doi.org/10.1109/IJCNN.2009.5178985
6. Sun, R., Guo, J., Gill, E.K.A.: Opportunities and challenges of wireless sensor networks in space. In: 61st International Astronautical Congress 2010, IAC 2010, vol. 6, no. 10, pp. 4993–5004 (2010)

7. Amin, S., Schwartz, G.A., Cardenas, A.A., Shankar Sastry, S.: Game-theoretic models of electricity theft detection in smart utility networks. IEEE Control Syst. Mag., no. February, 66–81 (2015)
8. Dos Angelos, E.W.S., Saavedra, O.R., Cortés, O.A.C., De Souza, A.N.: Detection and identification of abnormalities in customer consumptions in power distribution systems. IEEE Trans. Power Deliv. **26**(4), 2436–2442 (2011). https://doi.org/10.1109/TPWRD.2011.2161621
9. Di Martino, M., Decia, F., Molinelli, J., Ferńandez, A.: Improving electric fraud detection using class imbalance strategies. In: ICPRAM 2012—Proceedings of 1st International Conference Pattern Recognition Application and Methods, vol. 2, pp. 135–141 (2012). https://doi.org/10.5220/0003768401350141
10. Khalifa, T., Naik, K., Nayak, A.: A survey of communication protocols for automatic meter reading applications. IEEE Commun. Surv. Tutorials **13**(2), 168–182 (2011). https://doi.org/10.1109/SURV.2011.041110.00058
11. Nagi, J., Yap, K.S., Tiong, S.K., Ahmed, S.K., Nagi, F.: Improving SVM-based nontechnical loss detection in power utility using the fuzzy inference system. IEEE Trans. Power Deliv. **26**(2), 1284–1285 (2011). https://doi.org/10.1109/TPWRD.2010.2055670
12. Mashima, D., Cárdenas, A.A., Alvaro, A.C.: Evaluating electricity theft detectors in smart grid networks. Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 7462 LNCS, pp. 210–229 (2012)
13. McLaughlin, S., Holbert, B., Fawaz, A., Berthier, R., Zonouz, S.: A multi-sensor energy theft detection framework for advanced metering infrastructures. IEEE J. Sel. Areas Commun. **31**(7), 1319–1330 (2013). https://doi.org/10.1109/JSAC.2013.130714
14. Misra, S., Krishna, P., Saritha, V., Obaidat, M.: Learning automata as a utility for power management in smart grids. IEEE Commun. Mag. **51**(1), 98–104 (2013). https://doi.org/10.1109/MCOM.2013.6400445
15. https://blog.ipleaders.in/electricity-theft-punishments-india/
16. Jindal, A., Dua, A., Kaur, K., Singh, M., Kumar, N., Mishra, S.: Decision tree and SVM-based data analytics for theft detection in smart grid. IEEE Trans. Ind. Informatics **12**(3), 1005–1016 (2016). https://doi.org/10.1109/TII.2016.2543145
17. Ahmad, T., Chen, H., Wang, J., Guo, Y.: Review of various modeling techniques for the detection of electricity theft in smart grid environment. Renew. Sustain. Energy Rev. **82**(August), 2916–2933 (2018). https://doi.org/10.1016/j.rser.2017.10.040
18. Li, S., Han, Y., Yao, X., Yingchen, S., Wang, J., Zhao, Q.: Electricity theft detection in power grids with deep dearning and random forests. J. Electr. Comput. Eng. **2019** (2019). https://doi.org/10.1155/2019/4136874
19. Ahmad, S.: Detecting And Minimizing Electricity Theft : A Revie," vol. 1, no. 1, pp. 8–12 (2016)
20. Coma-Puig, B., Carmona, J., Gavalda, R., Alcoverro, S., Martin, V.: Fraud detection in energy consumption: a supervised approach. In: Proceedings—3rd IEEE International Conference on Data Science and Advanced Analytics (DSAA 2016,) pp. 120–129 (2016). https://doi.org/10.1109/DSAA.2016.19
21. Jokar, P., Arianpoo, N., Leung, V.C.M.: Electricity theft detection in AMI using customers' consumption patterns. IEEE Trans. Smart Grid **7**(1), 216–226 (2016). https://doi.org/10.1109/TSG.2015.2425222
22. Singh, S.K., Bose, R., Joshi, A.: Entropy-based electricity theft detection in AMI network. IET Cyber-Physical Syst. Theory Appl. **3**(2), 99–105 (2018). https://doi.org/10.1049/iet-cps.2017.0063
23. Li, W., Logenthiran, T., Phan, V.T., Woo, W.L.: A novel smart energy theft system (SETS) for IoT-based smart home. IEEE Internet Things J. **6**(3), 5531–5539 (2019). https://doi.org/10.1109/JIOT.2019.2903281
24. M. Nazmul Hasan, R. N. Toma, A. Al Nahid, M. M. Manjurul Islam, and J. M. Kim, "Electricity theft detection in smart grid systems: A CNN-LSTM based approach," Energies, vol. 12, no. 17, pp. 1–18, 2019, doi: https://doi.org/10.3390/en12173310.

25. Saeed,M.S., Mustafa, M.W., Sheikh, U.U., Jumani, T.A., Mirjat, N.H.: Ensemble bagged tree based classification for reducing non-technical losses in multan electric power company of Pakistan. Electronics **8**(8) (2019). https://doi.org/10.3390/electronics8080860
26. Mujeeb, S., Javaid, N.: ESAENARX and DE-RELM: novel schemes for big data predictive analytics of electricity load and price. Sustain. Cities Soc **51**(May), 101642 (2019). https://doi.org/10.1016/j.scs.2019.101642
27. Ismail, M., Shaaban, M.F., Naidu, M., Serpedin, E.: Deep learning detection of electricity theft cyber-attacks in renewable distributed generation. IEEE Trans. Smart Grid **11**(4), 3428–3437 (2020). https://doi.org/10.1109/TSG.2020.2973681
28. Chaudhari, A., Mulay, P.: A bibliometric survey on incremental clustering algorithm for electricity smart meter data analysis. Iran J. Comput. Sci. **2**(4), 197–206 (2019). https://doi.org/10.1007/s42044-019-00043-0
29. Adil, M., Javaid, N., Qasim, U., Ullah, I., Shafiq, M., Choi, J.G.: LSTM and bat-based rusboost approach for electricity theft detection. Appl. Sci. **10**(12), 1–21 (2020). https://doi.org/10.3390/app10124378
30. Zheng, K., Chen, Q., Wang, Y., Kang, C., Xia, Q.: A novel combined data-driven approach for electricity theft detection. IEEE Trans. Ind. Informatics **15**(3), 1809–1819 (2019). https://doi.org/10.1109/TII.2018.2873814
31. Talaat, M., Alsayyari, A.S., Alblawi, A., Hatata, A.Y.: Hybrid-cloud-based data processing for power system monitoring in smart grids. Sustain. Cities Soc. **55**(July 2019), 102049 (2020). https://doi.org/10.1016/j.scs.2020.102049

Chapter 44
Behavioral Analysis of Multitenant SaaS Applications

Poonam Mangwani and Vrinda Tokekar

Abstract Cloud Computing has received significant attention as it has changed the way of computation. Looking at the benefits of a cloud, more demand for cloud computing services and Multitenancy. Multitenancy offers various advantages including low infrastructure cost, low maintenance, and reducing the application hosting cost. This paper proposes a Multitenant Multiservice SaaS model for designing Multitenant SaaS applications. The model provides a configurable service accessed by different Tenants. The paper further evaluates the Performance implementing this model using data layer patterns as Tenant Specific and Tenant Shared and generates results using Apache JMeter. Our results show that under the Tenant Specific as a Single tenant, Response time is better than Tenant Shared for increasing load defined by the number of concurrent users. With Tenant Shared schema handling multiple tenants with a single instance, the Response time with an increasing number of tenants and concurrent users is almost stable. Since the Tenant Shared model is resource efficient as compared to Tenant Specific, its performance is affected due to shared instances instead of separate application instances for each tenant.

44.1 Introduction

Cloud computing has flourished in the computing world. A business model that delivers various services as required to multiple customers using a pay as you go basis through the internet. According to NIST [1], "a cloud model is composed of five characteristics, three service models and four deployment models". According to [2], One of the stepping stones of a cloud environment is sharing of resources. Sharing of resources at different levels provides benefits with minimum cost to organizations. The Infrastructure as a Service model provides access to hardware resources, Platform as a Service model provides an environment to develop, deploy and manage

P. Mangwani (✉)
International Institute of Professional Studies, DAVV, Indore, MP, India

V. Tokekar
Institute of Engineering & Technology, DAVV, Indore, MP, India

T. Senjyu et al. (eds.), *IOT with Smart Systems*, Smart Innovation, Systems and Technologies 251, https://doi.org/10.1007/978-981-16-3945-6_45

applications, and Software as a Service model delivers software applications to end users that can be accessed remotely via the Internet [3]. Customers need not have to purchase hardware and licensed software to access the SaaS application. They just have to subscribe the service and pay according to the use. Hence, SaaS is widely accepted by customers on a subscription basis as it can save 70–80% IT cost [4]. Hence many companies are interested in SaaS applications with multitenant architecture.

A multitenant SaaS application usually provides a solution for common organizational needs using a single shared application and DB instance. Multitenant SaaS supports tenant-level customization to accommodate individual tenant's requirements [5]. According to [6], a tenant is a set of users sharing a common view of Software as if they are the sole user of an application. It demonstrates the one-to-many relationship between an application and tenants. On the other hand, a single tenant SaaS serves each organization or tenant with dedicated application instances and infrastructure to fulfill the needs of just one organization. There is no sharing of resources. Multitenancy provides various advantages as "centralized version control, reduced software development time, lower maintenance cost, and much more" [7]. Multitenancy reduces the total cost of SaaS for both service providers as well as consumers. This results in more demand for developing a Multitenant SaaS application or enabling existing SaaS to Multitenant SaaS. Multitenancy can be implemented at different levels: infrastructure, database, and application level. Figure 44.1 shows the spectrum of Multitenancy in SaaS with different levels of sharing of resources among multiple tenants [8].

The two extremes of the spectrum are complete isolation as Tenant Specific and completely shared as Tenant Shared model. The former provides complete isolation that provides each tenant with a new instance with separate hardware, middleware, and database. One tenant cannot access another tenant's data and is completely secure and easily scalable. This model can support an average number of tenants.

The Tenant Shared model can support a large number of tenants where several tenants sharing a single instance of an application and database. Being a single shared

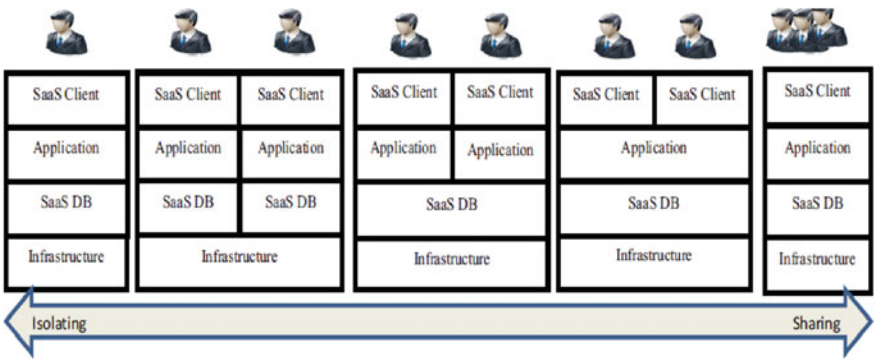


Fig. 44.1 Multitenancy SaaS spectarum [8] (*Source* IEEE, SOSE, p. 412)

instance, it is easy to maintain and deploy which saves effort and time and lowers the cost. This model is generally used where tenant data are almost the same and are willing to accept shared data with minimum cost. The selection of a Multitenancy model depends on the organization's requirement and specific application scenario. For example, a large organization may prefer to pay a higher amount with the isolated model without compromising the risk associated with resource sharing between tenants. While most small and medium enterprises (SME) would prefer lower costs with minimum isolation [9].

This paper contributes the empirical evaluation of the performance of the Tenant Specific and Tenant Shared data model under Multitenant Multiservice architecture using a real-life scenario. For this, SaaS online utility booking application with different services categorized as Travel, Stay and Local Conveyance is developed. Then the performance is evaluated in terms of Response time and Throughput for different load conditions with the different number of tenants and users, using Apache JMeter. Figure 44.2. shows the proposed model where a single instance of Multitenant SaaS application provides different services. Tenants can subscribe to multiple services and these services can be accessed by their users.

The next section describes related work that includes cloud computing, SaaS, and multitenant SaaS. Section 44.3 introduces the Research methodology and details about an online utility booking application for Performance Evaluation. Section 44.4 and 44.5 present Experimental Setup and Results. Section 44.6 concludes the paper.

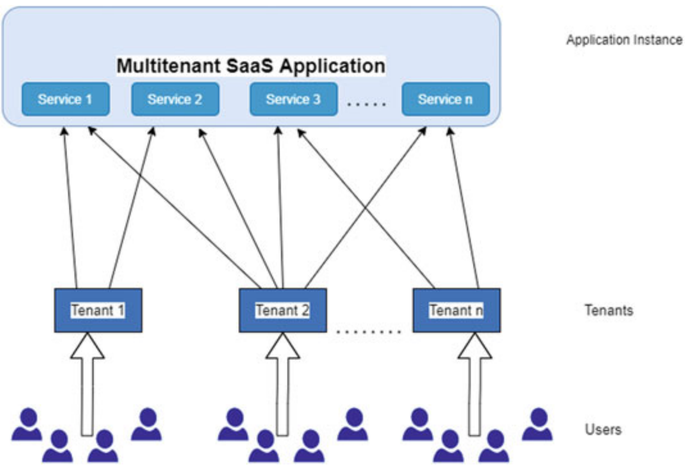


Fig. 44.2 A multitenant multiservice SaaS model

44.2 Related Work

In this section, previous research work, models, and different implementation patterns related to SaaS and analysis of Multitenant applications are presented.

Gao et al. in [10], focuses on scalability and performance evaluation for SaaS in clouds using graphic models. The results show how effective the proposed model is in evaluating SaaS performance and scalability. However, it lacks in performance evaluation and scalability in a Multitenant environment. In [11], the author analyzed different types of multitenant data tier implementation patterns. The results show the highest performance with Dedicate table/schema pattern and dedicated database pattern and the share table/schema pattern faces a decrease in performance. The authors in [12], have analyzed the behavior of cloud applications. Various experiments are performed using a single application executed in isolation for different machine configurations, multiple applications with conflicting resource demands, and Prediction of the application behavior based on resource usage. Krebs et al. [3], proposed a benchmark for a multitenant application for evaluating the maximum throughput and the number of tenants an application can support. This work motivated us to evaluate the performance of our Multitenant Multiservice SaaS application. The authors in [13], proposes an innovative quality model based on software quality and service quality, and few metrics of SaaS are derived. In this paper [7], the author proposes a Multitenant SaaS component-based scalable component model and is evaluated using a simulator. The results show that the proposed model supports increased application load distribution and provides better application response time. In [14], outlined the comparison between SaaS and Web application from customer behavior perspectives. The author proposes a tenant behavior analysis model and mining algorithm to classify customers into different segments as active tenants and loyal tenants. Experimentation is conducted on ERP SaaS. [15], introduces a middleware layer on top of Platform as a Service to minimize the gap between configuration and customization in multitenant applications. A prototype is implemented on top of the GoogleAppEngine and Guice dependency injection framework to support customizable multitenant applications. In [2], the author experimentally evaluates the performance in terms of latency of multitenant SaaS. The results show that the performance with isolated schema is better than the shared model.

44.3 Methodology

Our research employs the design and development of an online utility booking application using Multitenant Multiservice architecture to perform the behavior analysis in terms of performance. A Multitenant SaaS Online utility Booking application provides different services such as Travel, Stay, and Local conveyance which can be configured by tenants. End users can register with tenants and perform booking of these services. Different companies can register themselves as tenants whereas

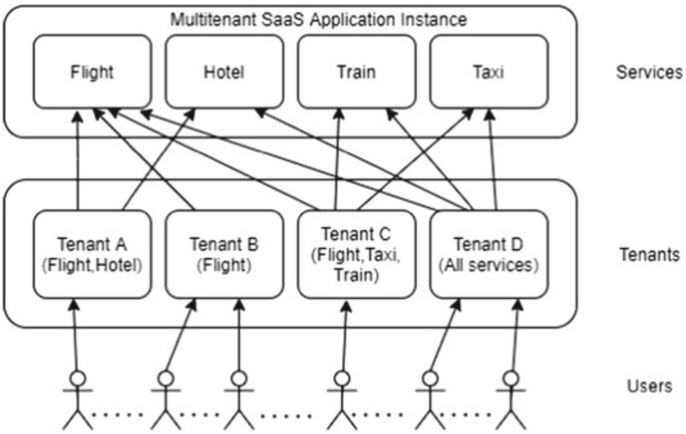


Fig. 44.3 A multitenant multiservice SaaS utility booking instance

customers are the end users performing online booking operations. A Tenant can login through URL and have their Configurable interface to select different services that correspond to their company. End users can register themselves under different tenants, can check availability, and perform online booking transactions. The application is developed using Python Flask Framework and deployed on Google App Engine. The view is implemented using HTML, CSS, and JavaScript at the presentation level while the Business logic is implemented using Flask framework, Python. For the model part of our application, the SQL database is used whose SQL instance is deployed on Cloud SQL. Features supported by the application are:

1. *Tenant Runtime Configuration*: Tenant can register using a specified URL. Tenant is offered a facility to select different services (Train, Hotel, Flight, and Taxi) they would like to serve. They can also view several users registered with them. Tenant data is isolated and is identified in the database using tenant-id.
2. *User Runtime Configuration*: The user can register with a specific tenant and perform booking transactions as per services offered by the tenant. When the user logins, different services available under the tenant can be viewed. Based on the availability of seats, users can perform booking operations. Figure 44.3. shows an application instance with four tenants, Tenant D serving all four services, Tenant C serves Flight, Taxi, and Train booking, Tenant B provides only Flight whereas Tenant A serves Flight and Hotel Booking.

44.4 Experimental Setup

There is a stringent requirement for analysis in terms of performance, as several concurrent users registered with different tenants will be accessing the application.

Performance is an important factor in understanding the behavior of Multitenant SaaS applications. As our application is a Multitenant SaaS application for online utility booking, it is deployed on Google App Engine, a cloud platform. Currently, the focus is on performance evaluation by experimenting with the SaaS online booking application under certain load conditions using Apache JMeter.

The objective of our research is to perform the following:

- To compare the performance of multitenant online utility applications under Tenant Specific model Vs Tenant Shared model in terms of Response time and Throughput.
- Relationship between tenants and users (varying tenants and users).
- The research comprises two different scenarios for evaluating the performance:

  Scenario 1: Tenant Specific Model: Here each tenant has a separate instance of VM, application instance, and database. An application is deployed on the application server running on vCPU (Intel ® Xeon ® CPU @2.20 GHZ with 2 CPU cores) with 3.75 GB memory and 10 GB SSD storage. Thus, the application uses a separate database schema hosted within a cloud. MYSQL 5.7 is executed on a database server.

  Scenario 2: In Tenant Shared Model, with a single instance of the application being deployed and shared among tenants, each tenant can access its data using a specific id. Thus, the experimental setup comprises of single Virtual Machine, shared application, and database schema. The specifications are the same as scenario I.

44.4.1 Testing Methodology

To achieve the above goals, several experiments were performed using both scenarios with varying numbers of tenants and users. Starting with one active tenant (Scenario I), and the number of users was fixed to 200, 400, 600, 800, and 1000. For Scenario II, Tenants were increased from one to four, eight, and twelve tenants with an above-mentioned number of concurrent users. Tenant and Users load is managed using Poisson Random Timer in JMeter to generate random delays between tenants and users. This timer is based on the Poisson Distribution Function. The Test is run for 20 min and at random times and subsequently, data is collected through Apache JMeter and displays performance metrics through various formats. The results were then averaged. The application response time is based on the number of concurrent tenants and users. The Average Response time is the average time taken by the concurrent booking transactions. Throughput is calculated as transactions/unit of time. It represents the load on the server generated by tenants and users performing transactions [16]. Apache JMeter application is used to perform load testing and measure average Response Time and Throughput of the multitenant online utility application. The concurrent workload of several Tenant and users are simulated by creating threads and thread groups.

44.5 Results

This section presents the results of several experiments performed under both the scenarios as Tenant Specific and Tenant Shared models. The various curves represent results with the average response time in shared multitenancy with four, eight, and twelve tenants while in Tenant specific model as a single tenant, for the different number of users. Figure 44.4 shows that the average response time of a single tenant is gradually increasing with the increasing number of users. While for multitenant applications the average response time with the increasing number of tenants and users with a tenant is almost the same. The figure also shows that there is a small increase in response time as we move from a single tenant to a multitenant.

Figure 44.5 shows Throughput based on the number of tenants and users. The Throughput is defined as the number of transactions per second that occurred at the server. It represents the load on the server. The throughput increases as the number of users increases. However, it does not change much with the increase in tenant size. Moreover, maximum throughput is achieved with a single tenant. As we move from a single tenant to multiple tenants, there is a very small degradation in throughput.

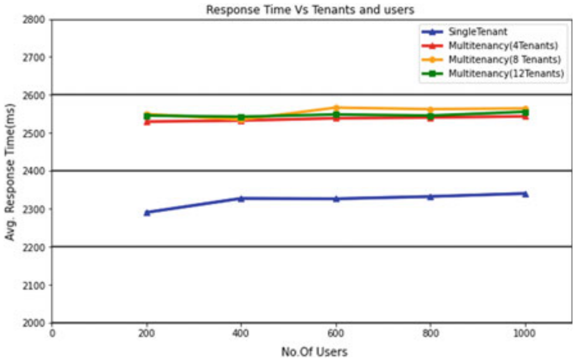


Fig. 44.4 Response time of tenants and users

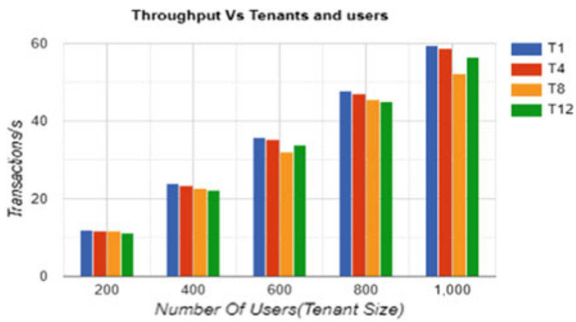


Fig. 44.5 Throughput versus users

44.6 Conclusion

Performance concerns the most important feature in the case of multitenant applications, where a single instance serves multiple tenants. The results show that response time with single tenant is better than multitenant architecture. As the number of tenants is increased to four, eight, and twelve, there is a small increase in response time. Also, the response time increases negligibly with an increasing number of users. Maximum throughput is achieved with single tenant than multitenant. Generally, a small decrease in throughput is observed as the number of tenants increases. Though multitenant SaaS being resource efficient, it does not provide these at the cost of performance degradation. Scalability analysis of Multitenant applications will be further evaluated.

References

1. Mell, P., Timothy, G.: The NIST Definition of Cloud Computing (2011)
2. Abdul, A.O., Bass, J., Ghavimi, H., MacRae, N., Adam, P.: Multi-tenancy design patterns in saas applications: a performance evaluation case study. Int. J. Digit. Soc. **9**, 1367–1375 (2018)
3. Krebs, R., Wert, A., Kounev, S.: Multi-tenancy performance benchmark for web application platforms. Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics) 7977 LNCS, pp. 424–438 (2013)
4. Su, W., Hu, J., Lin, C., Shen, S.: SLA-aware tenant placement and dynamic resource provision in SaaS. In: Proceedings—2015 IEEE International Conference on Web Services (ICWS), pp. 615–622 (2015). https://doi.org/10.1109/ICWS.2015.87
5. Tsai, W.T., Sun, X.: SaaS multi-tenant application customization. In: Proceedings of the 2013 IEEE Seventh International Symposium on Service-Oriented System Engineering (SOSE 2013), pp. 1–12 (2013). https://doi.org/10.1109/SOSE.2013.44
6. Tsai, W., Huang, Y., Shao, Q., Bai, X.: Data partitioning and redundancy management for robust multi-tenancy SaaS. Int. J. Softw. Informatics **4**, 437–471 (2010)
7. Samrajesh, M.D., Gopalan, N.P., Suresh, S.: A scalable component model for multi-tenant SaaS Application. Int. J. Adv. Intell. Paradig. **8**, 191–206 (2016)
8. Gao, J., Bai, X., Tsai, W.T., Uehara, T.: SaaS testing on clouds—issues, challenges, and needs. In: Proceedings of the 2013 IEEE Seventh International Symposium on Service-Oriented System Engineering (SOSE 2013), pp. 409–415 (2013). https://doi.org/10.1109/SOSE.2013.98
9. Chang, J.G., Sun, W., Huang, Y., Zhi, H.W., Gao, B.: A framework for native multi-tenancy application development and management. In: The 9th IEEE International Conference on E-Commerce Technology and The 4th IEEE International Conference on Enterprise Computing, E-Commerce and E-Services (CEC/EEE 2007), pp. 551–558 (2007). https://doi.org/10.1109/CEC-EEE.2007.4.
10. Pattabhiraman, P., Bai, X., Tsai, T.: SaaS Performance and Scalability in Clouds. In: Proceedings of 2011 IEEE 6th International Symposium on Service Oriented System, pp. 61–71 (2011)
11. Wang, Z.H., et al.: A study and performance evaluation of the multi-tenant data tier design patterns for service oriented computing. In: IEEE Int. Conf. E-bus. Eng. ICEBE'08 - Work. AiR'08, EM2I'08, SOAIC'08, SOKM'08, BIMA'08, DKEEE'08 94–101 (2008). https://doi.org/10.1109/ICEBE.2008.60.

12. Kaur, P.D., Chana, I.: Behavior analysis of multifarious cloud applications. In: Proceedings—2014 Fourth International Conference on Advanced Computing & Communication (ICACC 2014), pp. 175–178 (2014). https://doi.org/10.1109/ICACC.2014.49.
13. Jagli, D., Purohit, S., Subhash Chandra, N.: Saasqual: a quality model for evaluating Saas on the cloud computing environment. arXiv (2019). https://doi.org/10.1007/978-981-10-3818-1
14. Duan, N., Zhang, X., Sun, W., Li, L., Hu, K.: Tenant behavior analysis in software as a Service environment. In: Proceedings of 2011 International Conference on Service Operations and Logistics and Informatics (SOLI 2011), pp. 132–137 (2011). https://doi.org/10.1109/SOLI.2011.5986542
15. Walraven, S., Truyen, E., Joosen, W.: A middleware layer for flexible and cost-efficient multi-tenant applications. Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics) 7049 LNCS, pp. 370–389 (2011)
16. Users manual. https://jmeter.apache.org/usermanual/index.html

Chapter 45
Hackathon Methodology for E-Governance: Can We Get the Problem Solvers and Solution Seekers on One Common Platform?

Dipali D. Awasekar and Shashikant A. Halkude

Abstract The GOI has approved the 'Digital India' programme with the vision to transform India into a digitally empowered society and knowledge economy. E-Governance is among the nine pillars of the Digital India programme. At present E-Governance programme for Government of India has been designed & developed by the National Informatics Centre (NIC). To work towards our vision for Digital India and promote creativity the Ministry of HRD, AICTE, and UGC are undertaking multiple initiatives; one of the key initiatives collectively of MHRD is Smart India Hackathon (SIH). A direct outcome of this study was student's participation in the Smart India Hackathon 2020. In this paper, the students' informal learning gained through the design and development of 'TrackIT' an E-governance application to monitor all files movements for the Government of Sikkim. 90% of the team members reported having a very positive and new experience while developing a novel solution for the E-governance challenge faced by the Government of Sikkim.

45.1 Introduction

E-Governance has become the top priority of government efforts around the world. The Government of India (GOI) has approved the 'Digital India' programme with a vision to transform India into a digitally empowered society. The Digital India initiative aims to provide thrust to nine pillars identified as growth areas. These pillars include—broadband highways, everywhere mobile connectivity, public internet access program, e-Governance, e-Kranti, information for all, electronics manufacturing, IT for Jobs, and early harvest programs. E-Governance is among the nine pillars of the Digital India programme.

E-Governance Pillar focuses on reforming Government through technology to enable efficient government processes. E-Governance paradigm shift has brought

D. D. Awasekar (✉) · S. A. Halkude
Department of Information Technology, WIT, Solapur, India
e-mail: ddawasasekar@witsolapur.org

T. Senjyu et al. (eds.), *IOT with Smart Systems*, Smart Innovation, Systems and Technologies 251, https://doi.org/10.1007/978-981-16-3945-6_46

about a revolution in the value of service delivered to the citizens. At present E-Governance programme for Government of India has been designed & developed by the National Informatics Centre (NIC). If India has to progress and innovate, then the primary requirement is to create a large pool of good problem solvers. The Ministry of HRD, AICTE and UGC has realized this requirement and are undertaking multiple initiatives to work towards our vision for Digital India, promote creativity, critical thinking, and development of cognitive skills. One of the key initiatives collectively of MHRD is Smart India Hackathon (SIH).

45.1.1 SIH for E-Governance

SIH is a nationwide initiative to provide students with a platform to solve some of the vital and urgent problems faced by our country. It also aims to recognize disruptive digital innovations and thus inculcate a culture of product innovation and a mindset of problem solving. The philosophy behind MHRD's Smart India Hackathon is very simple. On one hand, we have millions of problems in this country and on the other hand, we have millions of students who are expected to be future problem solvers. Can we get these problem solvers and solution seekers on one common platform? This opportunity will not only help create good innovative products but will also give immense technical experience to students making them far more competent and market read. A direct outcome of this study was student's participation in the Smart India Hackathon 2020.

In this paper, investigates and reports how participation in Smart India Hackathon 2020 shifted student's focus from academic grades to the importance of application based practical knowledge and in turn contributing to nation building by being a solution provider to the pressing problems faced by the citizens. Secondly, we report the students' informal learning gained through the design and development of 'TrackIT' an E-governance application to monitor all file movements for the Government of Sikkim. The beginning of Hackathon is described in section two followed by the structure of SIH 2020 in section three. In section four we list the research question and r in section five we elaborate our 'TrackIT' case study. Finally, in section six answer the research question followed by conclusion in section seven.

45.2 Literature Review

45.2.1 Origin of Hackathon and Usefulness to Engineering Education

The term Hackathon was first exercised to describe an event aimed at idea generation and to create solution for predefined and identified problems within a specified timeframe. Literature reveals Hackathons serve a range of applications apart from idea generation and innovation. Of the existing literature available large part is non-academic and includes online articles from internet blogs [1–3]. As mentioned in [1, 2, 4].

The SIH primarily focuses on engaging engineering college students to propose innovative solutions to the problems listed by government ministries and private industries on the SIH website. The students registering for SIH are required to send an idea's presentation in the first round. If selected by the concerned ministries the student teams are expected to develop the product further. This problem solving approach is mutually beneficial to both the students and the government. Through part taking in SIH participating teams gets experience and hands-on experience while the government ministries get their E-Governance problems solved. Through this overall SIH activity, the GOI fosters a sense of competitive learning and also awards the best ideas with a cash prize.

Literature report very few papers related to the Hackathon Phenomenon have been written. The finding of those educational papers is restricted to the description of the events. However, there is no specific paper on SIH in India. Also, we find no account of literature reporting on the success achieved by SIH past edition apart from government reports. Also, it's only the government reports that mention the success of completion of the award winning projects, but no research paper reports the achievement. Thus, we have attempted to concentrate on this gap by inspiring final and pre final year students of our Institute to propose a novel solution for the E-governance challenge faced by the Government of Sikkim and presenting the effectiveness.

45.3 Smart India Hackathon 2020 (SIH 2020)

SIH is a non-stop product development competition. Each participating team comprises 6 students and 2 mentors. From a top perspective it has the 3 main components:

1. Organizers: The MHRD Innovation Cell, All India Council of Technical Education (AICTE), i4c, and Persistent System
2. Problem Statements Proposers: By GOI Ministries and Departments, Public Sector Undertakings and Non Government Organizations
3. Innovators: The bright minds across the nation who come up with a solution to the problem statements.

The overall SIH process includes problem statements submission. The finalized problem statements are and displayed on the website followed by registration of college Single Point of Contact (SPOC) SPOCs on SIH portal. Then the SPOC has to give a list of problem statements that the teams from the institute are interested to work on for the internal Hackathon as well as the main event. An internal Hackathon was conducted in the first week of February 2020 within each institute to select the top 5 whose ideas will be presented at the National Level. Team leaders of the shortlisted were registered in the SIH portal by the SPOC. Here, they had to submit their idea solution document for the problem statement they had chosen. For SIH Grand Finale the teams were announced. The list of the top 4–7 teams selected for each problem statement to compete in the SIH 2020 Grand Finale among all the colleges was published on the SIH website in the first week of March 2020.

45.4 Research Question

From an academic perspective we attempted to address the Smart India Hackathon phenomenon thus the research questions (RQ) addressed in this paper are:

- RQ 1: How successful was the participating team in proposing an E-governance solution works on a live project and in turn contributing to nation building?
- RQ 2: Did participation in SIH 2020 instill a mindset of problem solving culture of product innovation and among the team members?
- RQ 3: Did the participating team gain hands-on experience by creating an E-Governance prototype for providing a solution to monitor all file movements of every employee?
- RQ 4: How usable is 'TrackIT' as perceived by the evaluators and industry experts?

45.5 'TrackIT' for E-Governance Case Study

In this section, we present our functional prototype 'TrackIT' developed for the Rural Development Department (RDD) of Govt. of Sikkim to overcome the problem with the conventional HR Management System. One of the major drawbacks with traditional systems at RDD was to keep the track record of individual files. 'TrackIT' is a step in the direction of realizing the concept of the paperless office for enhanced e-governance. 'TrackIT', is a Quick Response (QR) based system aimed for highly efficient and transparent inter-government & intra-government transactions as shown in Fig. 45.1.

45.5.1 Workflow of the System

The application consists of the following modules: general desk module, admin module, employee module, office head module. The users of the developed system include super admin, office head of the respective department, general desk, employee, and HR. Figure 45.2 demonstrates the workflow of application. The file is submitted by HR to the general desk. For every incoming file in the department 'TrackIT' generates a unique QR code. The unique QR code for each file provides a

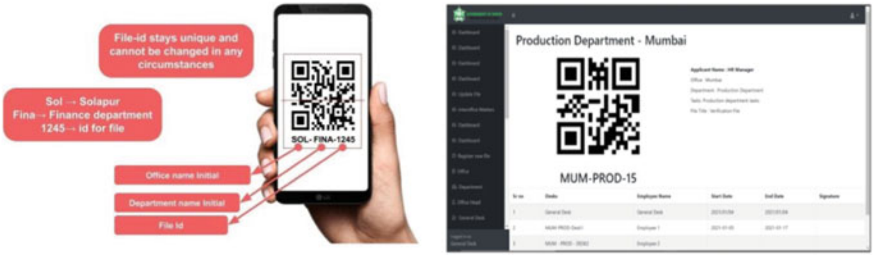


Fig. 45.1 QR Code structure and QR Generation by 'TrackIT' application

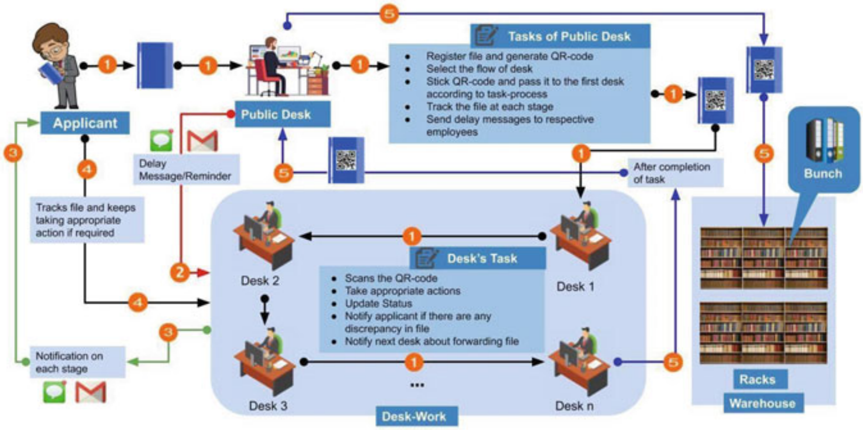


Fig. 45.2 TrackIT': Work flow diagram

special functionality to track the file. At the general desk, the new file is registered and the application/task details are filled. After registration QR code is generated through the inbuilt feature, further the generated QR code is pasted on the registered file. The government employees then directly scan the QR code and transfers file to the appropriate desk for further processing. Employees have to register themselves to the application using an IMEI number for security reasons. The application cannot be used after sharing through any sharing applications like Bluetooth and other similar applications. The employee module updates file status and perform required actions on file. Timely notifications are sent from time-to-time through email and SMS to each concerned person.

45.6 Discussion

As soon as the SIH 2020 Hackathon was completed, a survey link was emailed to all the six participating team members of SIH, and everyone (100%) submitted it. In this section, we answer our research question mentioned in Section four.

RQ 1: How successful was the participating team in proposing an E-governance solution working on a live project and in turn contributing to nation building?

A total of 21 ideas were submitted for this problem statement submitted by Govt. of Sikkim from all over India. All the 21 submissions were evaluated by experts from the industry and government on points principally focusing on novelty of the idea and practicability. Four teams were selected out of 21 teams for the grand finale to work on the problem statement given by Govt. of Sikkim. Our team 'Code Bluff' was selected out of the 4 teams for the final SIH 2020. The first author [1] was the mentor to the selected team and had a noteworthy role in mentoring the students for SIH 2020. The team's idea proposal selection at the grand finale demonstrates the team's potential in proposing an E-governance solution.

RQ 2: Did participation in SIH 2020 instill a mindset of problem solving culture of product innovation and among the team members?

> Of course the participation in SIH 2020 could instill a mindset of problem solving and inculcating product innovation culture among the participating teams.

The participating team during the entire process could closely investigate real world problems, provide an opportunity to apply the concepts and principles learned academically, and innovate. Further, it inculcates the value of teaming, provides opportunity to network with peers as well as other stakeholders, boosts self-confidence, and gets recognized by industry think tanks, bureaucrats, business decision makers, and investors as well. The participating experience (90%), learning new programming language, and get acquainted with latest technology stack were the main reasons for participating in the SIH. The Government of Sikkim had submitted 7 problem statements in the software category. The participating team members depending on their interest could select the problem statement of any ministry. The challenge faced by the Rural Development Department (RDD) of Gov. of Sikkim was with the conventional HR Management System and they expected to be solved using digital technology through SIH 2020 platform by participating teams. To systemize these operations and hence to keep track of those files, our team carried out a meticulous study of the existing system to come out with a web-enabled system and decided to work on 'Transparency and Safeguard measures in HR management' submitted by the Government of Sikkim. The developed web-enabled system not only accelerated file processing but also ensured that each and every file present in the database can be tracked using the QR code system. The functional prototype provided flexibility to government departments to maintain a huge amount of pending files. Most importantly for implementing the developed application in an organization, no radical change is required in its existing functioning.

RQ 3: Did the participating team gain hands-on experience by creating an E-Governance prototype for providing a solution to monitor file movements?

Apart from the pre Hackathon preparation, during the 36 h long Hackathon, the teams work on converting their ideas into a working prototype. There are 3 mentoring and evaluation sessions each spread across 3 days. The functional prototype designed and developed by the team 'CODEBLUFF' was abbreviated as 'TrackIT'. 'TrackIT' is a web and mobile application aimed to solve the problem identified by the Rural Development Department (RDD) of Govt. of Sikkim using digital technology as a part of SIH 2020. 'TrackIT' application was designed and developed by a team of six students under the mentorship of the first author of the paper. 'TrackIT', was developed using a Quick Response (QR) based system. The participating team explored the concepts of QR code generation, How to create a QR code for a mobile app, Design Your Code, Test Your QR Code and Share Your QR Code.

RQ 4: How usable is 'TrackIT' as perceived by the evaluators and industry experts?

Once the Problem Statements were displayed on the website, our team started working on them and then has come up with their own ideas as solutions to the chosen problem statement. In pre Hackathon phase the students have a thorough literature survey and feasibility study, a local visit to government offices group discussion with local stakeholders and industry experts. For the final evaluation session, the teams had to give a complete walkthrough of their application developed during this Hackathon and also submit a final presentation document. During the mentoring session the team members interacted with the judges and representatives from the organizations gaining valuable feedback on the progress of their work and how can it be improved. During The evaluation session in the evening, the teams gave a presentation of their application and answer the jury's questions. Each team had a scheduled time slot of approx. 10–20 min for each of these sessions conducted via web conference. As comments by the evaluating members:

Evaluator 1: "The applicant does not require visiting the office to track the file in person he will be notified via SMS."

Evaluator 2: The developed application enabled the employees to maintain a consistent watch over the movement of various files in the process of decision-making.

Evaluator 3: The system has been designed in such a manner that the controlling officer of the department can view the movement of the documents and could make appropriate decisions. Saves time and commuting cost of applicants. Greater accountability of each employee is calculated by the system to observe productivity.

During the Hackathon the 90% of participating team members found the mentorship of the faculty and industry experts to be extremely useful in completing the design and development of the application. In addition, 80% of them found the working environment and atmosphere of the competition to be very comfortable. 70% of the participating team found developing a product for an authentic client to be extremely useful from the academic point of view. In addition, some of the comments about the overall experience and knowledge gained are: "I extremely loved participating in the SIH, it has provided me with great exposure to some of the real time problems faced in the society and the environment. Also, the faculty mentors were highly helpful and guiding sources. One team member pointed out "My experience on a team was challenging. I got the opportunity to work with seniors." Second member mentioned, "Overall it was a fabulous experience".

45.7 Conclusion

The experiential learning through SIH 2020 has proven to be effective in bringing innovative, cost and time-efficient, and potentially more sustainable solutions to E-Governance than traditional product development. This is because the crafted solution and web application by students, in accordance with the challenge, seem to address the problems faced by the various stakeholders. After part taking in SIH event, we firmly mention and suggest the engineering institutes and graduates should participate proactively in Hackathon to gain a real time product design experience.

References

1. Backus, M.: E-Governance and Developing Countries Introduction and examples
2. Sachdeva, S.: E-Governance Strategy in India. CIT & DIT, Government of India
3. Halkude, S.A.: Learning By Competing
4. Chowdhury, J.: Hacking Health: Bottom-up Innovation for Healthcare. Technology (2012)
5. Innovation Management Review. Issue July 2012, pp. 31–35
6. Awasekar, D.: Effect of program visualization to teach computer programming in a resource constrained classroom. In: Technology for Education Conference 2013, IIT Khagragpur (2013)
7. Jordan, S.S.: Engineering Students Rapidly Learning at Hackathon Events American Society for Engineering Education (2017)
8. Rogers, W.O., Okot-Uma: Electronic Governance: A bridged Definitive Conceptual Framework. Commonwealth Secretariat London

Chapter 46
Design and Analysis of Blockchain-Based Resale Marketplace

Lalith Medury and Siddhartha Ghosh

Abstract Resale and exchange of items carry high risk and uncertainty. Fraudulent sellers and misleading item listings cause huge losses to consumers of digital resale marketplaces. Blockchain technology and smart contracts can help minimize the risk and uncertainty associated with resale marketplaces. The decentralized network of blockchain establishes trust by acting as a neutral third party. And the smart contracts guarantee that both parties will uphold the terms of sale. In this work, we propose the design of a blockchain-based decentralized and trustworthy marketplace. This design focuses on minimizing risk and uncertainty by using smart contracts, micro-escrows, and smart lock secured point-of-exchange. Additionally, zero-knowledge proofs are employed in the design enabling sellers to disclose the partial location of the point-of-exchange (PoE) while protecting their privacy.

46.1 Introduction

Resale and exchange of items carry high risk and uncertainty. Existing resale and exchange marketplaces like OLX and Quikr enable users to list products they intend to sell or exchange. They act as neutral third party that helps buyers find items for resale, and they define and impose the policy on sellers for posting items to sell. They also mediate dialogue between the buyer and seller before carrying out the exchange.

Few alternates additionally implement algorithmic trust and reputation of consumers for reducing buyer's skepticism. However, these platforms do not assure the quality of products, neither do they provide efficient return policies. Further, they offer little to no support in the event of fraudulent transactions, therefore, leaving buyers to proceed with risk and uncertainty. Ideally, a fair resale marketplace could be responsible for resolving conflicts between involved parties and reverts the transfer of funds in case of fraud. This would require the resale marketplace to additionally act as a financial escrow and deploy personnel to investigate each claim of fraud,

L. Medury (✉) · S. Ghosh
Vidya Jyothi Institute of Technology, Hyderabad, India
e-mail: siddhartha@vjit.ac.in

T. Senjyu et al. (eds.), *IOT with Smart Systems*, Smart Innovation,
Systems and Technologies 251, https://doi.org/10.1007/978-981-16-3945-6_47

which is not only impractical but also infeasible given the low-profit margins and high cost of legal fees.

Blockchain technology and smart contracts can help minimize the risk and uncertainty associated with resale marketplaces. The decentralized network of blockchain establishes trust by acting as a neutral third party. And the smart contracts guarantee that both parties will uphold the terms of sale. Implementing blockchain technology is advantageous as the ledger is public, ensuring transparency of all transactions, and the users are pseudoanonymous, thereby protecting their identity. Smart contracts are self-executing digital contracts stored on the blockchain that execute automatically when conditions enforced by the code are satisfied. By implementing the terms of exchange as a smart contract, a decentralized market place can avoid the buyer or the seller from dishonoring the terms of the agreement.

In this work, we propose the design of a blockchain-based decentralized and trustworthy marketplace. The scope of this marketplace is limited to the resale of items. The design focuses on minimizing risk and uncertainty by using smart contracts, micro-escrows, and smart lock secured point-of-exchange. We recommend Ethereum as the choice of blockchain platform due to its wide user-base and its support for turing-complete smart contracts. Additionally, zero-knowledge proofs are employed in the design enabling sellers to disclose the partial location of the point-of-exchange (PoE) while protecting their privacy. This also helps the buyers to purchase items based on their proximity to the PoE.

The remainder of this paper is organized as follows: Sect. 46.2 explains the problem that we are addressing, reviews previous work, smart contracts, decentralized finance (DeFi), and zero-knowledge proofs and its advantages to our solution. Section 46.3 presents the design of our blockchain-based secure sale and exchange of items wherein we describe the process, identify the threat vectors, and suggest potential solutions to address them. Section 46.4 identifies potential sources of centralization and analyzes the performance of our decentralized marketplace. Finally, we conclude with a brief summary in Sect. 46.5.

46.2 Background

Blockchain technology provides a public, decentralized, immutable, and a distributed ledger [1]. The peer-to-peer network of devices provides a trustworthy platform that is also resilient against tampering attempts. It was originally proposed to implement cryptocurrency solutions but has since evolved to support a wide range of applications. It also enabled the vision of decentralized marketplaces that is community-owned and operated.

The introduction of smart contract feature enabled programmable blockchains that could automate transactions by monitoring events. In a blockchain-based decentralized marketplace, neither the seller nor the buyer can alter the terms of sale retroactively as the blockchain ledger is immutable. Furthermore, blockchains inherently provide pseudoanonymity, thus ensuring to a certain extent the privacy of all users by protecting their identity. The P2P nature of a blockchain enables the community of sellers and buyers to democratically design the terms of sale without relying on an intermediary.

A decentralized digital marketplace is a platform where no single entity or organization has the authority to control the entire marketplace [2]. Typically, purchase of items can either take place directly between the buyer and seller (P2P) [3] or involve an intermediary [4]. It is beneficial to utilize a blockchain as the ledger is decentralized, preventing any entity from taking control of the entire network. The ledger is immutable, denying any attempts made to alter the information stored on the ledger. It is also public and available to all, ensuring transparency of all transactions and pseudoanonymous to protect the identity of the users [1].

46.2.1 Smart Contracts and Zero-Knowledge Proofs

A smart contract defines rules and automatically enforces them via code. The term smart contract was first coined by Nick Szabo in 1994 [5] but gained popularity after the introduction of blockchain and Ethereum. Smart contracts can perform turing-complete operations, and most blockchain frameworks and protocols support turing-complete smart contracts. An arbitrary logic sequence can be implemented as a smart contract and then deployed to a blockchain. They can then accept inputs and execute operations in the logic sequence. Ethereum first introduced the feature of smart contracts in 2015 [6].

Zero-knowledge proofs guarantee that the prover can convince the verifier of the statement, for example, $\phi(x)$ = true without revealing what x is [7]. Goldwasser, Micali, and Rackoff introduced the concept of zero-knowledge proofs (ZKP) in the 1980s [8]. A well-known application of ZKP is the cryptocurrency ZCash. Users' cryptocurrency addresses and balances in blockchains like bitcoin and Ethereum are public and can be viewed by looking up in a block explorer. ZCash implements zK-SNARKs [9] to prove that the conditions for a valid transaction are satisfied without revealing any crucial information about the addresses and balances of the user.

46.2.2 Related Work

Other research works [3, 10] have explored leveraging blockchain technology to design a decentralized marketplace. Kabi and Franqueira [3] identified the problems associated with a centralized marketplace and real-world escrow services, and they developed a proof-of-concept of a blockchain-based decentralized marketplace using Ethereum blockchain protocol. Their work applies to a general marketplace applying to a broad range of transactions, and we focus specifically on the resale market segment. Further, they focus on the operation of the marketplace and do not provide sufficient information to mitigate the risks associated with the exchange of items. We will be identifying the potential risks in a decentralized marketplace and suggesting suitable solutions to address them.

Prasad et al. [10] prototyped a decentralized marketplace wherein the seller can enlist the items on the marketplace, and interested buyers can submit their bid within a given time to purchase the item. Bids are entered and executed via smart contracts. Product images are stored on InterPlanetary File System (IPFS); their hash is submitted to the smart contract to be stored on the blockchain.

Their work does not mention how the buyer communicates with the seller to receive the product. In our work, we have proposed a design of a hardware secure point-of-exchange. It allows the buyer to purchase the item after viewing it, ensuring the buyer is satisfied with the purchase.

46.3 Design of Blockchain-Based Resale Marketplace

The operational logic involved in the sale of an item can be implemented as a smart contract. Buyer and seller can interact with this smart contract using their blockchain wallets. In this section, the design of a decentralized application using a smart contract is described and illustrated. Potential risks associated with such an application are listed, along with recommended mitigations.

The smart contract designed for the resale marketplace must handle the logic for listing, purchasing, and transferring ownership of the item. It must allow the seller to record item-specific details in a listing. The logic to accept inputs for a listing is defined by the policy for listing an item in the marketplace. The smart contract can securely automate policy compliance and validate all the required attributes before successfully listing the item.

46.3.1 Process

The process is explained from start to end in Fig. 46.1, and the lifecycle of each listing is shown in Fig. 46.2. The process is as follows:

- After successful validation of all attributes related to an item, the listing process is considered complete and the listing state is advanced to *Available* state. The front end of the application can monitor for listings and their attributes in *Available* state and display them according to the needs of interested buyers.
- Buyers can express intent to purchase items by initiating micro-escrow (simplified escrow, explained in more detail in Sect. 46.3.2) smart contract, which would require the buyer to deposit funds equivalent to the listing price of the item. After successful deposit, the state of the listing is updated to *Pending*.

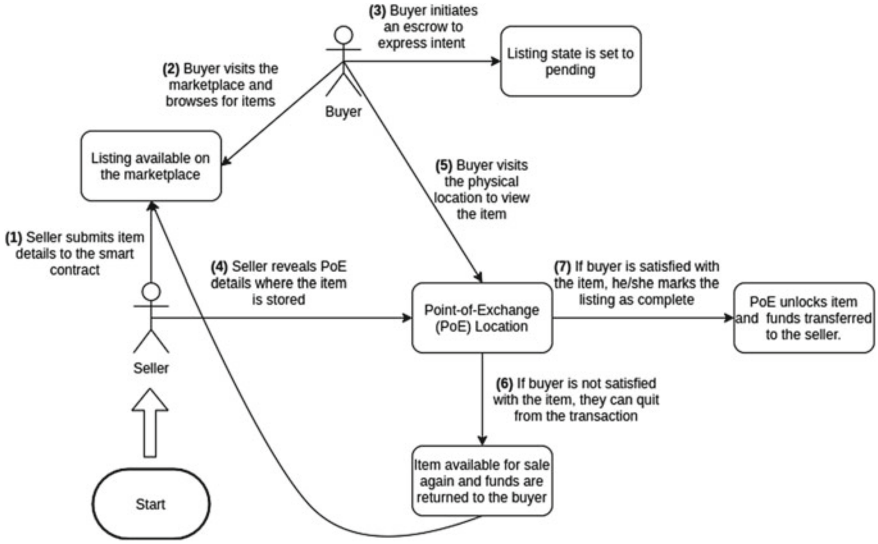


Fig. 46.1 Flow diagram of our decentralized marketplace

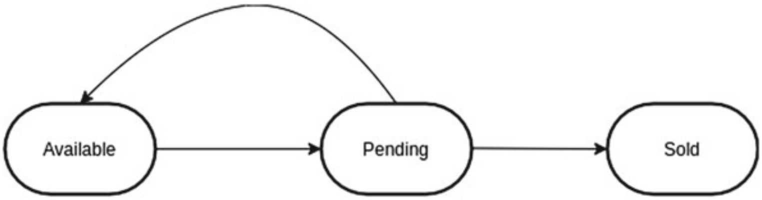


Fig. 46.2 Lifecycle of listing's state in our decentralized marketplace

- The seller may then provide the physical location of the smart lock secured point-of-exchange (PoE) (explained in detail in Sect. 46.3.3) containing the item. Alternately, the seller may also choose to outsource this procedure to smart contract. The PoE would be configured to unlock upon proving ownership of wallet address to the buyer's account in micro-escrow. Upon verification, the buyer would be allowed to view the item.
- If the buyer is satisfied with the quality of the product and wishes to proceed with the transaction, he/she can complete the transaction by invoking the micro-escrow smart contract method to release and transfer the funds to the seller. After the successful transfer of funds, the hardware box will allow the buyer to take possession of the item.
- If the buyer is not convinced with the item, he/she can quit the transaction, and the smart contract will automatically return the locked-in funds to the buyer's account. The item will then be available for display in the marketplace for other buyers.

Certain time-sensitive items like event tickets and perishable items may be suitable for bidding, in case the buyer claims that the product does not match the description. More details about item bidding are explained in Sect. 46.3.4.

With this approach, we have eliminated the following potential risks:

- The item stored in the PoE may not match the item description added by the seller in the listing. In case the exchange of funds had gone through, the buyer would have not gained any value out of this transaction and lost the entire amount.
- Fraudulent sellers may post fake listings to attract and trap buyers.
- In some occasions, interactions between buyer and seller during the exchange of item can be less than civil.

46.3.2 Micro-escrow

A micro-escrow (simplified escrow) smart contract ensures that the buyer and seller respect the exchange/resale policy. Both parties would require to lock-in cryptocurrency equivalent to the sale price. The lock-in would directly depend upon how soon the buyer can get to PoE, and this period may be higher if parties are distantly located. The full extent of terms and conditions of micro-escrow is beyond the scope of this research. Some of the conditions required in a micro-escrow smart contract are:

- Release funds to the buyer (after the buyer invokes the smart contract method) if they are not satisfied with the item they have seen.
- Release funds to the seller after the buyer invokes the smart contract method to authorize the transfer of funds in exchange for the product.

The micro-escrow for a particular item will be concluded after successfully transferring funds to the seller.

46.3.3 Smart Lock Secured Point-of-Exchange (PoE)

A smart lock secured point-of-exchange eliminates all risks associated with face-to-face interaction between the buyer and seller. Furthermore, it is designed to allow buyers to view the item before deciding to proceed with the purchase. We propose the design of a smart lock secured point-of-exchange for the decentralized resale marketplace. Conceptually, it is similar to a delivery mail-box equipped with a smart lock and two doors, one opaque and the other transparent.

The hardware box should be connected to the Internet and is expected to be capable of communicating with the smart contracts. The smart locks are supposed to:

- Request the buyer to provide the Ethereum address to verify if the same user has deposited funds to the micro-escrow for that item.

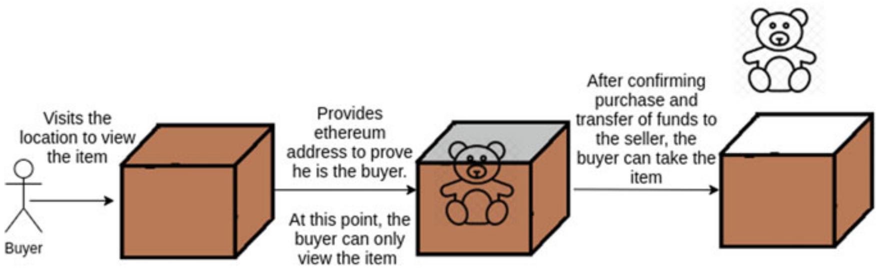


Fig. 46.3 Hardware secured point-of-exchange

- Post-verification allows the buyer to view the item (but not take possession of it) to verify the quality and accuracy of the item description listed in the decentralized marketplace.
- Allow the buyer to take possession of the item if the buyer proceeds to purchase the item by authorizing the transfer of funds to the seller (Fig. 46.3).

46.3.4 Bidding-Based Validation of Item

Certain time-sensitive items for resale, for ex: event tickets, can be put up for bidding in case the buyer claims a fraudulent product. Other interested buyers that may be within the vicinity of the point-of-exchange may enter bids to validate and buy the item. At a certain time T, the bid is supposed to close, and the highest bidder will then be authorized to view the item. If satisfied, the bidder will then proceed to buy the ticket at the bid price. Other bids will be invalidated, and the amount is released to respective accounts. In case even the highest bidder does not want to proceed with the purchase, it may be best to invalidate the listing. This would however expose the system to colluding attacks (which is highly unlikely given the escrow requirement) and mitigations against such attacks will be addressed in future work.

46.3.5 Zero-Knowledge Proofs for Sharing PoE Location

Existing marketplaces provide security and privacy for their customers by withholding certain attributes in listings and other communication. Airbnb, an online vacation and rental marketplace, provides an approximate location of the hosts in their listings but requires the hosts to upload proof-of-address documents. Uber, a P2P ride-sharing application, has recently introduced Voice over Internet Protocol (VoIP) [11] allowing the driver-partners to communicate with the riders in-app, avoiding the need for customers to share their phone numbers with drivers. These invaluable features that meant to preserve user-privacy are easier to implement in a centralized solution since customer data is not public. In an otherwise public platform, it is significantly difficult. Nevertheless, we attempt to enable these features in the design for a decentralized marketplace using zero-knowledge proofs (ZKPs).

During the process of listing an item, sellers can submit the location and unit details to the secured PoE using a ZKP. This would allow buyers to make more informed decisions based on how far the PoE is from their location. The smart contract verifies the proof before storing the item details on the blockchain.

46.4 Analysis

In this section, we offer our views on community governance in a decentralized marketplace. We also discuss costs and tradeoffs associated with our proposed design and mention potential limitations that need to be addressed for practical feasibility.

46.4.1 Governance

The decentralized marketplace is meant to be community-owned and operated. Therefore, consequential decisions, such as the attributes required in a listing and locations of PoE, need to be made carefully. The policy of a decentralized marketplace is the only component in our design that is susceptible to centralization. We recommend that even the policy be decided by the community, democratically. In keeping with the spirit of decentralization, we consider blockchain-based electronic voting or token-based voting to be ideal for this purpose.

46.4.2 Costs and Tradeoffs

This design incurs infrastructure costs including community-owned boxes of potentially different sizes and smart locks capable of communicating with the smart contracts. Other costs include developer fees, voting organizers, and interface hosting costs. The burden of bearing such costs will also be upon the customers, for which a small percentage of each completed transaction may be diverted to the community. The micro-escrow also forces the seller to lock-in amount equivalent to the listing price to ensure that he/the product is not fraudulent. This may limit the spending capacity of sellers for brief periods of time, which may not be desirable or acceptable for some sellers.

46.4.3 Limitations

Although the design reduces risk and uncertainty, the following limitations need to be addressed for a successful solution:

- Requires initial funds to bootstrap development and setup infrastructure
- Requires active community participation to keep the policy up to date, which is especially difficult to achieve on a utility platform.

46.4.4 Blockchain Analysis

Many blockchain-based distributed computing platforms exist that can support the design proposed in this research. Each platform has its unique benefits and tradeoffs, and choosing a blockchain platform to support a decentralized marketplace is an open design consideration. In this section, we highlight two choices of popular platforms and discuss their benefits and tradeoffs.

Bitcoin is the most popular cryptocurrency, and however, as shown in Table 46.1, it provides limited scalability and support for smart contracts. We highly recommend Ethereum as the choice of blockchain platform due to its wide user-base and its support for turing-complete smart contracts. However, the current proof-of-work (PoW) consensus algorithm in Ethereum limits the system's scalability to 7–15 transactions per second [12]. This scalability bottleneck can be overcome by deploying the smart contract on platforms with proof-of-stake consensus like Ethereum 2.0 or Algorand. Soon to be operational, Ethereum 2.0 employs an efficient proof-of-stake (PoS) algorithm wherein the users stake their cryptocurrency to validate the transactions. Ethereum 2.0 has also introduced sharding on their network, which further enhances scalability [13].

Algorand [14] blockchain, founded by Silvio Micali, implements a pure proof-of-stake consensus algorithm supporting over 1000 transactions per second, building on Algorand other benefits, including low transaction fees and fast processing. However, it supports limited storage for each application: 64 key-value pairs per smart contract and 16 key-value pairs per account in the smart contract. But approaches exist for overcoming this limitation to a certain extent, such as rekeying.

Table 46.1 Comparing different blockchain technologies.

Blockchain	Consensus algorithm	Block generation time (in seconds)	Transactions per second
Bitcoin	Proof-of-work	~ 600	~ 4.6
Ethereum	Proof-of-work	~ 13	~ 7–15
Algorand	Delegated proof-of-stake	~ 10	≥ 1000

46.5 Conclusion

We proposed the design of a blockchain-based decentralized and trustworthy marketplace. This design focuses on minimizing risk and uncertainty by using smart contracts, micro-escrows, and smart lock secured point-of-exchange. The operational logic involved in the exchange of an item can be implemented as a smart contract, and policy compliance can be automated by validating all listings before submission. Further, bidding of certain time-sensitive items was also proposed to minimize the loss in case the buyer tries to misbehave. Additionally, zero-knowledge proofs were employed in the design enabling sellers to disclose the partial location of the point-of-exchange (PoE) while protecting their privacy. Decentralized marketplaces are meant to be community-owned and operated and require active community participation in governance for successful operation. We also note other limitations, such as accumulating initial funds being a system prerequisite, to bootstrap operation and to set up infrastructure.

References

1. Yaga, D., Mell, P., Roby, N., Scarfone, K.: Blockchain technology overview. arXiv preprint arXiv:1906.11078 (2019)
2. Subramanian, H.: Decentralized blockchain-based electronic marketplaces. Commun. ACM **61**(1), 78–84 (2017)
3. Kabi, O., Franqueira, V.: Blockchain-based distributed marketplace (2018)
4. Banerjee, P., Ruj, S.: Blockchain enabled data marketplace-design and challenges. arXiv preprint arXiv:1811.11462 (2018)
5. Buterin, V., et al.: A next-generation smart contract and decentralized application platform. White Paper **3**(37) (2014)
6. On public and private blockchains | ethereum foundation blog. https://blog.ethereum.org/2015/08/07/on-public-and-private-blockchains/. Accessed 02 Feb 2021
7. Goldreich, O.: Foundations of Cryptography: Vol. 2: Basic Applications. Cambridge University Press (2009)
8. Goldreich, O., Micali, S., Wigderson, A.: Proofs that yield nothing but their validity or all languages in np have zero-knowledge proof systems. J. ACM (JACM) **38**(3), 690–728 (1991)
9. What are zk-snarks? | zcash. https://z.cash/technology/zksnarks/. Accessed 02 Feb 2021
10. Ranganthan, V.P., Dantu, R., Paul, A., Mears, P., Morozov, K.: A decentralized marketplace application on the ethereum blockchain. In 2018 IEEE 4th International Conference on Collaboration and Internet Computing (CIC), pp. 90–97. IEEE (2018)
11. Uber rolls out new VoIP feature for riders and drivers to communicate through the app—the verge. https://www.theverge.com/2018/10/18/17995540/uber-voip-data-call. Accessed 02 Jan 2021
12. Introduction to ethereum scaling—district0x education portal. https://education.district0x.io/general-topics/ethereum-scaling/introduction-to-ethereum-scaling/#:~:text=It. Accessed 02 Jan 2021
13. Ethereum 2.0 includes major changes that could end bitcoin's blockchain dominance | IEEE Computer Society. https://www.computer.org/publications/tech-news/trends/end-bitcoins-blockchain-dominance. Accessed 02 Mar 2021
14. Chen, J., Micali, S.: Algorand. arXiv preprint arXiv:1607.01341 (2016)

Chapter 47
An Analytical Approach for the Correction of Optical Readable Answer Sheets Using NLP

B. Sharookhan, A. Aneesh Kumar, Deepak Suresh, and C. V. Prasannakumar

Abstract We have concentrated more on the use of how to correct the answer sheets using a machine. In these days, correcting the mark sheets manually have become a burden and time-consuming methodology. In this era, we have come across identifying handwriting and converting them to text concept. So, we decided to dedicate to make up an improved version and accurate of this system which will be more helpful in evaluating the mark sheets, assignments, and notes of the students. We have gathered a lot of technology currently being used in today s' world. And thereby, taking an initiative to implement this in our educational system will be a great advantage and helpful in future. In this, we will be focusing on OCR, EPR method, applications of image processing, NLP, Quillionz platform, etc. We have managed to add more additional data into the work for an improved version of the system.

47.1 Introduction

Evaluating and identifying the handwriting of human are a current topic going around world. Finding the accuracy, evaluating how we can identify each and every individual's unique way of writing and converting them to text has been a great helpful discovery for mankind.

Evaluating them is not the only way but finding an application and fitting and folding them like a sheet in day today's life is of great importance. EPR acronym Exam Paper Reader helps in converting the hard copies into soft copies with the help of camera to take the pictures of the handwriting and scanning them with the help of VGA. And with the help of Adaptive Threshold for Color Detection (ATCD) helps to take the necessary part of the information and plays an important role in the

B. Sharookhan (✉) · A. A. Kumar · D. Suresh · C. V. Prasannakumar
Department of Computer Science and IT, Amrita School of Arts and Sciences, Amrita Vishwa Vidyapeetham, Kochi, India

C. V. Prasannakumar
e-mail: prasannakumarcv@asas.kh.amrita.edu

T. Senjyu et al. (eds.), *IOT with Smart Systems*, Smart Innovation, Systems and Technologies 251, https://doi.org/10.1007/978-981-16-3945-6_48

system. Optical character recognition (OCR) helps in recognizing the handwriting from papers and documents. Automatic evaluation of the answers using the keyword is also being implemented. Spelling-based evaluation is also tested with the help of deep learning technique.

Exam Paper Reader (EPR) image processing takes the columns to extract the useful information like our marks and roll number from the answer sheet and rest carried out using EPR image processing algorithm. Binarization and image enhancement are purely applied on documents based on thresholding and filtering hold together with image processing algorithms. Furthermore, using semantic analysis such as Information Retrieval and Extraction (IRE) and Natural Language Processing (NLP) used for deriving the necessary information are also used.

Let us see how they fit in evaluation of mark sheets with the current technology OCR, IBM, EPR image processing, ATCD algorithm, ILP, and NLP contributes into our proposed idea.

47.2 Background Study

47.2.1 Exam Paper Reader (EPR) Image Processing

EPR converts hard copy answer written papers into soft copies or digital text files were scanning is done by VGA camera. Image is then pre-processed which results in sufficient clarity or quality done by smoothening filter.

47.2.2 Hardware Part

With hack and pin mechanism and convey belt system the hardware for EPR, papers are processed one by one. Proper placement of paper on the conveyer is done with the help of sharp sensor. Soft copy of the photo that is the exam sheet must contain boxes having roll numbers and marks accurately with proper positioning. Webcam above the conveyor belt is used to take the images and to maintain the intensity of the light; a light is placed in an enclosed box. Serial communication between microcontroller and computer is attained by synchronization and interaction with webcam, and hardware is necessary.

47.2.3 Image Processing Part

By smoothening RGB the noises that is included in the picture while taking the photo of answer sheet with low-quality camera (which is easily available and affordable to

Fig. 47.1 Flow diagram of optical character recognition (OCR)

stimulate real world scenario, induces more noises), successfully reduces the noises produced from the low-quality camera (i.e., majorly the pixels of abrupt intensities). And thereby brings trade-off in sharpness of the image.

47.2.4 OCR—Optical Character Recognition

Mainly used for character recognition from the printed text. It will convert the region of interest from ATCD into machine editable text which is later updated to documents. Accuracy of system depends on the sample size of the OCR system. OCR algorithm and correlation are used for template matching to give accurate result. It involves several steps including segmentation, feature extraction, and classification. The template matching technique is used for recognition of characters.

Later then, binarization and image enhancement are carried out. Scanned document is converted into text file as shown in the Fig. 47.1 using OCR.

47.2.5 IRE (Information Retrieval and Action)

They play a major role in the query that is provided by the user for their need. These extract or retrieve the information in need.

47.2.6 NLP (Natural Language Processing)

They are a large volume of data or information. This makes easier for our machine to communicate with humans by reading text. NLP technique here is used to extract textual reference answers (TRA). These are then compared from the extracted words, i.e., from the image produced or the extracted text from the handwriting (Fig. 47.2).

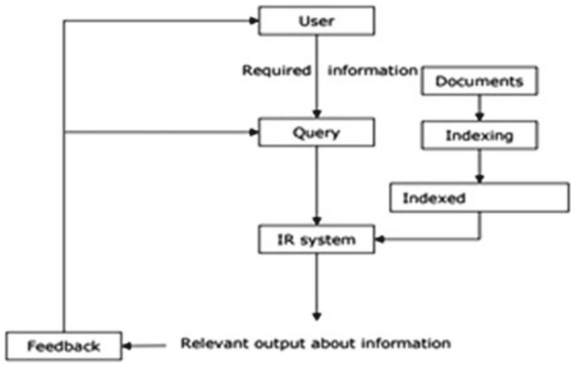


Fig. 47.2 Basic working of NLP

47.3 Related Works

In this work after identifying the handwriting or converting them into text, the next procedure that follows will be the way marking scheme is done. From the submitted answer sheets, we will identify the keywords through word spotting mechanism from the answer sheet were the keywords will be from the textual reference answers, Scoring could be performed from the reference [5].

Query Search Engine

Students usually have tendency to use synonyms and acronyms in the answers which could create difficulty in automatic evaluation for this query evaluation procedure is performed by querying the search engine. Therefore, all the solutions for this will performed with the help of the search engine. This will find all semantically related words from words in a query.

47.4 Discussion

47.4.1 Recognizing Unique Handwriting Challenge

It seems that using handwriting recognition system, we can recognize human's handwriting and could be evaluated using geometric character analysis. But still the question is could the machine learning technology could identify the more cursive complex human handwriting, could they extract them separately? The cursive handwriting of human beings sometimes confuses human brain system and still, could machine learning could identify that confuses human brain? Humans teaches the machine to work similar to human brain. But since human brain is a complex process, sometimes we fail to identify some of our friend's handwriting during taking notes. And when it comes to the machine learning part, human's handwriting varies from person to person, and still is a challenging task to identify the more cursive handwriting of students; because every individual in the world is unique and so do their handwriting. But still, there are chances to identify them and recognize them when it comes to a neat and predictable handwriting which is predicted through the geometric character analysis approach.

47.4.2 Understanding the Evaluation of Answer Sheets with Respect to NLP

A major part played in the evaluation of answer sheet is Natural Language Processing (NLP). We know that a human can master a language easily, while it is a human's duty to make it understand a machine to learn a language and perform like a human brain. And NLP became a difficult task for the machine to learn. Similarly in this process, the NLP is a great advantage and an approximate method for evaluating an answer sheet, after handwriting recognition.

Morphology

- Syntax
- Grammar
- Parsing
- Semantics
- Pragmatics.

Ambiguity

- Semantic ambiguity
- Referential ambiguity
- Local ambiguity.

These methods are adopted in NLP to check the written descriptive answers, and its role is mentioned in reference [11]. Let us see how these methods are connected while doing correction of answer sheets.

47.4.3 Grammatical Mistakes

Here grammatical mistakes are identified with the help of part of speech (POS) using inflectional morphology (Figs. 47.3 and 47.4).

If the sentence written by the student is a valid one, they will reward marks. If the sentence is not a valid one, then she will not be rewarded. In this method, the keyword in the sentence will also be checked with the answer in such a way that if it matches the answer, she will be given marks if any antonyms are present mark will be

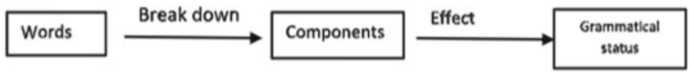
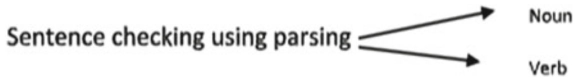


Fig. 47.3 Grammar checking work flow

Fig. 47.4 Checking the sentence is valid or not



reduced, if any synonyms are present then mark will be rewarded. The synonyms of the word in a sentence will be checked respect to keyword with the help of ambiguity in NLP. Similarly if antonyms or meaningless sentences are there, it will be given negative mark. Prediction of words is also applied.

Each and every step is evaluated using NLP reference [11]. Comparative understandable study is made with the keyword approach using NLP. We cannot find any better method to replace these methods of NLL, since it is performed neatly and accurately. All the NLP method performed have succeeded. These methods are able to detect the grammatical mistakes in sentences. And thus, NLP method is adapted in evaluation of mark sheets.

47.4.4 Time and Effort

In an organization during exam season, teacher faces a bulk amount of answer sheets to evaluate. And it takes them a lot of time and effort to devote into them. Since the grading answer sheet has come using CNN, it seems to reduce a lot of effort. Still the system has experimentally proved from the reference [7] that during the grading compared to a teacher and machine only a 0.5 difference in marking is noted; which is a great success. But we do not know if all the institutes and organizations will go to with this technology and adapt them instead of teacher's effort. Even after analyzing a mark sheet using CNN, we still make it clear that the machine performed by this task is correct or not. But still, an automated grading system is great method in educational system which reduces time and effort. In a survey, it is said that the teacher work 53 h per week on average in grading from home, stay late, answer phone calls, etc. But could this reduced by a machine?

There will be a bulk amount of answer sheets of student's in university. Adapting this method, we do not know how much time it will take to evaluate each answer sheets. We have to do an experimental study and time taken to evaluate each answer sheets compared to a teacher time taken to evaluate them. A comparison study or an experimental study should be carried out to prove this. If the machine takes less time than human with perfect grading, then the educational system can adapt them and would be a great advantage. If an experimental study based on the time taken by human v/s machine is carried out and if a favorable outcome is drawn out, then the whole educational system could adapt this.

47.4.5 Onscreen Evaluation for Answer Sheet Checking

It is an digital answer sheet evaluation method were the examiner can do the marking schemes with the help of keyword detection in this system, that makes it easier for evaluating the answer sheet more precisely with the reference [10]. Nowadays, the educational system moves toward online marking scheme due to pandemic, and this have succeeded in a great way of marching scheme. Still this method is done with the help of an examiner where he or she can evaluate the hard copies of answer sheet by converting to soft copies and evaluating on the desired document. This method is now widely adopted. In contrast to our automatic evaluating system, we are not using an examiner but evaluated automatically by the machine. But do the organizations will adapt our automatically correcting system for educational purpose? Still the proposed CNN method for evaluating answer sheet is of great advantage than the onscreen evaluation of answer sheet checking mechanism. But the currently as per to the knowledge the educational institutes or organizations have never make utilized or attempted of automatic evaluation of answer sheet generator technique. But if we are able to experiment on the time taken to do the following tasks maybe we can propose the system in every education field as mentioned in the above discussion in time and effort.

47.4.6 Estimation of Quillionz Platform

AI is used in Quillionz platform were questions are made automatically. This generates questions for the exams replacing teacher's time and effort devoted into it. Generating questions and answers will be a new breakthrough in the field if it is combined into one system. With the help of AI and machine learning, this makes it possible. Quillionz is platforms were we can create questions by giving the question and keywords as answers for evaluation, an article a website or a transcript of video in the textual format. Using these technologies adapting into one system make it easier for humans. This method could be adapted for question generation for teachers.

But we do not know the questions made are easy, medium, or tough! When it comes to textbook exercises and examination questions, the questions weight age varies. It actually tests the capability and potential of the students. We do not know how it varies when an AI performs this task. Every human beings IQ varies and so do their potential for creating easier and challenging questions. To estimate if human made question or AI made questions are challenging one, we have to make a survey or test to test the complexity of questions for the students. That's not all. We should estimate the time taken by a teacher to generate a question. The time increases proportional to the complexity of the question. But if we train our AI to do the complex methods a human brain do say making questions, then quillionz could be used.

47.5 Conclusion

The performance of the evaluating the mark sheet is deeply analyzed and reviewed which paves the path for the future development in this field that can perform in less time and effort. In this technology, the marking and identifying keywords are performed; in future, we will be able to train AI similar to human brain for creating questions and answers without the help of human. But from this review, marking and checking the marks are possible and would play a needful role in the future.

References

1. Sharma, D., Sharan, A., Sharma, H., Agharwal, A.: Data Extraction from Exam Answer Sheets using OCR with Adaptive Calibration of Environmental Threshold Parameters
2. Rowtula, V., Oota, S.R., Jawahar, C.V.: Towards Automated Evaluation of Handwritten Assessments (2019)
3. Yong, Z.Z., Yang, S.: The license plate image binarization based on Otsu algorithm and MATLAB realisze. In: International Conference on Industrial Control and Electronics Engineering, pp. 1657–1659 (2012)
4. Cuhadar, A.: Scalable parallel processing design for real time handwritten OCR: Pattern Recognition. In: Signal Processing, Proceedings of the 12th IAPR International Conference, vol. 3 (1994)
5. Sreehari, S.N., Sreehari, R.K., Babu, P., Sreenivasan, H.: On the automatic scoring of handwritten essays. IJCAI (2007)
6. Kacalack, ,W., Majewski, M.: Effective Handwriting Recognition System using Geometric Character Analysis Algorithm
7. Shaikh, E., Mohiuddin, I.A., Manzoor, A.: Automated grading for handwritten answer sheets using convolutional neural network. In: 2019 2nd ICTS
8. Rowtula, V., Oota, S.R., Jawahar, C.V. Automated Evaluation of Handwritten Assessments
9. Nanonets Automate Data Capture site by Anil Chandra Naidu Matcha (website)
10. Onlineexamhelp.eklavya website
11. Reshamwala, A., Pawar, P., Mishra, D.: Review on natural language processing

Chapter 48
A Comparison of 4-Parameter Mathematical Logistic Growth Model with Other SRGM Based on Bugs Appearing in the Software

Swati Singh, Monica Mehrotra, and Taran Singh Bharti

Abstract In this work, we present a comparison of 4-parameter logistic growth model with other SRGM models. The 4-parameter logistic growth model is a mathematical model in which the rate at which the function $m(t)$ is changing and the quantity of growth remaining by a time-dependent logistic function per quantity per unit time, have a directly proportional relation with each other. This model's result helps in determining how many failures can be expected in time t. The analysis of different existing software reliability logistic growth models is done on a data set of a software. We are considering three criteria for goodness-of-fit, such as MSE, SSE and R^2 for comparing various models. Our analysis with data set shows that among other existing growth models, the 4-parameter logistic model fits considerably better.

Notations

$m(t)$	Mean value function (count of failures expected in the software at time t).
N	Count of predictable faults that exists in the software initially.
$b(t)$	The rate of detecting faults, measured as fault in a unit of time.
a, b	Constants.

S. Singh (✉) · M. Mehrotra · T. S. Bharti
Jamia Millia Islamia, New Delhi, India

M. Mehrotra
e-mail: mmehrotra@jmi.ac.in

T. S. Bharti
e-mail: tbharti@jmi.ac.in

T. Senjyu et al. (eds.), *IOT with Smart Systems*, Smart Innovation, Systems and Technologies 251, https://doi.org/10.1007/978-981-16-3945-6_49

48.1 Introduction

As essentials of software quality improvement, the necessity for dependability on high-reliable systems is ever growing. It is extremely important to build highly stable and dependable software systems in order to provide our customers with high-quality services. However, it is a complex process of developing a software system which can be demanding a lot of time as well as cost. Improving the quality of software by enhancing software's reliability as well as keeping the expenses minimal is the strong focus of software industries [1]. Reliability is the probability which can either be one or zero. It means software will either be reliable or not [2]. In the interim, numerous development models for reliability (SRGM) have been proposed and examined for quite some time with the aim of obtaining a mathematical formula for mean value function [3].

In the course of recent decades, numerous SRGMs have been defined under non-homogeneous Poisson process (NHPP) system models. In non-homogenous Poisson process (NHPP) models, expected number of failures can vary with time. The mean value functions of various NHPP-SRGMs are used to predict future rates or the no. of residual defects in the code. In order to assess a software's reliability, various metrics such as number of faults remaining, rate of failure, etc., are used in various software reliability growth models [1, 4–14]. From the analysis of various experts, it is observed that NHPP based SRGMs captures the growth in reliability of a software at any time through testing and removing the bugs appearing. Goel and Okumoto's SRGM known as GO model predicts this rate of how often is the software encountering failures [4]. It can be seen that researchers have used this basic assumptions of the GO model for detecting and correcting faults integrating it in their developed models [11]. Ohba [5] in 1984 advanced the GO model keeping the basic assumptions same. Mostly, it was assumed that a software behavior is same as to where it is made as well as at the end user's. This rules out a lot of uncertainties. Which needs to be considered at the operating level [15].

As a growing area of application and development of software, numerous proposed SRGMs help to determine the credibility of the software [16]. Delay occurring between detecting a fault and its correction is well captured by Yamada et al. [6] in their delayed s-shaped model. Also described in their inflection s-shaped SRGM is how with advancement in software testing the ability to detect and correct errors is enhanced. For the most part, it is accepted when a fault is recognized in a software it tends to be corrected. But in reality, some identified faults cannot be corrected and they pass on to next versions as well [17]. Pham et al. proposed an SRGM called PNZ, assuming a time-dependent fault detection rate, introducing it in a debugging phase to capture the process of learning by testers of the software [8]. Talking about a dependent parameter model with imperfect debugging Pham et al. considered both two rate functions, one for detecting the fault and other for the content of fault [10]. Various reliability metrics have been estimated based on logistic growth function in different models [1, 4, 6, 8, 10]. Examining in this paper is a comparison between some of such parametric models along with 4-parameter logistic growth model on a data set, in order to find which model gives a best fit and best suitability.

48.2 A 4-Parameter Logistic Growth Reliability Model

Pham et al. [18] introduced a model with logistic growth having 4-parameters. It is a mathematical model in which the rate at which $m(t)$ is changing and the quantity of growth remaining by a time-dependent logistic function per quantity per unit time, have a directly proportional relation with each other. Basic assumption being $m(t)$ is directly proportional to the growth quantity remaining which is given by a three parameter logistic function $b(t)$ per quantity per unit time. In the form of an equation, it can be denoted as follows:

$$m(t) = \left[1 - e^{-\int_0^t b(x)dx} \right] \tag{48.1}$$

where it is assumed that initially at time zero (1) is also 0.

An "S-shaped" curve giving the rate of detecting faults with respect to time. Different values for $b(t)$ reflect upon various assumptions made in a growth process. Following is the equation for obtaining its value:

$$b(t) = \frac{a}{1 + \beta e^{-bt}}, \text{ where } a > 0,\ b > 0, \beta > 0 \tag{48.2}$$

Using (48.1) and substituting in (48.2), we obtain the following equation after simplification,

$$m(t) = N\left(1 - \frac{1}{\left(\frac{\beta + e^{bt}}{1+\beta}\right)^{\frac{a}{b}}} \right) \tag{48.3}$$

Represented by (48.3) is the mean value function. This 4-parameter logistic growth model is called the Pham inflection model. The model can be used to determine the number of software failures that in time t can be expected to be detected. The model can be used to provide a more realistic model for population growth. The model in fact can be used to describe the spreading of diseases or rumors, autocatalytic chemical reactions.

For analysis, we are going to apply the data set about the bugs appearing in the software application on few existing software growth models and have a comparative analysis of the results. An outline of the few existing SRGM used in this paper is given in Table 48.1.

Table 48.1 List of various models with their mean value functions

Model	*m(t)*
Goel-Okumoto	$a\left(1 - e^{-bt}\right)$
Delayed model S-shaped	$a\left(1 - (1 + bt)e^{-bt}\right)$
Dependent-parameter	$\alpha(1 + \Upsilon t)\left(\Upsilon t + e^{-\Upsilon t} - 1\right)$
Inflection model S-shaped	$\frac{a\left(1 - e^{-bt}\right)}{1 + \beta e^{-bt}}$
Yamada model with imperfect debugging	$a\left[1 - e^{-bt}\right]\left[1 - \frac{\alpha}{b}\right] + \alpha a t$
PNZ	$\frac{a}{1+\beta e^{-bt}}\left[\left(1 - e^{-bt}\right)\left(1 - \frac{\alpha}{b}\right) + \alpha t\right]$
Pham inflection	$N\left(1 - \frac{1}{\left(\frac{\beta + e^{bt}}{1+\beta}\right)^{\frac{a}{b}}}\right)$

48.3 Criteria to Compare Models

In order to compare the predicted values of different models, we try to analyze the goodness-of-fit for the models. For achieving this, we consider three of the most widely used and effective analysis criteria, which are:

(a) **SSE**: For evaluating effectiveness of prediction by a model this criteria of sum of squared error (SSE) is used. In this residuals of the observed data and mean value function obtained is first squared and then summed up.
The mathematical function to express SSE is as follows:

$$\mathrm{SSE} = \sum_{i=1}^{n} [y_i - \hat{m}(t_i)]^2 \tag{48.4}$$

where y_i and $m(t_i)$ are the total number of faults observed and estimated cumulative number of faults at time $t_{i,}$ respectively. Where $i = 1, 2, \ldots n$.

(b) **MSE**: For evaluating performance of the models, this criteria of mean squared error (MSE) is used. In this, we calculate estimation mean of the deviation of the value observed with respect to value predicted by analysis. The mathematical function to express MSE is as follows:

$$\mathrm{MSE} = \sum_{k=1}^{n} (\hat{m}(tk) - yk)2/n - N \tag{48.5}$$

where y_k and $m(t_k)$ are the total number of faults observed and estimated cumulative number of faults at time t_k, respectively. The number of observations is represented by n, and the number of parameters is represented by N. Where $k = 1, 2, 3 \ldots . n$

(c) **R^2**: *R* Square. It is simply a square of R. It gives a degree to which the input variable varies with output variable or predicted value of the variable. It is a record of the index of correlation of the equation of relapse curve. A better result is believed to be of the one whose value is as much close to 1.

48.4 Estimating Parameters and Data Set

Predicting reliability significantly requires estimating parameters for models. However, it is difficult to analyze these in order to have an optimal solution. It requires a lots of calculations and numerical methods. So, here in this paper, we use statistical package for social sciences (SPSS) to obtain an optimal result. We further perform nonlinear regression on our data set in SPSS. Results are obtained using Levenberg–Marquardt method and by an algorithm that iteratively estimates quadratic programming giving us an optimal solution. Models are validated on data set of project avro of apache open source software.

48.5 Model Results

The results are obtained using set of software failure data dependent on the bugs for existing reliability growth models. In our study, with the help of statistical package for social sciences (SPSS), we estimated the parameters for different models using regression modules of SPSS. Further, calculating MSE, SSE and R^2 Table 48.2 shows the value for different parameters obtained along with the value of MSE, SSE and R^2 for goodness-of-fit.

Further the graphs for these models are shown in Figs. 48.1, 48.2, 48.3, 48.4, 48.5, 48.6 and 48.7.

Discussion:

(1) Based on the results obtained, it is observed that 4-parameter logistic model gives the smallest value for MSE and SSE.
(2) R^2 being 0.995 is also one of the closest to value one, indicates a good result for fitness of this model.
(3) Among all the growth models considered in this paper, the best goodness of fit is achieved by the 4-parameter logistic model on the data set for all considered goodness-of-fit criteria.

48.6 Conclusion

This paper discusses about a 4-parameter dependent logistic growth model (Pham inflection model) that can at any time t, help in determining number of failures in

Table 48.2 Estimates of different parameters, MSE, SSE and R^2

Models used	Value obtained	MSE	SSE	R^2
Goel-Okumoto (GO)	$a = 977.279$ $b = 0.011$	114.978	7243.651	0.994
Delayed S-shaped	$a = 504.636$ $b = 0.063$	336.705	21,212.47	0.983
Dependent-parameter	$\alpha = 421$ $\gamma = 0.022$	531.803	33,503.64	0.424
Inflection S-shaped	$a = 650.349$ $b = 0.030$ $\beta = 1.00$	102.033	6428.133	0.995
Yamada Imperfect debugging	$a = 499.380$ $b = 0.022$ $\alpha = 0.008$	123.519	7781.737	0.994
PNZ	$a = 509.120$ $b = 0.037$ $\beta = 1.00$ $\gamma = 0.003$	102.939	6485.172	0.995
Pham inflection	$N = 642.11$ $a = 0.030$ $b = 0.034$ $\beta = 1.000$	101.821	6414.725	0.995

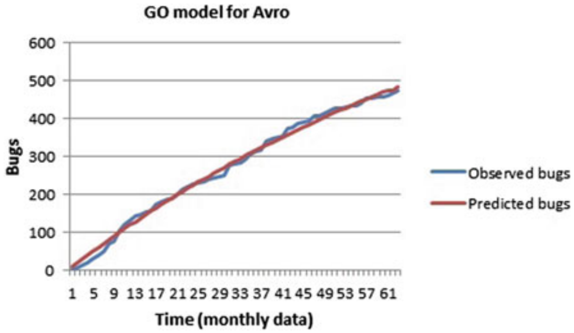


Fig. 48.1 Graph for GO

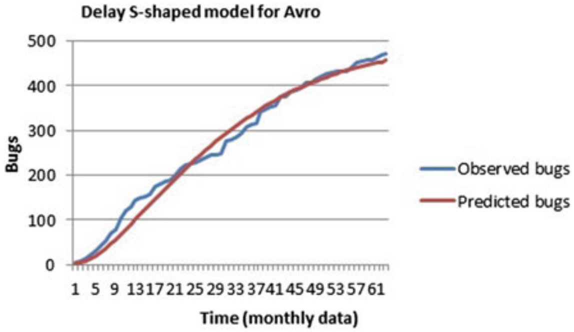


Fig. 48.2 Graph for delayed S-shaped

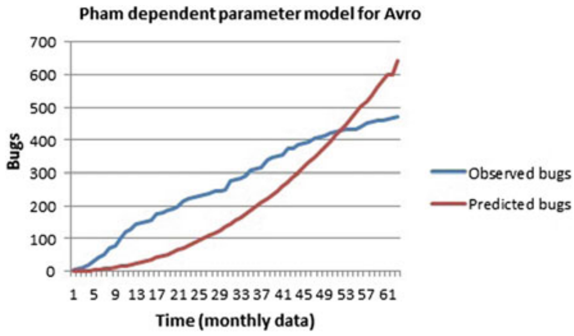


Fig. 48.3 Graph for dependent parameter

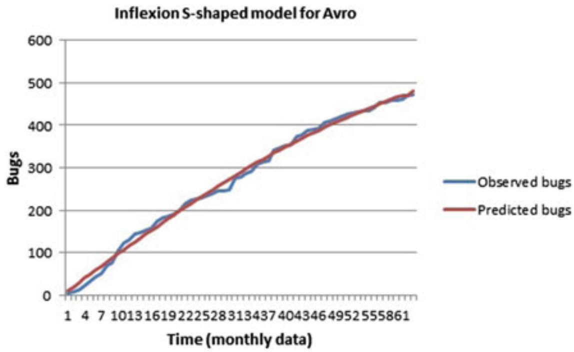


Fig. 48.4 Graph for inflection S-Shaped

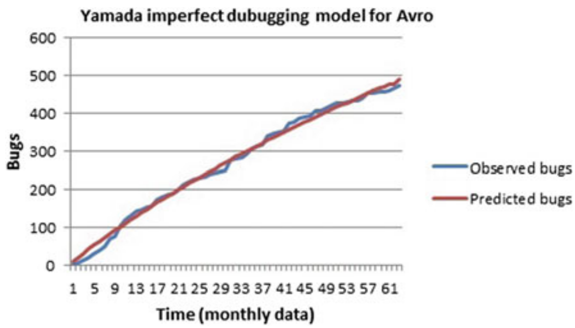


Fig. 48.5 Graph for Yamada imperfect debugging

the software. Illustrated in this paper on our data set is the validation of the Pham inflection model. This paper compares the results of the different existing software reliability growth models with different number of parameters. It is observed that Pham inflection model is fitting significantly better than the other SRGMs, giving the best goodness of fit on this data set. For further validation of this conclusion, more work could be done with other real life data sets. Other criteria to analyze the fitness of models are encouraged in future research.

Fig. 48.6 Graph for PNZ

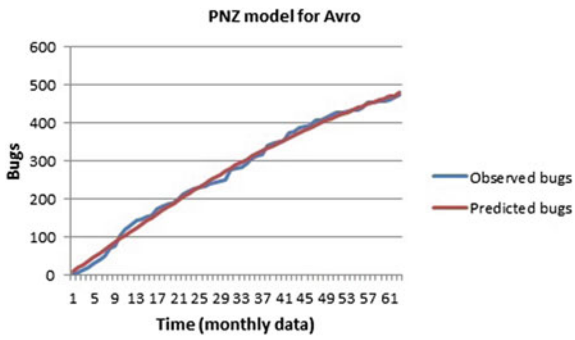
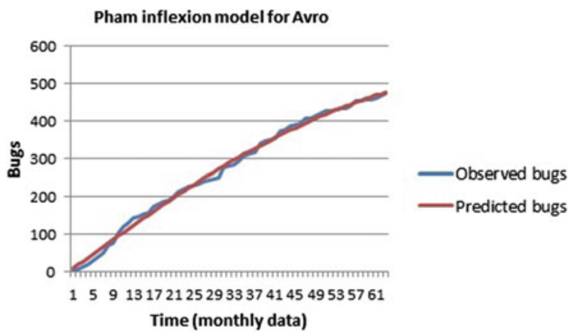


Fig. 48.7 Graph for Pham inflection



References

1. Garg, R.P., Sharma, K., Kumar, R., Garg, R.K..: Performance analysis of software reliability models using matrix method. World Acad. Sci. Eng. Technol. **4**, 31–38 (2010)
2. Sahu, K., Srivastava, R.K.: Revisiting software reliability. In: Balas, V., Sharma, N., Chakrabarti, A. (eds) Data Management, Analytics and Innovation. Advances in Intelligent Systems and Computing, vol. 808. Springer, Singapore (2019)
3. Song, K.Y., Chang, I.H., Pham, H.: Optimal release time and sensitivity analysis using a new NHPP software reliability model with probability of fault removal subject to operating environments. Appl. Sci. **8**(5) (2018)
4. Goel, A.L., Okumoto, K.: Time-dependent fault-detection rate model for software and other performance measures. IEEE Trans. Reliab. (28), 206–211 (1979)
5. Ohba, m.: Inflexion S-shaped software reliability growth models. In: Osaki, S., Hatoyama, Y. (eds.) Stochastic Models in Reliability Theory. Springer, Berlin, pp 144–162 (1984)
6. Ohba, M., Yamada, S.: S-shaped software reliability growth models. In: Proceedings of 4th International Conference on Reliability and Maintainability, pp. 430–436 (1984)
7. Yamada, S., Tokuno, K., Osaki, S.: Imperfect debugging models with fault introduction rate for software reliability assessment. Int. J. Syst. Sci. **23**(12), 2241–2252 (1992)
8. Pham, H., Nordmann, L., Zhang, X.: A general imperfect software debugging model with s-shaped fault detection rate. IEEE Trans. Reliab. **48**, 169–175 (1999)
9. Pham, L., Pham, H.: Software reliability models with time-dependent hazard function based on Bayesian approach. IEEE Trans. Syst. Man Cybern. Part A **30**(1), 25–35 (2000)
10. Pham, H.: An imperfect-debugging fault-detection dependent-parameter software. Int. J. Autom. Comput. **4**, 325–328 (2007)

11. Kapur, P.K., Singh, O., Garmabaki, A.S., Singh, J.: Multi up-gradation software reliability growth model with imperfect debugging. Int. J. Syst. Assurance Eng. Manage. **1**, 299–306 (2010)
12. Persona, A., Pham, H., Sgarbossa, F.: Age replacement policy in random environment using systemability. Int. J. Syst. Sci. **41**, 1383–1397 (2010)
13. Kapur, P.K., Pham, H., Anand, S., Yadav, K.: A unified approach for developing software reliability growth models in the presence of imperfect debugging and error generation. IEEE Trans. Reliab. **60**, 331–340 (2011)
14. Kapur, P.K., Pham, H., Aggarwal, A.G., Kaur, G.: Two dimensional multi-release software reliability modeling and optimal release planning. IEEE Trans. Reliab. **61**, 758–768 (2012)
15. Song, K.Y., Chang, I.H., Pham, H.: A three-parameter fault-detection software reliability model with the uncertainty of operating environments. J. Syst. Sci. Syst. Eng. **26**(1), 121–132 (2017)
16. Yang, J., Liu, Y., Xie, M., Zhao, M.: Modeling and analysis of reliability of multi-release open source software incorporating both fault detection and correction processes. J. Syst. Softw. **115**, 102–110 (2016)
17. Xu, R., Yuan, H., Wang, Q.: Software reliability growth model based on nonlinearity and test coverage. Syst. Eng. Electron. **42**(2), 473–479 (2020)
18. Pham, H., Pham, D.H., Pham, D., Jr.: A new mathematical logistic model and its applications. Int. J. Inf. Manage. Sci. **25**(2), 79–99 (2014)

Chapter 49
VPN Network Traffic Classification Using Entropy Estimation and Time-Related Features

Aswathi Balachandran and P. P. Amritha

Abstract The classification of Internet traffic is gaining prominence for the past few years due to the widening of the current Internet network and web-based applications. Many different approaches are being practiced based on the numerous studies conducted so far. The newer methodology can become imperative for service providers to offer a better quality of service to the users. Deep packet inspection is one such methodology used in the past for traffic characterization where network traffic is subsequently classified into different classes, but their accuracy has been declined due to the drastic changes in the Internet traffic, particularly the increase in encrypted traffic. Virtual private networks (VPNs) have become a preferred choice among the users for remote access communication over other public Internet or IP-based networks. In this paper, our proposed model is a combined approach of entropy estimation and machine learning algorithms, especially deep learning for the network traffic classification. We have used the time-related features to classify VPN or non-VPN network traffic and characterize encrypted traffic into different categories and application identification. We used random forest, KNN, and ANN to test the accuracy of our model. Our results show a good accuracy rate and performance and thus proved our proposed model with entropy as an additional feature, achieving accuracy levels above 90% for all the three algorithms to characterize VPN or non-VPN traffic and application identification.

49.1 Introduction

Network traffic characterization is one of the most important challenging tasks in today's security field. Due to the ever-increasing dynamic nature of wireless networks and mobiles and the fastest growth of the Internet have made the network traffic

A. Balachandran (✉) · P. P. Amritha
TIFAC CORE in Cyber Security, Amrita School of Engineering, Amrita Vishwa Vidyapeetham, Coimbatore, India

P. P. Amritha
e-mail: pp_amritha@cb.amrita.edu

T. Senjyu et al. (eds.), *IOT with Smart Systems*, Smart Innovation, Systems and Technologies 251, https://doi.org/10.1007/978-981-16-3945-6_50

categorization a major challenge. The continuous development and emergence of different types of applications and the development of encrypted communications make this task even more challenging. Due to the implementation of mechanisms such as design and engineering, accounting, network quality of service, identifying security threats or detection of anomalies in the firewall, etc., the traffic categorization technologies have gained increased attention throughout the last decade [1].

Many attempts have been made by the networking sector and the academic community to research these developments and have come up with several classification methods. In addition, in today's Internet, encryption and port obfuscation techniques are becoming widespread, acting as a foundation for safe and secure communication. This continuous development and growth make the classification of network traffic even more challenging to the research community. We can categorize traffic classification based on its fundamental purpose: encrypted traffic, protocol encapsulation like VPN; based on applications or based on the application type can be called as traffic characterization. Applications such as Facebook, Instagram, and Skype support different services like file transfer, voice calls, video calls, and each implementation involves understanding the required feature and the corresponding application. One of the crucial network monitoring and control tasks is the classification of network flows by identifying the corresponding application type. Most of the primary network management functions like the quality of service and billing are focused on the potential to identify Internet traffic into the correct application.

Classification of encrypted network traffic by machine learning has been widely explored. Especially deep learning models have gained attention for network traffic categorization because the results have proved that their accuracy and performance are better than most of the machine learning algorithms. Also, recent research shows that they can attain high accuracy as well as they do not require manual feature selection.

In this work, we emphasize the analysis of both normal and encrypted network traffic tunneled through a virtual private network and performed traffic characterization and application identification. The characterization of VPN traffic is a challenging task. VPN tunnels are mainly used to preserve data confidentiality exchanged over the physical network connection carrying packet-level encryption, rendering it very challenging to identify the applications going through these VPN services. We perform the classification based on entropy estimation as a new feature in addition to time-related features and the comparison of features with the existing approach.

49.2 Related Work

There are several research works that have concentrated on network traffic categorization. As emerging innovations continue to evolve, encrypted network traffic classification is a constant research subject. There are many methodologies for network traffic classification. The traffic categorization by port number is the oldest and has been very commonly used for this task. The port-based classification is the easiest and quickest approaches for network traffic classification. Although the port analysis classification approach works well and is useful, it appears not very practical because of its difficulty [2]. As port use is essentially unregulated, it cannot reliably classify network traffic. So to accurately classify modern network traffic, more advanced classification approaches are required.

Paxson [3] proposed the work associated with the Internet traffic development and about the current users find numerous methods to include the internet into their work patterns. Another proposal from Paxson is about the analytical models representing random variables related to various applications like SMTP, telnet, and FTP connections [4]. Several machine learning techniques have been shown to be accurate and efficient when applied to Internet traffic classification. Swetha [14] proposed an approach for analyzing the packets, where it processes the network packets in a network by real-time streaming computation system called spark streaming, which also helps detect the malicious packets. Sherry [5] proposed a new deep packet inspection platform that has the potential to inspect encrypted payloads without decrypting them, but it can only handle encrypted HTTPS traffic. Williams [6] proposed a different approach by using the "Bayesian network." The application classification by Bayesian network is by considering the flow-per statistics resulting from the independent features of payload and packet length [6]. Nguyen and Armitage developed a framework to identify internet traffic in real time [7]. They showed that network traffic classification is doable, even though we missed a flow's beginning or in between of a flow. Crotti [8] suggested the approach called protocol fingerprinting by considering the probability density function of packet's time-related features and normalized thresholds to develop an algorithm for network traffic classification.

As deep learning methods perform better than the other most used machine learning algorithms, most of the research later on network traffic classification was done using deep learning models. Vinaykumar [13] presented a method for network traffic prediction using the deep learning methodologies, where LSTM provides better output than recurrent neural networks and other classical approaches. Mohammad et al. [1] proposed a model which shows good performance for traffic classification where the network traffic is categorized into various kinds of classes and application identification. They have used stacked autoencoders and CNN for the classification.

Rahul [21] suggested a model using the deep learning concepts to identify network protocols and applications based on flow data and signatures. The model obtained a good accuracy rate of over 0.94. Bonfiglio et al. [9] proposed a method using two corresponding techniques based on the deep packet inspection and Bayes classifier for classifying real-time Skype traffic. Pratibha [10] proposed an efficient keyword matching technique explicitly for network traffic classification in one scan of payload. They chose the dataset which consists of the flows from ten various applications and achieved a reasonable classification accuracy rate. They have selected only time-related features to increase the performance and maintaining an encryption-independent traffic classifier.

49.3 Methodology

In this work, we proposed a model that comprises entropy estimation and machine learning methods like artificial neural network (ANN), random forest, and K-nearest neighbors' algorithm (KNN) for both traffic characterization and application identification tasks. Figure 49.1 demonstrates the architecture of our model. The network traffic data were first preprocessed before being input into the model. To ready the dataset for use, we first perform preprocessing steps on it. We trained and tested on 80% train and 20% test and accuracy are reported as the experiments average.

The details about implementation, dataset, architecture of proposed model, and design details of the preprocessing phase will be discussed in the next section.

49.3.1 Artificial Neural Networks

There are reports that neural network architectures like convolutional neural networks (CNN), autoencoders, multilayer perceptrons (MLP), generative adversarial networks (GAN), and recurrent neural networks (RNN) are all very useful for the traffic classification.

The study reveals the performance of artificial neural networks most recently accomplished using deep learning technologies [1]. In our model, we employed a traditional neural network consisting of input, computation, and output layers, to

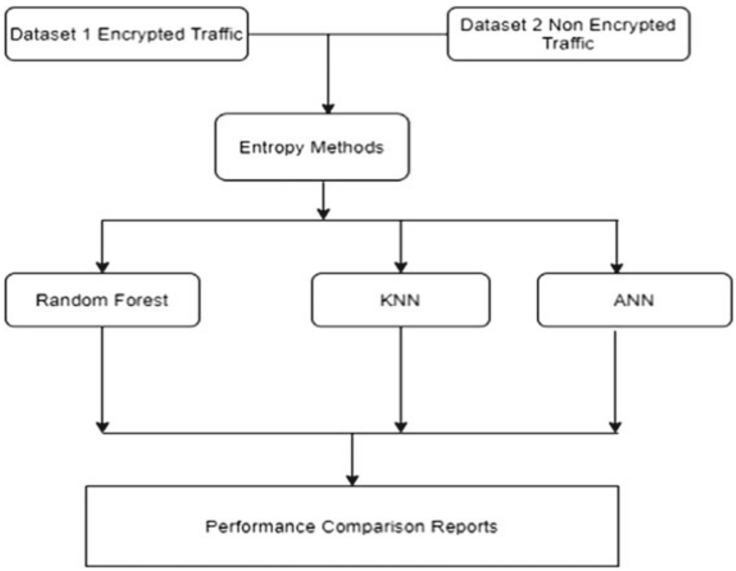


Fig. 49.1 Architecture of proposed model

assess our approach's efficacy. The idea behind each of the modules was to (a) calculate gradients, forward_propagation, backward_propagation, (b) combined forward and backpropagation, (c) calculate paragradients, and (d) apply after computation and iterate. The neural network uses cross-entropy (loss function), softmax (activation function), and modular backpropagation algorithms.

The neural networks comprise of a collection of three layers that consist of the input layer, which has 24 neurons after including entropy as an additional feature, then the hidden layer, which has around 100 neurons and the output layer, which has different classes such as audio streaming, VoIP, P2P, email, browsing, file transfer, video streaming, and chat.

49.3.2 Random Forest

Random forest is a commonly used decision tree-based ensemble approach with tremendous success compared to other supervised machine learning models [15]. Random forest forms many decision trees from a random subset of data and features. Random forests are considered to be stable and are quickly trained. However, ensemble approaches are a class of methods that incorporate base models to produce a meta-model with better results over the model regarding the precision, generalizability, accuracy, and robustness [15]. "The ensemble approaches such as random forest have demonstrated the capabilities to capture incredibly complex nonlinear patterns" and offer outstanding efficiency and robustness on organized datasets.

In our model, random forest showed good performance and accuracy rate of above 0.92 and also prevented the problem of overfitting.

49.3.3 KNN

K-nearest neighbor is one of the best-known machine learning algorithm for classification. KNN relies on certain similarity measures; hence, it can be calculated by the similarity measure that is used to achieve the distance between the samples. As a brief overview, KNN classification is known as the lazy approach that omits the training phase entirely and conducts classification using the highest number of votes [16]. The algorithm classifies as after obtaining the feature vector of a new flow, then the classification model selects the K-nearest neighboring flows in Euclidean distance in the training dataset.

Classification's output is determined according to the majority vote of its K-nearest neighbors. It depends on the fact that nearby datasets as a whole have the same label and high probability. In our proposed work, we studied and evaluated results on $k = 1$ and $k = 2$ neighbors and both showed a good accuracy rate of above 0.94.

49.3.4 Entropy

In the proposed model, we compute entropy as a new approach toward encrypted network traffic classification. After selecting the necessary features for our model after preprocessing will compute entropy as an additional feature. Entropy refers to the randomness collected by a system for algorithms that require random data. Therefore, having and using high entropy sources of data is critical to security. Equal probabilities mean it will have the maximum entropy value. When there is more unpredictable behavior or randomness such as encrypted traffic that indicates uncertainties and thus rises the entropy value.

Shannon's entropy theory measures information uncertainty. Given m possible events A, .., A_m with probabilities of occurrence p,.., p_m entropy H is defined by (49.1).

$$\text{Entropy(H)} = -\sum_{i=1}^{m} pi \log(pi) \tag{49.1}$$

A proper computation of entropy is required for small sample data; however, entropy estimation is harder for small datasets such as for $N < m$ [17]. An ideal approach by Olivain and Goubault-Larrecq is to calculating entropy by using the "N-truncated entropy regardless of the difficulties involved with estimating entropy for small lengths N" [18]. In addition to the contents from the experimental pcap files that are "ISCX-VPN-NonVPN2016" dataset, we calculate entropy as an additional feature for identifying the encrypted network traffic.

49.3.5 Dataset

We have used the well-known dataset from Canada Institute for Cybersecurity, which is considered as the commonly used dataset by the research community. The dataset features network traffic information for IDS, Tor–non-Tor, VPN–non-VPN, and android malware. In our work, we have used the "ISCX-VPN–non-VPN dataset." Dataset has both VPN and non-VPN of each traffic type (VoIP, P2P, etc.), and altogether, there is a total of 14 different traffic categories. For more information on the dataset and the traffic generation process of the captured network, refer to [11].

The network traffic data was captured using Tcpdump and Wireshark, and the size amounts to 28 GB of data. For the generated VPN dataset, they have made use of an external VPN provider and linked to it using Open VPN. To generate the file transfer protocols, they have used a third-party service provider and FileZilla as the client. We have analyzed each feature's importance in the dataset and have selected the necessary features for our network classification along with additional feature entropy. We evaluated the results using random sampling of training and testing sets.

We took 80% of the samples for training, and the other 20% for the testing set was used to assess the model's performance.

49.3.6 Data Preprocessing

Data cleaning and data reorganization were performed to improve the model's performance. There was the presence of duplicate and empty records in the dataset. Almost a total of more than 1300 duplicate records and 484 empty records were removed from the dataset. As the first stage in the preprocessing, the Ethernet header is discarded because the dataset is obtained at the data-link layer. The data-link header has the details like "Media Access Control address," which is not useful for either the application identification or traffic characterization tasks [1]. The packets that were not important to our purposes were discarded in our research output. There are a few domain name service segments in the dataset, and these details aren't needed for either traffic characterization or application identification so that they can be removed from the dataset. The segments, such as transmission control protocol or user datagram protocol, differ in the length of their header. There are chances that the neural network may attempt to learn the classification from the IP address. Hence, overfitting can be prevented by hiding the IP addresses in the IP header. In order to make these segments unvarying, the end of a UDP segment's headers is padded with zeros to make it equivalent to the TCP segment headers. Bits are then converted into bytes to decrease the size of the final input value. These packets are then converted into bytes, which decreases the input size of the model. The dataset includes few TCP segments which contain no payload. These are required for three-way handshaking but do not provide any data about the application generated them. Therefore, it is excluded from the dataset.

49.3.7 Data Reorganization

The pcap files of the dataset are named based on the tasks and applications that they were involved in. To perform traffic characterization and application identification tasks, we have to rename the labels for the task. For the traffic classification analysis, we combined both encrypted and non-VPN traffic captured by various applications engaged in the same task into a single file. After collecting all the datasets, it was noticeable that the number of samples differs across various groups, and all the datasets are substantially imbalanced. We know that if the training data is not balanced, it reduces classification efficiency. So to train the model, we used the "under-sampling" method. We removed the classes which are having more data until our selected classes are comparatively balanced; as we know, sampling is an easy but effective technique to overcome this challenge [19].

For every dataset, the VPN or non-VPN label was used as the new class label, and also, for our combined approach, we added entropy as an additional new class label.

49.4 Experimental Results

The findings from our proposed classification model are described below. Our analysis focuses on defining and classifying network traffic in VPNs and non-VPNs, and application identification. This problem is challenging because we cannot classify based on just analyzing the IP packet header. The main difference between the VPN and regular traffic is VPN traffic has better anonymity than non-VPN traffic because it holds both parties' identities confidential. Throughout VPN traffic, the intruder cannot identify the valid IP address. When we consider VPN traffic, we refer to the three protocols like point-to-point tunneling protocol, secure socket layer, and IPSec.

Here we provide the comparison results for our combined approach with entropy. The models compared are artificial neural network (ANN), random forest (RF), K-nearest neighbor (KNN). The experiment mainly contains four parts, and those are (a) reorganizing the selected data, (b) splitting our selected data randomly into training and testing sets, (c) building the proposed model, (d) training the model, and (e) testing the model. These processes such as splitting, training, and testing are performed many times to maintain accuracy and good efficiency in the results. The results are calculated in terms of accuracy, precision, recall, and F1 score.

49.4.1 Performance Metrics

An explicit and systematic way to present a machine learning model's prediction results is by the "confusion matrix." A confusion matrix will visualize the performance of classification models and show the number of correct as well as incorrect predictions made by the classification model, where rows and columns describe the class labels in the dataset and predicted class labels by the trained model.

However, the accuracy can be misleading when the number of application labels is imbalanced. We will convert the multi-class confusion matrix into the binary confusion matrix for every application class.

In the confusion matrix stated in Table 49.1, the true positive (TP) indicates the count of true positive classes that accurately identified that the actual label is positive.

Table 49.1 Confusion matrix

		Predictive	
		Positive	Negative
Actual	Positive	True positive: TP	False negative: FN
	Negative	False positive: FP	True negative: TN

False positive (FP) means the count of FP that is wrongly marked to a class where the actual label is negative. TN represents when the actual class and model prediction is accurately identified as negative. Finally, FN is the count that is inaccurately assigned to another class that is where the actual class is positive, but when it is predicted incorrectly as negative [12].

We have used accuracy, precision, recall, and F1 score in our work. Accuracy is the better metric to assess when both positive and negative are significant.

Sometimes when there is a significant imbalance, a model can predict the value of the majority class for all predictions and get the highest classification accuracy rate. This is called the accuracy paradox [19]. To overcome the accuracy paradox problem, we compute the additional measurements, like recall, precision, and F1-score, to evaluate our models [20].

Precision emphasis on when a good result of positive predictions is required. Precision is also called a "positive predictive value." Recall is effective when we expect our model to select many positive samples from actual positives. It is also called the "sensitivity." The F1 score shows the balance between recall and precision. It will reliably calculate the classification efficiency of a machine learning algorithm. All these metrics presume values varying from zero to one.

We evaluate the accuracy, recalls, precisions, and F1-score of all our data for performance comparison of our three models with both combined approaches, including entropy and the standard classification method.

These metrics are defined mathematically as follows: (49.2), (49.3), and (49.4)

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{49.2}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{49.3}$$

$$F1\text{ - score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{49.4}$$

We have trained and tested on 80% train, and 20% test and accuracy are stated as the experiments average for regular approach and combine approach by computing entropy as an additional feature in the dataset.

Table 49.2 presents the achieved performance of ANN, KNN, and random forest models. These were directly classified without computing the entropy. For traffic categorization, random forest shows better performance than others, and for application identification, random forest gives an accuracy rate of 0.88. Table 49.3 presents the performance of after computing entropy as an additional feature in the dataset for our proposed three models. We could find a good difference in their performance compared to the previous approach for all three models. Hence for the selected VPN and non-VPN dataset, computing entropy as an additional feature shows better accuracy and efficiency in both traffic classification and application identification tasks. The weighted average F1 score of above 0.90, respectively, implying that our models

Table 49.2 Metrics of our approach before adding entropy as a feature

Task	Algorithm	Accuracy
Traffic Characterization	ANN	0.81
	KNN	0.83
	RF	0.93
Application Identification	ANN	0.76
	KNN	0.81
	RF	0.88

Table 49.3 Metrics of our approach after adding entropy as a feature

Task	Algorithm	Accuracy	Precision	Recall	F1
Traffic characterization	ANN	0.98	0.961	0.958	0.952
	KNN	0.96	0.918	0.916	0.910
	RF	0.98	0.971	0.971	0.970
Application identification	ANN	0.97	0.956	0.952	0.954
	KNN	0.94	0.902	0.901	0.901
	RF	0.93	0.958	0.956	0.956

are capable of accurately classifying the VPN and non-VPN traffic categorization and application identification. Finally, in all experiments, using the base and new features delivered the best results for both tasks.

49.5 Conclusion

In this paper, we presented a network traffic classification model. Our approach is a fusion of information theory and machine learning, especially deep learning. We have combined the use of machine learning algorithm and computed entropy as an additional feature for each data for the proposed model. As a result, it showed a significant increase in the accuracy when the entropy parameter is considered. Entropy comes into the table when the network traffic's payload information is considered and gets added as an extra feature to the dataset. The proposed approach is found to be accurate and efficient. We have compared the results using various algorithms like KNN, ANN, and random forest for network traffic characterization, application identification, and segregation of VPN and non-VPN traffic. Our work emphasizes the importance of entropy for classifying the VPN traffic and the advantage of time-related features in categorizing encrypted traffic and application identification. Our results prove the proposed model with entropy as an additional feature, achieving accuracy levels above 90% for all three algorithms. Overall, random forest showed the best results in both normal and combined approaches of more than 90%. KNN and ANN gained more momentum and showed a significant increase in the accuracy from 0.81 to 0.94 and 0.81 to 0.98 when the entropy parameter is considered.

As future work, a detailed analysis of encryption protocols such as TLS 1.3 can also be performed along with this. The proposed model can be improved to be more efficient to perform more complicated tasks like multi-channel, and accurately classifying Tor's traffic.

References

1. Lotfollahi, M., Jafari Siavoshani, M., Shirali Hossein Zade, R., et al.: Deep packet: a novel approach for encrypted traffic classification using deep learning. Soft Comput. **24,** 1999–2012 (2020)
2. Schneider, P.: TCP/IP traffic classification based on port numbers, Division of Applied Sciences, Cambridge, MA 2138.5 (1996)
3. Paxson, V.: Growth trends in wide-area TCP connections. IEEE Network **8**(4), 8–17 (1994)
4. Paxson, V.: Empirically derived analytic models of wide-area TCP connections. IEEE/ACM Trans. Networking **2**(4), 316–336 (1994)
5. Sherry, J., Lan, C., Popa, R.A., Ratnasamy, S.: BlindBox: deep packet inspection over encrypted traffic. In: Proceedings of the 2015 ACM conference on special interest group on data communication, pp. 213–226 (2015)
6. Williams, N., Zander, S., Armitage, G.: A preliminary performance comparison of five machine learning algorithms for practical IP traffic flow classification. ACM SIGCOMM Comput. Commun. Rev. **36**(5), 5–16 (2006)
7. Nguyen, T.T., Armitage, G.: Training on multiple sub-flows to optimise the use of Machine Learning classifiers in real-world IP networks. In: Proceedings. 31st IEEE conference on local computer networks, pp. 369–376 (2006)
8. Crotti, M., Dusi, M., Gringoli, F., Salgarelli, L.: Traffic classification through simple statistical fingerprinting. ACM SIGCOMM Comput. Commun. Rev. **37**(1), 5 (2007)
9. Bonfiglio, D. Mellia, M., Meo, M., Rossi, D., Tofanelli, P.: Revealing skype traffic: when randomness plays with you. SIGCOMM, 37–48 (2007)
10. Khandait, P., Hubballi, N., Mazumdar, B.: Efficient keyword matching for deep packet inspection based network traffic classification. In: International conference on communication Systems & networks (COMSNETS), pp. 567–570 (2020)
11. Gil, G.D., Lashkari, A.H., Mamun, M., Ghorbani, A.A.: Characterization of encrypted and VPN traffic using time-related features. In: Proceedings of the 2nd international conference on information systems security and privacy, pp. 407–414 (2016)
12. Archanaa, R., Athulya, V., Rajasundari, T., Kiran, M.V.K.: A comparative performance analysis on network traffic classification using supervised learning algorithms. In: 4th international conference on advanced computing and communication systems, pp. 1–5 (2017)
13. Vinayakumar, R., Soman, K.P., Poornachandran, P.: Applying deep learning approaches for network traffic prediction. In: International conference on advances in computing, communications and informatics, pp. 2353–2358 (2017)
14. Swetha, K.V., Sathyadevan, S., Bilna, P.: Network data analysis using spark. In: Software engineering in intelligent systems, vol. 349 (2015)
15. Polikar, R.: Ensemble based systems in decision making. IEEE Circuits Syst. Mag. **6**(3), 21–45 (2006)
16. Altman, N.S.: An introduction to kernel and nearest-neighbor nonparametric regression. Am. Stat. **46**(3), 175–185 (1992)
17. Anderson, B., Paul, S., McGrew, D.: Deciphering malware's use of TLS, (without decryption). J. Comput. Virol. Hacking Tech. **14**(3), 195–211 (2018)

18. Bonachela, J.A., et al.: Entropy estimates of small data sets. J. Phys. A Math. Theoretical **41**(20) (2001)
19. Longadge, R., Dongre, S.: Class imbalance problem in data mining review, pp. 1305–1707 (2013)
20. Powers, D.M.W.: Evaluation: from precision, recall and f-measure to ROC, informedness, markedness & correlation. J. Mach. Learn. Technol. **2**, 37–63 (2011)
21. Rahul, R.K., Anjali, T., Menon, V.K., Soman, K.P.: Deep learning for network flow analysis and malware classification. In: Communications in computer and information science, vol. 746. Springer (2017)

Chapter 50
Aiding Team Leader Selection in Software Industry Using Fuzzy-TOPSIS Approach

Ajay Kumar and Kamaldeep Kaur

Abstract The objective of this paper is to address the problem of selecting the most suitable team leader in a software industry using multi-criteria decision-making (MCDM) method. This problem can be modelled as MCDM problem as selection of team leader may depend on various characteristics like technical competence, collaborative efficiency and innovativeness. The proposed approach is examined by an experimental study which is designed using one MCDM method fuzzy-TOPSIS (technique for order preference by similarity to ideal solution), four team leaders, three selection criteria and three decision-makers. This study will provide a ranking index for the evaluation of team leaders which will aid decision-makers in the selection of most suitable team leader in software industry.

50.1 Introduction

The selection of a team leader for leading successful software projects is a vital human resource issue in software industry. A team leader acts as an observer, navigator, initiator, external facilitator and a cultivator for a software development team in agile software development environments [1]. These job functions require high degree of technical competence, collaborative efficiency and innovativeness as desirable characteristics of a team leader. To fulfil the role of an observer, a team leader is required to continuously assess the technical challenges and impediments during software development process and propose solutions to technical problems faced by team members. This justifies technical competence as a desirable characteristic of a team leader. As a navigator and external facilitator, a team leader is required to efficiently carry out intra-team and inter-team coordination activities which justifies collaborative efficiency as a desirable characteristic. As an initiator, a high degree of innovativeness is required to introduce new tools and novel practices into the team

A. Kumar (✉) · K. Kaur
USICT, Guru Gobind Singh Indraprastha University, New Delhi, India

K. Kaur
e-mail: kdkaur99@ipu.ac.in

T. Senjyu et al. (eds.), *IOT with Smart Systems*, Smart Innovation,
Systems and Technologies 251, https://doi.org/10.1007/978-981-16-3945-6_51

for efficient, low cost and high-quality software development. As a cultivator, a team leader is required to handle conflicts and stimulate intellectual exploration within team [1] which map to the desirable characteristics of collaborative efficiency and innovativeness. These characteristics form the basis of selection criteria for decision-makers to recruit and appoint team leaders in their organizations. Decision-makers may assign different qualitative scores or ratings to potentially suitable candidates with respect to various selection criteria. This makes the selection of a team leader a non-trivial task that requires a multi-criteria decision support system. Conventional multi-criteria decision-making approaches are not suitable to deal with qualitative ratings which are fuzzy in nature. There is no precise boundary for the rating of a particular characteristic. This paper proposes a fuzzy-TOPSIS framework to aid selection of a team leader in software industry. This framework is based on technical competence, collaborative efficiency and innovativeness as selection criteria for a team leader.

The remaining section of this study is explained from Sects. 50.2 to 50.6. Related research is presented in Sect. 50.2, and Sect. 50.3 discusses the proposed method, selection criteria and MCDM method used in this study. Section 50.4 describes experimental study used in this paper. Section 50.5 describes results discussion and findings, and conclusion is presented in Sect. 50.6.

50.2 Related Work

There is not much previous intensive research work in the area of intelligent decision support systems for aiding the selection of most suitable team leader particularly for a software industry. Table 50.1 summarized some related research dissimilar yet related research work.

So far, only a few researchers have attempted to address the problem of selecting most suitable team leader. Although some researchers have attempted to address this issue by using MCDM approach, but none of them have emphasized to solve this problem using MCDM approach in fuzzy environment, where no precise boundary is fixed for the rating of a particular characteristic of a team leader. This research work proposes a fuzzy-MCDM-based approach to select best team leader in a software industry considering various different characteristics of a team leader as selection criteria.

Table 50.1 Overview of some related research

Author	Proposed method	Study purpose
Wi et al. [2]	A framework based on genetic algorithm	Selection of team manager and a team member in a research and development institute
Lele [3]	Hybrid model by integrating AHP and linear programming	Formation of an effective team for development of a software product
Muhisn et al. [4]	AHP-based model	Team leader selection for software engineering community
Sokolovska et al. [5]	Questionnaire survey	Emphasized on the importance of competency model for selecting best team leader
Sandhya et al. [6]	Euclidean distance-based approximation (EDBA)	Selection of a team leader in a software industry
Jaafar et al. [7]	Rough set theory and fuzzy set theory	To select best team leader in a software industry
Goyal and Gupta [8]	Intuitionistic fuzzy approach	Selecting efficient team for the development of a software product
Assavakamhaenghan et al. [9]	Model based on existing machine learning approach	Recommend a suitable software team member for a given task

50.3 Research Methods

50.3.1 Proposed Method

This paper proposes an MCDM-based method to generate ranking index of team leaders to select most suitable team leader (TL) in a software industry. For experimental study, this paper chooses four team leaders as the alternatives and three characteristics (technical competence, collaborative efficiency and innovativeness) as criteria as the input for MCDM method fuzzy-TOPSIS. The process of generating ranking index for team leaders is shown in Fig. 50.1.

50.3.2 Selection Criteria

This study chooses three characteristics of team leader as the selection criteria, namely technical competence (a person who has good command on technical knowledge for any type of software project), collaborative efficiency (ability to take a team

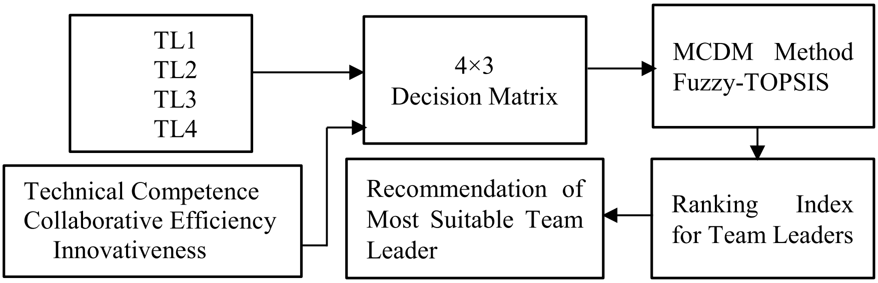


Fig. 50.1 Process of generating ranking index for team leader selection using MCDM

approach for solving a problem) and innovativeness (ability to do the things with new approaches other than the conventional methods).

50.3.3 MCDM Method Fuzzy-TOPSIS

This paper uses the following fuzzy-TOPSIS algorithm in the experimental study [10] for generating ranking index for team leaders for the selection of most suitable one in a software industry.

Step 1: Compute fuzzified decision matrix:

If there are K decision-makers, m alternatives and n criteria are involved. Then by using triangular membership function [11] (having triplet values p, q and r), fuzzy rating of alternative i for criteria j and fuzzified weightage of criteria j for the decision-maker k are given as follows:

$$\left.\begin{aligned}\tilde{x}_{ij}^{k} &= (p_{ij}^{k}, q_{ij}^{k}, r_{ij}^{k}) \text{ and}\\ \tilde{w}_{j}^{k} &= (p_{j}^{\prime k}, q_{j}^{\prime k}, r_{j}^{\prime k})\end{aligned}\right\} \tag{50.1}$$

where $\tilde{x}_{ij}^{k}$ and $\tilde{w}_{j}^{k}$ are fuzzy ratings of ith alternative for criteria j and fuzzified weightage of criteria j for the decision-maker k, respectively.

Further, total fuzzy rating of alternative i for criteria j and total fuzzified weightage of criteria j can be computed by calculating arithmetic mean of individual ratings and given as follows:

$$\left.\begin{aligned}\tilde{x}_{ij} &= (p_{ij}, q_{ij}, r_{ij}) \text{ and}\\ \tilde{w}_{j} &= (p_{j}^{\prime}, q_{j}^{\prime}, r_{j}^{\prime})\end{aligned}\right\} \tag{50.2}$$

where $\tilde{x}_{ij}$ and $\tilde{w}_{j}$ are the total fuzzy rating of alternative i for criteria j and total fuzzy rating of importance of criteria j, respectively.

Now total fuzzy ratings of alternatives can be represented as $\tilde{D}_{m\times n}$, fuzzified decision matrix, where each entry of this matrix represents value of $\tilde{x}_{ij}$. Similarly, total fuzzy ratings of criteria can be represented as $\tilde{W}_{1\times n}$, fuzzified weight matrix, where each entry of this matrix represents value of $\tilde{w}_j$.

Step 2: Normalization of fuzzified decision matrix using following equations.

$$\tilde{T} = \left[\tilde{t}_{ij}\right]_{m\times n} \tag{50.3}$$

$$\left.\begin{aligned} \tilde{t}_{ij} &= \left(\frac{p_{ij}}{r_j^+}, \frac{q_{ij}}{r_j^+}, \frac{r_{ij}}{r_j^+}\right) \\ r_j^+ &= \max_i r_{ij} \text{ (Benificial Criteria)} \end{aligned}\right\} \tag{50.4}$$

$$\left.\begin{aligned} \tilde{t}_{ij} &= \left(\frac{p_j^-}{r_{ij}}, \frac{p_j^-}{q_{ij}}, \frac{p_j^-}{p_{ij}}\right) \\ p_j^- &= \min_i p_{ij} \text{ (Non-Benificial Criteria)} \end{aligned}\right\} \tag{50.5}$$

Step 3: Use following equation to compute weighted normalized fuzzified decision matrix.

$$\left.\begin{aligned} \tilde{G} &= \left[\tilde{g}_{ij}\right]_{m\times n} \text{ where} \\ \tilde{g}_{ij} &= \tilde{t}_{ij}(\cdot)\tilde{w}_j = \left(p''_{ij}, q''_{ij}, r''_{ij}\right) \end{aligned}\right\} \tag{50.6}$$

Step 4: Find fuzzy positive ideal solution (FPIS) and fuzzy negative ideal solution (FNIS) denoted by F^+ and F^-, respectively. For each alternative i, F^+ and F^- are the best value and worst value, respectively, attained by a criterion j. F^+ and F^- are computed using following equations.

$$\left.\begin{aligned} F^+ &= \left(\tilde{g}_1^+, \tilde{g}_2^+, \ldots, \tilde{g}_n^+\right) \text{ where} \\ \tilde{g}_j^+ &= (r, r, r) \text{ such that:} \\ r &= \max_i \left\{r''_{ij}\right\} \end{aligned}\right\} \tag{50.7}$$

$$\left.\begin{aligned} F^- &= \left(\tilde{g}_1^-, \tilde{g}_2^-, \ldots, \tilde{g}_n^-\right) \text{ where} \\ \tilde{g}_j^- &= (p, p, p) \text{ such that:} \\ p &= \min_i \left\{p''_{ij}\right\} \end{aligned}\right\} \tag{50.8}$$

Step 5: For each alternative i, distance s_i^+ and s_i^- from FPIS and FNIS, respectively, can be computed as follows:

$$s_i^+ = \sum_{j=1}^{n} s_g\left(\tilde{g}_{ij}, \tilde{g}_j^+\right) \tag{50.9}$$

$$s_i^- = \sum_{j=1}^{n} s_g\left(\tilde{g}_{ij}, \tilde{g}_j^-\right) \tag{50.10}$$

where $s_g(\tilde{x}, \tilde{y})$ is the distance between two fuzzy numbers $\tilde{x}$ and $\tilde{y}$.

Step 6: Computation of closeness coefficient.

For each alternative, closeness coefficient (CC_i) represents the simultaneous distances F^+ and F^- from FPIS and FNIS, respectively. Use following equation to compute CC_i.

$$CC_i = \frac{s_i^-}{\left(s_i^- + s_i^+\right)} \tag{50.11}$$

Step 7: Assign rank to each alternative.

Assign rank to each alternative using closeness coefficient. Alternative having higher closeness coefficient value will be assigned higher rank.

50.4 Experimental Study

Following procedure is used for experimental design based on the fuzzy-TOPSIS method.

Step 1: Linguistic terms for the rating of team leaders and selection criteria with their fuzzy values using conversion scale of 1–9 are listed in Table 50.2. In this study, triangular membership function is used for calculating fuzzified value of ratings of team leaders and selection criteria.

Table 50.2 Fuzzified value for linguistic term

Fuzzified value	Linguistic term for team leader rating	Linguistic term for selecting criteria weightage
(1, 1, 3)	Poor (P)	Lowest weightage
(1, 3, 5)	Satisfactory (S)	Lower weightage
(3, 5, 7)	Good (G)	Average weightage
(5, 7, 9)	Very good (VG)	Higher weightage
(7, 9, 9)	Excellent (E)	Highest weightage

Step 2: Collect the rating of each team leader with respect to each selection criteria on the five-point scale described in step 1. Fuzzy ratings of team leaders are shown in Table 50.3.

Step 3: Prepare the combined fuzzified decision matrix (shown in Table 50.4) by calculating arithmetic mean of ratings given by all decision-makers.

Step 4: Get the normalized fuzzified decision matrix as shown in Table 50.5.

Step 5: Normalized fuzzified decision matrix is multiplied with fuzzified weights of selection criteria to get weighted normalized fuzzified matrix which is shown in Table 50.6. In this study, fuzzified weights for characteristics technical competence, collaborative efficiency and innovativeness are (5, 7, 9), (7, 9, 9) and (3, 5, 7), respectively.

Step 6: For each team leader, calculate the distance from FPIS and FNIS. This calculated distance is used to compute the closeness coefficient, which can be further used for assigning rank to each team leader considering higher the closeness coefficient higher the rank. Final results of this step are shown in Table 50.7.

50.5 Discussions and Findings

From Table 50.3, it can be observed that for selecting best team leader in a software industry, different decision-maker may have different subjective rating for a team leader with respect to a selection criterion. In this scenario, selection of best team leader is not an easy task. In this study, we have presented a MCDM-based approach using fuzzy-TOPSIS for selection of best team leader in software industry. From Table 50.7, it can be seen that team leader 4 is declared as the best team leader having maximum closeness coefficient (0.56).

50.6 Conclusion

This study provides an MCDM-based method to select the most suitable team leader for a software project in software industry. In this study, we have used fuzzy-TOPSIS approach to deal with qualitative ratings of team leaders given by different decision-makers. From the experimental study, it can be observed that four team leaders have different qualitative ratings given by three decision-makers for a particular characteristic (selection criteria). In this scenario, where different decision-makers have different opinion about a team leader with respect to a particular characteristic and there is no general agreement amongst decision-makers to select best team leader, a decision support system is needed to aid software managers for unbiased selection of best team leader. Further, our experimental study also shows that after applying fuzzy-TOPSIS approach, selection of best team leader can be done on the basis of

Table 50.3 Ratings obtained by four candidates based on opinion given by three decision-makers

Team leader (TL)/selection criteria		Technical competence		Collaborative efficiency		Innovativeness	
		Qualitative term	Fuzzified value	Qualitative term	Fuzzified value	Qualitative term	Fuzzified value
TL1	Decision maker 1	Excellent	(7, 9, 9)	Very good	(5, 7, 9)	Excellent	(7, 9, 9)
TL2		Good	(3, 5, 7)	Very good	(5, 7, 9)	Good	(3, 5, 7)
TL3		Good	(3, 5, 7)	Excellent	(7, 9, 9)	Poor	(1, 1, 3)
TL4		Very good	(5, 7, 9)	Good	(3, 5, 7)	Excellent	(7, 9, 9)
TL1	Decision maker 2	Very good	(5, 7, 9)	Satisfactory	(1, 3, 5)	Good	(3, 5, 7)
TL2		Good	(3, 5, 7)	Very good	(5, 7, 9)	Excellent	(7, 9, 9)
TL3		Poor	(1, 1, 3)	Excellent	(7, 9, 9)	Good	(3, 5, 7)
TL4		Excellent	(7, 9, 9)	Excellent	(7, 9, 9)	Excellent	(7, 9, 9)
TL1	Decision maker 3	Excellent	(7, 9, 9)	Good	(3, 5, 7)	Excellent	(7, 9, 9)
TL2		Very good	(5, 7, 9)	Very good	(5, 7, 9)	Good	(3, 5, 7)
TL3		Satisfactory	(1, 3, 5)	Excellent	(7, 9, 9)	Good	(3, 5, 7)
TL4		Good	(3, 5, 7)	Good	(3, 5, 7)	Excellent	(7, 9, 9)

Table 50.4 Combined fuzzified decision matrix

Team leader (TL)/selection criteria	Technical competence	Collaborative efficiency	Innovativeness
TL1	(6.33, 8.33, 9.00)	(3.00, 5.00, 7.00)	(5.67, 7.67, 8.33)
TL2	(3.67, 5.67, 7.67)	(5.00, 7.00, 9.00)	(4.33, 6.33, 7.67)
TL3	(1.67, 3.00, 5.00)	(7.00, 9.00, 9.00)	(2.33, 3.67, 5.67)
TL4	(5.00, 7.00, 8.33)	(4.33, 6.33, 7.67)	(7.00, 9.00, 9.00)

Table 50.5 Normalized fuzzified decision matrix

Team leader (TL)/selection criteria	Technical competence	Collaborative efficiency	Innovativeness
TL1	(0.70, 0.93, 1.00)	(0.33, 0.56, 0.78)	(0.63, 0.85, 0.93)
TL2	(0.41, 0.63, 0.85)	(0.56, 0.78, 1.00)	(0.48, 0.70, 0.85)
TL3	(0.19, 0.33, 0.56)	(0.78, 1.00, 1.00)	(0.26, 0.41, 0.63)
TL4	(0.56, 0.78, 0.93)	(0.48,0.70,0.85)	(0.78, 1.00, 1.00)

Table 50.6 Weighted normalized fuzzified decision matrix

Team leader (TL)/selection criteria	Technical competence	Collaborative efficiency	Innovativeness
TL1	(3.52, 6.48, 9.00)	(2.33, 5.00, 7.00)	(1.89, 4.26, 6.48)
TL2	(2.04, 4.41, 7.67)	(3.89, 7.00, 9.00)	(1.44, 3.52, 5.96)
TL3	(0.93, 2.33, 5.00)	(5.44, 9.00, 9.00)	(0.78, 2.04, 4.41)
TL4	(2.78, 5.44, 8.33)	(3.37, 6.33, 7.67)	(2.33, 5.00, 7.00)

Table 50.7 Distance of each alternative, closeness coefficient and rank of team leaders

Alternatives	Distance from FPIS	Distance from FNIS	Closeness coefficient	Rank
Team leader 1	11.48	12.86	0.53	2
Team leader 2	11.88	12.61	0.52	3
Team leader 3	13.36	10.44	0.44	4
Team leader 4	10.76	13.45	0.56	1

closeness coefficient, which forms a scientific basis of reaching consensus amongst experts in the presence of difference of opinion. In our study, fourth team leader is declared as best team leader on the basis of the highest closeness coefficient. This study can be further extended to large number of selection criteria. In our future work, we propose to develop hybrid techniques for team leader selection by using more than one MCDM methods.

References

1. Xu, P., Shen, Y.: The role of leadership in agile software development. In: International Research Workshop on IT Project Management (2018)
2. Wi, H., Oh, S., Mun, J., Jung, M.: A team formation model based on knowledge and collaboration. Expert Syst. Appl. **36**(5), 9121–9134 (2009). https://doi.org/10.1016/j.eswa.2008.12.031
3. Lele, A.: Formation of an efficient team by improvising employee selection process using AHP-LP for a software company in India. Manag. Labour Stud. **40**(1–2), 22–33 (2015). https://doi.org/10.1177/0258042X15601531
4. Muhisn, Z.A.A., Omar, M., Ahmad, M., Muhisn, S.A.: Team leader selection by using an analytic hierarchy process (AHP) technique. JSW **10**(10), 1216–1227 (2015). https://doi.org/10.17706/jsw.10.10.1216-1227
5. Sokolovska, B., Caganova, D., Kasnikova, K.: The requirements to develop a competency model for the position of a team leader in industrial enterprises in Slovakia. In: Smart City 2015, LNICST 166, pp. 572–583. Springer, Cham (2016)
6. Sandhya, S., Garg, R.: Implementation of multi-criteria decision-making approach for the team leader selection in IT sector. J. Proj. Manag. **1**(2), 67–75 (2016). http://doi.org/10.5267/j.jpm.2017.3.001
7. Jaafar, J., Gilal, A.R., Omar, M., Basri, S., Aziz, I.A., Hasan, H.M.: A rough-fuzzy inference system for selecting team leader for software development teams. In: Proceedings of the Computational Methods in Systems and Software, pp. 304–314. Springer, Cham (2017). http://doi.org/10.1007/978-3-319-67618-0_28
8. Goyal, M., Gupta, C.: Intuitionistic fuzzy decision making towards efficient team selection in global software development. J. Inf. Technol. Res. (JITR) **13**(2), 75–93 (2020). https://doi.org/10.4018/JITR.2020040105
9. Assavakamhaenghan, N., Suwanworaboon, P., Tanaphantaruk, W., Tuarob, S., Choetkiertikul, M.: Towards team formation in software development: a case study of moodle. In: 2020 17th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON), pp. 157–160. IEEE (2020)
10. Sodhi, B., Prabhakar, T.V.: A simplified description of fuzzy topsis. arXiv preprint arXiv:1205.5098 (2012)
11. Buckley, J.J.: Ranking alternatives using fuzzy numbers. Fuzzy Sets Syst. **15**(1), 21–31 (1985)

Chapter 51
Multicore Embedded Worst-Case Task Design Issues and Analysis Using Machine Learning Logic

Sumalatha Aradhya, S. Thejaswini, and V. Nagaveni

Abstract In critical embedded systems especially hard real time systems, worst-case execution time analysis is considered rather than expensive testing tools. In real time, most of the optimization methods fail because software or code performance is optimized only up to 10–20% but not more. A strategy to analyze the tasks across multicores during worst-case scenario when event occurs is needed as the solution. The worst-case execution time is derived as a solution for scheduling event-related tasks. Estimations are needed to verify the response time of unit of code under test, to understand the finishing time of interrupt handlers and to set the sample rate of control loop. The worst-case execution time analysis is required to understand the program flow to analyze dynamic behavior of the program in execution and also to understand machine timing effects. In this paper, the machine learning logic is used to analyze the design space issues.

51.1 Introduction

The transactions across core are puzzled due to atomicity. The other aspects of parallel programming problems are identifying concurrency, partitioning work, and ordering actions. To overcome design issues, the methodology such as auto-tuning is followed in most of the applications. Auto-tuning though is successful as library approach such as FFTW [1], Atlas [2], and OSKI [3], a much work is needed to apply to applications rather than kernels as a huge search space is needed to progress ineffective truncated searches [4]. Depending upon which location written by one processor and used by another processor and data dependence analysis, a location-based communication code is generated [5].

S. Aradhya (✉) · S. Thejaswini
Siddaganga Institute of Technology, Tumkur, Karnataka, India
e-mail: sumalatha@sit.ac.in

V. Nagaveni
Acharya Institute of Technology, Soladevanahalli, Karnataka, India

T. Senjyu et al. (eds.), *IOT with Smart Systems*, Smart Innovation, Systems and Technologies 251, https://doi.org/10.1007/978-981-16-3945-6_52

The automated runtime debuggers such as Intel's thread checker [6], VTune [7], and double checker [8] check for storage conflicts and search the places where data stalled. Most of the time, these dynamic analysis tools are dataset dependent. When the application has a greater number of individual tasks, functional decomposition is used [9]. The best way to do the functional decomposition is explicit threading. If the dataset has huge set of independent data, then the threaded model needs to adapt data decomposition [10]. The data parallelism is achieved through compiler directed methods such as OpenMP [11], OpenCL [12], and Cilk [13]. Thus, to exploit multicore, programmer needs to identify and analyze suitable parallelism for the application and then create the threads to run multiple tasks concurrently. When it comes to the work sharing across heterogeneous, the performance trades off with work distribution done following the data dependencies [14]. The communication is handled by exchanging the data and control messages across core through the prefetching model [15]. However, the defect pattern arises because of performance scalability, and unnecessary waiting time issue occurs due to in effective load balancing. The way to prevent the defect in such case is to ascertain that all processors are very well balanced or efficiently occupied to handle unbalanced tasks. The bug-related symptoms occur during compile time, and some defects surface only under specific conditions such as processor issues, input value issues, and alignment issues. But, the identification of design space issues and analyzing the behavior of the program are time consuming and may need expensive tool kits. Hence, an effective and easy approach to carry out WCET analysis is needed. In our paper, we proposed WCET analysis by using machine learning approach. To carry out the research work, we used emulators (ReneasesRX232T), time-accurate simulators, static analysis tools (absInt), software profiling tools, and machine learning-based failure prediction logics [16].

51.2 Literature Survey

51.2.1 Existing Work on Design Space Issues in Multicore Embedded Systems

The parallel programming by pattern can lead to high-quality solutions in embedded domain. Each pattern provides common vocabulary for obtaining the solutions to the programming community [17]. If not written by pattern, it is too difficult for programmers and software to exploit parallelism to full extent. With respect to parallelizing programs, the design space patterns such as finding concurrency for designing algorithm expression and for building parallel software construction are considered. The defect patterns are common sort of bugs in debugging and are obtained due to inappropriate choice of function calls. The common design space pattern issues are found very frequently in embedded multicore systems [18]. The most difficult defect pattern is synchronization. Task synchronization enforces preconditions and postconditions of logical processes [19]. However, the parallelism techniques used to schedule and parallelize subtask do not guarantee parallelization of subtasks. A semipartitioned approach proposed by Bado et al. [20] with fork-join task model and Dziurzanski [21] suggested task level-migration and have used semipartitioned scheduling algorithms. If the dataset does not provide the data path, the tools fail to evaluate code portions. The defect arises due to the improper coordination between the processors, and common defects that occur in parallel programming are deadlocks and race conditions [22]. The symptoms of such errors are program hangs and incorrect, non-deterministic outputs. One way to prevent the synchronization defect is to make sure that all communications are correctly coordinated. Hence, in embedded applications, it is relevant to create a dataset that simulates relevant asynchronous processes [23].

One more kind of defect pattern observed is due to side effect of parallelization. The way to prevent such errors is to just not to focus on parallel codes alone, and checking that defect may surface only in parallel context [23]. The parallelism techniques used to schedule and parallelize subtask do not guarantee parallelization of subtasks. As described by Kim et al. [24], scheduling through thread clustering provides better system throughput and fair scheduling especially in heterogeneous systems. In heterogeneous cores, the different architectural parameters and embedded application-specific constraints can be addressed using machine learning-based configurable parameters [25] but not on parallel programming patterns. Thus, as per literature survey carried out, an effective WCET analysis for critical design space issues using machine learning logic is not found in the existing work. In this paper, we proposed the WCET analysis by applying machine learning logic on critical design space issues. We have considered the critical design space issues such as jitter, incorrect priority assignment, priority inversion, deadlock, memory leak, high coupling and data dependency, segmentation fault, incorrect outputs, program hangs, non-deterministic outputs, correctness and performance problems, sublinear scalability, and waiting time for WCET analysis in multicore embedded systems.

51.3 Derivation of Mathematical Model for Worst-Case Task Execution

51.3.1 Derivation of Optimal Schedule for Worst-Case Task Execution

Let μ be the vector size allocated by a task v during the computation phase. For all $i = 0 \ldots n$ tasks, let T_i represents execution time of task i. Then, the mean duration, μ, of n tasks is computed as shown in (51.1).

$$\mu = \sum_{i=0}^{n} T_i \tag{51.1}$$

Here, μ represents mean duration of n tasks. For the worst-case, let, Δt is the latest task with longer time duration. The final task is scheduled if the condition at (51.2) is met.

$$\left\lceil \frac{n}{P_i} \right\rceil \sum_{i=1}^{n-1} T_i \leq \left\lceil \frac{n-1}{P_i} \right\rceil . \mu + \Delta t \tag{51.2}$$

Case 1: Finding lower limit of the execution time that is no sharing of resources as per (51.3) [26]

$$\text{mean time}_i^{\text{low}} = \left\lceil \frac{T_i}{P_i} \right\rceil \times \text{totalavg}(T_i)\ \forall i \in n \text{ tasks} \tag{51.3}$$

Here, totalavg is the average execution time of the tasks. The $\lceil T_i \div P_i \rceil$ provides the execution time rounds when task T_i allocated to P processor. When resources are not shared, then meantime evaluates to 0. This means no overhead on the execution time of tasks.

Case 2: Finding upper limit of the execution time that when the resources are shared and is computed by (51.4) [26].

$$\text{mean time}_i^{\text{up}} = \left\lceil \frac{T_i}{P_i - 1} \right\rceil \times \text{total avg of } T_i + t_{\max} \text{ of } P_i \tag{51.4}$$

Here, $i \in n$ tasks and P_i is the maximum time observed in the event ting upper and lower bounds for the task; we can compute the end-to-end estimation time of the job. Then, $t_{\max}$ represents the maximum task execution time at P_i. The completion of task i requires the condition given in (51.5).

$$x_i(C_i) = \int_0^{C_i} f_i u_i(t)\mathrm{d}t = w_i \tag{51.5}$$

The status of job i is resented by x_i, and C_i represents the completion time. The capacity of the task is represented by w_i. It is easy to verify that $f(e_i, t_i) = 0$ if C_i is not satisfied, and $f(e_i, t_i) = 1$, if C_i is satisfied under almost e_i scenarios. Furthermore, if $f(e_i, t_i + \Delta t) = m$, where m is medium instance, then e_i must belong to almost m scenarios. Otherwise, $f(e_i, t_i + \Delta t) \geq 2m$. Thus, schedule e_i such that occurrence of event in the worst-case scenario is represented in (51.6) [27]

$$\text{eventOccured} \leq (m + e_i) \div (m + 1) \tag{51.6}$$

Relating objective measure of the work related to the function event sequence $f(e_i, t_i)$, we can derive the scheduling algorithm. The complex dependency of the tasks to get scheduled across the multiple processors is given by considering the parameters such as size w_i of the job i, execution time, and resource allocation time. We can compute the optimal scheduling, represented by OptimSchedule as per the Formula (51.7).

$$\text{OptimSchedule} = w_i \times f(e_i, t_i) \times \text{mean time}_i^{\text{low}} \times t_{\min} \text{ of } P_i) \tag{51.7}$$

The classification and training datasets are randomly sampled for machine learning. The worst-case training data is determined by the sample ratio of worst-case defect pattern class and its associated processor with mean time execution time with upper bound. The sampled points are mapped with upper bound of all design space issues mentioned in the literature survey section.

The assessment for accuracy is done by using the confusion matrix [28]. We propose, a better measured accuracy, OptimizedAccuracy, as per the formula represented in (51.8)

$$\text{OptimizedAccuracy} = \text{TP} \div (\text{TP} + \text{FN} + \text{FN}) \tag{51.8}$$

Here, TP represents TruePositive class and FN represents FalseNegative class. Further improvement is suggested by using Kappa coefficient estimation [29]. The decision tree generated is considered for the training and validation of the inputs. The ReLu function model [30] is used to predict WCET analysis further by capturing the interactions and nonlinear inputs.

51.3.2 *Functional Block Representation to Analyze WCET Using ML*

The functional model using ML logic is represented through the block diagram shown in Fig. 51.1. The WCET analysis is performed on the executable of unit under test. It is ensured that the execution of test is uninterrupted by the exceptions and other background activities. The dynamic behavior of the unit under test is determined by the program flow analysis. Both the control flow and data flow analyses provide the

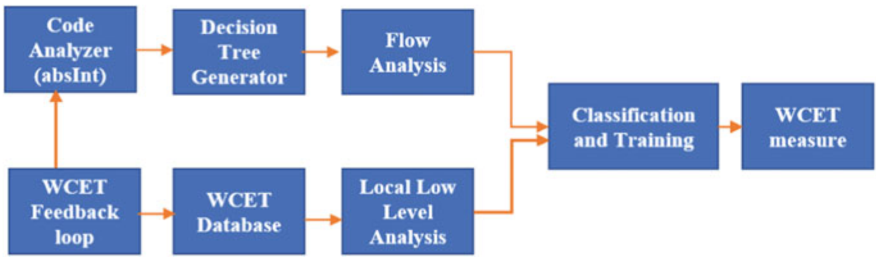


Fig. 51.1 Worst-case execution time functional block diagram using ML

information related to the function task in use, function task called, iteration count of loops, and inter dependencies of tasks.

WCET database is updated after the analysis of local low-level codes and control flow. Local low-level analysis module will internally analyze logical calculation, mathematical calculation, and function flow of a program based on which core/function taking more execution time. If programmer used extra arithmetical operation and incorrect logical flow, then it led to a more execution time of program. The scheduler handles the critical timing issue caused by memory conflicts. The worst-case scheduling algorithm for obtaining the schedule for all critical tasks submitted is shown in Fig. 51.2.

The classification and training ML model is represented in Fig. 51.3. ReLu function model [30] is used because it avoids vanishing gradient problem. With the usage of worst-case task execution input data, further classification, training, and validation of WCET analysis are accomplished.

```
for T: = {1, n}
for all i belongs to O // where, O is Optimized minimum position among all schedule
//Obtain the processing time for all scheduled tasks
procTime(eventOccurred): = sigma procTimeO(eventOccurredi)//O belongs to Task
end for
for jobs r ranging down from n to 1
  //find task j belong to T, which has no successor in T and has the minimum value of function Fj(T)
  Fj(T) = max {i belong O} worstj(eventOccurredi)(procTime(eventOccurredi)-otherSetj(eventOccurredi))
  schedule(r): =j
  T: =Tj
  for all i belongs to O
    procTime(eventOccurredi): = procTime(eventOccurredi)-procTimej(eventOccurredi)
  end for
end for
return schedule
end for
```

Fig. 51.2 Worst-case task scheduling algorithm for all critical tasks submitted

```
def build_model():
  model = Sequential()
  # Input Layer
  model.add(Dense(units = 128, activation = 'relu', input_shape = [len(X.keys())]))
  # Hidden Layer - I
  model.add(Dense(units = 256, activation = 'relu'))
  # Hidden Layer - II
  model.add(Dense(units = 512, activation='relu'))
  # Output
  model.add(Dense(units= 1))
  # optimizers
  # error function - mean_squared_error , optimizers = RMSProp with learning rate
  optimizers = keras.optimizers.RMSprop(learning_rate=0.001)
  # Model Compile - Only for Regression
  model.compile(loss = 'mean_squared_error', optimizer = optimizers, metrics = ['mean_squared_error', 'mean_absolute_error'])
  return model
model = build_model()
model.summary()
history = model.fit(X_train, Y_train, epochs = 600, batch_size=25, validation_split=0.20)
predictions = model.predict(X_test)
predictions
yhat = np.round(predictions)
from sklearn.metrics import confusion_matrix
confusion_matrix(Y_test, yhat)
from sklearn.metrics import classification_report
print(classification_report(Y_test, yhat))
```

Fig. 51.3 ReLu activation model for classification and training

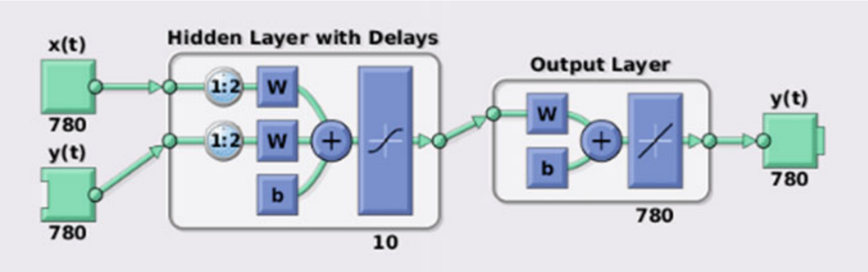


Fig. 51.4 Worst-case task execution ML model

51.4 Results and Analysis

51.4.1 Solution Representation of Worst-Case Task Execution Model Using Machine Learning Logic

The worst-case execution time adds the complexity to the core-balancing model. One such scenario is represented in Fig. 51.4. Here, the defect pattern issues are sampled for 780 input data values. The dynamic data has 11 time steps of 780 elements.

The output parameter $y(t)$ represents the outcome of ReLu activation function [30]. The training, validation, and testing values for the defect classes are randomly divided into 8580 steps such that 70% training dataset with 6006 target time steps, 15% of validation dataset with 1287 target time steps, and 15% of testing dataset with 1287 target time steps. The ML decision tree learning result for design space issue class is shown in Fig. 51.5. The optimized accuracy result depicts that for critical design space issues such as sublinear scalability, program hangs, non-deterministic outputs, incorrect outputs, high coupling and data dependency, correctness and performance problems, and jitter have occurred at the samples x, and their best and worst execution time in microseconds across cores have to be sampled to get the optimized accuracy results.

The worst-case dynamic scheduling algorithm when deployed on the sampled data, the result obtained for one of the design spaces-correctness and performance problems, is shown in Fig. 51.6.

The OptimSchedule value shows that the WCET analysis of correctness and performance issue has helped the programmer to fix the issue before hand, and the optimum result is achieved from 1.2335 upper bound to OptimSchedule value 0.6847. Thus, the prediction helped to improve the result by 55%.

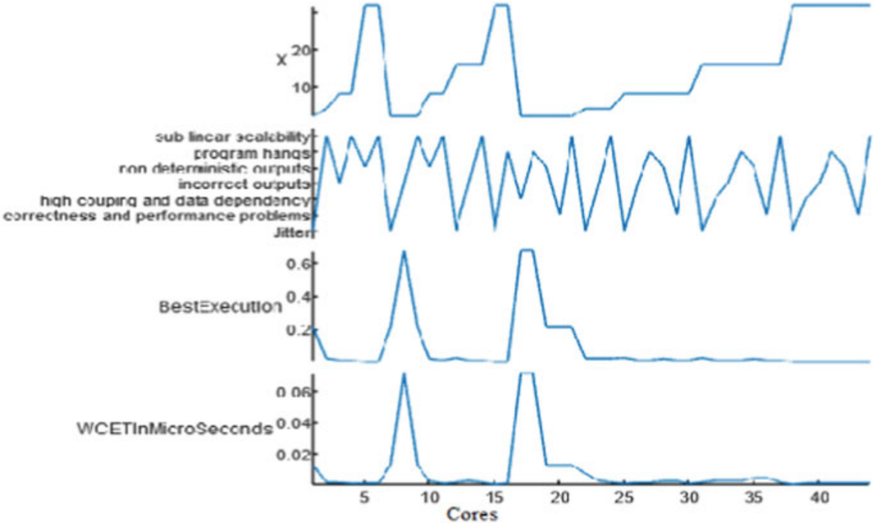
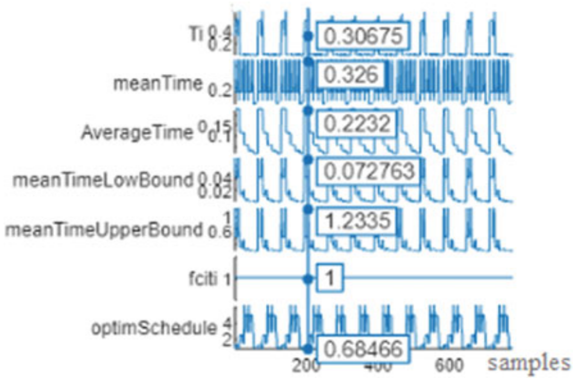


Fig. 51.5 ML decision tree learning result for design space issue

Fig. 51.6 WCET analysis for correctness and performance issue



51.5 Conclusion

To work with real-time multicore embedded systems, an embedded developer has to focus on critical challenges such as core balancing, scheduling, critical safety measures, functionalities, limited resource utilization, and performance. An effective real-time program flow analysis especially worst-case execution time analysis shall ease the job of developer. There exists very minimal efficient framework to understand the worst-case analysis in critical systems. In this paper, the design defect scenarios are deeply analyzed using machine learning algorithms. The worst-case execution time analysis is done by considering a real-time use case and the design defect pattern issues. Estimations of WCET are needed to verify the response time of unit of code under test, to understand the finishing time of interrupt handlers and to set the sample rate of control loop. Using the machine learning-based validation data, the programmer shall have an option to verify the calling sequence and to correct the design and code based on WCET analysis. WCET analysis shall help the embedded system developers to overcome the dynamic execution challenges as they will get the prior information regarding the dynamic design space issues.

References

1. Popovici, D.T., Low, T.M., Franchetti, F.: Large bandwidth-efficient FFTs on multicore and multi-socket systems. In: IPDPS (2018)
2. Öhman, H., Panitkin, S., Hendrix, V.: Using puppet to contextualize computing resources for ATLAS analysis on Google compute engine. J. Phys. Conf. Ser. **513** (2014). http://doi.org/10.1088/1742-6596/513/3/03207
3. Ibrahim, K.Z., Williams, S.W., Epifanovsky, E., Krylov, A.I.: Analysis and tuning of libtensor framework on multicore architectures. In: High Performance Computing Conference (HIPC) (2014)
4. Tang, H., Byna, S., Bailey, S., Lukic, Z., Liu, J., Koziol, Q., Dong, B.: Tuning object-centric data management systems for large scale scientific applications. In: 26th IEEE International Conference on High Performance Computing, Data, and Analytics (HiPC 2019), 18 Dec 2019
5. Chaimov, N., Ibrahim, K., Williams, S., Iancu, C.: Exploiting communication concurrency on high performance computing systems. In: IJHPCA, 17 Apr 2015
6. James, R.: Intel Threading Building Blocks: Outfitting C++ for Multi-core Processor Parallelism, p. 2007. O'Reilly Media, Inc., Sebastopol, CA (2007)
7. Intel (R) VTune (TM) Performance Analyzer 9.0 by Intel Corporation. http://intel-r-vtune-tmperformance-analyzer.software.informer.com/9.0/
8. DoubleCheck—Integrated Static Analysis Tool by Green Hills Software. https://www.ghs.com/products/doublecheck.html
9. Zhao, T., Hall, M., Basu, P., Williams, S., Johansen, H.: SIMD code generation for stencils on brick decompositions. In: Proceedings of the 23rd ACM SIGPLAN Symposium on Principles and Practice of Parallel Programming (PPoPP), Feb 2018
10. Preissl, R., Shalf, J., Wichmann, N., Ethier, S., Long, B., Koniges, A.: Multithreaded global address space communication techniques for gyrokinetic fusion applications on ultra-scale platforms. In: SC'11, p. 15 (2011)
11. Butko, A., Bruguier, F., Gamatié, A., Sassatelli, G.: Efficient programming for multicore processor heterogeneity: OpenMP versus OmpSs. In: Open Source Supercomputing Workshop. Springer's Lecture Notes in Computer Science (LNCS). Frankfurt, Germany, 22 June 2017
12. Koshiba, A., Sakamoto, R., Namiki, M.: OpenCL runtime for OS-driven task pipelining on heterogeneous accelerators. In: 2018 IEEE 24th International Conference on Embedded and Real-Time Computing Systems and Applications (RTCSA), Hakodate, pp. 236–237 (2018). http://doi.org/10.1109/RTCSA.2018.00038
13. Agrawal, K., Lee, I.A., Li, J., Lu, K., Moseley, B.: Practically efficient scheduler for minimizing average flow time of parallel jobs. In: 2019 IEEE International Parallel and Distributed Processing Symposium (IPDPS), Rio de Janeiro, Brazil, pp. 134–144 (2019). http://doi.org/10.1109/IPDPS.2019.00024
14. Zhang, C., Peleato, B.: On the average rate for coded caching with heterogeneous user profiles. In: ICC 2020—2020 IEEE International Conference on Communications (ICC), Dublin, Ireland, pp. 1–6 (2020). http://doi.org/10.1109/ICC40277.2020.9148779

15. Aradhya, S., Srinath, N.K.: Proliferation framework on input data set to improve memory latency in multicore for optimization. In: Proceedings of IACC. IEEE Publications, June 2015. 978-1-4799-8047-5. http://doi.org/10.1109/IADCC.2015.7154712
16. Mohammed, B., Awan, I.U., Ugail, H., Mohammad, Y.: Failure prediction using machine learning in a virtualized HPC system and application. Cluster Comput. J. Netw. Softw. Tools Appl. 471–485 (2020)
17. Li, Z., et al.: Time and space-efficient write parallelism in PCM by exploiting data patterns. IEEE Trans. Comput. **66**(9), 1629–1644 (2017). http://doi.org/10.1109/TC.2017.2677903
18. Bertauld, T., Dagenais, M.R.: Low-level trace correlation on heterogeneous embedded systems. EURASIP J. Embed. Syst. (2017). http://doi.org/10.1186/s13639-016-0067-1
19. Chung, M.T., Weidendorfer, J., Samfass, P., Fuerlinger, K., Kranzlmüller, D.: Scheduling across multiple applications using task-based programming models. In: 2020 IEEE/ACM Fourth Annual Workshop on Emerging Parallel and Distributed Runtime Systems and Middleware (IPDRM), GA, USA, pp. 1–8 (2020). http://doi.org/10.1109/IPDRM51949.2020.00005
20. Bado, B., George, L., Courbin, P., Goossens, J.: A semi-partitioned approach for parallel real-time scheduling. In: Proceedings of the 20th International Conference on Real-Time and Network Systems, At Pont à Mousson, France, RTNS'12, ACM. 978-1-4503-1409-1
21. Dziurzanski, P., Singh, A.K., Indrusiak, L.S.: Multi-criteria resource allocation in modal hard real-time systems. EURASIP J. Embed. Syst, pp. 1–16. (2017)
22. Chiu Y.-C., et al.: Data race avoidance and replay scheme for developing and debugging parallel programs on distributed shared memory systems. Parallel Comput. **37**(1) (2011)
23. Kast, B., Schmitt, P.S., Albrecht, S., Feiten, W., Zhang, J.: Hierarchical planner with composable action models for asynchronous parallelization of tasks and motions. In: 2020 Fourth IEEE International Conference on Robotic Computing (IRC), Taichung, Taiwan, pp. 143–150 (2020). http://doi.org/10.1109/IRC.2020.00029
24. Kim, Y., Papamichael, M., Mutlu, O., Harchol-Balter, M.: Thread cluster memory scheduling: exploiting differences in memory access behavior. In: 43rd International Symposium on Microarchitecture (MICRO), pp. 65–76, Atlanta, GA (2010)
25. Vazquez, R., Gordon-Ross, A., Stitt, G.: Machine learning-based prediction for dynamic architectural optimizations. In: 2019 Tenth International Green and Sustainable Computing Conference (IGSC), Alexandria, VA, USA, pp. 1–6. (2019). http://doi.org/10.1109/IGSC48788.2019.8957207
26. Zacheilas, N., Kalogeraki, V.: A pareto-based scheduler for exploring cost-performance tradeoffs for MapReduce workloads. EURASIP J. Embed. Syst, pp. 1–24. (2017)
27. Różycki, R., Waligóra, G., Węglarz, J.: Scheduling preemtable jobs on identical processors under varying availability of an additional continuous resource. Int. J. Appl. Math. Comput. Sci. **26**(3), 693–706 (2016). http://doi.org/10.1515/amcs-2016-0048
28. Congalton, R.G., Green, K.: Assessing the Accuracy of Remotely Sensed Data—Principles and Practices, 2nd ed. CRC Press, Taylor & Francis Group, Boca Raton, FL (2009). 978-1-4200-5512-2
29. Mchugh, M.L.: Interrater reliability: the kappa statistic importance of measuring interrater reliability. Biochemia Medica **22**(3), 276–282 (2012)
30. BrauBe, F., Khasidashvili, Z., Korovin, K.: Selecting stable safe configurations for systems modelled by neural networks with ReLU activation. In: 2020 Formal Methods in Computer Aided Design (FMCAD), Haifa, Israel, pp. 119–127 (2020). http://doi.org/10.34727/2020/isbn.978-3-85448-042-6_19

Chapter 52
Exploration of Technology-Driven Income Sources for an Agricultural Community in West Bengal, India

P. S. Vishnu Priya, L. M. Frey, R. Chinmayi, Renjith Mohan, E. Lalith Prakash, S. Anne Rose, S. R. Shiva, G. Muthusurya, Dongala Abihas Balaji, and S. Vidhya

Abstract Global climate changes are having a severe impact on India's agricultural sector in impoverished rural areas. Addressed in this study, many who have depended upon agriculture are experiencing climatic changes, leading to problems in sustaining secure livelihoods. To mitigate this risk, other sources of income need to be introduced to supplement agricultural income. This paper examines alternative income sources that provide an option for a sustainable income in a rural village of West Bengal, India. The prevalent technologies available in the village were assessed, besides those introduced, to realize sustainable augmented income goals. Studies revealed that village resources comprised fishponds, a poultry farm for producing and selling eggs, and small-scale kitchen gardens. Results also showed a lack of knowledge of other ways of additional income generation among villagers. Close observation of the

P. S. Vishnu Priya · S. Vidhya
Department of Computer Science and Engineering, Amrita Vishwa Vidyapeetham, Coimbatore, India
e-mail: s_vidhya@cb.amrita.edu

L. M. Frey · R. Mohan (✉)
Amrita Center for International Programs, Amrita Vishwa Vidyapeetham, Kollam, India
e-mail: renjithmohan@am.amrita.edu

L. M. Frey
e-mail: lisamfrey@am.amrita.edu

R. Chinmayi
Department of Electronics and Communication Engineering, Amrita Vishwa Vidyapeetham, Kollam, India
e-mail: chinmayir@am.amrita.edu

E. Lalith Prakash
Department of Civil Engineering, Amrita Vishwa Vidyapeetham, Coimbatore, India
e-mail: e_lalitprakash@cb.amrita.edu

S. Anne Rose · S. R. Shiva · G. Muthusurya
Department of Electrical and Electronics Engineering, Amrita Vishwa Vidyapeetham, Coimbatore, India

D. A. Balaji
Department of Aerospace Engineering, Amrita Vishwa Vidyapeetham, Coimbatore, India

T. Senjyu et al. (eds.), *IOT with Smart Systems*, Smart Innovation, Systems and Technologies 251, https://doi.org/10.1007/978-981-16-3945-6_53

current occupational practices determined that technological innovations could be introduced to maximize available village resources for improved sustainable income, concurrent with traditional agricultural activities. For example, an innovative engineering design can increase vegetable gardening and sales through a sophisticated system of maximizing outcomes in a small space, introducing watering technologies and temperature modifiers.

52.1 Introduction

Globally, more than half of the 7.6 billion people in the world live on less than $2.50 a day. Among them, at least 1.3 billion people endure extreme poverty conditions, living on less than $1.25 a day [1]. Poverty disrupts human lives in numerous ways; the less educated are more likely to remain ensnared in poverty. Studies have shown that about 263 million children worldwide were not enrolled in schools for 2016 [2]. As a single country, India has one of the world's largest populations and an equally large number of impoverished citizens. As of 2013, about 30% of its population survived on less than $1.90 per day [3]. As a developing country, India has put in steady efforts to help millions of its impoverished citizens. However, despite significant progress in reducing poverty across the country, the rural poor have not realized or experienced these gains, many of whom continue to experience extreme poverty [1]. Researchers find that some states show notable improvement even in rural India, whereas others lag in relentless poverty [1, 4]. People living on multiple levels in West Bengal, India, have reported adverse effects of climate change impeding their livelihood; in the last two decades, the following areas were notably affected: agriculture, fishing, forestry, and human health [5]. Aggravated unremitting poverty contributes to urban migration [6].

Under the aegis of Amrita Vishwa Vidyapeetham's Live-in-Labs® program, a study was conducted at Kalinagar, a village in West Bengal, to explore the poverty alleviation opportunities in the rural community [7]. The village is 26.6 km away from Kolkata, the state capital; Pailan is the nearest city to the village. There are around 294 families in the village with a population of 1264 [data collected from Amrita Self Reliant Villages (SeRVe)]. The study focused on appraising the impediments to generating adequate income and exploring sustainable income generation from village resources.

52.2 Methodology

This study aims to map potential resources that can lead the aspirations of the community to be self-reliant. A naturalistic study was performed to understand a clear rationale of the present activities in Kalinagar, its history, culture, and other perspectives. Inaccessible remote terrains and the village's geographical expanse were the

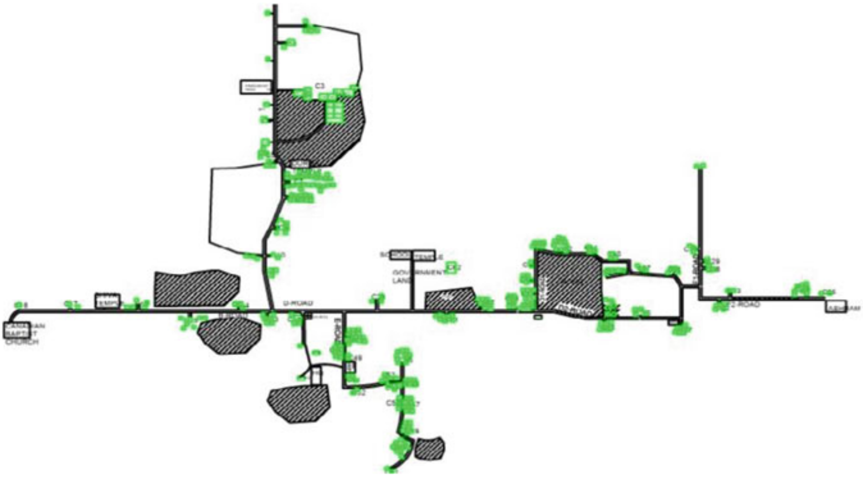


Fig. 52.1 Kalinagar village, West Bengal

team's most significant challenges to conducting a naturalistic survey. The village's geographical spread is shown in Fig. 52.1, where the ponds, roads, and houses are mapped. The ponds are mapped as the gray shaded area, where the houses are mapped in green dots. The team deployed various data collection methods such as observation analysis, personal interviews, and discussions with focus groups. The observations were planned in 45 clusters by resources available in those clusters.

The research team spent six days in the village, observing its residents' daily routine of beautiful Kalinagar, surrounded by ponds and the ecosystem. The team focused on data collection from women [primarily from self-help groups (SHGs)] and men (primarily farmers) by in-person interviews. Observational studies consisted of probing available village resources, the ponds and water-related issues, farming, access to technology available in the village, potential sources of supplementing income, and occupational opportunities. Fact-finding meetings were also held with village heads and activist community members. The village coordinator provided primary, secondary data on the village, collected by Amrita SeRVe. The SeRVe data mainly involved demographic information of the community. Interviews were conducted with semi-structured questionnaires focused on learning about villagers' subsidies, any debts or loans taken by farmers, and farmer credits. Additional interviews probed the Anganwadi teachers to understand students' behavior and inquire if they continued their education in schools. The questions focused on the villagers' preferred occupations, the crops they grew, whether they owned any assets such as land, and finally, the money earned by all of these activities. Observational and field studies included learning the villagers' agricultural practices, the farmers' crops, the climatic and weather conditions that favor the crop yields, the market value farmers get, and soil testing camps. Other field studies were with small-scale business people, who sold duck eggs and water cans, engaged in fish farming, the fish food they consumed, the market value of the fish they cultivated, and reviewed mill owners' economic gains by grinding flour of various grains. Generally, we spoke to them to familiarize ourselves with the hardships experienced by them in income generation.

52.3 Results

The village's socio-economic status was mapped and analyzed using key observations and data collected over six days. The data were classified according to various socio-economic parameters and are discussed in this section.

52.3.1 Education Status

The educational qualifications of the villagers are summarized in Fig. 52.2. Villagers mentioned the tuition center, which runs a few years back in the villages. However, after seeing that the educated are unemployed, the villagers are no longer interested to send their children to school or higher education.

Fig. 52.2 Number of people having completed up to a certain educational standard

52.3.2 Economic Status

Although most of the villagers are school dropouts, the Kalinagar people seem to have better financial literacy than similar villagers. As noted in Fig. 52.3, more than 90% of the villagers have bank accounts. The economic status of the Kalinagar village can be better represented by the type of ration card they hold. Three different ration cards are actively used in this village: Antyodaya Anna Yojana (AAY), State Priority Households (SPHH), and Priority Households (PHH) cards.

AAY cardholders have a daily income of less than Rs. 100, whereas PHH and SPHH cardholders earn less than Rs. 250 and Rs. 300 per day, respectively. Figure 52.4 shows the categorization of villagers according to the type of ration card.

All these three cardholders are considered below the poverty line (BPL) by the state government. The above poverty line (APL) cards SGRY1 and SGRY2 are not issued in the village. Data collection indicated several communal activities—some were for the generation of income, and others primarily for supplying food to needy families.

All these three cardholders are considered below interviews with the selected people, and post-interview interactions with their family members revealed the various modes of income generation in the village. Kalinagar is beautified by 60 ponds distributed throughout the village. Each family owns at least one pond, used to cultivate local fish breeds for their self-consumption. Some villagers were professional fish farmers who purchased international breeds from outside, cultivated them in their ponds, and sold them. These fishes are sold in large amounts, as they are produced in the ponds.

Fig. 52.3 Distribution of Kalinagar villagers, with and without bank accounts

Fig. 52.4 Number of persons per each ration card type

According to the villagers, some of them die due to malnutrition, pathogen attacks, and diseases. Few farmers sell the fish at young ages to mitigate the burden of breeding them (because stealing fish is a problem during the fishing season); they are happy even if it returns smaller profits. However, most farmers sell fish only after they are fully grown so that they gain decent profits.

With a plurality of ponds in this village, integrated farming can help them augment their revenue. Poultry–fish culture can be practiced, wherein poultry excreta is recycled and reused as feed for fishes in ponds, as shown in Fig. 52.5. Poultry excreta rich in minerals like nitrogen, phosphate, and potash serves as a good fish feed. The concept of recycling and reuse forecloses the application of extra fertilizer for

Fig. 52.5 Poultry house over a fish pond in the village

fish breeding. One of the farmers pursued a sustainable system, constructing a chicken coop that stands on stilts over a pond. Chicken eggs are sold for income. The chicken droppings become food for the fish so that the farmer gets double benefits. Further, ducks are grown, swim in the ponds, and sell the duck eggs. This duck–fish culture has solid remunerative potential. Ducks are called living manuring machines. Duck droppings, rich in calcium, nitrogen, potassium, carbon, and other nutrients, are a bonus for fishponds, as they remove potentially harmful snails and insects (with their larvae) that are known to carry disease-causing microorganisms that can infect fish and thereby humans as well. Ducks can bring additional income by farmers selling the eggs [8].

52.3.3 Key Activities of the Villagers

1. Agriculture: The villagers' staple diet is constituted of roti (Indian-style bread made from wheat and rice). Field visits to the paddy fields, which cover 60 villages, are used for paddy and jute cultivation. A critical irrigation facility is a canal that runs along the boundary of the village. During the rainy season, the fields often get flooded. The majority of the harvest is kept for self-consumption, with the remaining sold in the nearby market.
2. Kitchen gardening: Kitchen gardening is a cost-effective task activity that provides a good nutrition source for village residents, especially women, children, and the elderly. Although many are not practicing kitchen gardening, the prime vegetables grown are lady's finger (okra), brinjal (eggplant), and cabbage. Cow dung used as manure serves as the primary source of fertilizer.
3. Some of the villagers work as daily wage laborers in fields owned by others. A few of them work as laborers in industrial companies outside Kalinagar.
4. Women are very active in the village. They run a food-prep machine used for making papad. Papad (widely known as papadam) is a thin, crisp, circular-shaped savory Indian food, based on a seasoned dough, commonly made from peeled gram flour and fried in oil. Although the group receives not many outside orders, SHGs have sufficient demand to run with a merger profit unit.
5. Tailoring classes are also taught but attended by very few women due to the lack of demand. Active occupations are water can sales, pen packaging, flour mill and small shops, driving auto rickshaws, and tuition classes to students.

In the state of West Bengal, majority of the girls get married at a very early age. The major reasons for early marriages, as mentioned by villagers, are family's poor financial status, atmosphere (many people get married at very young age), customs, family background, and domestic problems.

52.4 Discussion

The aquifers in the village were identified with the highest potential to serve as sources of income generation avenues for hundreds of families. This attributes to many big and small ponds, the hesitation from the women to go out and work in the town or nearby villages. However, fish farming is considered to be riskier at the moment, as any change in water quality affects the fish's health. A large number of pollutants are currently in the ponds due to improper waste disposal mechanisms. Also, the yearly monsoon flooding in the plains washes away the fish or changes the water quality. In order to maintain the quality of the water, a four-step process was proposed [9].

- Establish a netting structure within the ponds such that fishes are confined to one pond section. The nets should extend from the water's surface to the depth of the pond, attached to poles (possibly bamboo). This would keep debris thrown into the pond from affecting the fish. For example, plastics are terrible for water and the fish; they would be protected from such pollutants if they were in ponds. This is an ancient system that has been used in rural fishing villages of Kerala with success.
- A submersible sensor will be devised, augmenting and upgrading existing technology to appraise various water quality parameters such as pH, temperature, and oxygen. pH is critical as an indicator of the level of alkalinity or acidity of the water. The average span of pH for a fish is 6.5–9. Levels higher or lower than this range can cause illness and death to fish. In addition to pH, sensors will include measures of contaminants and impurities, nutrients, chlorophyll, dissolved oxygen, and water temperature, all of which could influence the fish's health and life. The sensor should be placed just inside of the netting and need to be checked only once daily.
- Create an overflow channel from the pond to a holding area or a naturally flowing drain or canal area. Thus, in times of excessive rain, the ponds' overflow could put the fish at risk of going out of the pond and ultimately ending up in shallow areas with too little water to survive. Thus, the netting helps to keep them within the pond.
- Establish an experimental duck farm above ponds with willing farmers to test the benefit of their droppings on the fish [10, 11]. Some research indicates that duck droppings are healthier for fish and influence fish's more considerable growth [12].

An architectural diagram of the proposed model based on the four-step process is shown in Fig. 52.6. The experiences from north Indian villages and Kerala in participatory technology innovation implementation are considered for the model's design [13, 14]. The ponds are connected by a pipeline network regulated by an automated valve shown by the figure in red lines. Sensors are deployed to monitor the water's pH levels and temperature, levels of biological dissolved oxygen, and the presence of a pathogen, dissolved contaminants, and suspended particles. A water

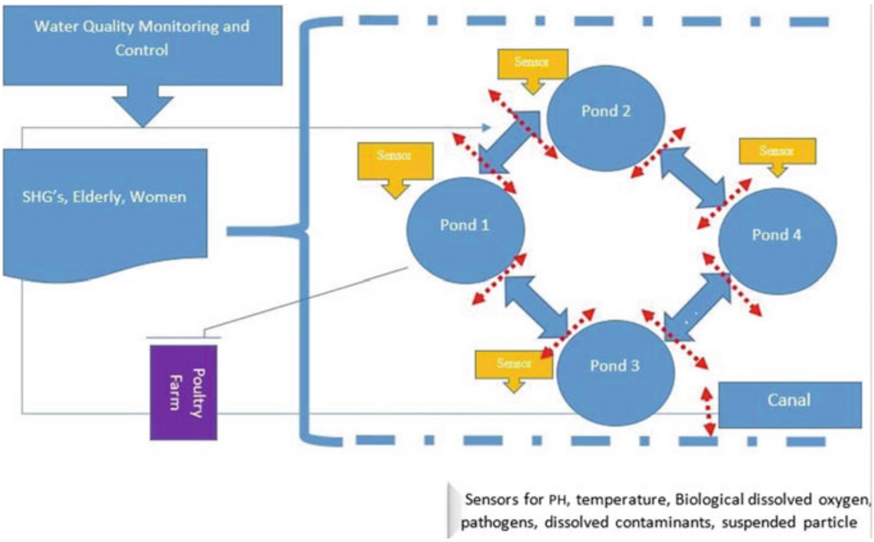


Fig. 52.6 Architecture diagram on technology-driven integrated fish and poultry farming

channel is created to facilitate water overflow during the rainy season into a nearby canal. This implementation is anticipated to be highly useful for the operability of integrated poultry–fish farming. Farming would help the women SHGs, homemakers, and older women contribute to the family income by working in the village, averting commutes back and forth to workplaces outside the village.

52.5 Conclusion

The study's purpose was to assess the village resources' potential, their sustainable utilization, and villagers' aspirations to be a self-resilient community. Life in the agrarian village is no longer self-reliant owing to the unprofitable farming job. Although the villagers sought other jobs by utilizing local resources, the traditional methods were riskier, and the villagers were skeptical about choosing them as alternative revenue opportunities.

The paper discussed leveraging advances in technology to develop sustainable income generation opportunities from available resources in the village.

Acknowledgements This research was funded by Amrita Vishwa Vidyapeetham's Live-in-Labs® program, and the field implementation was funded by the UN-recognized NGO, MA Math. The study was conducted as part of the UNESCO Chair for Experiential Learning for Sustainable Innovation Development. The authors express their immense gratitude to Sri. Mata Amritanandamayi Devi, Chancellor of Amrita Vishwa Vidyapeetham, who has inspired them in performing selfless service to society. The authors would also like to express their sincere appreciation to Mr. Gopalaiengar Sheshadri for his extensive support in enhancing the quality of the paper within the limited time frame. The authors also thank the faculty, staff, and students of Amrita Vishwa Vidyapeetham, the government authorities of Kalinagar, West Bengal, and also the Amrita SeRVe program staff.

References

1. 11 Facts About Global Poverty—DoSomething.org—Volunteer for. https://www.dosometing.org/us/facts/11-facts-about-global-poverty. Accessed 5 Feb 2018
2. Out-of-School Children—Global Partnership for Education. https://www.globalpartnersip.org/focus-areas/out-of-school-children. Accessed 5 Feb 2018
3. Alkire, S., Seth, S.: Multidimensional poverty reduction in India between 1999 and 2006: Where and how? World Dev. **72**, 93–108 (2015)
4. Chauhan, R.K., Mohanty, S.K., Subramanian, S.V., Parida, J.K., Padhi, B.: Regional estimates of poverty and inequality in India, 1993–2012. Soc. Ind. Res. **127**(3), 1249–1296 (2016)
5. Sarkar, S., Padaria, R.N.: Farmers' awareness and risk perception about climate change in coastal ecosystem of West Bengal. Indian Res. J. Extension Educ. **10**(2), 32–38 (2016)
6. Viswanathan, B., Kumar, K.K.: Weather, agriculture and rural migration: evidence from state and district level migration in India. Environ. Dev. Econ. **20**(4), 469–492 (2015)
7. Ramesh, M.V., Mohan, R., Menon, S.: Live-in-labs: rapid translational research and implementation-based program for rural development in India. In: GHTC 2016—IEEE Global Humanitarian Technology Conference: Technology for the Benefit of Humanity, Conference Proceedings (2016)
8. Rajeshkumar, A., Balusamy, M.: The effect of different fish feeding methods on growth performance and fish yield in composite fish culture system (2017)
9. Ramesh, M.V., Nibi, K.V., Kurup, A., Mohan, R., Aiswarya, A., Arsha, A., Sarang, P.R.: Water quality monitoring and waste management using IoT. In: 2017 IEEE Global Humanitarian Technology Conference (GHTC), pp. 1–7. IEEE (2017)
10. Handbook on Fisheries and Aquaculture Technology. https://books.google.com/books. Accessed 5 Feb 2018
11. Wohlfarth, G.W., Schroeder, G.L.: Use of manure in fish farming—a review. Agric. Wastes **1**(4), 279–299 (1979)
12. Kalita, H.K., Talukdar, B., Baishya, R.A., Basumatary, S., Das, J., Dutta, A., Sarma, D.: Role of duck droppings on pond. Productivity through fish-cum-duck integrated farming system in agro climatic condition, Assam. Nat. Environ. Pollut. Technol. **14**(3), 541 (2015)
13. Kadiveti, H., et al.: Water management through integrated technologies, a sustainable approach for village Pandori, India. In: 2019 IEEE R10 Humanitarian Technology Conference (R10-HTC) (47129). IEEE (2019)
14. Nair, R., et al.: Community development and energy equality: experiences from micro hydro implementation in a tribal settlement in India. In: 2016 IEEE Region 10 Humanitarian Technology Conference (R10-HTC), Agra, India, pp. 1–7 (2016). http://doi.org/10.1109/R10-HTC.2016.7906794

Chapter 53
Design and Development of LDPE Plastic Bricks Through Triangulation Methodology

Harish T. Mohan, Renjith Mohan, Francesca Whitaker, Daniel Gaskell, and Gaspard Gindt

Abstract Annual consumption of plastic materials in India is predicted to surge from 12×10^6 T to 20×10^6 T by 2020, posing a severe environmental threat. Most of the plastics are either incinerated or discarded due to India's substandard waste management infrastructure. The lack of adequate sanitation facilities portents paucity of personal hygiene through open defecation in rural communities. There is an urgent need for the utilization of the available human resources to address these challenges. The paper includes case studies on waste management measures in Ranchi (the capital of the Indian state of Jharkhand) and Bhubaneswar (the capital of the Indian state of Odisha), along with sanitation challenges in the villages of Dewgain and Gupdapada. These case studies accentuate the need for a single solution for confronting waste management and sanitation infrastructure implementation. A potential remedy has been explored by developing interlocking plastic bricks composed of low-density polyethylene (LDPE) and raw construction materials. The structured system was proposed to incorporate state government initiatives in collaborative partnerships with the urban and rural sectors, to establish an open defecation free (ODF) hygienic environment.

53.1 Introduction

Sanitation is key to national development. Enormous amounts of money have been spent on improving access to sanitary facilities. Nevertheless, the world's economies

H. T. Mohan
Ammachi Labs, Amrita Vishwa Vidyapeetham, Amritapuri, India

R. Mohan (✉)
Amrita School for Sustainable Development, Amrita Vishwa Vidyapeetham, Amritapuri, India
e-mail: renjithmohan@am.amrita.edu

F. Whitaker · D. Gaskell
Materials Science and Engineering, The University of Manchester, Manchester, UK

G. Gindt
Industrial Engineering and Management INP Grenoble, Grenoble, France

T. Senjyu et al. (eds.), *IOT with Smart Systems*, Smart Innovation, Systems and Technologies 251, https://doi.org/10.1007/978-981-16-3945-6_54

have lost noticeably large sums of money due to deficient or ineffective sanitary mechanisms such as improper waste management, open defecation, and unhygienic living standards. As the world's sixth-largest economy, India has to bear its share of the unenviable socio-economic debacle. A nation's vulnerable rural and urban population hamstrings its achievement of sustainable survivability. This paper discusses plastic waste deployment to developing open defecation free (ODF) rural Indian communities.

As urbanization and consumption maintain an unconstrained growth trajectory, solid waste management becomes a significant matter of the first order. In 2012, the total waste generation by cities worldwide reached 1.2×10^9 T per year, and this rate is expected to reach 2.2×10^9 T by 2025 [1]. Developed nations, which contribute 44 face the momentous task to mitigate the fallout of unsatisfactory waste management. Mixed outcomes were observed in sanitation projects due to meager financial allowance and inadequacies in the technical, institutional, and social sectors [1]. Thus, the vulnerability of the developing nations is a matter of even more significant concern, where social, economic, and physical infrastructure lack the cost-effective guidelines set by the developed nations.

An article in New York Times reported [2] that as of April 2017, India is the world's sixth-largest economy per the IMF's World Economic Outlook Database, and the second-most populous country in the world. India is experiencing rapid economic growth; however, as the economy grows, India's impact on the environment regarding waste and greenhouse gases. In 2011, the rate of waste generation in urban India was estimated at 47.3×10^6 T per year. In 2021, it should be about 71.15×10^6 T per year; considering the spiraling urbanization growth in India, this rate could reach 160.96×10^6 T per year by 2041 [3]. According to the Central Pollution Control Board, the average waste collection ranges from 50 to 90% of the production, but 94% of this collected waste is disposed of outdoors, in landfill or unrevealing areas [3]. The lack of segregation of waste collected always increases the risk of failure of the mitigation plan. The rate of urban migration population adds to the challenge of developing a waste management model [4].

Like most countries in the world, urbanization is an increasing phenomenon in India. In 2001, the Indian urban population was $286{,}110 \times 10^6$, representing 27.81% of the population. Ten years later, this number reached $377{,}160 \times 10^6$. Alternatively, 32.16% of India's population accounts for half of the population by 2050 [5]. This rapidly changing magnitude of urban waste generation disturbs the nation's development plans and policies. Although programs like the government developed the MAHATMA GANDHI NATIONAL RURAL EMPLOYMENT GUARANTEE ACT to control urban migration; the livelihood generation remains a challenge for rural development. The role and impact of sanitation in rural and urban development are still a matter of discussion.

Open defecation is a crucial challenge for rural sanitation. Barely, 16% [6] of rural households have access to basic sanitation facilities, and over 7×10^9 children die each year across the globe due to diarrhea and other sanitation-related diseases. Open defecation persists in the rural communities, which ensues in a lack of personal hygiene and contaminated water supplies. As per the $\$4 \times 10^6$ campaign launched by the Government of India, over 50×10^6 toilets are anticipated to be built by 2019 [2]. The use of toilets to store cattle food, farming equipment, and tools are show spoilers. Since 2014, the community-driven sanitation model's introduction by building toilets through women groups has successfully mitigated this challenge; a community-driven model aimed at building toilets in Dewgain village has shown remarkable results. A group of women, employed by a collaboration of the Government and Amrita University, was launched to build toilets in the Amrita Serve adopted village using molding techniques.

Albeit successful, these community-driven models have not scaled satisfactorily. The concerned communities have not participated enough. Although awareness was generated across the nation through the Swatch Bharat scheme, it has not touched its daily chores. The majority of the Swatch Bharat scheme stake-holders are still government agents or NGO members, not the community of interest. Lack of livelihood opportunities is a deterrent to community inclusion and sustained participation. Community-driven models should devise strategies to widen the stakeholder base. There should be a change in roles for these communities from laborers to skilled workers.

We investigated three significant challenges for sanitation: reducing the amount of plastic waste produced by urban populations, improving rural communities' sanitation facilities, and trying to bring all three on the same page by proposing a single solution. The study has been carried out as a part of the Live-in-Labs® program [7–9]. A significant gap in understanding the cross-linkages, while addressing the challenges through mitigation. Amrita Vishwa Vidyapeetham developed a waste management model based on the LDPE-based plastic waste to take on the challenges mentioned above. For the design and development of the LDPE plastic brick-based composite, we used the triangulation process [10–12]. This year-long study was limited to two rural communities, two nearby urban communities, and the stakeholder base on the two rural communities. Data triangulation approach, deployed in the study, involved students' perception of social presence, method consist of questionnaires, interviews, followed by class observations, online discussion, and lab observation [13].

53.2 Methodology

Multiple triangulations were done to ensure the design and development of plastic bricks would have a positive impact on government efforts to mitigate the waste management challenges for both urban and rural sectors. This paper investigated the possibility of improving the stakeholder participation of urban and rural sector small groups, while interviewed, were unaware of their exact ages, so an educated estimate was made. Furthermore, ten percentage of the villagers interviewed was completely unsure of their age. This project is aimed at both men and women of working age, 18–60 years. The intention was to involve women, as much as possible, to encourage them to seek careers of their own and not solely rely on men to support themselves and their families. Older community members were also interviewed to understand waste,

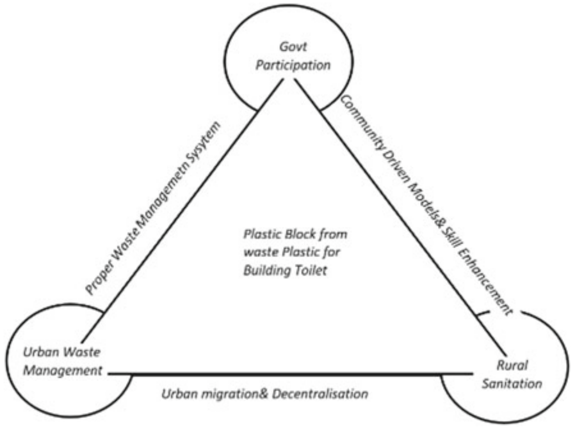


Fig. 53.1 Triangulation methodology

waste management, and sanitation in the village. Selection of urban community and associated rural community in the vicinity was of utmost priority in understanding the challenge (Fig. 53.1).

In Jharkhand, the number of towns had grown from 119 in 2001 to 152 in 2011. Urbanization in the state is 24.05%, versus 22.24%, ten years ago [14]. Compared to India's national rate, Jharkhand's urbanization stands behind right now but tends to increase in the coming future. Although Ranchi district is not the most urbanized, and its rate growth is the highest in the state [14]. In 2012, Ranchi population was 1090 $\times$ 10^6 and estimated to reach 1340 $\times$ 10^6 [15]. Rapid urbanization and growth in Bhubaneswar city's population, which increased overall waste generation in recent years. The solid waste management system's performance was far from required due to outdated technology and inefficient methods. Other barriers like a shortage of financial and human resources, lack of modern technology. Analysis of the physical waste showed, 100 g of plastic out of every 1000 g of waste, which gives the plastic waste composition 10 of the total waste generated. The village of Dewgain is located in the state of Jharkhand, 22 km from the city of Ranchi, in East India. The village's living population comprised around one thousand people, but this number could not be verified. The village is composed of three or four sectors, mainly depending on religious and family gatherings. It is surrounded by wheat and rice fields, where most of the villagers work during the day. The other location of this study was the rural village of Gupdapada, located on the eastern coast of India, in the state of Odisha. The village is approximately 30 km from the state capital, making the plastic waste from the capital easily accessible. Gupdapada is surrounded by thick forestry, and villagers experience many dangerous encounters with wildlife such as elephants and snakes. The village is now surrounded by a large ditch, protecting them from the elephants; however, snakes still prevail.

A survey was carried out, targeting both men and women of working age, 18–60 years. Particular emphasis was devoted to women in this study, which included equal gender distribution of different ages. This helped gain a broad understanding of waste production, waste management, and sanitation in the village, reflecting their employment conditions.

Educational demographics were also surveyed, as education is vital for income generation and ensuring better sanitation practices.

In order to get a better understanding of the user groups to aid in effective analysis, personas were created based on the data obtained from the interviews. Though fictitious, creating personas would help in understanding the user demography more broadly, as it is a representation of a collective user group with shared goals. Scenarios were then made, to deeply understand the pain points, enabling the researchers to focus on the magnitude of the pain endured by the users.

Since the data obtained from the study was more qualitative in nature, data analysis was done mainly through thematic analysis. Here, the analysis was focused around the themes—goals, tasks, and pain points.

53.3 Results

In Dewgain, the villagers, older than 35 years who had not studied past the 10th grade, were interviewed separately. All those who had reached the 12th grade of education were under 25. Except for two villagers who were occupied as drivers, and all others were agricultural farmers (Fig. 53.2).

Waste management in Ranchi, one of the main cities in central India has been carried out at five waste collection centers. Each center collects around 40 truckloads per day by going door to door. The collected waste is weighed and transferred to a compactor; the compacted waste is transported to landfilling sites like Jhiri, located 25 km away. The center reported that around of waste is dumped into landfill sites every day. This approach to waste disposal has raised serious concern among the nearby residents for fear of contamination of their water resources.

Fig. 53.2 Compactor at one of the waste collection unit

Triggered by the Jhiri field visit experience, research moved toward plastic waste, challenging to handle. LDPE plastic materials are one of the significant components in the generated waste. This type of plastic is not easily recyclable. It has a recyclability number of 4, as per the United States Government's Environmental Protection Agency (EPA).

Construction blocks from the LDPE plastic block research were carried out to overcome the demand for construction materials.

The Dewgain and Gupdapada villages have 17 toilets, which equates to one toilet per every 18 persons. It would be a significant improvement to improve this ratio to 5 people per toilet. Among the survey respondents, 28% answered positively to the sanitation issues. There is a lack of infrastructure; however, the villagers want the situation to improve (Fig. 53.3).

Villagers without access to a personal/family sanitation structure were surveyed to understand their experiences with open defecation. Villagers especially experienced struggles when going to defecation sites. Some of the fascinating responses came from the female group. They explained their fear when commuting to open defecations sites, especially at night. They felt threatened due to the dangerous animals such as elephants and snakes and inebriated men—as alcoholism is prevalent in many rural communities, including this village.

A survey showed that barely 46% of the Gupdapada villagers are employed in the nearby private organizations as daily waged workers. The remaining 54% are unemployed.

Bhubaneswar Municipal Corporation is the government organization responsible for the solid waste management of Bhubaneswar city. The waste is collected in Bhubaneswar and across 1500 locations. 600–700 T of waste are being collected and dumped. Plastic constitutes 30% of the waste being collected. LDPE plastic is being used as a landfill for the local industries. A local official confirmed that LDPE plastic waste is causing significant landfills due to its low-recyclability properties.

The rag pickers employed at the waste management center can earn an income of INR. 200 to INR 400, through the sales of glass and thick plastic. They are not interested in segregating the waste collected since they can get significant monetary benefit by selling it (Fig. 53.4).

Fig. 53.3 Apprentice availability chart

Fig. 53.4 Interlocking blocks made from LDPE sand

53.4 Discussion

The need for income generation opportunity has been confirmed by discussing in the village with unemployed women, who have shown great interest in participating. Further discussion with a field officer from a waste management center and field observations suggests that LDPE plastics being dumped for the land-fill every day can be easily procured and reused for some other purpose. So the team decided to use LDPE as the primary material for making bricks. A specific design was developed and tested.

The double-cross interlocking brick design has 15 cm × 30 cm × 10 cm as width, length, and height, respectively. Two circular holes of approximately 30-mm diameter go directly through each brick cross's center to allow for rebar to be fitted into a walled structure. Thus, the design gives the structure to a wall in the brick bond pattern, with rebar running through the hole, which adds structural strength. The compression strength of the construction block is around 14 N/mm^2, which exceeds the minimum standard requirement of 10 N/mm^2.

In the subsequent triangulation, we study the use of plastic bricks for sanitation in a rural environment. There are some key factors to be assessed for this process. The first one is to see if there is any interest in the villagers for the sanitation process, and the next is the number of resources required. The resources and cost of procurement have been assessed from visits to the urban waste management facility in Bhuasuni. Using interlocking blocks to build a structure required a semi-skill level of human resources. The minimum cost for LDPE plastic is another factor that needs to be considered. Like sand, the other materials are readily available throughout the country, which is an advantage when deploying it in a rural community.

53.5 Conclusion

The use of triangulation methodology based on the stakeholder base expansion improves the design and development process of LDPE-based plastic bricks. Thus, the use of LDPE-based plastic brick as a building resource for the unit is being studied. The study benefited the development of LDPE-based plastic bricks to curb the urban and rural sanitation challenges. The triangulation of the urban and rural sanitation challenges is being analyzed through non-statistical methods. Further statically, studies to be done quantify the impact of triangulation in the design process. Also, the brick being developed is to be tested for technical feasibility as the triangulation does not cover this topic.

Acknowledgements This research was funded by Amrita Vishwa Vidyapeetham's Live-in-Labs® program and the field implementation was funded by the UN-recognized NGO, MA Math. The study was conducted as part of the UNESCO Chair for Experiential Learning for Sustainable Innovation and Development. The authors express their immense gratitude to Sri. Mata Amritanandamayi Devi, Chancellor of Amrita Vishwa Vidyapeetham, who has inspired them in performing selfless service to society. The authors would also like to express their sincere appreciation to Mr. Gopalaiengar Sheshadri for his extensive support in enhancing the quality of the paper within the limited time frame. The authors also thank the faculty, staff, and students of Amrita Vishwa Vidyapeetham, the government authorities of Kangra district, Himachal Pradesh, and also the Amrita SeRVe program staff.

References

1. Shawn Burke, C., Salas, E., Smith-Jentsch, K., Rosen, M.A.: Measuring macrocognition in teams: some insights for navigating the complexities. In: Macrocognition Metrics Scenarios, pp. 29–44 (2018)
2. Padhi, B.K., et al.: Risk of adverse pregnancy outcomes among women practicing poor sanitation in rural India: a population-based prospective cohort study. PLOS Med. **12**(7), e1001851 (2015)
3. Annepu, R.K.: Sustainable solid waste management in India, MS dissertation, pp. 1–189 (2012)
4. Asnani, P.U., Zurbrugg, C.: Improving Municipal Solid Waste Management in India. The World Bank (2007)
5. IOM, International Organization for Migration (IOM): World Migration Report. Migrants and Cities: New Partnerships to Manage Mobility (2015)
6. The World Bank: Sanitation Overview. [Online]. Available: http://www.worldbank.org/en/topic/sanitation3. Accessed: 17-Mar-2019
7. Office of the Registrar General Census Commissioner, Census of India Website: Office of the Registrar General and Census Commissioner, India (2011). [Online]. Available: http://www.censusindia.gov.in/2011census/population_enumeration.html. Accessed: 17-Mar-2019
8. Varma, D.S., Nandanan, K., Vishakh Raja, P.C., Soundharajan, B., Pérez, M.L., Sidharth, K.A., Ramesh, M.V.: Participatory design approach to address water crisis in the village of Karkatta, Jharkhand, India. Technol. Forecast. Soc. Change **172**(2021), 121002 (2021)
9. Goris, L.M., Harish, M.T., Bhavani, R.R.: A system design for solid waste management: a case study of an implementation in Kerala. In: 2017 IEEE Region 10 Symposium (TENSYMP), pp. 1–5. IEEE (2017)

10. Mohan, H.T., et al.: Transforming urban waste into construction blocks for a sanitation infrastructure: a step towards addressing rural open defecation. In: 2017 IEEE Global Humanitarian Technology Conference (GHTC), pp. 1–9 (2017)
11. Bhavani, R.R., Sheshadri, S., Maciuika, L.A.: Addressing the first teachers: education and sustainable development for children, families and communities through vocational education, technology and life skills training for women. In: Children and Sustainable Development, pp. 319–334. Springer, Cham (2017)
12. Balu, S., Sampath, K.S., Hari, L.R., Ramprasath, M., Aishwarya, N.K., Aiswarya, M.S., Frey, L.M., Nandanan, K.: Assessment of waste management practices and its impact on the village ecosystem in Himachal Pradesh, India. In: 2020 IEEE 8th R10 Humanitarian Technology Conference (R10-HTC), pp. 1–6. IEEE (2020)
13. Tu, C.-H., McIsaac, M.: The relationship of social presence and interaction in online classes. Am. J. Distance Educ. **16**(3), 131–150 (2002)
14. Kumar, N.: Urban continuum in Jharkhand: beyond the state formation. Int. J. Res. Appl. Nat. Soc. Sci. **3**(1), 57–66 (2015)
15. Kadiveti, H., Eleshwaram, S., Mohan, R., Ariprasath, S., Nandanan, K., Divya Sharma, S.G., Siddharth, B.: Water management through integrated technologies, a sustainable approach for village Pandori, India. In: 2019 IEEE R10 Humanitarian Technology Conference (R10-HTC) (47129), pp. 180–185. IEEE (2019)
16. Ogawa, H.: Sustainable solid waste management in developing countries. In: 7th ISWA International Congress and Exhibition (1996) [Online]. Available: https://www.gdrc.org/uem/waste/swm-fogawa1.htm. Accessed: 17 Mar 2019
17. Ramesh, M.V., Mohan, R., Menon, S.: Live-in-labs: rapid translational research and implementation-based program for rural development in India. In: GHTC 2016—IEEE Global Humanitarian Technology Conference: Technology for the Benefit of Humanity, Conference Proceedings (2016)

Chapter 54
A Review of Algorithms for Mental Stress Analysis Using EEG Signal

Sherly Maria, J. Chandra, Bonny Banerjee, and Madhavi Rangaswamy

Abstract Mental stress is an enduring problem in human life. The level of stress increases exponentially with an increase in the complexity of work life. Hence, it is imperative to understand the causes of stress, a prerequisite of which is the ability to determine the level of stress. Electroencephalography (EEG) has been the most widely used signal for understanding stress levels. However, EEG signal is useful only when appropriate algorithms can be used to extract the properties relevant to stress analysis. This paper reviews algorithms for preprocessing, feature extraction and learning, and classification of EEG, and reports on their advantages and disadvantages for stress analysis. This review will help researchers to choose the most effective pipeline of algorithms for stress analysis using EEG signals.

54.1 Introduction

Stress is the level at which a person begins to feel overwhelmed and is unable to function efficiently. Hans Selye came up with a theory called General Adaptation Syndrome (GAS). He classified stress into three stages called alarm, resistance, and exhaustion based on various factors that contributed to it. These stages could have neurological responses which could be measured using the electroencephalogram

S. Maria (✉) · J. Chandra
Department of Computer Science, CHRIST (Deemed to be University), Hosur Road, Bangalore, India
e-mail: sherly.maria@res.christuniversity.in

J. Chandra
e-mail: chandra.j@christuniversity.in

B. Banerjee
Institute for Intelligent Systems, and Department of Electrical and Computer Engineering, The University of Memphis, Memphis, TN, USA
e-mail: bbnerjee@memphis.edu

M. Rangaswamy
Department of Psychology, CHRIST (Deemed to be University), Hosur Road, Bangalore, India
e-mail: madhavi.rangaswamy@christuniversity.in

T. Senjyu et al. (eds.), *IOT with Smart Systems*, Smart Innovation, Systems and Technologies 251, https://doi.org/10.1007/978-981-16-3945-6_55

(EEG). EEG is a widely used medical technique to analyze brain signals, and it is one of the most trusted modalities used when it comes to the field of psychology, medicine, and computer science. Several new instruments can acquire EEG data that is being developed for ease of access, portability, and affordability. The traditional EEG machines are stationary and lab-based and very expensive. Inventors have come up with alternatives and built EEG machines that are smaller in size, easy to carry around, and affordable [1]. EEG has been preferred over other methods of analyzing brain signals, such as Positron Emission Tomography (PET) and functional magnetic resonance imaging (fMRI), because of its excellent temporal resolution and non-invasive properties. EEG has been utilized for predicting clinical outcomes in conditions such as epilepsy, autism, and Alzheimer's disease among others [2–6].

54.1.1 Stress Detection Devices

There have been various methods used to detect stress using psychological, physiological, and behavioral signals. A few of the most commonly used stress analysis signals include questionnaires, ECG and GSR, EEG, keyboard and screen time, and wearable sensor signals. Questionnaires are designed to analyze the different conditions of the mind. The set of questions have a couple of choices for answers and the participant chooses the answer that suits him/her the best. In most cases, each answer would have a particular weight attached to it. Based on the total weight, the level of stress is identified. Most researchers developed questionnaires using internationally established guidelines. Ahuja and Banga [7] used a Holmes-Rahe scale, and their main aim was to categorize the participants based on the source of origin of stress and the severity of the stress. Pozos-Radilloa et al. [8] made use of the PSS test that included 14 questions. Each answer has a weight assigned and based on the total weight, and the students were categorized into various categories like stressed or normal. Suresh et al. [9] have used benchmark questionnaires like the stress symptom inventory that was developed by Lipp and Guevara which contains a list of 42 psychophysiological symptoms that characterize chronic stress. Mohammad et al. [10] have used questionnaires for participants with diseases such as bipolar disorder as well. The researchers made use of an android device to record the answers for the questionnaire. Abdulhamit and Ismail [11] have used the Stroop color test, while the questionnaires were being answered. Electrocardiograph (ECG) and the galvanic skin response (GSR) techniques are popular stress detecting methods. The symptoms detected for the ECG would be the discrepancies in the electrical activity of the heart, whereas in the GSR, the different conductance levels of the skin would be tested. If a person is stressed, the nervous system would release sweat in response [12]. After the desired features are extracted from the ECG and GSR, they are typically fed to the system as inputs to classifiers to be further classified into stressed/non-stressed categories. Panagiotis and Leontios [13] put the participants in certain stressful situations, and then assess their performance and their reactions using the GSR and ECG machine. Rajdeep and Tathagata [14] use non-invasive methods for extracting different features in the time domain using GSR and ECG. Raja and Hyo [15] put the participant in a certain conditional situation to assess stress. The authors in this paper collected signals from automobile drivers who were made to drive on different types of roads. Many researchers [16, 17] have used smartphones and wearable sensors to detect stress. The wearable sensors would typically be skin conductance devices or accelerometers. Mobile sensors would typically be based on the usage of the mobile and are associated with the stress level. EEG is known to be one of the most reliable sources that help detect brain activity. The fluctuations in the voltage are associated with the neurons in the brain. EEG is used to identify various brain disorders like epilepsy, sleep disorders, coma, Parkinson's disease to name a few. EEG records milliseconds of data and for long durations, even during sleep activity and is mostly used because of its reliable and non-invasive properties [18–21].

54.1.2 EEG Signal Classification

Spectral analysis generates the most well-known classification of EEG signals into physiologically significant frequency bands. Typically, most studies focus on five spectral bands in the EEG—α (8–12 Hz), β (13–30 Hz), ϒ (30 Hz and above), Δ (0.5–4 Hz) and Θ (4–8 Hz) [22]. Studies have discussed the significance of these spectral bands in the context of neural activities associated with cognition [23]. Delta (∂) waves (0.5–4 Hz) are usually observed in individuals that are in deep sleep. It is not produced in adults who are awake. If present in awake state, it points toward cerebral dysfunction. However, in infants, it is normal for delta waves to be present during the awake state. Delta wave has the lowest frequency range and the highest amplitude. It is also the slowest wave and typically accompanies the release of growth hormone that helps in healing the body. Theta (Θ) waves (4–8 Hz) are associated with deep relaxation, meditation, and subconscious activity. It is not typically found in adults in their wake state but is normal for children below the age of 13. Theta waves are also associated with the production of human growth hormones and serotonin. These hormones help in relieving stress and memory recollection. Alpha (α) waves (8–12 Hz) are usually observed in humans during the awake state. It is observed in all age groups. Alpha waves are also considered the ideal situation for creative activities, deep reflection, and problem-solving capabilities. These waves are recorded from the occipital and parietal lobe of the brain. Beta (β) waves (13–30 Hz) are the most commonly observed waves in humans. Beta waves are associated with concentration and alertness. In this state, the brain is easily able to analyze the problem, process the information, and generate solutions. Beta waves are extremely advantageous for productivity. Gamma (ϒ) waves (>30 Hz) (30–100 Hz) are linked to wakefulness and consciousness. They are also often associated with extreme alert activity and active sensory inputs. Gamma waves usually combine senses to promote learning.

54.2 BioSignal Classification Using EEG

Figure 54.1 depicts the workflow of the brain signal classification technique. An EEG machine typically consists of a cap-like structure that has electrodes attached to it. The electrodes conduct the EEG signals to a preamplifier, and this is fed into the EEG acquisition system and can be viewed on a screen. The electrodes are attached to the scalp in a specific montage based on the 10-20 system of electrode location and nomenclature. The electrodes are typically dipped in saltwater or conductive gel for a while and then carefully placed on the head of the participant. Once all the electrodes are in place, the impedance of each electrode is checked by the system and the experimenter has to ensure a value that is less than 5k ohms. The participants would be asked to remain relaxed and to sit in a comfortable position that promotes minimal movement since any slight movement would be picked up as noise by the EEG. Video cameras may be used to capture facial expressions during the experiment. During data collection, there could be a lot of noise that can be recorded along with the signal. Since the EEG is an extremely sensitive device, there could be a lot of reasons why noise is generated due to electrodes not being able to conduct sometimes or impedance, physiological artifacts that are caused due to other parts of the body, etc. These artifacts have to be removed as it renders the data useless. Noise removal could be done using specific filters. There are many methods of artifact correction based on the type of artifacts like bandpass filtering, low-pass filter, and high-pass filter [24]. The second phase is the post-processing that includes feature extraction and classification. The most common techniques used for feature extraction include Fourier Transformation, Wave Transformation, Independent component analysis, and Principal Component Analysis [21, 26–28]. Feature selection/extraction is a very important step as it saves time and reduces computational cost and decreases complexity. Classification is done after feature extraction. The most common classification techniques are Artificial Neural Network (ANN), Support Vector Machine (SVM), K Nearest Neighbor (KNN), and Naïve Bayes. It is recommended to have the largest size of training data as compared to the feature vectors to get better results. In the field of research for EEG Data, Linear discriminant Analysis (LDA), Support Vector Machine (SVM), Naïve Bayes, Multi-Layer Perceptron (MLP) have been the most prominent classifiers. Apart from them, quite recently researchers have been coming up with novel classification techniques like the hybrid method. LDA has been used to separate data that is represented by different classes [29, 30]. In some cases [31], LDA has been used to create a new predictor variable that is observed to be the combination of the original predictors. Their main aim was to come up with a discriminant score that is a combination of predictor scores to form

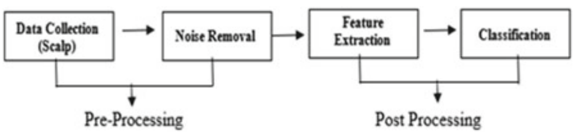


Fig. 54.1 The workflow of brain signal classification

a single-composite variable. They consider the degree of overlap between the scores to be a measure of the success of this technique. The LDA is useful for multiclass domain and binary classification. LDA also has a really good assumption value of multivariate normality, and all the matrices are observed to have equal covariance matrices. SVM has been widely used to classify EEG data. It is observed to have the most accuracy rate as compared to other classifiers [32, 33]. SVM is most commonly used as a multiclass classifier. In some cases, the various classifiers were compared to each other depending on various factors and research goals, in most cases, SVM seems superior [34] but in some cases, other classification algorithms seemed superior [35] and in some cases, the models are trained using multiple classifiers [36]. ANN has been used widely in detecting diseases like Parkinson's, Huntington's with the use of an EEG [32, 37]. The data was divided into sets to train the ANN and were validated later. In many other works [38], the EEG data is recorded on random patients to test ANN practically. ANN has also been used to study the depth of anesthesia in patients [29], which has been very useful in understanding the same and is proven to make things easier or feasible for monitoring patients. Plenty of research related to EEG signal classification has been done by making use of the MLP property of generalization [39]. MLP has been used for classification to detect epilepsy as a diagnostic decision support mechanism. When it comes to smaller training set samples, MLP is quite useful. It is also one of the most common feed-forward neural networks because it is very easy to implement. MLP is also used for emotion detection in targeting various emotions like joy, anger, sadness, etc. It was found that the accuracy of using MLP is way more than the other classification techniques [40]. The main aim of combining classifiers is to achieve a better classification that surpasses all the limitations of the existing classifier. One of the main reasons that the hybrid classifier would work is because of the diversity of various classifiers which would make up for the errors caused by the previous classifiers [41]. Although however, using highly diverse classifiers do not guarantee accuracy either. There are different ways to combine classifiers. The ensemble learning method is to combine various classifiers after training them. In a few cases, researchers make use of ensemble learning by missing various classifiers made from using the same classification algorithm. This does not necessarily become a hybrid classifier. In most cases, hybrid ensembles work better than non-hybrid ones. An ensemble method helps in decreasing variance, bias, and in improving predictions. In boosting, the weakest model is chosen and is run through a couple of iterations. In this method, the model tries to build a strong classifier by combining all the weak classifiers. In simple terms, a model is created with the help of training data. This model would certainly have some errors. A second model is then created, trying to fix the errors of the first model. Likewise, the second model would have errors too. This cycle repeats until the model can classify the training set perfectly or the number of models is exhausted. Figure 54.2 describes the workflow of brain signal classification and how the perfect model is built by learning the mistakes of the previous model and avoiding them. The reason the iteration is repeated is that each dataset has rules to classify. If a model misses one rule, it would not be able to classify it. All the rules together would help classify

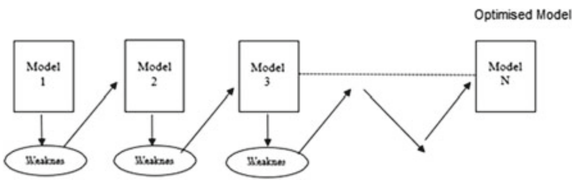


Fig. 54.2 The workflow of brain signal classification

the data. To ensure the accuracy of the predictors, the outcome of each of the weak learners is used.

Hybrid models combine the strengths of various classification models. They improve prediction accuracy and have applications in various fields like e-commerce, medical fields, and finance. Based on the use, different hybrid models could be developed.

54.3 Result Analysis

There have been significant advancements in the field of neurocognitive science and machine learning. EEG has been effectively used in analyzing how the brain works. From the above literature review, it may be observed that SVM and hybrid models are the most widely-used algorithms. Hybrid models allow the best features of many algorithms to be exploited as per the need, thereby enhancing accuracy and efficiency of classification. It is observed that the hybrid classifier has the highest accuracy, at around 96.69%, followed by SVM with an accuracy of 95.72%. NB and ANN are the next best performing algorithms with accuracy of 82.86% and 78.3%, respectively. This is followed by LDA, MLP, and KNN yielding 74.5%, 65.08%, and 61.03%, respectively, with KNN being the lowest. Therefore, it an be concluded that a hybrid classifier is most likely to yield the best result, followed by SVM with around a 2% accuracy difference. It is note worthy that SVM is the most widely-used classifier, followed by a hybrid classifier of various kinds [42].

54.4 Conclusion

Inability to cope with mental stress is an ongoing issue all around the world. It is the need of the hour to automate the stress analysis process which requires a pipeline or model consisting of algorithms that exploits their strengths. This review concludes that a hybrid model is superior than other popular machine learning algorithms based on the accuracy of stress classification and is also cost-effective and efficient.

References

1. Sateesh, K., Bhuvaneswari, P.: Analysis of electroencephalography (EEG) signals and its categorization—a study. Procedia Eng. **38**(2012), 2525–2536 (2012)
2. Ames, F.R.: Self induction in photosensitive epilepsy. Brain **94**(4), 781–798 (1983)
3. Shanbao, T., Nitish, V.: Quantitative EEG Analysis Methods and Clinical Applications. ISBN: 9781596932043 (2009)
4. Ali, S.: Human emotion detection using brainwave signal: a challenging. Eur. J. Sci. Res. **44**(4), 640–659 (2010)
5. David, L.: Applicability of brainwave biofeedback to substance use disorder in adolescents. Child Adolesc. Psychiatr. Clin. **14**, 125–136 (2005)
6. Classen, J., Lawrence, H., Stephan, A.: Treatment of status epilepticus: a survey of neurologists. J. Neurol. Sci. **211**, 37–41 (2003)
7. Ahujaa, R., Banga, A.: Mental stress detection in university students using machine learning algorithms. In: International Conference on Pervasive Computing Advances and Applications—PerCAA 2019
8. Pozos-Radilloa, B.E., de Lourdes Preciado-Serranoa, M., Acosta-Fernándeza, M., de los Ángeles Aguilera-Velascoa, M., Delgado-Garcíab, D.D.: Academic stress as a predictor of chronic stress in university students. Psicología Educativa **20**, 47–52 (2014)
9. Suresh, B., Aravind, Joseph, P.: Linear discriminant analysis for signal processing problems. In: Proceedings of IEEE (1999)
10. Mohammad, R., Muhammad, I., Motakabber, S.M.A., Shahjahan, S.: Classification of Multi-channel EEG Signal by Linear Discriminant Analysis. Springer International Publishing, Switzerland (2015)
11. Abdulhamit, S., Ismail, M.: EEG signal classification using PCA, ICA, LDA and support vector machines. Expert Syst. Appl. **37**, 8659–8666 (2010)
12. Fares, A., Tong, B., Nasreen, B., Masashi, K.: Towards multilevel mental stress assessment using SVM with ECOC: an EEG approach. Med. Biol. Eng. Comput. **56**(1), 125–136 (2017). https://doi.org/10.1007/s11517-017-1733-8. Epub 2017 Oct 18
13. Panagiotis, S., Leontios, J.: Emotion recognition from EEG using higher order crossings. IEEE Trans. Inf. Technol. Biomed. **14**(2) (2010)
14. Rajdeep, C., Tathagata, B.: EEG based motor imagery classification using SVM and MLP. 2375-5822/16 IEEE (2016)
15. Raja, M., Hyo, J.: Emotion classification of EEG brain signal using SVM and KNN. In: IEEE International Conference on Multimedia & Expo Workshops (ICMEW) (2015)
16. Nazari, K., Asadi, G., Gharaviri, A., Teshnehlab, M., Aliyari, M.: Classification of alcoholics and non-alcoholics via EEG using SVM and neural networks. In: 3rd International Conference on Bioinformatics and Biomedical Engineering (2009)
17. Rekha, G.: Neural networks in signal processing. In: Studies in Fuzziness and Soft Computing Book Series (STUDFUZZ), vol. 38 (2000)
18. Rodrigues, P.M., Teixeira, J.P.: Classification of electroencephalogram signals using artificial neural networks. In: 3rd International Conference on Biomedical Engineering and Informatics. 978-1-4244-6498-2/10 (2010)
19. Nikola, M., Aleksandar, M., Natasa, J.: Artificial neural network-based approach to EEG signal simulation. Int. J. Neural Syst. **22**(3), 1250008 (2012)
20. Yue, G., Zhenhu, L., Satoshi, H.: Use of multiple EEG features and artificial neural network to monitor the depth of anesthesia. Sensors **19**, 2499 (2019)
21. Mishra, B., Mehta, S.C., Sinha, N.D., Shukla, S.K., Ahmed, N., Kawatra, A.: Evaluation of work place stress in health university workers: a study from rural India. Indian J. Commun. Med. **36**(1) (2011)
22. Lopes da Silva, F.H.: Neurocognitive processes and the EEG/MEG. In: Schomer, D.L., da Silva, F.L. (eds.) Electroencephalography: Basic Principles, Clinical Applications, and Related Fields, 6th ed. Lippincott Williams & Wilkins (LWW) (2011)

23. Wendling, F., Congendo, M., Lopes da Silva, F.H.: EEG analysis: theory and practice. In: Schomer, D.L., Lopes da Silva, F.H. (eds.) Niedermeyer's Electroencephalography: Basic Principles, Clinical Applications, and Related Fields. Oxford University Press, Oxford, UK (2017)
24. Gupta, R., Afshar Alam, M., Agarwal, P.: Modified support vector machine for detecting stress level using EEG signals. Comput. Intell. Neurosci. (2020). http://doi.org/10.1155/2020/8860841
25. Subasi, A., Gursoy, M.I.: EEG signal classification using PCA, ICA, LDA and support vector machines. http://doi.org/10.1016/j.eswa.2010.06.065
26. Lee, H., Choi, S.: PCA+HMM+SVM for EEG pattern classification. 0-7803-7946-2/03/$17.00 02003 IEEE
27. Upadhyay, P.K., Nagpal, C.: PCA-aided FCM identifies stressful events of sleep EEG under hot environment. Journal homepage: https://www.tandfonline.com/loi/tijr20
28. Barua, S., Begum, S.: A review on machine learning algorithms in handling EEG artifacts. In: The Swedish AI Society (SAIS) Workshop SAIS, 14, 22–23 May 2014, Stockholm, Sweden
29. Alfred Lim, C.-K., Chia, W.C.: Analysis of single-electrode EEG rhythms using MATLAB to elicit correlation with cognitive stress. Int. J. Comput. Theor. Eng. **7**(2) (2015)
30. Srirampraskash, S., Prasanna, V.D., Murthy, O.R.: Stress detection in working people. In: 7th International Conference on Advances in Computing & Communications, ICACC 2017, Aug 22–24, 2017, Cochin, India
31. Lu, H., Rabbi, M., Chittaranjan, G.T., Frauendorfer, D., Mast, M.S., Campbell, A.T., Gatica-Perez, D., Choudhury, T.: Stress sense: detecting stress in unconstrained acoustic environments using smartphones. In: UbiComp '12, Sept 5–Sept 8, 2012, Pittsburgh, USA
32. Arsalan, A., Majid, M., Anwar, S.M., Bagci, U.: Classification of Perceived Human Stress Using Physiological Signals. arXiv:1905.06384v1 [eess.SP], 13 May 2019
33. Dayalin, S.J., Anuja, H.S.: Detection of automobile drivers stress from physiological signals. IJCSMC **3**(4), 454–458 (2014)
34. Sano, A., Picard, R.W.: Stress recognition using wearable sensors and mobile phones. In: 2013 Humaine Association Conference on Affective Computing and Intelligent Interaction
35. Maxhuni, A., Muñoz-Meléndez, A., Osmani, V., Perez, H., Mayora, O., Morales, E.F.: Classification of bipolar disorder episodes based on analysis of voice and motor activity of patients. 1574-1192/© 2016 Published by Elsevier B.V.
36. Sulaiman, N., Taib, M.N., Lias, S., Murat, Z.H., Aris, S.A.M., Hamid, N.H.A.: Novel Methods for Stress Features Identification using EEG Signals. ISSN: 1473-804x 27 online, 1473-8031 print
37. Halim, Z., Rehan, M.: On identification of driving-induced stress using electroencephalogram signals: a framework based on wearable safety-critical scheme and machine learning. http://doi.org/10.1016/j.inffus.2019.06.006
38. Tsoi, A., So, D., Sergejew, A.: Advances in Neural Information Processing Systems 6 (NIPS 1993) (1983)
39. Umut, O., Mahmut, H., Mahmut, O.: EEG signals classification using the K-means clustering and a multilayer perceptron neural network model. Expert Syst. Appl. **38**(2011), 13475–13481 (2011)
40. Kuo, W.: A theoretical analysis of why hybrid ensembles work. Comput. Intell. Neurosci. **2017**(Article ID 1930702) (2017). https://doi.org/10.1155/2017/1930702
41. Rahman, Md.A., Hossain, Md.F., Hossain, M., Ahmmed, R.: Employing PCA and t-statistical approach for feature extraction and classification of emotion from multichannel EEG signal. http://doi.org/10.1016/j.eij.2019.10.002
42. Priya, T.H., Mahalakshmi, P., Naidu, V., Srinivas, M.: Stress detection from EEG using power ratio. In: 2020 International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE), Vellore, India, pp. 1–6 (2020). http://doi.org/10.1109/ic-ETITE47903.2020.401

Chapter 55
Case Study on Water Management Through Sustainable Smart Irrigation

P. B. Abhinaya, T. Adarsh, Prasanthi Vanga, S. Sivanesh, Yisanaka Vishnuvardhan, N. Radhika, and A. S. Reshma

Abstract In India, more than 70% of people in rural areas depend on agriculture and thus depend on large quantities of water for irrigation purposes. Sustainable irrigation methods are required to reduce the wastage of water and effectively use it. Drip irrigation method is an effective way to reduce the water usage significantly without losing crop yield. This study conducted in Gudipadu Cheruvu Village was focused on the challenges faced by the people due to water scarcity. Challenges faced by the people were analysed using participatory rural appraisal approach and human-centred design methods. Also, these approaches helped to design and derive an effective method to overcome the water scarcity issues prevalent in the village.

P. B. Abhinaya · N. Radhika
Department of Computer Science and Engineering, Amrita Vishwa Vidyapeetham, Coimbatore, India
e-mail: n_radhika@cb.amrita.edu

T. Adarsh
Department of Civil Engineering, Amrita Vishwa Vidyapeetham, Coimbatore, India

P. Vanga
Department of Electronics and Communications Engineering, Amrita Vishwa Vidyapeetham, Coimbatore, India

S. Sivanesh
Department of Chemical Engineering, Amrita Vishwa Vidyapeetham, Coimbatore, India

Y. Vishnuvardhan
Department of Mechanical Engineering, Amrita Vishwa Vidyapeetham, Coimbatore, India

A. S. Reshma (✉)
Amrita Center for International Programs, Amrita Vishwa Vidyapeetham, Amritapuri, Kollam, India
e-mail: reshmaas@am.amrita.edu

T. Senjyu et al. (eds.), *IOT with Smart Systems*, Smart Innovation, Systems and Technologies 251, https://doi.org/10.1007/978-981-16-3945-6_56

55.1 Introduction

Nature has provided us with 1.4 billion cubic metres (m^3) of water globally [1], out of which the quantity of potable water is 200,000 km^3. Present world population is 7.7 billion people. The amount of freshwater available is insufficient to satisfy the global population. More than one out of six people encounter water scarcity in the world [2]. India stands at the thirteenth position in the list of water-stressed countries, with a human population of 1.392 × 1013 billion [3].

The water usage by the Indian population is shown in Fig. 55.1. Agricultural sector consumes the major portion of water—about 80% [4]. This water requirement for agriculture is inevitable, since feeding a country with a large population, such as India, requires an enormous amount of water. The Indian population is expected to rise to 2 billion by 2050, exacerbating water scarcity [5].

According to the latest census, the majority of the Indian population resides in rural areas in which people mainly depend on agriculture. Due to the lack of proper education and awareness, people in rural areas are unaware of sustainable irrigation methods that enhance agricultural yields. This paper presents a case study conducted in Gudipadu Cheruvu, a village located in India's state of Andhra Pradesh (AP). This research work was conducted as part of Amrita Vishwa Vidyapeetham's Live-in Labs® program [6]. Next session introduces the study area which includes details such as the main occupation, traditional practices, climatic variations and problems they face. Section 55.3 describes the methods used in our study and their findings. In Sect. 55.4, we have discussed the solution proposed by the participants and conclusion and future aspects in Sect. 55.5.

Gudipadu Cheruvu, a village situated in the southern Indian state of AP, attained the status of village, in the past decade. Inhabited by 547 people in 173 households, this village belongs to Kandlakunta Panchayat, Veldurthi Mandal, Guntur District

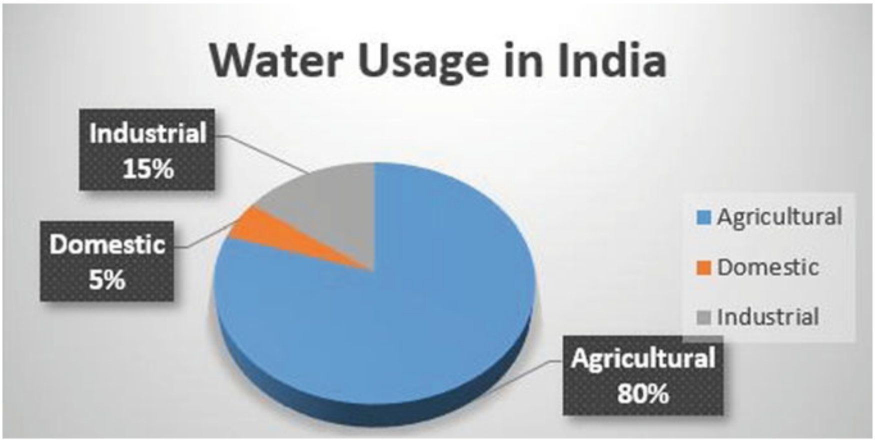


Fig. 55.1 Water usage in India

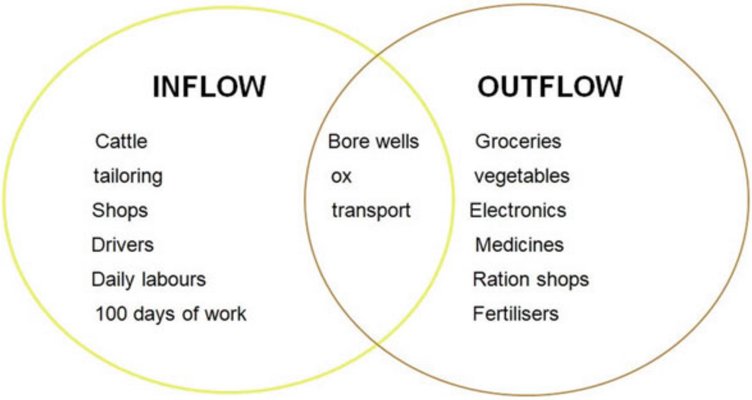


Fig. 55.2 Inflow–outflow diagram

in AP (Fig. 55.2). The village is about 45 min (24 km) of travel from Macherla, the nearest town. The Gudipadu Cheruvu Village is known for its varying geographical terrains and varying climates. The village is situated in one of the hottest regions of the state having an average temperature of about 45 °C. Major crops cultivated here are paddy, chilli and cotton. In order to ensure water accessibility to all members of the community, a micro-water distribution system was built by researchers as Amrita Vishwa Vidyapeetham [7]. This was part of similar implementations done by the University in villages in other Indian states [7]. Borewells have been exploited by the villagers for agricultural activities. Due to the lack of rain, the water table has failed to be replenished in Gudipadu Cheruvu. The traditional method of irrigation followed by the people in the village is through the canal system. Because of climatic variations caused rainfall shortage which forced the farmers to drill deeper borewells to draw subsurface water. The villagers have been practicing conventional irrigation such as canal method of irrigation for many years. For the past six years, there has been shortage of rainfall that impelled the farmers to drill deeper borewells to draw subsurface water. Due to the climatic change and failure in monsoons, low rainfall has prevailed for the past few years.

55.2 Methodology and Results

Ethnographic research was incorporated into the study where the research team visited the village and lived along with the community to understand their problems. Data was collected using human-centred design (HCD) and participatory rural appraisal (PRA) which has been detailed in this section.

55.2.1 *Participatory Rural Approach (PRA)*

PRA is a participatory approach which aids in identifying the resources and the socio-economic conditions of the village; data from which is used to focus on a challenge that is significantly affecting the community [8–10]. Based on the initial field visit, it was understood that there is only one primary school in that village which have class only upto 5th grade. Most of the children discontinue studies after that and thus have only a basic level of education. So most men are farmers and women engage in their household activities. We interacted with both men and women to understand the challenges faced by both the groups.

PRA tools utilized in this study are delated below:

Resource map A resource map is used to obtain a better insight into the village's resources and their distribution of the village's resources [11]. Resources pertaining to various thematic areas (water, agriculture, sanitation, health, education, income generation, skill development and waste management). With the help of certain community members, the research team went around the village to identify various resources in the village.

Seasons and activities This tool enabled the research team to identify various activities conducted by the village communities in different seasons of the year [12]. Table 55.1 summarizes the occupation carried out by the villagers in each month. Analysing the table, we can clearly understand that their work pattern changes according to the climatic variations.

Paddy being a water intense crop, cannot be nurtured during the hot months of the year (February to June). But cotton and chilli are more sustainable and can be

Table 55.1 Summary of crops grown by the people in different seasons

Month/crops grown	Paddy	Chilli	Cotton	MGNREGA
January	5	3, 4, 5	3, 4, 5	
February		4	4	
March		5	5	6
April		5	5	6
May		1	1	6
June		1, 2	1, 3	6
July		2, 3	2, 4	
August	1	3	3	
September	2, 3	4	4	
October	2, 3	3, 4	3, 5	
November	3	3, 4	3, 4	
December	5	3, 4	3, 4	

where the numbers denote the following
1—ploughing, 2—sowing, 3—weeds, 4—fertilizers, 5—harvest, 6—labour

grown throughout the year. And due to harsh climatic conditions from March to June, cultivation is hard for people and they take up labour jobs via the 100 days labour government scheme to balance the income flow (Table 55.1).

Inflow–outflow diagram The inflow–outflow diagram helped to identify the resources that are exported from the village, the resources which are imported into the village, and the resources which are produced in the village and utilized within the village itself (Fig. 55.2). This enabled us to further understand the resources that are abundantly available in the village (leading to outflow) and the resources that are scarce (needing inflow).

Problem tree It is through the problem tree method that we understand the core problems along with their causes and effects. Problem tree mapping helps to make sense of all the data collected and organize them in a meaningful way to identify the main challenges [13]. In this village, agriculture and water shortages were the main problems.

Brainstorming Session Brainstorming is a way to generate ideas within a group setting. Villagers were invited for group discussions in which the main challenges faced by the people were identified [14]. Problems which were raised by the majority of the people were selected for our study purpose. Their ideas and solutions were documented.

55.2.2 Human-Centred Design

Human-centred design approach was used to understand the challenges faced by the village residents [15–17]. The villagers' requirements, their pain points and daily experiences were documented. Additional details were retrieved by human-centric tools such as participant observation, interviews, personas and scenarios, as explained below.

Participant Observation The diverse activities of the villagers were observed over a period of seven days. These observations enabled determination of the activities, environment, interactions, objects and users in a particular boundary by using the AEIOU observation framework [18]. Using this method, we could clearly appreciate that the villagers were completely dependent on agriculture. Major crops grown are chilli, paddy and cotton. Villagers either own land or work as farm labourers in nearby villages. Open water sources were utilized for irrigation. Dried lands indicated a shortage of rainwater which was confirmed during interview session.

Interviews Semi-structured interviews were conducted with 60 families in the village during field visits. The questions asked were mainly open-ended with emphasis on details pertaining to demography, technography, psychography, a day in one's life, goals, tasks and pain points.

Fig. 55.3 Research team conducting interviews with the village residents

In a few instances, the default questions were tailored to each household. The interviews were semi-structured so that the participants can freely express their experiences and emotions. Most of the participants who attended the interview were dependent on agriculture for their livelihood (Fig. 55.3). The villagers opined that the scarcity of water faced during the summer season and its ill-effects on agriculture affected their lives. Participants reported that rainfall had reduced since the past few years because of seasonal variations that had a negative impact on agriculture.

Interviews with the participants revealed that, due to the unavailability of water and low rainfall, villagers had to adapt their farming methods. They started cultivating cotton and chilli instead of paddy as they required much less water.

Personas Personas were developed from the interviews conducted. Understanding the user groups as well as their challenges is possible from the personas [19]. The villagers confirmed that the absence of water for the past three years had worsened the situation by further depleting the water table.

Scenarios Scenarios enabled us to deeply understand and analyse the pain points experienced by the village community. Scenarios highlighted the adverse impacts on their livelihoods due to the minimal agriculture that was practised due to water scarcity. Due to a lack of knowledge of scientific agricultural practices, the available water is not properly utilized. This added to drastic seasonal variations and inadequate rainfall causes the groundwater to dry up easily.

55.3 Pathway for Development

Based on the study, it was clear that people of Gudipadu Cheruvu were highly affected by water scarcity. Drip irrigation is one of the types of micro-irrigation systems that save nutrients and use water efficiently. It directly supplies water drops to plant roots, at a slow rate, either from below or above the surface [20]. Through this method, water loss caused by evaporation can be reduced effectively. Drip irrigation system delivers water to crops, via a network of primary lines, secondary lines and lateral lines, with dripping points spaced along their lengths [21]. The dripping space was determined by the crop sown in the field. Each dripper/emitter orifice supplies a calculated, precisely controlled uniform flow of water, along with other nutrients required by the plants. Water and nutrients enter the soil from the emitters which are fed directly into the root zone of the plants. The combined forces of gravity and capillary action lead the water from the orifice towards the root nodes [22]. The plants retrieve the moisture and nutrients almost instantly so that the plant never gets water-stressed, thus enhancing quality, its ability to achieve optimum growth and high yield.

By implementing smart irrigation, the crop yield could be augmented by enhancing the proposed system with the application of IoT. The deployed sensors identify or read the physical and chemical parameters of the farmlands such as weather, temperature, humidity, plant health, soil moisture, pH of the soil, soil nutrients [14]. The outputs from the sensors are analysed to notify the farmers about the essential remedial measures to be undertaken. The analysis helps in the determination of the apt fertilizers and their quantities, besides the amount of water required for irrigation. Such data help in automated motor/pump for irrigation.

Smart irrigation can be implemented by the following five methods.

- Local agricultural parameters should be sensed.
- Location of sensors should be identified, and data should be collected.
- Collected data should be transferred from fields and should be used for decision making.
- By sensing soil water conditions, it will either turn on and off the motor based on the requirement.
- Early warning and support of decision-making based on data analysis, image processing and domain knowledge.

55.4 Design of Proposition

DHT11, a temperature and humidity sensor, is fitted with a separate negative temperature coefficient (NTC) to calibrate temperature, and an 8-bit microcontroller that outputs the values of temperature and humidity as serial data. DHT11 sensor is calibrated to industry standards and interoperable with other standard microcontrollers.

The sensor can measure temperature in the range of (0–50) °C ± 1 °C and humidity in the range of (20–90).

The moisture sensor is used to detect the water content (moisture) present in the soil. When the soil is dry due to water shortage, the sensor module output is at high level, or else the output is at low level. Therefore, the sensor reminds the user to water the plants and monitor the moisture content of soil [15]. Based on the outputs of these sensors, automated On/Off mechanisms can be employed. Also borewells, operated by three-phase power, are used in the farm, eight hours a day for irrigation. Villagers sometimes adopted an all-nighter field irrigation schedule. At times, even though the farmland is well-irrigated, failure to switch off power results in the wastage of large amounts of water. From a sustainability perspective, it is crucial to conserve water, which can be done by the practice of drip irrigation. Such a system would help villagers save energy as well as water.

55.5 Discussion and Conclusion

The study conducted in Gudipadu Cheruvu clearly reveals that the major challenge faced by the people is water scarcity. Participatory tools utilized in this study helped to share their experiences and pinpoints. Most of the people were farmers, and thus, a huge amount of water was utilized for irrigation purposes. Groundwater and rainwater were the two sources used for irrigation. Due to the important climatic variations in recent years, we could understand that availability of rainwater reduced and groundwater also reduced. During the interactive session participants revealed that, for all crops equal amounts of water was used which caused huge amount of water loss. So we could understand that a more sustainable method of irrigation for the village was necessary to overcome the water scarcity issue that existed there. We suggest drip irrigation as it takes the water directly to the roots of the plant. This method of irrigation ensures sustainability by saving 70% of water.

The challenge for developing the system in the area is low literacy rate and also lack of proper signals for the sensors to work. Increase in depth of groundwater level is also an additional challenge for drip irrigation that needs to be confronted. The proposed solution helps to reduce the water required and the sensor network, which is going to be deployed at Gudipadu Cheruvu, prompts the farmers to assess the condition of the soil and take required actions to keep the soil fertile and useful for cultivation. Even though the proposed solution seems to be effective, as a next step, the effectiveness of the solution should be monitored and evaluated at an equal interval of time. This will help to modify the solution based on the obtained results and also help to implement in other villages facing similar issues. The proffered solution also contributes to the United Nations' Sustainable Development Goals (SDGs)—no poverty (SDG 1) and Clean Water and Sanitation (SDG 6). This method could be adopted by other communities tackling similar problems.

Acknowledgements This research was funded by Amrita Vishwa Vidyapeetham's Live-in Labs® program. The study was conducted as part of the UNESCO Chair for Experiential Learning for Sustainable Innovation Development. The authors express their immense gratitude to our Chancellor and renowned humanitarian leader Sri. (Dr.) Mata Amritanandamayi Devi (Amma) for her inspiration and support towards working on interdisciplinary research that has direct societal benefit. The authors would also like to express their sincere appreciation to Mr. Gopalaiengar Sheshadri for his extensive support in enhancing the quality of the paper within the limited time frame. The authors also thank the faculty, staff and students of Amrita Vishwa Vidyapeetham, the government authorities of Guntur District, Andhra Pradesh, and staff of the Amrita Self-Reliant Villages (SeRVe) program.

References

1. Gleick, P.H., Palaniappan, M.: Peak water limits to freshwater withdrawal and use. Proc. Natl. Acad. Sci. **107**(25), 11155–11162 (2010)
2. Wang, D., Hubacek, K., Shan, Y., Gerbens-Leenes, W., Liu, J.: A review of water stress and water footprint accounting. Water **13**(2), 201 (2021)
3. Forney, S.Z., Sadar, A.J.: Environmental Risk Communication: Principles and Practices for Industry. CRC Press, Boca Raton (2021)
4. Zhang, D., Sial, M.S., Ahmad, N., Filipe, A.J., Thu, P.A., Zia-Ud-Din, M., Caleiro, A.B.: Water scarcity and sustainability in an emerging economy: a management perspective for future. Sustainability **13**(1), 144 (2021)
5. Subramanian, S.: The fast lane to India's water management. World Water Policy **6**(2), 279–285 (2020)
6. Ramesh, M.V., Mohan, R., Menon, S.: Live-in-labs: rapid translational research and implementation-based program for rural development in India. In: 2016 IEEE Global Humanitarian Technology Conference (GHTC), pp. 164–171. IEEE (2016)
7. Ramesh, M.V., Mohan, R., Nitin Kumar, M., Brahmanandan, D., Prakash, C., Lalith, P., Ananth Kumar, M., Ramkrishnan, R.: Micro water distribution networks: a participatory method of sustainable water distribution in rural communities. In: 2016 IEEE Global Humanitarian Technology Conference (GHTC), pp. 797–804. IEEE (2016)
8. Kadiveti, H., Eleshwaram, S., Mohan, R., Ariprasath, S., Nandanan, K., Divya Sharma, S.G., Siddharth, B.: Water management through integrated technologies, a sustainable approach for village Pandori, India. In: 2019 IEEE R10 Humanitarian Technology Conference (R10-HTC) (47129), pp. 180–185. IEEE (2019)
9. Chambers, R.: The origins and practice of participatory rural appraisal. World Dev. **22**(7), 953–969 (1994)
10. Chambers, R.: Participatory rural appraisal (PRA): analysis of experience. World Dev. **22**(9) 1253–1268 (1994)
11. Kumar, P.S., Pranay, N., Ilam Bharathi, G., Suchit, Y., Saravanan, R., Nandanan, K., Mohan, R.: Technology enabled pathways for a waste free village. In: ICDSMLA 2019, pp. 1273–1283. Springer, Singapore (2020)
12. Chandra, G.: Participatory rural appraisal. In: Katiha, P.K., Vaas, K.K., Sharma, A.P., Bhaumik, U., Chandra, G. (eds.) Issues and Tools for Social Science Research in Inland Fisheries, vol. 163, pp. 286–302. Central Inland Fisheries Research Institute, Barrackpore, Kolkata, India (2010)
13. Snowdon, W., Schultz, J., Swinburn, B.: Problem and solution trees: a practical approach for identifying potential interventions to improve population nutrition. Health Promot. Int. **23**(4), 345–353 (2008)
14. Patton, M.Q.: Qualitative evaluation methods, p. 1 (1980)

15. Vechakul, J., Agogino, A.: A comparison of two transdisciplinary human-centered design approaches for poverty alleviation. In: The Future of Trans-disciplinary Design, pp. 179–192. Springer, Cham (2021)
16. Brown, T., Wyatt, J.: Design thinking for social innovation. Dev. Outreach **12**(1), 29–43 (2010)
17. Ganesh, V., Sharanya, B., Sri Vyshnavi, U., Pradeep, K.R., Kumar, A.G., Purna Chaitanya, G., Frey, L.M., Cornet, S., Chinmayi, R.: Technology for addressing income insufficiency in rural India. In: 2020 IEEE 8th R10 Humanitarian Technology Conference (R10-HTC), pp. 1–6. IEEE (2020)
18. Lee, M.J., Wang, Y., Been-Lirn Duh, H.: AR UX design: applying AEIOU to handheld augmented reality browser. In: 2012 IEEE International Symposium on Mixed and Augmented Reality-Arts, Media, and Humanities (ISMAR-AMH), pp. 99–100. IEEE (2012)
19. Miaskiewicz, T., Kozar, K.A.: Personas and user-centered design: How can personas benefit product design processes? Des. Stud. **32**(5), 417–430 (2011)
20. Rekha, V.P.R., Ramesh, M.V., Nibi, K.V.: High yield groundnut agronomy: an IoT based precision farming framework. In: 2017 IEEE Global Humanitarian Technology Conference (GHTC), San Jose, CA, pp. 1–5 (2017). http://doi.org/10.1109/GHTC.2017.8239287
21. Prathibha, R., Hongal, A., Jyothi, M.P.: IOT based monitoring system in smart agriculture. In: 2017 International Conference on Recent Advances in Electronics and Communication Technology (ICRAECT), Bangalore, pp. 81–84 (2017). http://doi.org/10.1109/ICRAECT.2017.52
22. Bhanu, B.B., Hussain, M.A., Ande, P.: Monitoring of soil parameters for effective irrigation using wireless sensor networks. In: 2014 Sixth International Conference on Advanced Computing (ICoAC), pp. 211–215 (2014)

Chapter 56
A Survey on Sentiment Lexicon Creation and Analysis

Ashish R. Lahase, Mahesh Shelke, Rajkumar Jagdale, and Sachin Deshmukh

Abstract This is a survey paper that presents sentiment lexicons and discusses the state of the art in the field of methods of producing sentiment lexicons (Creation). Like websites, social networks, blogs, online portals, reviews, thoughts, suggestions, ratings, and feedback are generated by writers with accelerated Internet development. Sentiment material can be about books, people, hotels, products, surveys, activities, etc., and these sentiments are very useful to businesses, governments, and individuals. Sentiment analysis is the method for processing natural language and text analysis to define and extract subjective knowledge from text. This paper provides a survey on the problems of sentiment analysis related to their methods and techniques. There are different types of lexicons covered, differing in aspects such as coverage, development methods, lexical unit, and granularity. It aims to provide a representative sample of the field of lexicons of sentiment.

56.1 Introduction

Sentiment analysis is one of the most important applications of natural language processing. It refers to the study of the extraction from the perspective of texts. There are two key approaches to sentiment analysis: a classifying approach that considers sentiment analysis to be a special case of text classification and uses conventional machine learning methods to solve a problem or approach based on a lexicon, using emotion lexicons dictionaries of words with labels representing their feelings to define the feelings of the text. Both methods have their benefits and disadvantages [1].

Sentiment analysis has become an active field of ongoing research since its introduction by IBM. This was in 1961. Findings were produced to understand feelings at the document level, sentence, and aspect level. Even so, the use of opinionated terms to accomplish the task is typical among many research results. It is because the

A. R. Lahase (✉) · M. Shelke · R. Jagdale · S. Deshmukh
Dr. Babasaheb Ambedkar Marathwada University, Aurangabad, India

T. Senjyu et al. (eds.), *IOT with Smart Systems*, Smart Innovation, Systems and Technologies 251, https://doi.org/10.1007/978-981-16-3945-6_57

sentiment lexicon recreates the visual information present in the human in devices. With a lexicon, knowing emotion becomes easier [2].

The development of a lexicon containing such details comes with its own set of difficulties. A lexicon like that is usually a specific domain. Typically, this is because words play various roles in different ways, and it is difficult for a lexicon to possess all domains equally. For example, the word 'large' can be seen as negative in the term 'big house.' Contradictory, the same word 'big' can be seen as positive in 'a big engine.' Another thing to note is that these programs vary based on language and culture. For example, Gandhi, for instance, could be seen in the Hindi SentiWordNet as 'positive' as 'neutral' in the English SentiWordNet. There are also words in languages that in other languages do not have a literal translation. Capturing such words is critical in a holistic sense since such a resource depends on the language [3].

Sentiment analysis is a method of analyzing a predefined feeling concerning a given subject from online texts written in a natural language. The need for an analysis of sentiment is the product of a sudden spike in blogs, reviews, and discussions of opinions or sentimental texts [4].

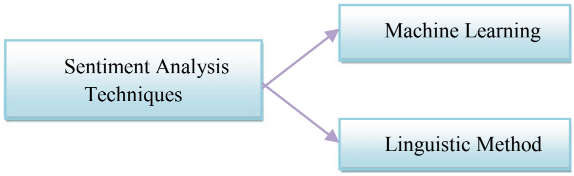
The task can be seen as a classification question in which, according to the opinions found therein, documents must be classified into either positive or negative categories. When a large amount of annotated data is available, a good technique is to create classifiers based on supervised machine learning that relies on document terms. But if the size of the annotated one is correct, this method may be incorrect. The corpus is incomplete since it has a low frequency of word types. Sentiment lexicons can help to solve the problem of lexical sparsity, containing a predefined set of positive and negative words. In addition to n-grams, this information can be used to extract various features. There are many general-purpose emotion lexicons available for English, but there are far fewer for Hungarian [5].

56.2 Related Work

A good starting point may be an established lexical resource when constructing a sentiment lexicon automatically, the contents of which could then be associated with a sentiment ranking. WordNet, whose existence is influenced by the psycholinguistic theories of human lexical memory, is a widely used online (semantic) lexical resource. WordNet is intended to be used under software supervision and allows for the distinction between word types and definitions of words. WordNet is organized into {synset {collections of synonyms}} that can be distinguished based on their Part-of-Speech (POS) type. Each synset represents a distinct concept and is related by various types of relationships to other synsets [6].

Four key POS forms are included in WordNet: verbs, nouns, adjectives, and adverbs. Usually, WordNet's semantic network of English verbs is considerably more complex than the network of nouns, which means that verb meanings are more usable than noun meanings. Based on distinguishing characteristics, nouns in WordNet are grouped hierarchically. Rarely does this hierarchy surpass more than a dozen levels?

Fig. 56.1 Sentiment analysis techniques [9]



In WordNet, adjectives are split into two classes: Descriptive and relational can have multiple kinds of relationships. Of all POS forms, adverbs have the least complex structure. Adverbs not derived from an adjective only have occasional antonym relationships, and derived adverbs are semantically related to their base adjectives. Key techniques and methods of sentiment analysis are available, such as machine learning, lexical dictionaries, natural language processing, psychometric scale, imaging, and cloud-based techniques. Machine learning requires a large data resource because of the training portion. The verbal approach to action and understanding is much easier than machine learning. Typically, these two techniques are now combined [7].

Many linguistic rules and constraints provide an unsupervised and supervised machine learning system that could enhance the accuracy of calculations and classifications. A more basic field is the psychometric scale system. It primarily analyzes people's moods and presents as a formalized indicator of societal happiness and depression the new smile or cry index. It is also combined with lexical dictionaries. To some extent, the lexical dictionary approach is the development of lexical affinity and linguistic technique. If you're a beginner, the simple process can be easy to run. It is not going to take too many measurements or data resources. The processing of natural language is a tool that can enforce the relationship between humans and computers. It will assist us to evaluate texts' polarity. The techniques are the basis of the sentient. It is an approach that categorizes texts as positive or negative [8].

Figure 56.1 demonstrates the techniques of sentiment analysis. Hindi uses SentiWordNet in English and links to WordNets in both English and Hindi. Polarity scores have been copied from the English SentiWordNet terms to the corresponding Hindi SentiWordNet translated terms. Further work on Hindi has also been done to create a sentiment lexicon [9].

56.3 Lexicon

Lexicon is an important instrument that plays a role in the study of emotions. SentiWordNet is the most prevalent and well known of current lexicons. For each word of opinion, SentiWordNet has three degrees of emotion: positivity, negativity, and objectivity. SentiWordNet has been created from version 1.0 to version 3.0.0. There are some variations between versions 1.0 and 3.0 of SentiWordNet: WordNet versions and algorithms used to automatically annotate WordNet, which can now be used to optimize scores randomly. SentiWordNet 3.0 is working to make the part better [10].

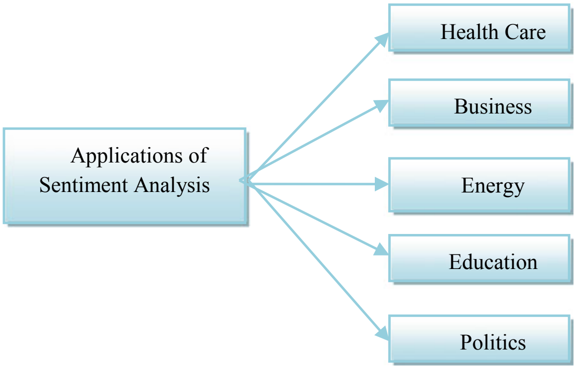


Fig. 56.2 Applications of sentiment analysis [12]

56.4 Applications of Sentiment Analysis

Sentiment analysis originates from the products and services of customers. Amazon.com is a representative example of this. Twitter and Facebook are both popular and popular sites for many sentiment analysis applications.

There are various uses for the analysis of sentiment. Thousands of text documents can be processed in minutes through sentiment analysis, compared to the hours that a team of individuals will need to complete manually. The data can be sentences, phrases, or paragraphs. Sentiment analysis is directly linked to the analysis of feelings in China. It means that it is possible to determine what feelings or moods people have. On the other hand, digital statistics do not tell us what people feel. There are different applications for sentiment analysis. Thousands of text documents can be processed in minutes through sentiment analysis, compared to the hours that a team of individuals would need to complete manually. The data can be words, phrases, or paragraphs. Sentiment research is specifically related to the research of emotions in China. It means that it is possible to decide what emotions or moods people have. Digital figures, on the other hand, do not tell us what individuals feel [11] (Fig. 56.2).

56.5 Proposed Methodology

Propose a method to build a Marathi sentiment lexicon using a SenticNet and a WordNet. Instead of a translation-based approach, this method was primarily used to take into account language variations. A word may mean different things in different languages [13].

56.5.1 Translation-Based Approach

Marathi, apart from being a language that is commonly spoken and understood, but not highly researched. It also has a wide variety of publicly available data on the Internet, such as film reviews, survey responses, drama scripts, Marathi translation collections, and parallel corpus. This promotes the idea of creating a sentiment lexicon in the Marathi language to further advance research in sentiment analysis [14].

56.5.2 Subjectivity Lexicon

The word is labeled as either strong or weakly subjective by the subjectivity parameter. The four distinct types are type, part of speech, stemming, and prior polarity [15].

56.5.3 Translation Method

Google Translate5 will be used to translate the final list of words from English to Marathi. To exclude multiword entries, incorrect translations, and other unclear terms, the final list of translated terms was cross-reviewed by Marathi annotators. English SentiWordNet, but when translated, it loses its subjectivity in the target language. Google Translate is a fast, free, accurate translation service. The translation may also be achieved using a bilingual dictionary. However, the Google Translate API is simple to use and has a wide coverage. Using a bilingual dictionary is a painful, laborious process to carry out [16].

56.5.4 IndoWordNet

Google Translate is a fast, free, accurate translation service. The translation may also be achieved using a bilingual dictionary. However, the Google Translate API is simple to use and has a wide coverage. Using a bilingual dictionary is a painful, laborious task to carry out [17] (Fig. 56.3).

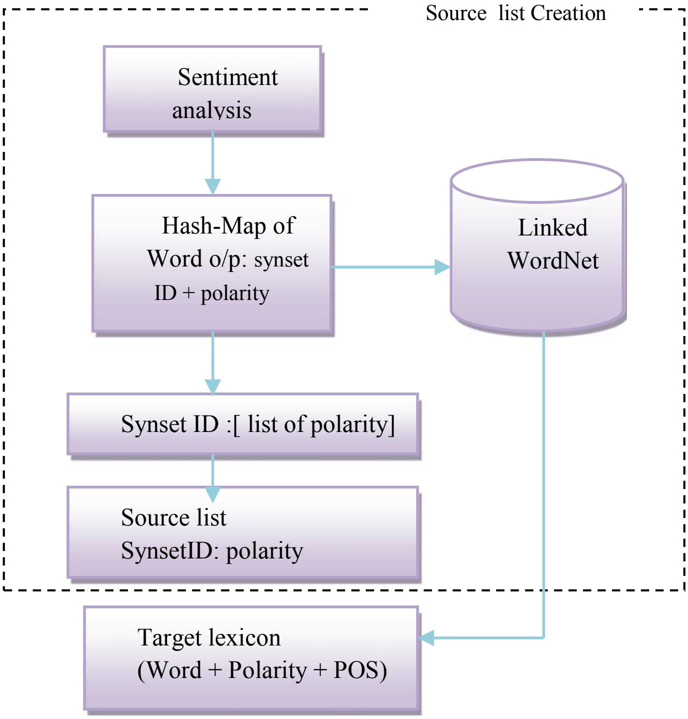


Fig. 56.3 WordNet model for sentiment lexicon creation [18]

56.5.5 SenticNet

SenticNet5 is a publicly available platform that deals with concept-level sentiment analysis. SenticNet uses common sense reasoning tools and domain-specific ontology for sentiment analysis, SenticNet deals with statistical learning models. This common-sense/taken-for-granted data is taken into account by SenticNet. It uses the lexical representation of affective information to assist in the task of opinion mining [19] (Fig. 56.4).

SenticNet4 has 50,000 entries that are a combination of phrases and phrases. These entries are graded on a scale of [−1, 1]. But many of SenticNet's entries are unique to both domain and language. It would also be helpful to expand the English language to other European languages. If we were to use this resource in Indian languages, we would have phrase-level translation followed by human error correction for multiword entries [21].

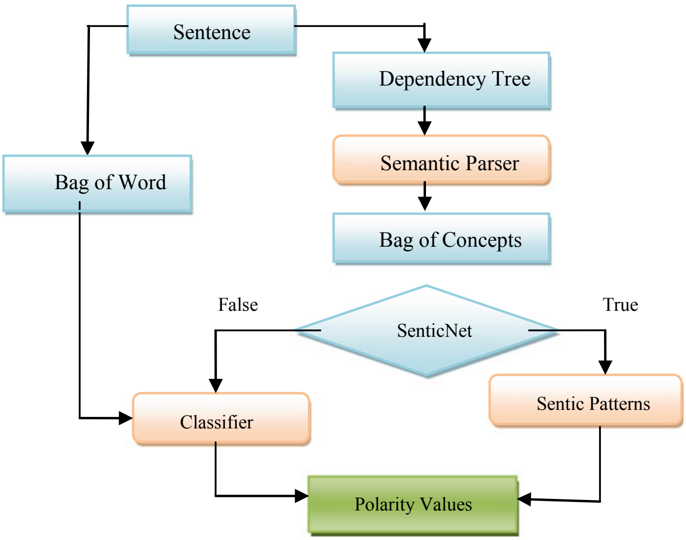


Fig. 56.4 SenticNet model for sentiment lexicon [20]

56.5.6 Comparison of Sentiment Analysis Methods

See Table 56.1.

56.6 Conclusion

The most popular positive and negative words databases that can help to perform sentiment analysis were described: Liu and Hu opinion lexicon, SentiWordNet, Senti-Words, AFINN, WordStat Sentiment Dictionary, SenticNet, the Affective Norms for English Words, the Whissell's Dictionary of Affect in Language, Pattern, Sentiment140 Lexicon, Linguistic Inquiry, and Word Count, the MPQA Subjectivity Lexicon. A systematic analysis of existing literature on sentiment lexicon for languages. In that survey, what is the initial application of sentiment lexicon, what is the methodology used? In future work, it will be helpful to investigate how sentiment lexicon creation for low-resource languages can be developed. simplistic negation handling algorithm implemented as part of the bespoke sentiment analysis system may have affected the degree to which negation was detected is the limitation of sentiment analysis, and how accurately it was detected. In the future work, we will develop sentiment lexicon creation and analysis for Marathi language.

Table 56.1 Table of comparison of sentiment analysis methods

Name	Description	L	ML
Sentiment140 lexicon [22]	A dictionary of lexicons based on the same dataset was used for the Sentiment140 process training. To identify the tweet as positive or negative, the lexicon was constructed to emoticons. Then the n-gram score in each class of tweets is based on the frequency of occurrence	Yes	
SentiStrength [23]	Builds a human-annotated lexicon dictionary enhanced with the use of machine learning	Yes	Yes
SenticNet [24]	Uses dimensionality reduction to infer the polarity of concepts of common sense and thus provide a resource at a semantic, rather than just syntactic level	Yes	
SentiWordNet [25, 26]	Construction of a WordNet-based lexical resource for opinion mining, associating three polarity scores (positive, negative, and neutral) for each	Yes	Yes
Opinion lexicon [27]	Build a Lexicon to predict the polarity of phrases of product features that are summarized to provide the product feature with an overall score	Yes	
Stanford recursive deep model [28]	RNTN is proposed that processes all sentences dealing with their structures and measures the interactions between them	Yes	Yes

References

1. Medhat, W., Hassan, A., Korashay, H.: Sentiment analysis algorithms and applications: a survey. Ain Shams Eng. J. **5**(4), 1093–1113 (2014)
2. Yadollahi, A., Shahraki, A.G., Zaiane, O.R.: Current state of text sentiment analysis from opinion to emotion mining. ACM Comput. Surv. **50**(2) (2017)
3. Boguraev, B., Pustejovsky, J.: Corpus Processing for Lexical Acquisition. The MIT Press, Cambridge (1996)
4. Liu, B.: Sentiment analysis and subjectivity. In: Indurkhya, N., Damerau, F.J. (eds.) Handbook of Natural Language Processing, 2nd ed. (2010)
5. Kowsari, K., Meimandi, K.J., et al.: Text Classification Algorithms: A Survey. arXiv, 23 Apr 2019
6. Neviarouskaya, A., Prendinger, H.: SentiFul: a lexicon for sentiment analysis. iEEE Trans. Affect. Comput. **2**, 22–36 (2011)
7. Miller, G.A., Beckwith, R., Fellbaum, C.: WordNet: an on-line lexical database. Commun. ACM **38** (2008)
8. Trotzek, M., Koitka, S., Friedrich, C.M.: Utilizing neural networks and linguistic metadata for early detection of depression indications in text sequences. IEEE Trans. Knowl. Data Eng. arXiv: 1804.07000 [cs.CL] (2018)
9. Pandey, P., Govilkar, S.: A framework for sentiment analysis in Hindi using HSWN. Int. J. Comput. Appl. (2015)

10. Chauhan, U.A., Afzal, M.T., et al.: A comprehensive analysis of adverb types for mining user sentiments on amazon product reviews (2020)
11. Fang, X., Zhan, J.: Sentiment analysis using product review data (2015)
12. Korashy, W.M.A.H.H.: Sentiment analysis algorithms and applications: a survey. Ain Shams Eng. J. **5**(4), 1093–1113 (2014)
13. Vilares, D., Peng, H., Satapathy, R., Cambria, E.: BabelSenticNet: a commonsense reasoning framework for multilingual sentiment analysis. In: IEEE Symposium Series on Computational Intelligence SSCI (2018)
14. Callison-Burch, C., Osborne, M.: Constructing parallel corpora for six Indian languages via crowdsourcing. In: Proceedings of the Seventh Workshop on Statistical Machine Translation, June 2012
15. Wilson, T., Wiebe, J., Hoffmann, P.: Recognizing contextual polarity: an exploration of features for phrase-level sentiment analysis. Comput. Linguist. **35**(3), 16 (2008)
16. Mishra, D., Venugopalan, M., Gupta, D.: Context-specific lexicon for Hindi reviews. In: 6th International Conference on Advances in Computing & Communications, ICACC 2016, 6–8 Sept 2016, Cochin, India
17. Bird, S., Klein, E., loper, E.: Natural Language Processing Using Python. Oreilly, USA (2009)
18. Yousef, A.H., Medhat, W., Mohamed, H.K.: Sentiment analysis algorithms and applications: a survey. Ain Shams Eng. J. **5**(4) (2014)
19. Cambria, E., Hussain, A.: A Common-Sense-Based Framework for Concept-Level Sentiment Analysis. Springer, Switzerland (2015)
20. Cambria, E., Poria, S., Bajpai, R., Schuller, B.: SenticNet 4: a semantic resource for sentiment analysis based on conceptual primitives. In: Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers, pp. 2666–2677, Osaka, Japan, 11–17 Dec 2016
21. Cambria, E., Poria, S., Hazarika, D., Kwok, K.: SenticNet 5: discovering conceptual primitives for sentiment analysis by means of context embeddings. In: The Thirty-Second AAAI Conference on Artificial Intelligence (AAAI-18) (2018)
22. Go, A., Bhayani, R., Huang, L.: Twitter sentiment classification using distant supervision (2009)
23. Thelwall, M.: Heart and soul: sentiment strength detection in the social web with SentiStrength. https://sci-hub.se/, http://sentistrength.wlv.ac.uk/documentation/SentiStrengthChapter.pdf (2013)
24. Cambria, E., Olsher, D., Rajagopal, D.: SenticNet 3: a common and common-sense knowledge base for cognition-driven sentiment analysis. In: 28th AAAI Conference on Artificial Intelligence, pp. 1515–1521 (2014)
25. Esuli, A., Sebastiani, F.: SentiWordNet: a publicly available lexical resource for opinion mining. In: 5th International Conference on Language Resources and Evaluation (LREC), pp. 417–422 (2006)
26. Baccianella, S., Esuli, A., Sebastiani, F.: SentiWordNet 3.0: an enhanced lexical resource for sentiment analysis and opinion mining. In: 7th International Conference on Language Resources and Evaluation (LREC), pp. 2200–2204 (2010)
27. Hu, M., Liu, B.: Mining and summarizing customer reviews. In: KDD'04, pp. 168–177 (2004). http://doi.org/10.1145/1014052.1014073
28. Socher, R., Perelygin, A., Wu, J., Chuang, J., Manning, C.D., Ng, A.Y., Potts, C.: Recursive deep models for semantic compositionality over a sentiment treebank. In: Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing (EMNLP '13), pp. 1631–1642 (2013)

Chapter 57
A Survey of Recent Advances in Recommendation Systems

Kanika Pasrija and Kavita Mittal

Abstract Present recommendation schemes such as content grounded filtering and collaborative filtering practice dissimilar databases to create references. Content-based filtering creates recommendations built on customer favorites for product types. Collaborative filtering mimics user-to-user recommendations. It forecasts customer's favorites as a linear, weighted grouping of other user preferences. Both approaches have limits. Content-based filtering can recommend a new item, but requests more data of customer preference in order to include the finest match. Like, collaborative filtering wants huge dataset with lively customers who valued a product before in order to create precise predictions. Arrangement of these dissimilar recommendation schemes is named hybrid systems. These schemes can blend the topographies of the item itself and the favorites of other customers. This paper reviews recent advances in recommendation approaches and their findings.

57.1 Introduction

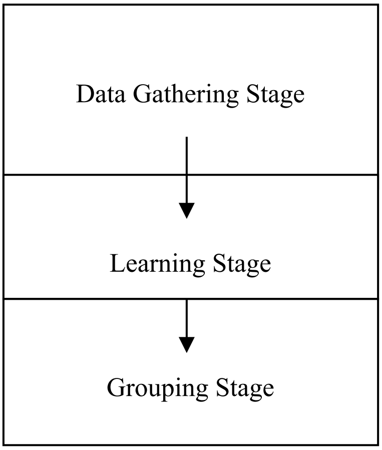
For healthcare field, machine learning is useful and the fastest developing technologies. Machine learning offers superior assistance in better disease analyses, investigates, and avoidance. Lots of machine learning-based classifications have been considered to deliver adapted daily life recommendation/mediation. The explosive progress in the work of digital data and large count of people on Internet has produced an impending task of data burden that obstructs well-timed entrance to things of importance on the Internet. Recommender schemes are facts purifying arrangements which contract with the difficulty of data excess by clarifying vibrant evidence portion out of huge volume of vigorously produced data allowing to customer's favorites, awareness, or practical activities around point [1] (Figs. 57.1 and 57.2).

Recommender schemes create recommendations and references to help their customers in various supervisory practices. Customers are further expected to access suitable produces and facilities by means of the recommender classifications.

K. Pasrija (✉) · K. Mittal
Jagannath University, Bahadurgarh, India

T. Senjyu et al. (eds.), *IOT with Smart Systems*, Smart Innovation, Systems and Technologies 251, https://doi.org/10.1007/978-981-16-3945-6_58

Fig. 57.1 Typical recommendation system [1]



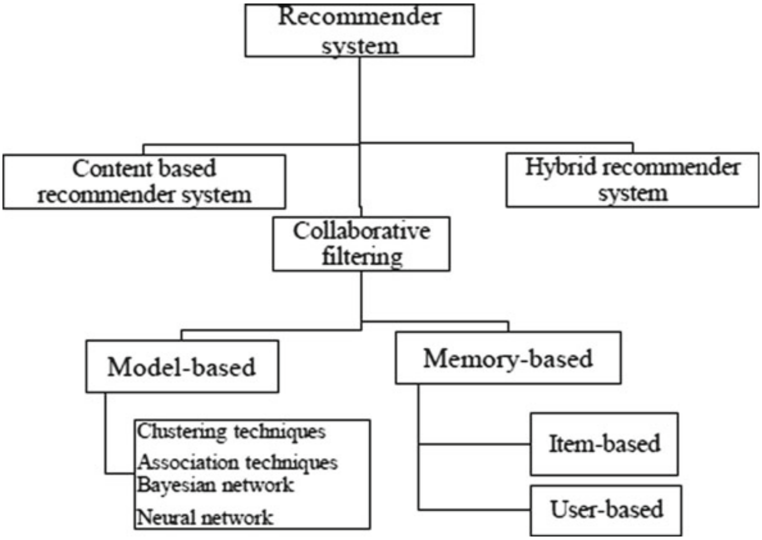


Fig. 57.2 Summary of recommendation techniques [1]

According to Adomavicius and Tuzhilin [1], recommender schemes can be categorized into three main types:

Collaborative Filtering Recommender

CF recommender schemes create suggestions to its customers built on the likings of other consumers with comparable perceptions. It is area self-governing expectation procedure for content that cannot simply and effectively be designated by metadata such as pictures and composition. This method works by the construction of a catalog (user–item matrix) of inclinations for things by customers. Collaborative filtering

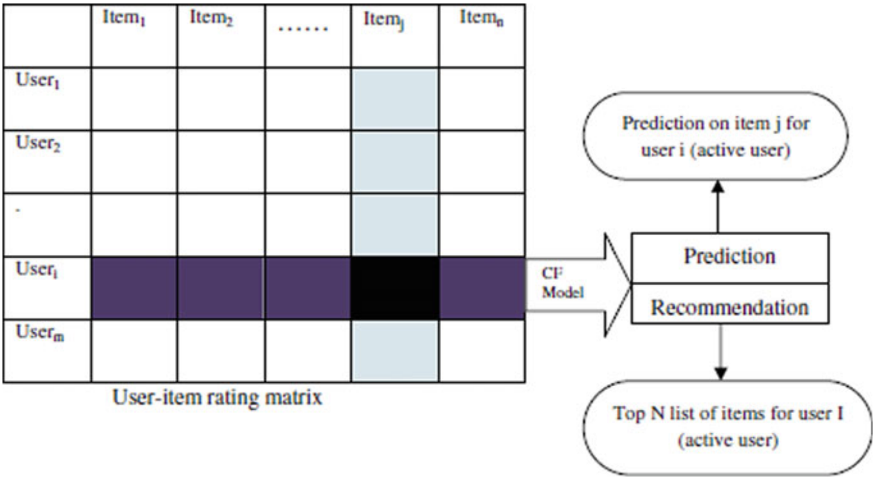


Fig. 57.3 Collaborative filtering

then contests customers with related importance and inclinations through scheming resemblances among their profiles to create references [2] (Fig. 57.3).

Content-Based Recommender

Content-based recommender schemes create suggestions founded on resemblances of fresh articles to those that the customer be fond of in the previous by manipulating the imaginative features of objects. It is an area reliant procedure and highlights further on the examination of the characteristics of objects in direction to make guesses, where official forms such as network sheets, newspapers, and newsflash are to be suggested, contented created filtering system is best effective [3].

Hybrid Recommender Systems

Hybrid recommender schemes employ several methodologies together, and they overwhelmed the drawbacks of certain methods by manipulating recompenses of the other. Hybrid filtering system pools dissimilar recommendation procedures in direction to advance improved classification optimization to evade certain limits and difficulties of clean reference schemes.

Section 57.1 presents the introduction part, Sect. 57.2 presents what is deep learning, basic terminologies of deep learning and findings of deep learning-based recommender systems, Sect. 57.3 presents the literature review, and Sect. 57.4 concludes the paper with future work.

57.2 What Is Deep Learning?

- It comes under class of machine learning systems
- It uses hierarchy of nonlinear processing layers and complex model structures
- Layers learn to represent different representations of facts
- Advanced stage features are constructed from minor stage abstract features
- Trendy name for "neural networks with deep layers".

Deep learning has become increasingly more famous all through subfields of software engineering, for example, natural language preparing, picture and video handling, PC visualization, and information withdrawal because there has not been such a typical way to deal with taking care of various types of figuring issues previously.

Deep learning strategies also assist these subfields to team up with one another wherever this was somewhat questionable earlier because of the assorted variety and unpredictability of used systems [4, 5]. For whatever length of time that the personalization pattern stays famous, the recommender frameworks research will assume a basic job in data examining. In spite of the fact that the use of profound learning into recommender frameworks field guarantees noteworthy and empowering results, difficulties, for example, the precision and marketability are as yet open for upgrades and warrant future work.

Why Deep Learning for Search and Recommender System?

- Direct content feature extraction instead of metadata
- Text, image, audio
- Better representation of users and items
- Hybrid algorithms and heterogeneous data can be used
- Better suited to model dynamic behavioral patterns and complex feature interactions.

Deep Learning-Based Recommender Systems: Findings

- Deep learning practices are not specific for a single reference technique. These can be employed for many type of recommendation approaches for dissimilar purposes. In content-based filtering deep learning-based practices usually extract features to produce user/item profiles from mixed data foundations. Whereas in collaborative filtering these are usually employed as a model-centered methodology to mine hidden features on consumer detail background. Deep learning approaches in hybrid recommender schemes are usually employed to extract features from secondary data and mixing them into the recommendation process.
- DBNs are typically employed to extract features construction from the writing, auditory, and graphical contributions. These mined features can be castoff in either content-built filtering procedures or as lateral facts in collaborative filtering.
- For dimensions reduction of top level and sparse structures into small level and deeper topographies, deep learning-based approaches are useful.

- Deep learning grounded methods are used in circumstance responsive recommender arrangements to perfect appropriate data or catching both customer inclinations and environments.
- Deep learning built methods are also castoff in recommendation arrangements for the resolution of supervision of massive scales data through dimension reduction. Along with dimensions lessening, current deep knowledge systems can also be implemented into a further ascendable system to compact with big scales statistics.
- Deep learning prototypes return further precise references compared to superficial ones. Additionally, using deep learning approaches alongside both consumers grounded and item grounded methods deliver references with greater accurateness, as well.

57.3 Literature Review

See Tables 57.1 and 57.2.

Table 57.1 Accuracy of various classifiers

Authors	Year	Dataset	Classifier	Accuracy
Jenni A. M. Sidey-Gibbons, Chris J. Sidey-Gibbons	2019	Breast cancer Wisconsin diagnostic data set	ANN SVM GLM	94% 96% 95%
Gulshan	2016	EyePACS-1 and Messidor-2 dataset (retinal images)	DCNN	94–96%
Quinn	2016	Microscopic image	DCNN and shaped features like moment and morphological	100% for malaria; 99% for tuberculosis and hookworm
Korsuk	2016	ADNI dataset Alzheimer disease using dataset of PET, MRI, PET, and MRI	DBM	92.38% 92.20% 95.35%
Sarraf	2016	fMRI and MRI images	CNN	97.77% and 100% for fMRI and MRI
Iidaka T.	2015	fMRI data	Deep learning classifiers	90%

Table 57.2 Findings of various recommendation systems

Author (s)	Year	Approach	Findings
Joseph A. Konsta John Riedl	2012	Enhancing user experience by including algorithm in recommender system	Reported advances in collaborative filtering recommender systems, progression from research concentrating on the ironic set of queries concerning the consumer involvement with the recommender
Xun Zhoua, Jing Heb, Guangyan Huang, Yanchun Zhang	2015	Proposal of singular value decomposition-based incremental algorithm	Error analysis for demonstrating the efficiency of the performance of incremental ApproSVD algorithm
Ashwin Belle, Raghuram Thiagarajan	2015	Three encouraging and impending capacities of medicinal investigation: image, genomics and gesture-centered Big Data analytics in healthcare	Enormous capacity of medicinal statistics directed can be developed by joining multimodal records from distinct foundations
Lidong Wang and Cheryl Ann Alexander	2015	Conversion of innovative analyses for Big Data	Big data has unlimited prospective to develop medicine; monitor clinicians in supplying value-based carefulness
Carlos Luis Sánchez Bocanegra	2015	Used SNOMED-CT and bio-ontology semantic skills to endorse healthiness websites	Mostly all websites recommending health videos were relevant provided through semantic skills
Alejandro Baldominos, Fernando De Rada, Yago Sae	2017	Apache Spark, an open-source collection workout structure	Proposed framework is intelligent to recover facts from devices which are set up in the healthcare center places and form appropriate info
Mohamed Hussein Abdi	2017	A study on context-aware healthcare recommender systems point out that a socio-technical issue of privacy, security, and trust is emerging	The incorporation of contextual information is limited even though it is suggested as a key ingredient to improving the quality of the recommendations and the accuracy of the predications

(continued)

Table 57.2 (continued)

Author (s)	Year	Approach	Findings
Hanna Schäfer	2017	Personalization of recommender systems 1. On the basis of customers trust and 2. Evaluation approaches and actions (consumer fulfillment) in HRS	Helping users with customized, complex medical support or interventions with precautionary healthcare measures
Donghui Wanga	2018	Web Crawler approach to regularly keep informed the keeping fit set and the knowledge prototypical	Proposed hybrid model centered on softmax regression and chi-square feature collection provides efficient interactive responses online
Gourav Bathla, Himanshu Aggarwal	2018	Recommendation approach centered on deep knowledge and outsized scale graph splitting	Deep knowledge and outsized scale graph splitting approaches provide better recommendation accuracy for large-scale social data
Jun Yi LIU	2018	A recommendation approach to extract customers and objects topographies centered on deep learning neural system	Proposed recommendation approach focuses on put on in the real-world situations for better accuracy, cost competence, and low resource depletion in real-world use
Hanafi, Nanna Suryana	2018	A recommendation approach centered on application field arrangement using deep learning	Provides enhanced performance to capture appropriate consciousness concerning examination, feedback, theoretical and product description
K. U. Kala, M. Nandhini	2018	Applicability of deep knowledge practices in recommender schemes	Deep learning practices are very effective in the area of recommendation schemes
Hanaa El Fazazi, Mohammed Qbadou	2018	A recommendation approach centered on grouping, collecting, and relationship rules	Proposed recommendation approach based on three modules, a learner module, a domain module, and a recommendation module delivers good results
Shadi Alian, Juan Li	2018	Personalized recommendation system for diabetic American Indians	Provides personalized recommendations such as food intake and physical workout for AI patients based on the extraordinary socioeconomic, educational, and environmental status

(continued)

Table 57.2 (continued)

Author (s)	Year	Approach	Findings
Zeynep Batmaz, Ali Yurekli,	2018	A approach to examine the relation among deep knowledge classifications and purposive things of recommendation schemes	Efficiently categorizes the journals into the conforming recommender arrangement category
Guanjie Zheng, Fuzheng Zhang, Zihan Zheng, Yang Xiang	2018	News recommendation framework using deep Q-learning	Efficiently captures consumer response data by customer arrival pattern by click/no click tag
Puja Deshmukh, Dr. Geetanjali Kale	2018	A music recommendation system built on content, collaborative, metadata, and emotion-based recommendation	Compared to other recommendation systems, emotion and context grounded model give high quality of recommendation by seeing the community data
Liao liang Jiang	2018	A reliance constructed Ecommerce recommendation scheme using collaborative filtering scheme	Deployment of slope one algorithm in various recommender systems based on the combination of user similarity and trusted data
OlegRybakov, Vijai Mohan	2018	Formulating recommendation problem to predict the future behavior by encoding historical behavior using soft data split, combining predictor and auto-encoder models	Learning the significance (time decay) of the purchases depending on the purchase date by introducing convolutional layer and indicating that the shape of the time decay function can be approximated by a parametrical function
Georgia Koutrika	2018	Focus on multi-armed bandits, matrix factorization, and methods for amalgamation of recommendations	Matrix factorization provides good performance in terms of recommendation excellence and scalability
Sundus AyyazID, Usman Qamar	2018	A hybrid recommendation scheme built on filtering system with fuzzy extrapolation system	Proposed hybrid recommendation scheme applied in the area of cinemas provides worth recommendations to customers with a sureness level and a better accurateness

(continued)

Table 57.2 (continued)

Author (s)	Year	Approach	Findings
Qiwei Han	2019	Temporal dynamics of patient–doctor relationships using consultation histories	Proposed model displays greater extrapolative accurateness than both a collaborative filtering method and a heuristic model

57.4 Conclusion

Collaborative filtering is reflected as more superior to other tree approaches (contents-centered, knowledge-centered, and demographic filtering). A deep learning methodology for collaborative and content-centered methods will permit the traditional model to acquire dissimilar topographies of customers and things routinely to progress the accuracy of recommendations. Collaborative filtering recommendation procedure with deep learning technology provides good accuracy. This model uses a feature extraction scheme built on a quadric polynomial regression model, which gets the hidden topographies further precisely by old-style matrix factorization procedure. Then, these hidden features are considered as the input of the deep neural system. Implementation of hybrid recommendation procedure using deep learning methods to lower time complexity and to increase the accurateness of formed expectations is our immediate future work.

References

1. Adomavicius, G., Tuzhilin, A.: Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions. IEEE Trans. Knowl. Data Eng. **17**(6), 734–749 (2005). http://doi.org/10.1109/TKDE
2. mySugr Bolus Calculator. Available in Diabetes Healthcare. Available: https://mysugr.com/mysugr-boluscalculator/ (2018)
3. Vlahu-Gjorgievska, E., Trajkovik, V.: Personal healthcare system model using collaborative filtering techniques. Int. J. Adv. Inf. Sci. Serv. Sci. **3**(3), 64–74 (2011)
4. Konstan, J.A., Riedl, J.: Recommender Systems: From Algorithms to User Experience. Springer Science+Business Media B.V., Berlin (2012). http://doi.org/10.1007/s11257-011-9112-x
5. Sahu, S.P., Nautiyal, A., Prasad, M.: Machine learning algorithms for recommender system—a comparative analysis. Int. J. Comput. Appl. Technol. Res. **6**(2), 97–100 (2017). ISSN: 2319-8656

Chapter 58
New Multiphase Encryption Scheme for Better Security Enhancement

Rajalaxmi Mishra, Jibendu Kumar Mantri, and Sipali Pradhan

Abstract With the progression of technology, nowadays, it becomes essential to communicate through unsecure open network. Several researchers have proposed various methods to conceal the messages in order to securely transfer them. In this paper, we propose a new multiphase encryption method which combines the modified Caesar cipher, Atbash Cipher, and XOR cipher, where the encryption is done in multiple phases. It also utilizes the polyalphabetic substitution concept to enrich the encryption process.

58.1 Introduction

With the progression of technology, nowadays, it becomes an integral part of life to communicate through the Internet. All e-commerce and banking transactions are taking place digitally. Hence, in the modern digital world, it becomes utmost important to protect the sensitive information. The transmitted information can easily be acquired by the intruders from the vulnerable and open networks. These types of attacks can be reduced by communicating the information in an encrypted form. To conceal the confidential information from the outside world, it becomes necessary to keep inventing novel, efficient cryptographic methods which can protect the sensitive information while transmitting through open and unsecure public networks. Several researchers have presented remarkably secure and strong cryptographic schemes like data encryption standard (DES) [1, 2], 3DES [3, 4], AES [5], blowfish [6], and RSA [2]. The implementation complexity of these ciphers is quite high, and also they consume more resources.

There are several cryptographic schemes available. One of the important, simplest, and monoalphabetic substitution ciphers is the additive cipher. It is also called as shift cipher [7]. Julius Caesar communicated secret messages with his troops by using an

R. Mishra · J. K. Mantri (✉)
Department of Computer Application, North Orissa University, Baripada, Orissa, India

S. Pradhan
Department of Computer Science, RBVRR Women's College, Hyderabad, India

T. Senjyu et al. (eds.), *IOT with Smart Systems*, Smart Innovation, Systems and Technologies 251, https://doi.org/10.1007/978-981-16-3945-6_59

additive cipher with 3 as the key value. Hence, additive cipher is also called as Caesar cipher. But these ciphers are vulnerable to cipher text only attack and brute force attack because the domain of the key is very small. The key can take any of 25 (1–25) values, so it is possible for the intruder to get the value of the key [8]. Many researchers have proposed several encryption schemes to improve the security of Caesar Cipher. The algorithm [9] is proposed in a modified way to utilize the basic concepts of affine ciphers, transposition ciphers, and randomized substitution techniques to generate a cipher text which is quite difficult to decipher. To ensure enhanced security, a complex technique of key generation is used to generate two keys from a single key. The encryption scheme [10] combines the concept of Caesar cipher and rail fence cipher. It employs an automatic key generation technique to generate the key for the Caesar cipher. It provides security for the data containing case sensitive alphabets, numbers, and special characters. The modified method of the Caesar cipher [11] produces readable cipher text. The encryption process is done by dividing the alphabet into two parts; the vowels and consonants. The vowels were replaced with the vowels, and the consonant alphabets were substituted with consonants. The text is readable, and the cryptanalyst would not be suspicious of the cipher text. An encryption scheme [12] combines the modified Caesar cipher with the transposition cipher. The encryption operation is performed three times. The modified Caesar cipher is first applied on the plain text, the resultant cipher text is then enciphered by the transposition cipher, and finally, the encrypted text is again encrypted with Caesar modification cipher. The Caesar cipher if modified by including the characters of the ASCII set (32–126 characters). It generates the key dynamically by choosing the key as the ASCII value of one plaintext character. For different plaintext, the key value is different. An encryption system [13] encrypts and decrypts the plain text by using the values of the index and value of the position of the corresponding characters as the key. The scheme is a multistage encryption scheme which improves the security of the plain text and protects it from brute force attack. An encryption scheme in [14] proposes a modification to the classical Caesar cipher scheme where the key is fixed to one. It checks the alphabet index; if it is even then the character value is increased by one; otherwise, the character value is decreased by one. The symmetric key encryption scheme, based on the modified Caesar cipher [15], encodes the actual message into secret message by using mathematical principles. The proposed technique produces the cipher text which includes special characters. The algorithm [16] is a modification to traditional Caesar cipher which uses variable key for each character. It uses a variable length key whose value depends on the length of the plain text string, the place of character in the string and time of file. The receiver gets the time along with name of file, and if the receiver has appropriate key for decryption, the message can be deciphered. An encryption scheme [17] utilizes the basic techniques of substitution and transposition to encrypt. To present a more secure and stronger cipher, it uses a single columnar transposition followed by a double substitution based on Caesar cipher.

The Caesar cipher method may be modified to add more complexity to the system. The keys can be generated in such a way that the value of the key depends on the plain text. The key values are different for different plain text. Hence by using brute force attack, the intruder cannot get the keys. To make the encryption scheme robust, the Caesar cipher may be combined with other ciphers. Some researchers have proposed encryption schemes which combine Caesar cipher with other ciphers to improve security of the system. The scheme [18] proposed an enhanced encryption method by combining modified Vigenere with Caesar cipher. The paper [19] proposed a secure communication program which involves three design stages, i.e., the encryption, serial transmission, and encoding stages. The encryption technique adopts a combination of Caesar cipher and XOR encryptions.

In this paper, we proposed a multiphase encryption method. It combines some of the enhanced concept of Caesar cipher described in [10, 13, 14] with XOR operation and Atbash cipher implemented in different phases. It is a polyalphabetic substitution scheme. The key values are generated from the plain text. This paper proposes the modification to Caesar cipher by combining its different enhancements to make the system robust.

58.2 Our Proposed Scheme

58.2.1 Theoretical Background

The symmetric key cryptography uses the same key to encrypt and decrypt. The key needs to be shared by the sender and the receiver through a secured channel. P is the plain text, C is the ciphered text, and K is the key. The sender uses the encryption algorithm to produce the ciphered text by applying the key on the plain text. Then the ciphered text is sent to the intended receiver who shares the key with the sender through the unsecured open network. The receiver then uses the decryption algorithm and by applying the key on the ciphered text can produce the plain text.

Encryption Algorithm:

C = Encption(P,key).

Decryption Algorithm:

P = Decryption(C,key) = Decryption(Encption(P,key),key)).

The traditional symmetric key ciphers are broadly classified into substitution ciphers and transposition ciphers. The substitution cipher is again classified as monoalphabetic substitution cipher and polyalphabetic substitution cipher.

The additive cipher is one of the monoalphabetic substitution cipher also called as shift cipher. It uses modular arithmetic in the encryption and decryption process.

If P is the plain text, C is the cipher text and the shared secret key is K, then

Encryption process is to get the ciphered text as $C = (P + K) \bmod 26$ and

Decryption process is to get the plain text as $P = (C - K) \bmod 26$.

Caesar cipher is a type of additive cipher where the key value is 3

Encryption process in Caesar cipher is to get the ciphered text as $C = (P + 3)$ mod 26 and

Decryption process in Caesar cipher is to get the plain text as $P = (C - 3)$ mod 26

For example, the word 'cryptography' is encrypted as 'fubswrjdskb.'

58.2.2 The Scheme

We are proposing one encryption scheme where the encryption process occurs in multiple phases:

Phase 1 is using Atbash cipher in a modified form.
Phase 2 is using XOR cipher
Phase 3 is using Caesar cipher in an enhanced way.

The operations at sender side:

1. Accept the message that has to be encrypted and send to the receiver.
2. Generation of the keys
3. Share the keys
4. Encryption process by using
   - i. Atbash cipher
   - ii. XOR cipher and
   - iii. Modified Caesar cipher.
5. Send the encryption message to the receiver.

The operation at the receiver side:

Receive the shared keys.
Receive the encrypted text.
Decrypt the text by using
Modified Caesar cipher

XOR cipher and finally
Atbash cipher.

The Key generation Process

After getting the message which has to be send to the receiver, the ASCII values of the characters of the message have to be determined. Then the keys are being generated as:

sum = summation of the ASCII values of all characters in the message

The key $k1$ = (int)sum/number of characters in the message
The key $k2$ = sum%256

The key $k3 = \text{sum}$.

The encryption process is carried out in multiple phases:

Phase I uses the Atbash Cipher

About 500–600 BC, the Hebrew scholars used a cipher based on simple substitution method known as Atbash cipher [18]. It works by reversing the alphabets, i.e., the letter 'a' is replaced by the letter 'z,' 'b' the second letter is replaced by the letter 'y,' the second letter from the end, and so on. For example, the word 'cryptography' is replaced by 'xibkgltizksb.'

The Atbash cipher is a special form of the affine cipher. The character of an alphabet having m letters is mapped to the numbers 0, 1, …, $m - 1$. (The Hebrew alphabet has $m = 22$, and the standard Latin alphabet has $m = 26$). The encryption and decryption function of Atbash cipher in the form of an affine cipher can be written by setting $a = b = (m - 1)$;

$$E(x) = D(x) = ((m-1)x + (m-1)) \mod m = (m-1)(x+1) \mod m = -(x+1) \mod m$$

In the proposed scheme, we are using the ASCII codes, where the value of $m = 256$. Hence, the encryption and decryption function for Atbash cipher is

$E(x) = D(x) = -(x+1) \mod 256$.

Hence, $C1 = E(P) = -(P + 1) \bmod 256$

Phase II uses XOR cipher with key value k2

Hence, $C2 = C1 \oplus k2$.

Then in phase III, a modified version of Caesar cipher is used,

$C2 = C2 - 32$.

$C3 = (C2 + k1) \bmod 256$.

$C4 = (\text{sum} - C3) \bmod 256$.

$C5 = (C4 + \text{position of the character in the message}) \bmod 256$.

$C5 = C5 + 32$.

To avoid non-printable ASCII characters, 32 is being subtracted.

$C5$ is the final ciphered text which has to be communicated from the sender to the receiver.

The decryption process is just the reverse of encryption process:

The modified Caesar cipher is applied in decryption phase I.

C is the ciphered text.

$C = C - 32$

$T1 = (C - \text{position of the character in the message}) \bmod 256$

$T2 = (\text{sum} - T1) \bmod 256$

$T3 = (T2 - k1) \bmod 256$

$T3 = T3 + 32$.

In decryption phase II, XOR cipher is applied.

$T4 = T3 \oplus k2$.

Then, finally, the decryption phase III applies Atbash cipher.
Plain text = − ($T4$ + 1) mod 256.

58.3 Result and Discussions

The message, secret message has to be sent to the receiver. Before sending the message through the unsecured network, it needs to be encrypted.

The ASCII values for the secret message are:

83, 101, 99, 114, 101, 116, 32, 77, 101, 115, 115, 97, 103, 101.

The sum = 1355.

$K1$ = avg = (int) sum/count of characters = 96.

$K2$ = sum%256 = 75.

The encryption process (Table 58.1):

The decryption process is the reverse of the encryption process (Table 58.2).

This proposed scheme is secured from brute force attack at the key values which are dependent on the plain text values. So for different plain text, the key would be different. It is also secured from chosen cipher text attack, because the scheme uses a polyalphabetic substitution where the same character can be encrypted to a different character. As it is mentioned in the above example, e is encrypted to [, ˆ, b and g,

Table 58.1 Encryption process

Plain text	S	e	c	r	e	t	Space	M	e	s	s	a	g	e
ASCII value, P	83	101	99	114	101	116	32	77	101	115	115	97	103	101
Atbash cipher, C_1	172	154	156	141	154	139	223	178	154	140	140	158	152	154
XOR cipher, C_2	231	209	215	198	209	192	148	249	209	199	199	213	211	209
$C_2 = C_2 - 32$	199	177	183	166	177	160	116	217	177	167	167	181	179	177
$C_3 = (C_2 + k_1)$mod 256	39	17	23	6	17	0	212	57	17	7	7	21	19	17
$C_4 =$ (sum $- C_3$)mod 256	36	58	52	69	58	75	119	18	58	68	68	54	56	58
$C_5 = (C_4$ + pos) mod 256	36	59	54	72	62	80	125	25	66	77	78	65	68	71
$C_5 = C_5 + 32 = C$	68	91	86	104	94	112	157	57	98	109	110	97	100	103
Cipher text	D	[	V	h	ˆ	p	¥	9	b	m	n	a	d	G

Table 58.2 Decryption process

Cipher text	D	[	V	h	ˆ	p	¥	9	b	m	n	a	d	G
ASCII value, C	68	91	86	104	94	112	157	57	98	109	110	97	100	103
$C = C - 32$	36	59	54	72	62	80	125	25	66	77	78	65	68	71
$T_1 = (C - \text{pos}) \bmod 256$	36	58	52	69	58	75	119	18	58	68	68	54	56	58
$T_2 = (\text{sum} - T_1) \bmod 256$	39	17	23	6	17	0	212	57	17	7	7	21	19	17
$T_3 = (T_2\text{-}k_1) \bmod 256$	199	177	183	166	177	160	116	217	177	167	167	181	179	177
$T_3 = T_3 + 32$	231	209	215	198	209	192	148	249	209	199	199	213	211	209
XOR cipher, T_4	172	154	156	141	154	139	223	178	154	140	140	158	152	154
Atbash cipher (P)	83	101	99	114	101	116	32	77	101	115	115	97	103	101
Plain text	S	E	C	r	e	t	space	M	E	s	s	a	g	E

similarly, s is encrypted to m and n. Hence, it would be difficult for an intruder to get the plain text message from the cipher text.

58.4 Conclusion and Future Work

Our proposed scheme uses multiple phases to encrypt the message by employing polyalphabetic substitution scheme. In the first phase, it used Atbash cipher, second phase used XOR cipher, and in the last phase, modified Caesar cipher is used. The value of the key for Caesar cipher is generated from the plain text value. Hence, the proposed scheme is providing security against brute force and chosen cipher text attacks. The key exchange scheme can also be incorporated to it as the future work.

References

1. Murali Krishna, B., Khan, H., Madhumati, G.L., Praveen Kumar, K., Tejaswini, G., Srikanth, M., Ravali, P.: FPGA implementation of DES algorithm using DNA cryptography. J. Theor. Appl. Inf. Technol. **95**(10) (2017). ISSN: 1992-8645

2. Singh, G.: A study of encryption algorithms (RSA, DES, 3DES and AES) for information security. Int. J. Comput. Appl. **67**(19) (2013)
3. Abdelwahab, M.M.: A compact cryptosystem design of triple-DES. UKEJ **8**(1), 25–29 (2018)
4. Abood, O.G., Guirguis, S.K.: A survey on cryptography algorithms. Int. J. Sci. Res. Publ. **8**(7) (2018). ISSN 2250-3153
5. Chowdhury, Z.J., Pishva, D., Nishantha, G.G.D.: AES and confidentiality from the inside out. In: Advanced Communication Technology (ICACT), The 12th International Conference on IEEE, vol. 2, pp. 1587–1591 (2010)
6. Singh, G., Kumar, A., Sandha, K.S.: A study of new trends in blowfish algorithm. Int. J. Eng. Res. Appl. (2011)
7. Ferozan, B.A., Mukhopadhyay, D.: Cryptography and Network Security, 2nd edn. McGrawHill Education (India) Private Limited (2013)
8. Stallings, W.: Cryptography and network security, 4th edn. Prentice Hall, Pearson (2000)
9. Jain, A., Dedhia, R., Patil, A.: Enhancing the security of Caesar cipher substitution method using a randomized approach for more secure communication. Int. J. Comput. Appl. **129**(13), 0975–8887 (2015)
10. Bazith Mohammed, B.: Automatic key generation of Caesar cipher. Int. J. Eng. Trends Technol. (IJETT) **6** (2013). ISSN: 2231-538
11. Purnama, B., Hetty Rohayani, A.H.: A new modified Caesar cipher cryptography method with legible ciphertext from a message to be encrypted. Procedia Comput. Sci. **59**, 195–204 (2015)
12. Lubis, F.I., Simbolon, H.F.S., Batubara, T.P., Sembiring, R.W.: Combination of Caesar cipher modification with transposition cipher. Adv. Sci. Technol. Eng. Syst. J. **2**(5), 22–25 (2017)
13. Pinheiro, G., Saraf, S.: IJCST **7**(1) (2016). ISSN: 0976-8491 (Online) | ISSN: 2229-4333 (Print)
14. Goyal, K., Kinger, S.: Modified Caesar cipher for better security enhancement. Int. J. Comput. Appl. **73**(3), 0975–8887 (2013)
15. Ilayaraja, M., Shankar, K., Devika, G.: A modified symmetric key cryptography method for secure data transmission. Int. J. Pure Appl. Math. **116**(10), 301–308 (2017). ISSN: 1311-8080 (printed version); ISSN: 1314-3395 (on-line version), Special Issue
16. Patni, P.: A poly-alphabetic approach to Caesar cipher algorithm. Int. J. Comput. Sci. Inf. Technol. **4**(6), 954–959 (2013). ISSN: 0975-9646
17. Dar, S.B.: Enhancing the security of Caesar cipher using double substitution method. Int. J. Comput. Sci. Eng. Technol. **5**(07) (2014). ISSN: 2229-3345
18. Gautam, D., Agrawal, C., Sharma, P., Saini, P., Mehta, M.: An enhanced Cipher technique using Vigenere and modified Caesar cipher. In: 2nd International Conference on Trends in Electronics and Informatics (ICOEI) (2018)
19. Han, L.C., Mahyuddin, N.M.: An implementation of Caesar cipher and XOR encryption technique in a secure wireless communication. In: 2nd International Conference on Electronic Design (ICED), 19–21 Aug, Penang, Malaysia (2014)

Chapter 59
A Novel Deep Ensemble Learning Framework for Classifying Imbalanced Data Stream

Monika Arya and G. Hanumat Sastry

Abstract Machine learning has emerged as one of the most indispensable fields of this era in mining stream data. The stream data have the tendency to change their characteristics over time. Mining imbalanced stream is a research demanding subfield of this area. In imbalanced data, one of the target classes has much less instances than other class. The imbalanced data may differ either in their ratio between majority and minority class or dimension or the number of classes. The performance of the classifier is affected by variety of imbalanced data set used for training the classifier. Because, the learning of classifier is different for different types of data sets. Imbalanced data can result due to rare events which can have negative impact on society. Most traditional data mining algorithm misclassifies the minority class of the imbalance data sets or considers it as noise. Therefore, the decision is biased toward majority class and hence reduces the accuracy and overall performance of the algorithm. The algorithms for classifying imbalanced data sets thus demand high adaptability to changes in the majority and minority class ratios. The performance of traditional machine learning algorithms is enhanced by applying ensemble method and deep learning approaches. *Purpose*: This paper proposes a framework based on deep ensemble learning for classifying imbalanced stream of data. *Methodology*: The deep ensemble methods ensemble multiple base learners. While deep learning approach is applied to improve the performance by extracting lower-level features and feeding them forward for the next layer to identify higher level features. *Results*: In this method, the effect of highly imbalanced classes is reduced by uniting the ensemble method with deep learning. The accuracy of the classifier is improved in terms of not only accuracy but also categorical accuracy. *Conclusion*: In addition to accuracy, other performance measures like categorical prediction accuracy, training accuracy and prediction accuracy have also been compared which are crucial metrics for evaluating imbalance stream of data.

M. Arya (✉)
Bhilai Institute of Technology, Durg, Chhattisgarh, India

G. Hanumat Sastry
University of Petroleum and Energy Studies, Dehradun, India

T. Senjyu et al. (eds.), *IOT with Smart Systems*, Smart Innovation, Systems and Technologies 251, https://doi.org/10.1007/978-981-16-3945-6_60

59.1 Introduction

Imbalanced data are the data set where the number of observations in different classes is not evenly distributed, i.e., observations across the classes are biased or skewed. The class having a greater number of observations is referred as majority class, while the class having fewer observations is called minority class. This can occur not only in binary class but also multi-class data. Real-world data set like fraud detection, anomaly detection, medical diagnosis, oil spillage detection, facial recognition, stock market data, etc., are the examples of imbalanced data where the rare events such as fraudulent transactions in credit card transaction are the minority class as most of the transactions will be 'not fraud'. Machine learning algorithms were generally designed by assuming that the data are balanced and the examples are evenly distributed among the available classes. Thus, classifying imbalanced data is posing significantly new challenges for traditional machine learning classifiers as they result in poor predictive performance, specifically for minority classes. There can be various causes of class imbalance and perhaps can be grouped as data sampling and properties of the domain. The imbalance in the distribution of observations across each class can be caused by the way they are sampled. Like data collected from small geographical region or slice of time or the way of collecting the data. This might introduce bias in collected data or errors might be made during data collection. The imbalance caused by a data sampling bias can be corrected by using improved data sampling methods and/or correcting the measurement error. Secondly, the problem domain might possess the property that naturally the occurrence of one class may dominate the other classes. This can be due to various reasons like: (1) the cost of generating the observations of minority class can be expensive or more time consuming, (2) it may require more computation, (3) it may be infeasible, etc. As mentioned earlier, most of the traditional machine learning algorithms assume data to be balanced. These models focus on the majority class characteristics and neglect the minority class examples. But minority class examples are, in fact, of more interest and whose predictions are more valuable in certain critical areas like medical, credit card frauds, etc. As a result, the traditional machine learning algorithms may cause bad classification of less frequent class and reduce the overall performance and accuracy of the classifiers. For sensitive data like credit card fraud detection, medical diagnosis or stock market predictions, the misclassification might be expensive or might have cause serious consequences.

59.2 Ensemble Learning Methods for Imbalanced Data Stream

This section analyzes a number of ensembles learning methods for classifying imbalanced data streams. For each surveyed work, we have summarized the concepts and the methodology implemented by the authors to improve the performance of the classifier. In the later part of the section, challenges are highlighted, and open issues for designing the new algorithm for learning from imbalanced data stream are discussed.

Zyblewski et al. [1] proposed a classifier which updates on new arrival of data. The proposed approach proved to be useful in non-stationary data streams scenario. Klikowski et al. [2] in their work introduced a novel *multi-sampling random subspace ensemble (*MSRS), an ensemble method for imbalanced non-stationary data stream classification and was having an appropriate diversity of the classifiers in ensemble. Sun et al. [3] in their work developed an under-sampling bagging framework in which a new fitness function is designed, and it can generate an ensemble of classifiers. The performance is evaluated, and superiority of the classifier is demonstrated on the basis of not only accuracy but also recall, geometric mean and AUC. Zhang et al. [4] in their work employed differential evolution algorithm to optimize the weight of base classifiers in the ensemble. The approach showed performance improvement as compared with the simple vote-based ensemble method and non-ensemble method. Authors (Yijing et al. [5]) in their work proposed a classifier AMCS, which is an adaptive classifier and is used for multi-class classification of imbalanced dat. Arabmakki et al. [6] used partially labeled samples to develop reduced labeled samples-ensemble (RLS-E) for imbalanced data stream. This method predicts the labels either by combined output of all the classifiers of the ensemble or by individual classifiers of the ensemble, and the more accurate predictions among the two are selected as the final prediction. The method has been tested on many real-world data sets and proved to produce better performance. Wang et al. [7] in their work developed a multi-objective ensemble method (MOSOB). Experimental results on five real-world data applications show that MOSOB performs well in both static and dynamic data streams. Wang et al. [8] proposed two new ensemble methods that maintain both OOB and UOB with adaptive weights for final predictions and showed improved performance. Ghazikhani et al. [9] worked with an online ensemble of neural network (NN) classifiers consisting of two layers. One layer is of learning which is embedded into training phase of NNs, and the second layer is the ensemble of classifiers. This approach was used for handling the class imbalance of non-stationary data streams.

We have experimentally compared several states-of-art ensemble classifiers on synthetic imbalanced data for binary classification problem. We have performed the study for 10,000 samples with imbalance ratio of 99:1. The ensemble classifiers are compared on basis of ROC AUC score (Table 59.1 and Fig. 59.1).

Challenges in classifying imbalanced data with standard machine learning techniques

Traditional machine learning algorithms like decision tree, logistic regression, etc., are biased toward majority class, and often, they treat minority class features as noise and ignore them. Therefore, the probability of misclassification of minority classes is high.

1. The traditional machine learning techniques evaluate the performance of a classifier by confusion matrix which only gives the information about actual and

Table 59.1 Comparison of ensembler by ROC AUC score

Ensembler used	Algorithm	ROC AUC score
Bagging	Standard bagging	0.871
	With random under-sampling	0.962
Random forest	Standard RF	0.869
	RF with class weighting	0.871
	RF with bootstrap class weighting	0.884
	RF with random under-sampling	0.970
Easy ensemble	Easy ensemble	0.968

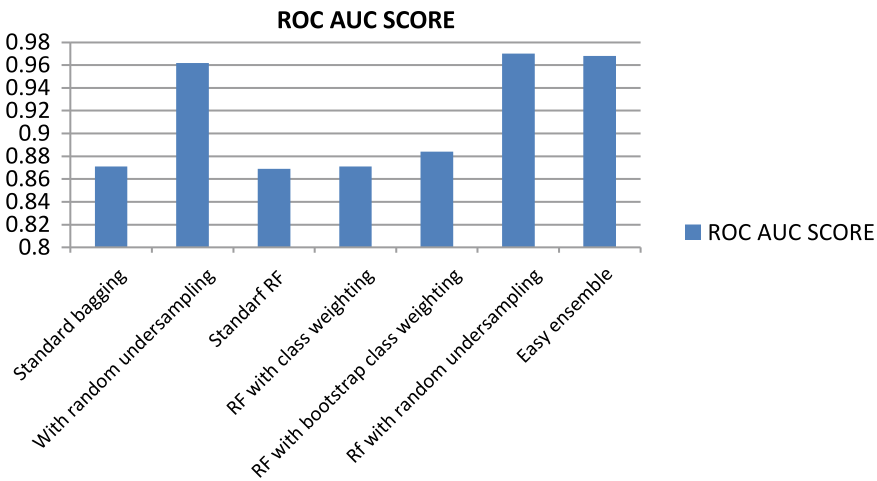


Fig. 59.1 Graph of ROC AUC score

predicted class. However, for imbalanced data set, accuracy is not enough to measure the performance of the classifier.

2. The imbalance of the class may vary across problem domain. It may range from slight imbalance to severe imbalance. The traditional machine learning techniques concentrates on imbalance ratio ranging between 1:4 and 1:100. However, in real-life application like fraud detection, the imbalance ratio may range from 1:1000 to 1:5000. Such severe imbalance is challenging to model and requires specialized techniques.

Open issues need to be addressed when designing new algorithms for classifying imbalanced data streams

Following are the open issues which are important to be addressed when designing new algorithms for classifying imbalanced data streams:

1. One of the most challenging issues while designing new algorithms for imbalanced data is dealing with new emerging class and/or old fading class. In both the cases, naturally the class which is newly emerging or the old class which is disappearing will be the minority class.
2. Another vital challenge in classifying imbalanced data stream is the availability of class labels. A major drawback with the traditional machine learning algorithms is that it assumes availability of class label as soon as new sample is classified. However, in the real-life scenario, it is not true and also imposes labeling cost on the system. So, while designing the new algorithm for imbalanced data stream, new labeling strategies must be developed.
3. A very common recurring drift problem in data stream caused due to re-appearing sources can also be the reason of class imbalance. As such drifts occur again and again over time. Developing the dedicated methods for such scenario and storing general solutions for re-appearing class imbalance are required.

Real-life applications face the data imbalance problem and therefore motivate the learning from imbalanced data. Due to a smaller number of available examples, the minority classes are harder to predict. As the traditional machine learning algorithms assume classes to be evenly distributed and focus on the learning the characteristics of majority class only, it is a challenge for standard machine learning techniques to learn from the characteristics of the minority class and to differentiate it from the majority class.

Deep learning strategies for handling imbalanced data streams

Deep learning employs two strategies for handling imbalanced data sets. It can be either by improving the classification algorithms or by balancing the imbalance data set.

I. **Balancing techniques for imbalanced data stream**

Deep learning uses two techniques for balancing the imbalanced data set:

(1) Balancing the weight
(2) Sampling.

In weight balancing, the weight of the training example is when computing the loss. Normally, while calculating the loss function, it is considered that each example and class will carry equal weight. But sometimes certain classes or certain examples are more important for example in credit card fraud detection data set, and the examples of fraud class are of more importance than that of non-fraud class. So, we might want that these examples hold more weight. Selecting the class weight makes the process complicated. Another way to balance our data is via sampling. In under-sampling, only *some* of the data from the majority class will be selected, i.e., only using as many examples as the minority class has. In oversampling, *copies* of minority class data are created in order to have the same number of examples as in the majority class.

We have experimentally compared base model, weight balancing model and sampling model based on the performance. For the experiment, we have used credit card fraud detection data set hosted on Kaggle having 284,807 instances with 492

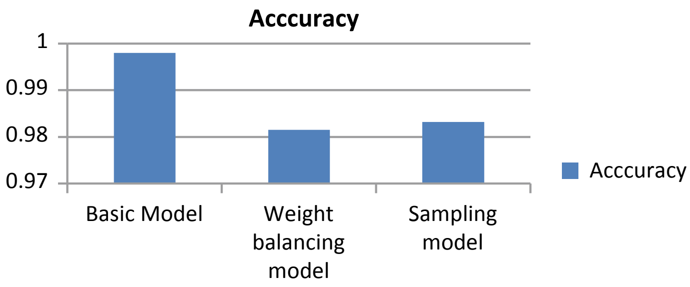


Fig. 59.2 Graph comparing accuracy

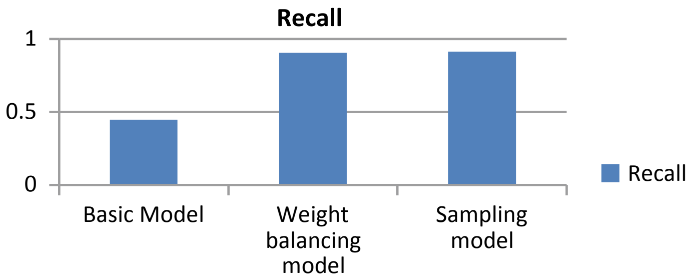


Fig. 59.3 Graph comparing recall

Table 59.2 Comparison of base model, weight balancing model and sampling model

Model	Loss	Accuracy	Precision	Recall	AUC	tp	fp	tn	Fn
Basic model	0.00553	0.998	0.8793	0.4473	0.9161	51	7	56,841	63
Weight balancing model	0.07716	0.9815	0.0911	0.9035	0.9790	103	1040	55,808	11
Sampling model	0.0764	0.9832	0.0990	0.9122	0.9780	104	946	55,902	10

fraudulent examples, i.e., only 0.17% of total instances. Thus, the ratio of imbalance is 492:284,807 (Figs. 59.2, 59.3 and Table 59.2).

Here, it is clearly visible that basic model has more accuracy than weight balancing model and sampling model, but the recall value is less. This clearly indicates that the ability of weight balancing and sampling model to detect the false positive case is more than the basic model which is very crucial in imbalance class.

II. **Algorithmic approaches for imbalanced data stream**
Table 59.3 summarizes the deep learning approaches proposed for classifying imbalanced data stream along with the type of data set in which the proposed algorithm is tested and the advantages and disadvantages of the algorithm.

Table 59.3 Algorithmic approaches for imbalanced data streams

S. No.	Title	Proposed algorithm	Description	Data set	Adv/Disadv
1	An efficient feature generation approach based on deep learning and feature selection techniques for traffic classification [10]	Feature optimization approach based on deep learning and feature selection (FS) techniques	used feature generation model based on deep weighted symmetric uncertainty (WSU) methods	Internet traffic data stream	High classification and runtime performance
2	Ensemble of online neural networks for non-stationary and imbalanced data streams [9]	An online ensemble of neural network (NN) classifiers	Used a two-layer approach for handling class imbalance and non-stationarity	Real-world data set	Improved performance as compared to online ensemble methods with similar features
3	Hybrid geometric sampling and AdaBoost-based deep learning approach for data imbalance in E-commerce [11]	AdaBoost-based deep learning classification approach	It used diverse solution to provide a balance among prediction, accuracy, precision, specificity, sensitivity and usability of data in E-commerce	E-commerce data	Reduce the data imbalance problem
4	An imbalance modified deep neural network with dynamical incremental learning for chemical fault diagnosis [12]	A novel incremental imbalance modified deep neural network (incremental-IMDNN)	It employed an imbalance modified method combined with active learning for the extraction and generation of the most valuable information	Tennessee Eastman (TE) data set	Better than existing methods and possesses significant robustness and adaptability
5	Ensemble classification for skewed data streams based on neural network [13]	Novel ensemble classification method (ECSDS) for classifying data streams with skewed class distributions	Back-propagation neural network used as the base classifier	UCI machine learning repository	Deals better with classification problems of imbalanced data streams

(continued)

Table 59.3 (continued)

S. No.	Title	Proposed algorithm	Description	Data set	Adv/Disadv
6	Dynamic sampling in convolutional neural networks for imbalanced data classification [14]	A novel model based on the convolutional neural networks (CNNs)	Used a dynamic sampling technique	Real-time visual data from public network cameras	Handles imbalanced and heterogeneous data

59.3 A Framework for Classification of Imbalanced Data Stream Based on Deep Ensemble Learning

In this work, a predictive (DEAL)/DEL framework with novel deep ensemble learning model is proposed. In this approach, we have combined deep learning and ensemble approach for classifying imbalanced data stream for classification in real-time data stream such as credit card transaction data and stock data. The objective in CCFD data is to classify the transactions as fraudulent and genuine, while in stock data, the objective is to obtain stock buy/sell decision for the next day to earn profits. The proposed framework is adaptive and robust.

The methodology for proposed framework is presented below:

1. **Feature Generation**: Depending on the type of model applications, suitable features are generated by applying various methods such as technical analysis for stock prediction.
2. **Feature Extraction**: In this phase, the best features are extracted using correlation-based analysis. The current best subset of features is chosen and sent for further processing and tensor transformation. Initially, variable size subsets of features are supplied.
3. **Tensor Transformation**: Before transformation, suitable preprocessing such as deletion of rows with empty values and data normalization is required. The data normalization step is necessary to fit the data on common scale. Deep learning considers the categorical and numerical data as numbers, and every neuron executes arithmetic operations on inputs and weights.
4. **Model Building**: We utilized a five-layer DEL model having three dense layers, an input and an output layer. The dense layers 1 and 2 utilize rectified linear unit (ReLU) as activation, and dense layer 3 utilizes sigmoid as activation function.

The model utilized in proposed framework is utilizing deep learning algorithm and extra-tree ensemble optimization for better and efficient prediction. The framework outcomes are predicted trends or decisions. The cost or loss function, to minimize, chosen is binary cross entropy/log loss represented using E(W). Other conventions are presented in

$$[E(W)] = -\frac{1}{m}\sum_{i=1}^{m} y_i \log(\gamma) + (1 - y_i)\log(1 - \gamma)$$

5. ***Extra-Tree* Ensemble Learning Optimization**
   The proposed method utilizes the number of features to optimize the probability scores. The majority vote is final prediction or classification and thus the best subset of features. The score function/cost function used is binary cross entropy/log loss (Fig. 59.4).

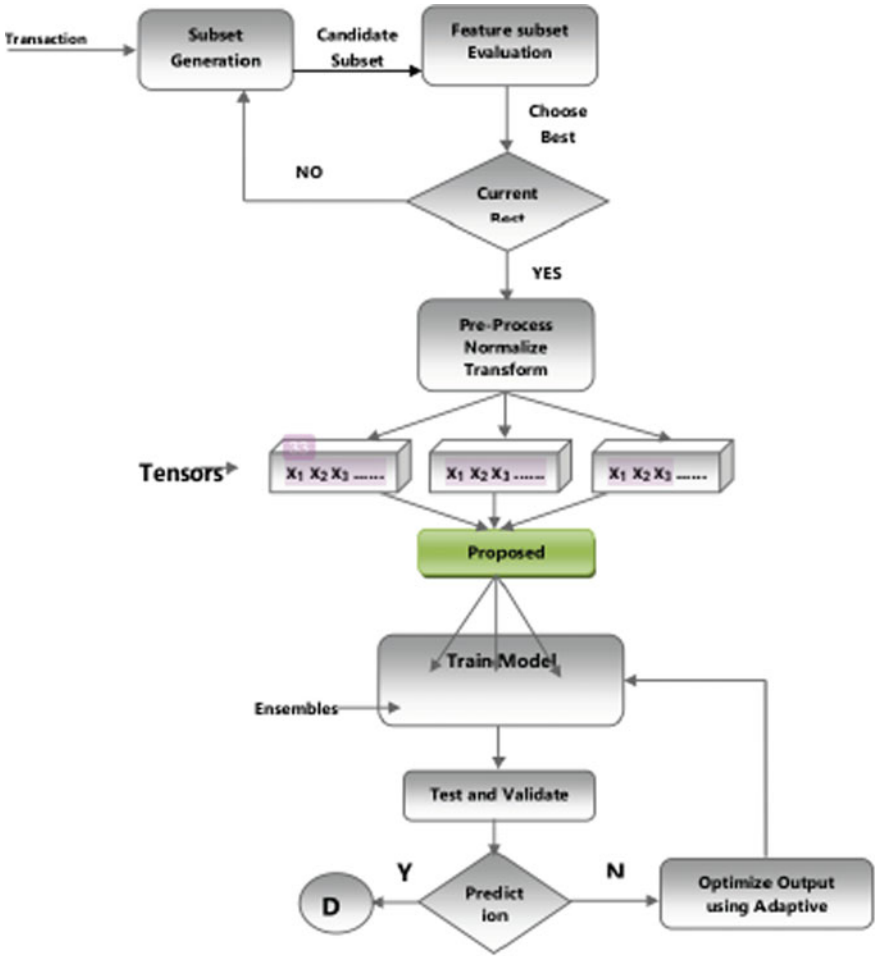


Fig. 59.4 Flowchart showing steps of DEAL

Table 59.4 Comparison of training accuracy, prediction accuracy and categorical prediction accuracy of CNN, MLP and DEAL

Model	Training accuracy (%)	Prediction accuracy (%)	Categorical prediction accuracy (%)	
			Legitimate	Fraudulent
CNN	98.80	97.87	97.81	82
MLP	97.93	96.94	96.79	81
DEAL	99.81	99.78	99.79	99.78

59.4 Experiments and Results

This section experimentally compares DEAL framework, CNN and MLP. For the experiment, we have used credit card fraud detection data set hosted on Kaggle having 284,807 instances with 492 fraudulent examples, i.e., only 0.17% of total instances. Thus, the ratio of imbalance is 492:284,807 (Table 59.4).

59.5 Conclusion

In this work, the performance of traditional ensemble algorithms for classifying imbalance data stream has been compared. We have summarized the issues for designing the new algorithms and future challenges in classifying imbalance data stream with traditional machine learning approaches. We have also studied the deep learning strategies for imbalance data classification and concluded that using balancing technique and algorithmic technique the imbalanced data streams can be handled. A novel deep ensemble learning framework for classifying imbalanced data stream has been proposed, and performance has been evaluated. In addition to accuracy, other performance measures like categorical prediction accuracy, training accuracy and prediction accuracy are also compared.

References

1. Zyblewski, P., Ksieniewicz, P., Woźniak, M.: Classifier selection for highly imbalanced data streams with minority driven ensemble. In: International Conference on Artificial Intelligence and Soft Computing, pp. 626–635. Springer, Cham (2019)
2. Klikowski, J., Woźniak, M.: Multi sampling random subspace ensemble for imbalanced data stream classification. In: International Conference on Computer Recognition Systems, pp. 360–369. Springer, Cham (2019)
3. Sun, B., Chen, H., Wang, J., Xie, H.: Evolutionary under-sampling based bagging ensemble method for imbalanced data classification. Front. Comp. Sci. **12**(2), 331–350 (2018)
4. Zhang, Y., Liu, B., Cai, J., Zhang, S.: Ensemble weighted extreme learning machine for imbalanced data classification based on differential evolution. Neural Comput. Appl. **28**(1), 259–267 (2017)

5. Yijing, L., Haixiang, G., Xiao, L., Yanan, L., Jinling, L.: Adapted ensemble classification algorithm based on multiple classifier system and feature selection for classifying multi-class imbalanced data. Knowl. Based Syst. **94**, 88–104 (2016)
6. Arabmakki, E., Kantardzic, M., Sethi, T.S.: Ensemble classifier for imbalanced streaming data using partial labeling. In: 2016 IEEE 17th International Conference on Information Reuse and Integration (IRI), pp. 257–260. IEEE (2016)
7. Wang, S., Minku, L.L., Yao, X.: A multi-objective ensemble method for online class imbalance learning. In: 2014 International Joint Conference on Neural Networks (IJCNN), pp. 3311–3318. IEEE (2014)
8. Wang, S., Minku, L.L., Yao, X.: Resampling-based ensemble methods for online class imbalance learning. IEEE Trans. Knowl. Data Eng. **27**(5), 1356–1368 (2014)
9. Ghazikhani, A., Monsefi, R., Yazdi, H.S.: Ensemble of online neural networks for non-stationary and imbalanced data streams. Neurocomputing **122**, 535–544 (2013)
10. Shi, H., Li, H., Zhang, D., Cheng, C., Cao, X.: An efficient feature generation approach based on deep learning and feature selection techniques for traffic classification. Comput. Netw. **132**, 81–98 (2018)
11. Dhote, S., Vichoray, C., Pais, R., Baskar, S., Shakeel, P.M.: Hybrid geometric sampling and AdaBoost based deep learning approach for data imbalance in E-commerce. Electron. Commerce Res. 1–16 (2019)
12. Hu, Z., Jiang, P.: An imbalance modified deep neural network with dynamical incremental learning for chemical fault diagnosis. IEEE Trans. Industr. Electron. **66**(1), 540–550 (2018)
13. Zhang, Y., Yu, J., Liu, W., Ota, K.: Ensemble classification for skewed data streams based on neural network. Int. J. Uncertain. Fuzziness Knowl. Based Syst. **26**(05), 839–853 (2018)
14. Pouyanfar, S., Tao, Y., Mohan, A., Tian, H., Kaseb, A.S., Gauen, K., et al.: Dynamic sampling in convolutional neural networks for imbalanced data classification. In: 2018 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR), pp. 112–117. IEEE (2018)

Chapter 60
Location-Based Crime Prediction Using Multiclass Classification Data Mining Techniques

Vishva Upadhyay and Dushyantsinh Rathod

Abstract A wrongdoing is an activity which represents a culpable offense by rules. It is destructive for community so as to anticipate the criminal movement, it is critical to recognize crime Information driven inquires about are valuable to anticipate and fathom wrongdoing. Up to date research appears that 50% of the wrongdoings are committed by as it were modest bunch of criminals. The law requirement executive requires quick data approximately the criminal movement to response and illuminate the spatio-temporal criminal movement. In this inquire about, supervised learning calculations are utilized to anticipate criminal movement. The proposed facts driven framework predicts wrongdoings by analyzing San Francisco city criminal activity dataset for 12 a long time. Decision tree and k-nearest neighbor (KNN) calculations are tried to anticipate wrongdoing. But these two calculations are given precision in prediction. Then, arbitrary woodland is connected as gathering strategies, and optimized XG-BOOST algorithm is used as a boosting strategy to extend the precision of expectation. Be that as it may, log-loss is used to degree the execution of classifiers by penalizing untrue classifications. As the dataset contains exceedingly course awkwardness issues, an arbitrary undersampling method for arbitrary woodland calculation gives the finest precision.

60.1 Introduction

A huge number of wrongdoings are perpetrated each day, and most likely hundreds are happening right now on the planet. Somebody some place will be executed, ransacked, and assaulted until you read this. Unmistakably, wrongdoing can be viewed as the plague of society. Regardless of what is done, nothing can truly stop it. In 2010, specialists recommended that it was conceivable to anticipate certain wrongdoings, similar as researchers estimate quake delayed repercussions. The primary

V. Upadhyay (✉)
Alpha College of Engineering, Khatraj Kalol Road, Thirumazhisai, Gujarat 382721, India

D. Rathod
Sankalchand Patel College of Engineering, Visnagar, Gujarat 384315, India

T. Senjyu et al. (eds.), *IOT with Smart Systems*, Smart Innovation, Systems and Technologies 251, https://doi.org/10.1007/978-981-16-3945-6_61

motivation behind wrongdoing forecast is to help forestall repeating violations in a territory by following the examples of wrongdoings just as the most well-known kinds of wrongdoing in an area. Using the idea of data mining, we can remove beforehand obscure, valuable data from an unstructured information. Data mining is the patterns for analyzing information and the process to extract the interesting knowledge. In data mining, various data mining tools are available which are used to analyze different types of data.

60.1.1 Mining Methods

To find the better accuracy and better prediction of the dataset, we have used four algorithms. They are decision tree, random forest, K-nearest neighbor and optimized XG-BOOST.

Decision Tree Algorithm

A decision tree is a formation that incorporates a root hub, branches and leaf junctions. Each inner junction indicates a test on a quality, each branch signifies the result of a test, and each leaf junction holds a class label. The highest junction in the tree is the root junction.

Random Forest Algorithm

Random forest is a directed learning calculation which is utilized for both classifications just as relapse. But however, it is mainly utilized for classification problems. We realize that a forest is comprised of trees and more trees imply more hearty forest. Additionally, random forest calculation makes choice trees on information tests, afterward gets the expectation from every one of them and lastly chooses the best arrangement by methods of voting.

K-Nearest Neighbor Algorithm

K-nearest neighbors (KNN) calculation utilizes ‘include closeness’ to anticipate the estimations of new data points which further imply that the new information point will be doled out a worth dependent on how intently it coordinates the focuses in the preparation set.

Optimized XG-BOOST

The working of optimized XG-BOOST is dependent on a tree structure. In this algorithm, there will be multiple tree structure, and each and every tree is connected to each other. Unlike the random forest, here the main difference is this only. In first tree, the attributes will be evaluated. If there are any attributes which are not evaluated will go in the second tree and further it will continue to the last tree.

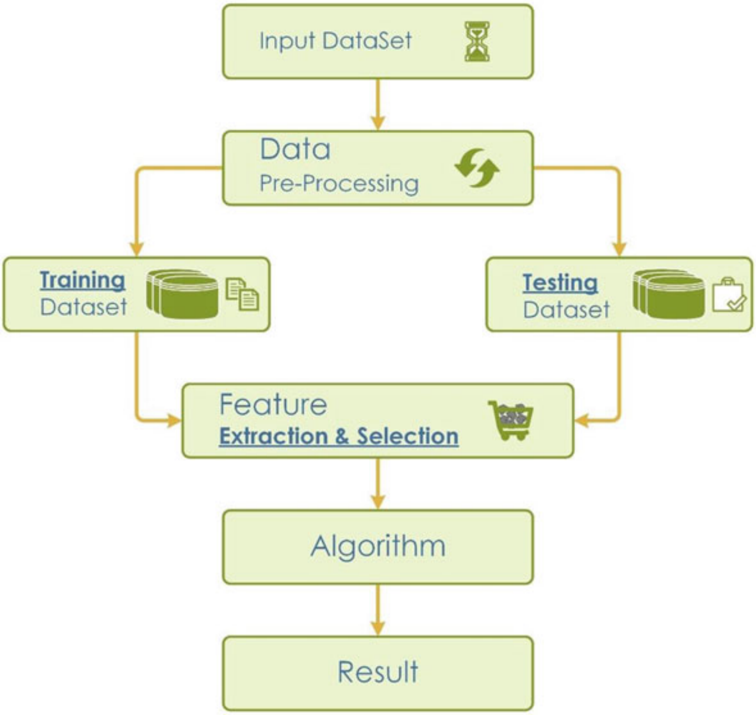


Fig. 60.1 Existing architecture

60.2 Existing Architecture

In terms of crime prediction, the author worked on the dataset of the 12 years of Chicago city. After applying the data preprocessing, the dataset is divided in the testing and training part. They have applied DS and got the accuracy of 31.17%, after that, KNN was applied and got 28.50% as an accuracy, but the accuracy was very low, and after that, they have applied RF to get the better accuracy (i.e., 31.71). To get the final result, oversampling and under sampling are performed. The final accuracy is 68.03% (Fig. 60.1).

60.3 Proposed Architecture

Step 1 We propose that post data preprocessing we will apply the min–max standardization that will scale all the attributes to the 0–1. So that we can clearly know the contribution of the attribute in the output.

Step 2 The data will be divided into two parts, one is training (i.e., 80%), and other is training (i.e., 20%). The partition can be changed according to the needs.

Step 3 Post that we will train four models (i.e., DT, RF, KNN and optimized XG-BOOST). After that, we will do the model evaluation. Here, the main benefit of this algorithm is that it will train multiple model parallelly. From the results, we can understand which model has performed better by voting approach (Fig. 60.2).

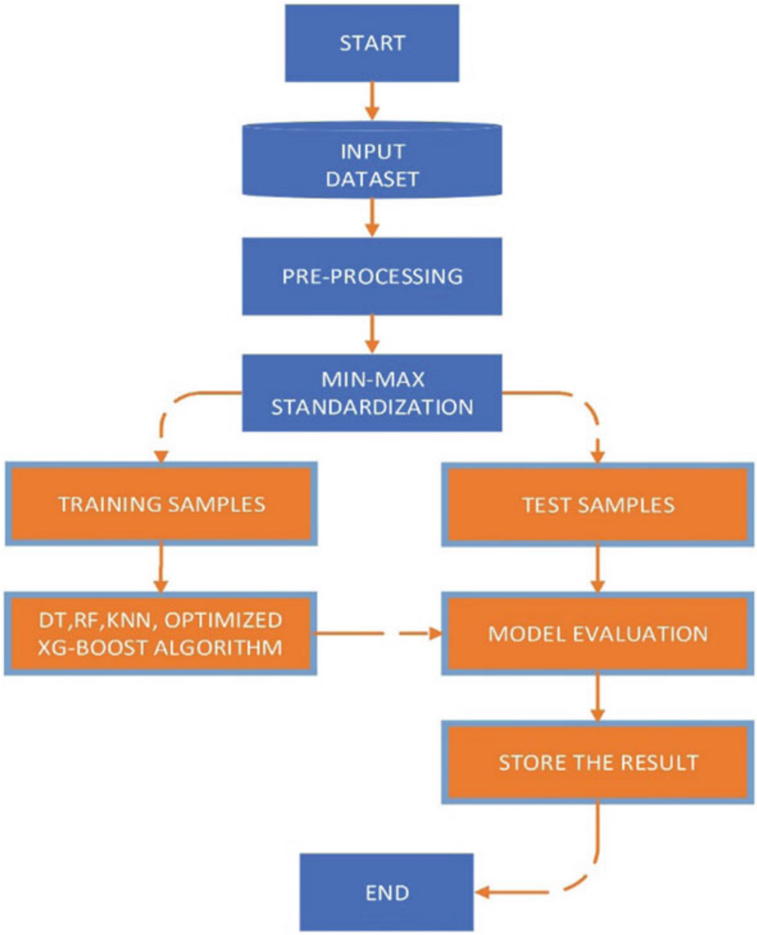


Fig. 60.2 Proposed architecture

```
from sklearn.ensemble import RandomForestClassifier
clf = RandomForestClassifier(max_depth=3, random_state=0,n_estimators=100)
clf.fit(X_train_new,yyy)
clf.score(x_test_temp,y_test_temp)
```

```
0.4454
```

The Final Accuracy of Random Forest is 44%

Fig. 60.3 Random forest accuracy

60.4 Proposed Algorithm

Here, to get the better accuracy, we are using the following algorithm steps.

Steps of Proposed Algorithm

(1) START
(2) **Take input** from dataset
(3) **Data preprocessing** from dataset
(4) **MIN–MAX standardization** on dataset
(5) Divide training (80%) and testing (20%) data from dataset.
(6) Train model using DT, RF, KNN, optimized XG-BOOST algorithm.
(7) **Model evaluation**
(8) Store result.
(9) TERMINATE/END.

60.5 Result

60.5.1 Random Forest Code

See Fig. 60.3.

60.5.2 Decision Tree Code

See Fig. 60.4.

```
from sklearn import tree
clf1 = tree.DecisionTreeClassifier(max_depth=3,random_state=0)
clf1.fit(X_train_new,yyy)
clf1.score(x_test_temp,y_test_temp)
```

```
0.4374
```

The Final Accuracy of Decision Tree is 43%

Fig. 60.4 Decision tree accuracy

```
from xgboost.sklearn import XGBClassifier as Optimized_XGBOOST
model=Optimized_XGBOOST(max_depth=3,n_estimators=10,random_state=0)
model.fit(X_train_new,yyy)
model.score(x_test_temp,y_test_temp)
```

```
0.9561
```

The Final Accuracy of Optimized XG-BOOST is 95%

Fig. 60.5 Optimized XG-BOOST accuracy

60.5.3 Optimized XG-BOOST Code

See Fig. 60.5.

60.5.4 KNN Code

See Fig. 60.6.

60.5.5 Comparison Chart

See Fig. 60.7.

60.5.6 Accuracy Table

See Fig. 60.8.

```
from sklearn.neighbors import KNeighborsClassifier
clf2 = KNeighborsClassifier(n_neighbors=3)
clf2.fit(X_train_new,yyy)
clf2.score(x_test_temp,y_test_temp)
```

```
0.2892
```

The Final Accuracy of KNN is 28%

Fig. 60.6 K neighbor accuracy

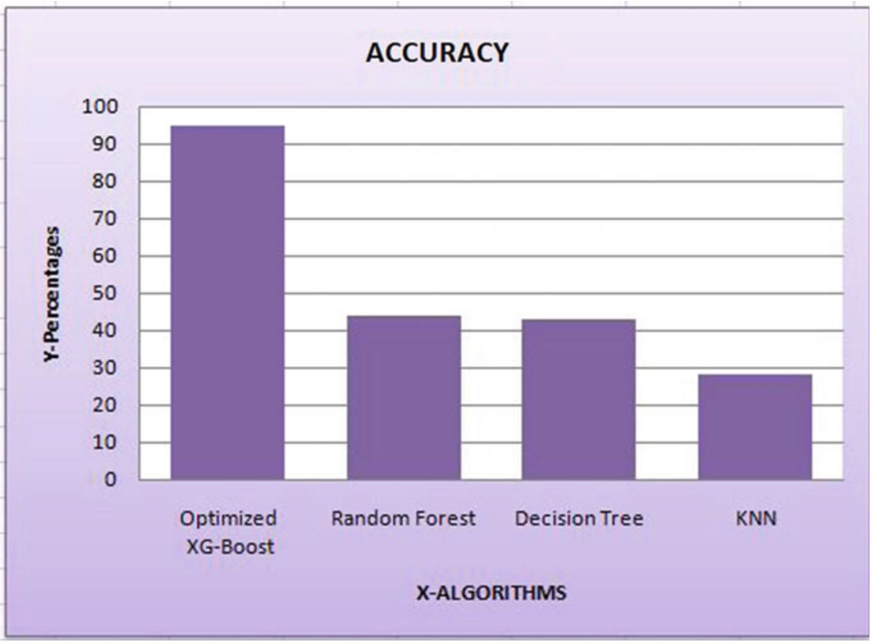
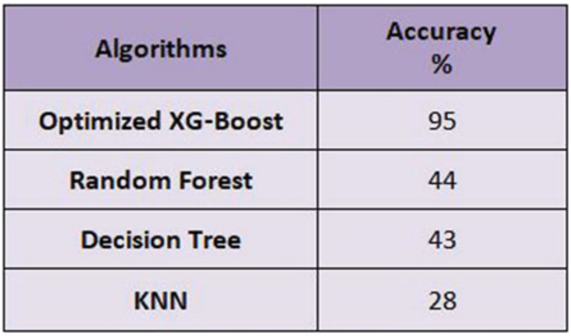


Fig. 60.7 Comparison chart

Here, we can see that accuracy of optimized XG-BOOST is the maximum among four algorithms which is my proposed algorithm.

Fig. 60.8 Accuracy table



Algorithms	Accuracy %
Optimized XG-Boost	95
Random Forest	44
Decision Tree	43
KNN	28

TABLE-1 . ACCURACY TABLE

60.6 Implementation Tools and Technology

Here, in this work, I have used Python 3.8 Anaconda in which I have used IDE Jupyter as well as Google Colab as my implementation tool.

60.7 Conclusion

After implementing the above code, I have got the results for all the four algorithms. The accuracy of decision tree is 43%. In random forest, accuracy is 44%. KNN algorithm got 28% accuracy, and my proposed algorithm optimized XG-BOOST got 95% accuracy, which is the highest accuracy among all the algorithms. So here, we can say that optimized XG-BOOST algorithm performed with better accuracy within a time. We can use this algorithm to prevent crimes and make the better changes for the society. In future, I will work with other algorithms to get the better results.

Chapter 61
Overview of Augmented Reality and Its Trends in Agriculture Industry

Simran Garg, Priya Sinha, and Archana Singh

Abstract Augmented reality (AR) is an intelligently real-world involvement in which computer-generated perceptual data, regularly through different tactile modalities, such as visual, sound-related, haptic, somatosensory, and olfactory, improves the objects that happen within the world. Augmented reality is recent technology that shows a knowledge base examples in various industries. AR remains rising within the scientific eventualities as well as agriculture areas. AI in agribusiness assists ranchers with mechanizing their cultivating as well as movements to exact development for higher harvest yield and better quality while utilizing fewer assets. This paper features importance of augmented reality in the ongoing farming industry. The ongoing patterns of cultivating innovations are examined individually. Augmented reality will bring modernization to agriculture industry.

61.1 Introduction of Augmented Reality

Augmented reality (AR) is an intelligently real-world involvement in which computer-generated perceptual data, regularly through different tactile modalities, such as visual, sound-related, haptic, somatosensory, and olfactory, improves the objects that happen within the genuine world [1]. Augmented reality is the item of utilizing innovation to superimpose information around the physical world, such as sounds, pictures, and content. Not at all like virtual reality (VR), rather than building the total advanced world to replace original with virtual, AR shows up straightforwardly in see of a real environment and includes sounds, pictures, and design to it [1]. In fundamental terms, AR could be a see-through of this present reality actual world with superimposed PC produced pictures, thus changing the acknowledgement of the real world [2]. In 1990, the actual term was instituted, and one of the essential business applications was on TV and military organizations. AR propelled its moment wave with the rise of the Web and smartphones and is presently primarily related to correlative objective. 3D models are anticipated specifically onto real things or

S. Garg · P. Sinha · A. Singh (✉)
Amity University, Noida, Uttar Pradesh, India
e-mail: asingh27@amity.edu

T. Senjyu et al. (eds.), *IOT with Smart Systems*, Smart Innovation, Systems and Technologies 251, https://doi.org/10.1007/978-981-16-3945-6_62

melded together in real time, influencing our schedules, social life, and the excitement industry through various virtual reality applications. The overlaid material information can be significant or risky [3]. This phenomenon is reliably gotten along with the existent world so much that it is seen as a clear point of the existent world. Along these lines, virtual reality changes one's propelling affirmation of the existent world, while augmented reality completely replaces the customer's actual surroundings with a new clone one. Augmented reality [1] is utilized to acknowledge characteristic conditions and give improved trials. With the help of progressed AR progresses (e.g. counting PC vision, uniting AR cameras into cell phone applications and protest acknowledgement), the information enveloping genuine universe of the customer will be naturally and deliberately controlled. Information around the climate and its items is overlaid on this present reality (Fig. 61.1).

The augmented reality is continuously being explored in cognitive sciences [4]. The principle regard of augmented reality or the truth is the manner by which the segments of the computerized world are mixed inside the insight that an individual has of the existent world, not as an essential demonstration of data, but rather through the reconciliation of vivid sensations, which are viewed as characteristic pieces of the climate. The essential valuable AR frameworks which gave enormous mixed reality experiences to customers were planned inside the mid-1990s, beginning with the virtual establishments system made at the Armstrong Exploration Office of the U.N. Examine Drive. Inside 1992, [5] augmented reality applications were spread to business organizations, for example, schooling, interchanges, drug, and delight. There are four types of increased reality today, namely markerless AR, marker-based AR, projection-based AR, and superimposition-based AR. How this is consistently [6] overlaid depends upon the character of the experience and besides the gear you

Fig. 61.1 AR gadgets enhancements [3]

are seeing the experience on. The most ideal way is utilizing your cell phones—any spot what you see through the camera has modernized portions extra to that. This can feel abnormal, [5] and therefore, for a lot of striking experience, you will have the choice to wear a headset like Microsoft's HoloLens—that lets free you from passing on up a phone and makes the combo of this present reality and CG (PC delivered) fragments feel tons even more veritable [5]. Currently, augmented reality applications and technologies have evolved a lot in agriculture industry. It has been yet to be implemented properly due to lack of awareness as well as knowledge of farmers, but augmented will revolutionize farming totally.

61.2 Evolution of Augmented Reality

The advancement of augmented reality technology [7] began when Ivan Sutherland, first individual to create the main head-mounted presentation in 1968 and positions it as a window into a virtual world. His fantasy was to extend a shape out of his creative mind and have the option to collaborate with it, and he attempted to make it genuine. Ivan Edward Sutherland was broadly viewed as the "father of PC illustrations".

- Myron Krueger made video place which permitted the used to communicate with virtual articles. This prompted [2] the making of the primary wearable PC and PC vision framework and graphical overlays in 1980 by Steve Mann.
- In 1992, Louis Rosenberg [2, 7] pushed one of the primary working AR structures, suggested as Virtual Installations, at the US Air Exploration Research center. Also, Steven Feiner, Blair MacIntyre, and Doree Seligmann presented AR machine model, KARMA, on the designs interface show.
- In 1999, [2] the US Maritime Exploration Research facility connects on 10 years broad assessments program called the B.A.R.S. program to show a portion of recent wearable systems for fighters working in metropolitan/urban condition and areas for situation focus and preparing.
- In 2004, outdoor protective cap built up AR device confirmed by method for Trimble Navigation and the Human Interface Technology Laboratory (HIT lab). Microsoft reported the world Holographic and the HoloLens increased reality headset in 2015 [7].
- In July 2016, Niantic discharged the game Pokémon Go which turned into the most prevalent cell phone application and quickest application to achieve 500 million downloads.
- Today, [8] the AgriCare industry and all cutting edge ranchers are going to computer-based intelligence advances to help yield better harvests, supervise pests, screen soil and developing conditions, arrange data for farmers, encourage with the work, and improve an enormous amount just as the nature of horticulture-related undertakings inside the whole food production network.

61.3 Devices of Augmented Reality

In scorn of the advancement take-off and rapidly creating improvement instruments update [6, 7]. From cells and tablets to gadgets like Google Glass or hand-held contraptions, it got advanced. For dealing with and projection, AR devices and gear, from the start, have necessities like sensors, cameras, accelerometer, spinner, progressed compass, GPS, focal processor, features, and things we have recently referred to.

References [2, 8] devices suitable for augmented reality fall into the following categories.

- Special AR gadgets spread out by and large and just developed reality encounters. One portrayal is head-up introductions (HUD), sending information directly into the customer's view. From the start, it was acquainted with plan military's pilots, as of now such gadgets have applications in flying, vehicle industry, making, sports, etc.
- AR glasses (or present-day glasses)—Google Glasses, Meta 2 Glasses, Laster Transparent, Laforge AR eyewear, etc. These units show popups in your cell phone, gathering with coating workers, access information without hands, and so forth.
- AR contact central focuses, making extended reality one step without a doubt more distant. Creators like Samsung and Sony have articulated the improvement of AR central focuses. Autonomously, Samsung is managing central focuses as the assistant to mobile phones, while Sony is organizing central focuses as only AR gadgets (with highlights like eliminating photographs or putting information).
- Virtual retinal display (VRD), making pictures by anticipating laser light into the common eye. Pointing at shinning, high separation, and significant standard pictures, such designs, regardless, stay to be made for viable use (Fig. 61.2).

61.4 Augmented Reality Working

AR likely could be appeared on shifted contraptions, screens, glasses, hand-held devices, cells, and head-mounted grandstands. It incorporates progresses like S.L.A.M. (Simultaneous, localization, mapping), significance following (quickly, locater information critical thinking the power source to the things), and in this way [2] the identified with parts,

- Cameras and sensors. The cameras inside the contraptions check the ecological components, and with this information, a comfort finds real things and makes 3D models.
- Handling. AR devices demonstration is like miniature PCs, things that present-day cell phones currently do. Additionally, they need a PC chip, a GPU, streak memory, Hammer, Bluetooth/Wi-Fi, a GPS, so forward to claim the decision to gauge speed, point, heading, bearing in house, etc.

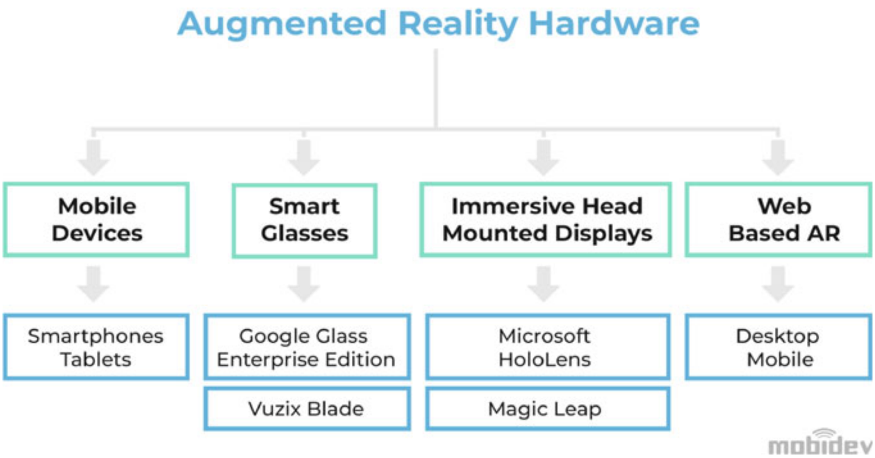


Fig. 61.2 AR devices overview [7]

- Projection. It is a cut back projector on AR headsets that takes information from sensors and adventures automated data (the eventual outcome of preparation) onto a surface to ascertain.
- Reflection. A few AR tools have mirrors to help basic eyes by watching imaginary images. Many have accomplice "collection of most negligible bowed mirrors", and a couple have a twofold sided mirror to mirror light-weight to a camera and to a client's eye.

Another key commitment [7] of augmented reality is in agriculture industry. The innovation permits learners to adjust to ranch or farm hardware while not controlling it inside the real sense. With this, farmers will encourage instructing their workers while not accruing costs with taking care of actual tools and hardware. Cultivating is one of the reaches that enormously affect environmental conditions. The enthusiastic idea of environment prompts wastage of troublesome work, time, and resources of farmers. AR being, all things considered, will offer help to agriculturists in making a couple of proactive strides.

61.5 Present AR Trends in Agriculture Fields,

Augmented reality in agribusiness is new trend dealing with the developing food interest as well as its demand. Also, expanding food request prompts intense trouble among farmers. Ranchers live with a desire for fine days. The need to satisfy developing food request could have genuine repercussions [9]. Starvation, mortality, and political turmoil are only some of them. Thus, a temporary modification in farming turns into a need. This tech can also offer extra supportive applications to the current area serving to the world contend with food production issues for the developing

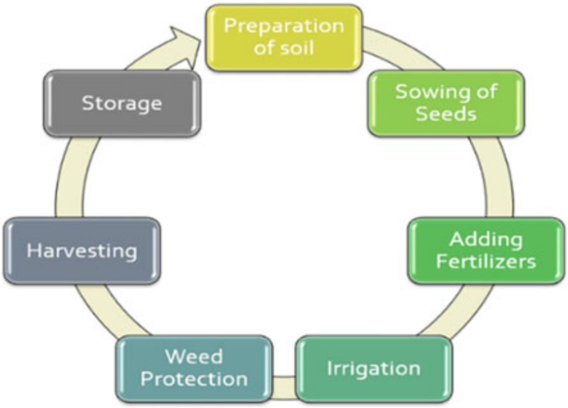


Fig. 61.3 Life cycle of agriculture [10]

populace [3]. Utilization of AR in agriculture fields will offer help with pushing ahead farm produce. The inventive influx of AR in cultivating will reform developing. Along these lines, the meaning of augmented reality in cultivating will be boosted. A couple of the manners by which AR will offer help to our farmers are as per the following (Fig. 61.3).

61.5.1 Augmented Reality for Field Assessment

Reference [9] do you think farming is a two-stage interaction of planting seeds and gathering? No, agribusiness is significantly more than this. Farmers perform manual richness control on each plot of land, cultivating a complex and tedious interaction. Augmented reality in horticulture assumes a transformative part in the field, control, and cycle location of pests/creepy crawlies and then decides the properties of the dirt and the harvest to be planted on a specific package of land [7]. Cultivating involves exact estimations, planning, and a logical way to deal with yield the premier plentiful gather. With a few investigations investigating augmented reality in agribusiness, potential applications flourish. Truth be told, by exploring soil properties, agriculturists can pick which yields would thrive in a specific land. There are many examples of various augmented reality farming applications.

61.5.2 Expanded Reality in Horticulture for Mimicked Preparing

The utilization of augmented reality is unmistakable [3] in assembling, medical care, and land, yet its utilization in farming still cannot seem to be investigated. It is an intelligent and more secure type of approach for preparing farmers. Farm gear tasks are achievable with a tablet or cell phone! [3]. In this manner, presently ranchers can catch and share data about apparatus, yields, and animals in the most ideal way.

61.5.3 Augmented Reality in Horticulture for Climatic or Daily Weather Updates

In addition, [8] AR is regularly utilized for receiving the environment figures as well! Augmented reality in cultivating gives ranchers right and steady environment-related conditions. AR, for this situation, [1] goes about as help for ranchers for acquiring ongoing weather experiences. Constant, exact, and area explicit climate refreshes help with relieving the dangers identified with cultivating. The powerful idea of climate brings about wastage of ingenuity, time, and assets of ranchers. AR being, all things considered, can encourage ranchers in making some proactive strides. Combined with weather information [8], they could approach farming during an additional informed method. This can doubtless minimize crop losses and guarantee healthy harvests.

61.5.4 AR Wearable Modern Glasses

These glasses advancement [1, 6] seems a sparkling future in AI industry. Currently, these modish glasses are a sort of wearable advancement that coordinates virtual information with information inside the viewpoint of the field to the ranchers. The reformist augmented reality glasses permit farmers to put on PCs and examine their fields. The showcase [7] highlights the isolating highlights of the environment. These modish frames have a splendid display as there is fundamental headway inside the quantity of its shipments. While only 150,000 AR modish glasses dispatched around 2016, various packages are anticipated to climb as much as 22.8 B consistently in 2022. The veil got from AR [1] which is largely supportive to farmers incorporates flow soil sogginess content, surface porosity, and soil water ingestion limit. The once-over of the upsides of using these cutting edge glasses in cultivating is extended (Fig. 61.4).

61.5.5 AR Agricultural Games

Reference [1] cultivating Games! Fascinating … No? The AR cultivating games are one among the great methods for ranchers to acknowledge legitimate cultivating encounters. This beguiling experience utilizing AR can satisfy the need of ranchers of getting their fantasy ranch. AR joined with subtle and current illustrations [4] prompts

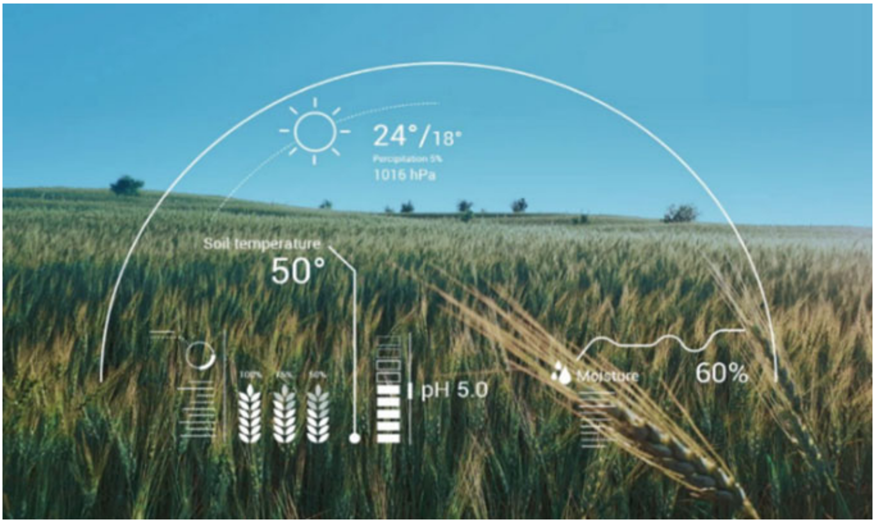


Fig. 61.4 Augmented reality in horticulture for climatic or daily weather updates [10]

the best insight for ranchers. For instance, AR games can allow ranchers to drive cultivating vehicles of their fantasy [11]. AR developing diversions will offer help to ranchers with the equivalent. For delineation, ranchers can with these redirections build their craziest type of cultivations and have a great time. Farmville is one of the many developing games inside the feature. Accordingly, these entertainments will progress as a reformist learning instrument in developing (Fig. 61.5).

Fig. 61.5 Farming games [1]

61.6 Future Scope

With our expanding populace, decreasing common assets, and the impacts of environmental change making agribusiness seriously testing, putting resources into front line innovation could be a suitable advance for ranchers around the globe [3]. Augmented reality is hoping to additional its place in cultivating measures, as it keeps on giving advantageous answers for the business. Augmented reality has a great potential to turn farming 360°. Thus, it is greatly benefiting our farmers as well as economy. Currently, augmented reality can sense soil, air moisture, and full profile of any plant and weather forecast. But in near future, it will totally revolutionize farming equipment and farming techniques for our farmers. It will assist in more holistic and precise approach to cultivate the best harvest every time. Also, so far augmented reality devices have been a little explored in all aspects of farming, and it has proven its significance. But just imagine if it gets further explored as well as implemented properly, it would become few steps work for our farmers. And further, it will boost our economy.

61.7 Conclusion

Augmented reality is the item of utilizing innovation to superimpose information around the physical world, such as sounds, pictures, and content. This technology began near 1968, and till now, it has evolved tremendously. Alongside its enormous developed history, its simple and convenient working and devices have surely proved its further growth in near future. Unquestionably, augmented reality in farming has a colossal extension. An absence of information, assets, and foundation, notwithstanding, brings about its complex execution. Ranchers are the mainstays of an economy and consequently should be ensured. Brilliant cultivating instruments of augmented reality will improve the current circumstance of our ranchers.

Rural advancement is the key to the financial turn of events. Augmented advancements can possibly uphold shrewd agribusiness for expanded profitability. Further investigation will educate, control, and will be executed in not so distant future exploration in the utilization of vivid innovation to farming. There is a wide scope of augmented reality in agriculture industry, but it is yet to be explored. It will surely assist in more holistic and precise approach to cultivate the best harvest every time. Augmented reality will be the “God of Farming”.

References

1. http://www.queppelin.com/augmented-reality-in-agriculture/
2. https://thinkmobiles.com/blog/what-is-augmented-reality/

3. https://arpost.co/2019/01/18/how-augmented-reality-could-revolutionize-farming/#,~,text=These%20include%20an%20AR%20system,presence%20of%20pests%20and%20insects.&text=Aside%20from%20data%20about%20insects,flourish%20in%20a%20particular%20land
4. https://www.digitalartsonline.co.uk/features/hacking-maker/everything-you-need-know-about-designing-with-augmented-reality/
5. https://psu.pb.unizin.org/ist110/chapter/6-2-augmented-reality/
6. https://www.livescience.com/34843-augmented-reality.html
7. https://mobidev.biz/blog/augmented-reality-development-guide
8. https://www.frontiersin.org/articles/10.3389/fpsyg.2018.02086/full
9. Silva, R., Oliveira, J.C., Giraldi, G.A.: Introduction to Augmented Reality, pp. 3–6 (2014)
10. https://www.analyticsvidhya.com/blog/2020/11/artificial-intelligence-in-agriculture-using-modern-day-ai-to-solve-traditional-farming-problems/
11. Azuma, R.: Baillot, Y., Behringer, R., Feiner, S., Julier, S., MacIntyre, B.: Recent Advances in Augmented Reality, pp. 3–4 (2011)

Chapter 62
Comparative Analysis of Single-Stage YOLO Algorithms for Vehicle Detection Under Extreme Weather Conditions

Udaya Mouni Boppana, Aida Mustapha, Kavikumar Jacob, and Nagarajan Deivanayagampillai

Abstract Efficiency of vehicle detection algorithms in Computer Vision is predominantly calculated under clear weather condition; good weather with favorable lighting. Benchmarking of top-performing algorithms still reports a relatively low performance for vehicle detection under extreme circumstances. This is because vision-based detection algorithms face challenges in dealing with low quality images with background noise, bad lighting, and weather-caused distortions. The literature has reported substantial work on restoration of images prior to object detection. Nonetheless, the measures affecting the vision based algorithms and their effectiveness as well as how much degradation of the quality of the input image reduces the detectors output are less investigated. This paper focuses primarily on single-shot detector, which is You Look Once (YOLO) algorithm along its variations to detect vehicles real-time. The comparative experiments are set to evaluate their efficiencies based on precision, mean average precision (mAP), and recall rate using distorted vehicle images from the AAU Rain Snow Dataset. Based on our experimental analysis, YOLOv3 performed better as compared to other variants of YOLO. These findings will be used as the benchmarking results for improvement of vehicle detection algorithms under extreme weather conditions.

62.1 Introduction

Adverse weather conditions have a remarkable impact on people's health, transport management, and autonomous driving vehicles due to its various influences on human behavior across the globe. Traffic management always suffers from congestion, low visibility, and unsafe driving situations due to its abnormality in nature, and drivers are failing to adjust their speed, often leads to crashes [1]. Improving traffic control countermeasures under adverse weather conditions is challenging due to the degrada-

U. M. Boppana (✉) · A. Mustapha · K. Jacob · N. Deivanayagampillai
Universiti Tun Hussein Onn Malaysia, Batu Pahat, Johor 86400, Malaysia
e-mail: aidam@uthm.edu.my

K. Jacob
e-mail: kavi@uthm.edu.my

T. Senjyu et al. (eds.), *IOT with Smart Systems*, Smart Innovation, Systems and Technologies 251, https://doi.org/10.1007/978-981-16-3945-6_63

tion of the quality of images, down to the point they are not useful for further analysis. According to [2], a computer vision-based traffic monitoring surveillance system is one of the most effective alternatives for sensors. It extracts high-level information from image and video-based systems. Due to its less disruptive and more affordable nature, the results from video camera are far better in tracking due to the strategic placement of the cameras. Vision-based detection algorithms can perform various tasks such as vehicle detection segmentation and tracking with better outcomes than modern, sensor-based technologies. Traffic surveillance system, for example, is able to cover several hundred square meters with adequate capacity to draw out premium information like an exact path, color, shape, and dimensions.

Image analysis in automatic traffic surveillance generally employs Computer Vision approach. This approach works well in daylight when the road users are visible to the camera but often struggles when the scene's visibility is impaired due to insufficient lighting or distortions from extreme conditions like rain, snow, haze, and fog [3]. Reference [4] stated that adverse weather, with occasional adverse events such as tornadoes, twisters, and hurricanes, is categorized as unusual occasions a opposed to rainfall, snow, and wind, which are considered usual events. Both occasions have a substantial effect on the transportation system, due to its strong influence on human driving behaviors that can cause traffic congestions as well as road collisions. Therefore, the traffic surveillance systems must be appropriately regulated to prevent accidents under extreme weather conditions.

At present, deep learning models have taken extensively in fundamental object detection as well as domain-specific object detection in Computer Vision [5]. This is credited to the rapid advancement in detection strategies began with the progress of Graphics Processing Units (GPUs) and Deep Convolution Neural Networks (DCNN). Object detection is the process where it handles the semantic objects and classifies into different classes in various images and videos. It involves the procedure of both localization and classification, and some of its applications involve posture detection, crowd analysis and detection, face detection, and detecting pedestrians as well as fast-moving vehicles. In the transportation domain, detection of automated license plate and pedestrian plays a vital role but are impacted by the extreme weather situations the most [6].

This paper is set to quantify the degradation level of state of art detectors due to rain and snow weather conditions, which are YOLOv1 [7], YOLOv2 [8], and YOLOv3 [9]. Section 62.2 describes the Convolutional-based Detection Models, followed by single-stage YOLO algorithm along with its variants in Section 62.3. Section 62.4 presents the comparative analysis of YOLO variants in terms of precision, mAP, and recall rate based on the AAU rain-snow dataset. Finally, Sect. 62.5 concludes the paper with some indication for future work.

62.2 Convolutional-based Detection Models

Deep learning has substantially enhanced the efficiency of many tasks, such as speech recognition, object detection, and various other hefty detection tasks. After the arrival of Convolutional Neural Networks (CNNs), one of the critical groups of neural networks, the area of Computer Vision has been impacted the most. In terms of handling various challenges like detection, classification, semantic segmentation, and deepness evaluation, CNNs have been proven to be highly effective. The algorithm is able to deal with spatial dependences with no requirement for a typical feed-forward neural networks. In recent years, vision-based detection systems have seen significant growth and have produced excellent results under favorable conditions. Convolutional Neural Networks (CNN) consider the depictive design for deep learning. In a standard CNN design, every nerve cell is in touch with nearby neurons from the previous layer. Each layer stands for a feature map representing a three-dimensional matrix of numerous pixel intensities of colour channels for the input layer. Based on [10], a feature map is an induced multichannel picture in every internal layer, where pixel is deemed a specific attribute. Various sorts of transformations carried out on feature maps, like pooling and filtering.

To produce more robust feature descriptions, the convolution layer considers neurons' values using the convoluted matrix. It uses nonlinear functions like ReLU and sigmoid to achieve final responses for a more robust feature description. By adding fully connected layers between the pooling and filtered layer, a feature hierarchy layer tunes in a supervised manner for adapting different visual tasks. To get the specific conditional probability for each output neuron depends on the task in the final layer, different activation functions added to optimize the network objective function like cross-entropy loss or mean squared error uses Stochastic Gradient Descent (SGD) method. A traditional CNN contains 13 convolutional layers, three max-pooling layers, three fully connected layers, and a SoftMax classification layer. It convolutes the 3 * 3 filter window for producing the feature maps and reduces two strides. If there are any variations in sizes, rescaling or cropping can be performed.

To date, CNN has been widely used in different research study areas, such as classification of image retrieval of images, construction of image resolution, video analysis and pedestrian detection. The main advantages of CNN against traditional methods as follows.

1. CNN represents features in a hierarchical manner, where multilevel feature representations from pixel to high-semantic level are learned by a multistage framework. Therefore, data can be disentangled via multilevel nonlinear mappings.
2. CNN enhance several relevant tasks like rapid R-CNNby collectively incorporating category and bounding boxregression right into a multitasking discovery manner.
3. CNN mostly solves many high-dimensional transformation problems. Its deeper architecture also tremendously enhances the expressive capacity.

In general, CNN has been widely used for classification and detection. Nonetheless, very few research concentrates on rain, snow, and haze under extreme weather conditions. Some work focused on adherent raindrops instead of streaks of rain and how much quantity of rain or snow degradation was added to influence the image quality and vision-based detectors performance. Rain streaks that are applied to images will cause degradation to occur and leads to poor detection efficiency. Alternatively, this method can be used to predict how much image quality enhancement is required to achieve a proper detection level for a given object detector.

62.3 You Only Look Once (YOLO)

YOLO (you only look once), a one-stage object detector mainly to provide real-time image detection and well with surveillance. YOLO's inspiration came from Google Net, it contains 24 convolutional layers for extracting the features and two dense layers for predictions. YOLO works upon Darknet architecture. It predicts less than 100 bounding boxes per image while Quick R-CNN using careful search anticipates 2000 area propositions per image. YOLO detection takes as a regression issue, so a unified architecture can extract features from input images directly to forecast bounding boxes as well as class likelihoods.

In contrast with other object detection algorithms, YOLO has the following advantages: In real-time, YOLO acts very fast with a detection rate of 45 frames per second. High accuracy for real-world prediction due to generalized representation.

State-of-the-art, one-shot detectors such as YOLOv1, YOLOv2, and YOLOv3 take the cell-to-box prediction right into account and permit multiple objects to be found in a solitary cell. Detection is considered by YOLO, as well as its variations as a regression problem. These regions are categorized into object classes by specific areas where the image pixels arrange the bounding boxes into corresponding probabilities of classes.

62.4 Comparison of YOLO-Based Algorithms

The adverse weather dataset used for comparing YOLO's performance, and its variants detection algorithms is AAU Extreme Weather Dataset created by [11]. It contains seven different traffic surveillance intersection videos in extreme weather conditions, mainly considering adverse rainfall and hazy images in both day and night vision. It is also having the reflections from vehicles and distorted rainy images. AAU dataset is the mixed combination of both RGB and thermal infrared cameras for robust vision-based detection and classification.

By analysing quantitative as well as visual contrast of YOLO variations with severe weather conditions, including snowfall and even significant rain photos, based on AAU extreme weather conditions as seen in Fig. 62.1.

Fig. 62.1 Vehicle detection in a blurry image under adverse weather condition

Based on the Evaluation of the AAU Weather dataset, extreme weather endures the performance efficiency of the detectors', and the outcome is distinct in consecutive frames, even though the detected objects are the same.

Detection results are contrasted with ground truth images. If the intersection over union (IoU) score of the observed bounding box corresponds to more than fifty percent of the ground truth bounding box, it is called True Positive. The detector returns the bounding box coordinates and the corresponding class labels of the detected objects, and a confidence value indicates how certain the detector detects the detection.

Precision is the positive predictive Value and even represents as True Positive. When Compared with a ground truth image, the number of objects classified and correctly detected in a tested frame is True Positive. False Positive is the case where it cannot be able to detect accurately in the tested frame.

The precision curve is comparitivly low for YOLOV1 when contrasted with YOLOv2 and YOLOv3 seen in Fig. 62.2 due to more localization errors in version1. YOLOv2 measures the classes conditional probabilities based on SoftMax to conclude that each box has one label that makes threshold value more, than compared to version1 but still due to the overlapping problem the threshold value is average when contrasted with YOLOv3. Version 3 shows better performance in dealing multi-label problem and uses individual logistic classification for each class and binary cross entropy for each class to measure the loss.It makes use of multi-label classification in adapting to different extreme weather complex dataset, and contain lots of overlapping tags leads to increase the confidence level and minmize the threshold value which raises the precision rate.

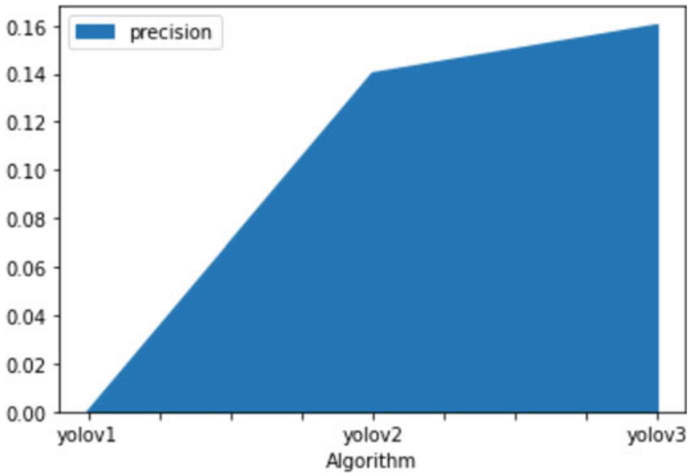


Fig. 62.2 Comparison of precision for YOLO, YOLOv2, and YOLOv3

Recall is represented as Sensitivity, and it is defined as the proportion of specific instances from the overall cases. When compared with the ground truth, if the object is not found, it is False Negative.

YOLOv3 is having the min threshold value due to its layout is similar to the pyramid network features for detecting obstacles at various sizes and predicting the boxes as a logistic classifier at two additional scales instead of using the SoftMax classifier. In order to avoid the loss of low-level characteristics and attributed to pooling, it uses some residual blocks to extract the features of convolutional layerphase to sample the feature maps. At each scale, it forecasts three boxes and calculates the three-scale function maps. It leads to high recall rate where as YOLOv1 and YOLOv2 is having low precision which will raise the threshold and decreases the recall rate seen in Fig. 62.3.

Mean Average Precision is the mean of average precision. It calculates the mean of each class and average them. Average precision is the region in which both accuracy and recall lie in the cases tested. Due to YOLOv1 has a relatively low recall rate due to more localization errors, and only two bounding boxes can be suggested for each grid.

YOLOv2 achieved enormous mean average precision than contrasted to YOLO due to anchor boxes and batch normalization as well as fine-grated features with multi-scale training to improve YOLO speed and precision. YOLOv3 utilizes three various range attribute maps to predict the bounding box. The last convolutional layer anticipates a 3D tensor inscribing class forecasts, objectness, and also bounding boxes. Due to the advantages of multi-scale predictions, YOLOv3 can find smaller objects but achieved relatively low efficiency on larger sized objects runs within the darknet system and handles more frames per second with less inference time and more processing speed leads to more mean average precision seen in Fig. 62.4

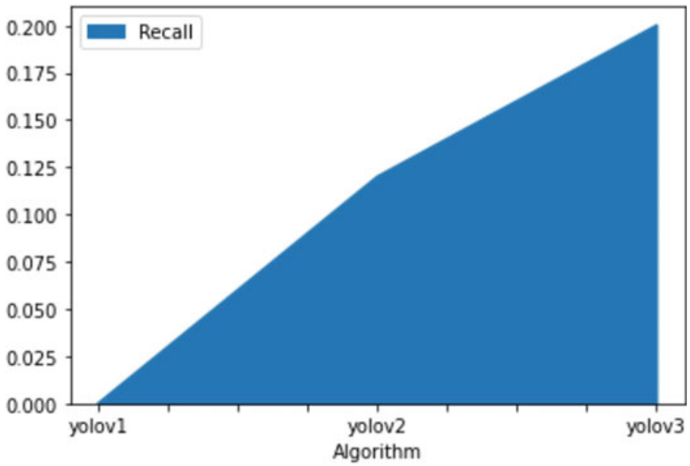


Fig. 62.3 Comparison of recall for YOLO, YOLOv2, and YOLOv3

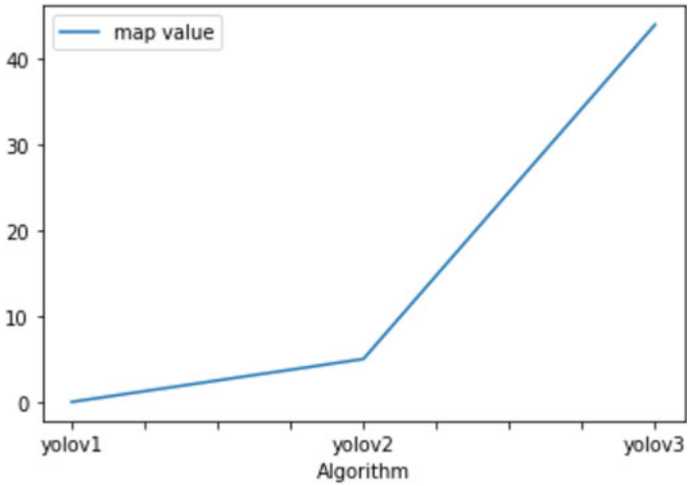


Fig. 62.4 Comparison of mAP for YOLO, YOLOv2, and YOLOv3

The precision and recall curve value for a class varies by measuring at each point, from YOLO and its varients depend on the confidence level from min to max. Still the performance of the detectors lacks due to its degradation of the quality of the input image in extreme weather conditions.

Precision, mAP and Recall are going positively, and near to 1.0 reveals it strongly impacts the false positive rate. By using COCO labels (only vehicles are considered) and repeating the same experiment with different YOLO variants in extreme weather conditions and comparing it with ground truth images, more objects are detected by YOLOv3 with higher confidence than compared to other variants of YOLO. The object detection experiments were measured based on precision, mean Average Precision (mAP), and recall rate. The experimental results show that YOLOv3 produced better results in mAP, precision, and recall under extreme weather conditions. Parameters like threshold and confidence, as well as non-maximum suppression, are set to 0.4.

62.5 Conclusion and Future Work

Vision-based detection methods face challenges from multiple perspectives; background noise, bad lighting, and weather conditions, which all have influence on the quality of image detection. This paper quantitatively measured and compared the performance based of vehicle detection model for three variants of the You Only Look Once (YOLO) object detection algorithm. YOLO is a vision based detection algorithm designed to detect in good weather conditions where these algorithms were facing the problems in detecting extreme weather conditions and impacts the quality of the image. In future, this work will be extended to focus on how much degradation occurs in a vehicle detection model and how much it suffers for a given input image deterioration. These information can be used to evaluate the quality of images and to gauge how much improvement is needed in order to achieve an acceptable level of detection rate in a vehicle detection model. The correlation between the quality of images and the performance of the vehicle detection model is imperative to ensure a robust algorithm that can work even under severe weather conditions.

Acknowledgements The research was supported by Ministry of Higher Education (MOHE) through Fundamental Grant Scheme (FRGS/1/2020/ICT02/UTHM/02/4).

References

1. Peng, Y., Jiang, Y., Lu, J., Zou, Y.: Examining the effect of adverse weather on road transportation using weather and traffic sensors. PLoS One **13**(10) (2018). ISSN 19326203. https://doi.org/10.1371/journal.pone.0205409
2. Zhang, X., Chen, G., Saruta, K., Terata, Y.: Deep convolutional neural networks for all-day pedestrian detection. Lect. Notes Electr. Eng. **424**, 171–178 (2017)
3. Huang, D.A., Kang, L.W., Wang, Y.C.F., Lin, C.W.: Self-learning based image decomposition with applications to single image denoising. IEEE Trans. Multimed. **16**(1), 83–93 (2014)
4. Oquab, M., Bottou, L., Laptev, I., Sivic, J.: Weakly supervised object recognition with convolutional neural networks. Technical report, INRIA, HAL-01015140, pp. 1–11,"Â hal.inria.fr (2014)
5. Han, X., Lu, J., Zhao, C., You, S., Li, H.: Semisupervised and weakly supervised road detection based on generative adversarial networks. IEEE Signal Process. Lett. **25**(4), 551–555 (2018)
6. Wang, Z., Chen, Z., Wu, F.: Thermal to visible facial image translation using generative adversarial networks. IEEE Signal Process. Lett. **25**(8), 1161–1165 (2018)

7. Redmon, J., Divvala, S., Girshick, R., Farhadi, A.: You only look once: unified, real-time object detection. In: Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 779–788, December 2016
8. Redmon, J., Farhadi, A.: YOLO9000: better, faster, stronger. In: Proceedings of 30th IEEE Conference on Computer Vision and Pattern Recognitionition (CVPR 2017), vol 2017, pp. 6517–6525. Institute of Electrical and Electronics Engineers Inc., Nov 2017
9. Redmon, J., Farhadi, A.: YOLOv3: an incremental improvement (2018). arxiv.org
10. Krizhevsky, A., Sutskever, I., Hinton, G.E.: ImageNet classification with deep convolutional neural networks. ACM Commun. **60**(6), 84–90 (2017)
11. Bahnsen, C.H., Moeslund, T.B.: Rain removal in traffic surveillance: does it matter? IEEE Trans. Intell. Transp. Syst. **20**(8), 2802–2819 (2019)

Chapter 63
IoT-Based Smart Irrigation System—A Hardware Review

Anamika Chauhan, Ravi Ranjan Sah, and Ronit Khatri

Abstract India is a subcontinent with a population of about 1.38 billion. In 2020, India's population is almost 18% of the human population on earth. It is a worrisome and concerning matter. In the event that the populace continues expanding in light of current circumstances, at that point, shortly, there will be a scarcity of food. Today's farming area is enduring because of the helpless yielding of harvests in light of an absence of rain and poor watering framework. This research paper aims to introduce a computerized water system framework that is not time-consuming, cost-efficient, and reduces the farmer's heavy workload. The conventional watering procedures like the dribble water system, the dekhli water system, the manual water system, and the sprinkler water system require many working hours and tedious work. Additionally, water wastage is a great matter of concern with these systems, as they waste thousands of liter of water. With automated frameworks, farmer's heavy work can be diminished. It can detect any adjustment in their environmental factors and adjust according to the climate and send the signal to its controllers about any significant change that can harm the crops or farmland.

63.1 Introduction

Before the time of Indus Valley Civilization, India is cultivating different types of crops. Still, in the twenty-first century, India is on the second position in the world for growing crops, fruits, and vegetables. As per the Niti Aayog survey of 2018–2019, 40% of India's population are directly or indirectly bent on farming- and agricultural-based works. According to the R.B.I., farming and agriculture contributed 15.4% in 2019 toward the GDP of India. Framer situation can become better only when the new mechanism and techniques will emerge in the field of irrigation and fertilizing crops.

A. Chauhan · R. R. Sah (✉) · R. Khatri
Delhi Technological University, Shahbad Daulatpur Main Bawana Road, New Delhi, Delhi 110042, India
e-mail: raviranjansah_2k17it96@dtu.ac.in

A. Chauhan
e-mail: anamika@dce.ac.in

T. Senjyu et al. (eds.), *IOT with Smart Systems*, Smart Innovation, Systems and Technologies 251, https://doi.org/10.1007/978-981-16-3945-6_64

However, that requires immense research and development in this field. Farming and its other associated sectors related to farming and agriculture are the major sector in India that provides a livelihood to the people of India, especially in the rural area.

In India, there have been many revolutions in agriculture and its associated fields, and major one is the green revolution, which made India stable in agriculture and food production. Similarly, there were other revolution like White Revolution and blue in the field of milk production and fisheries. In the 1960s, 65% of the people were involved in farming or related work. Nevertheless, recently, this number has seen a significant decline. According to the Policy Commission, in 2016, only 37% of people are involved in farming activities. Slowly, people are disillusioned with the work related to farming. There are many reasons for this water problem, drought, flooding, hard work, fewer profits, lack of facilities, lack of knowledge and modern technologies, etc. Farming consistently stays in the dangers of flood, dry season, animal grazing, and substantial or surprising precipitation. There are numerous strategies for aiding in cultivating for developing yields. The fundamental goal of these procedures is to guarantee more water conservation and less water utilization, less tedious work, but rather more profit to farmers and help expand the harvests' profitability. When contrasted with agribusiness's standard technique, it utilizes Indian farmers still use ox or other animals rather than big technologically updated machines. New Technology and innovation cause agribusiness to build profitability, save-time when contrasted with the conventional technique of farming also helps in keeping up and upgrading soil quality and exact utilization of water. After the machine revolution, hard labor work has significantly reduced, but environmental change has seriously influenced farming and agriculture-related jobs in India. India is very diverse in climatic conditions. In some places, it may flood, and in some places, it may be drought. Hence, water management for irrigation is necessary for India.

63.2 Literature Review

K. Aggarwal in 2018 stated that a mechanized framework anticipates the framework's activity to have at least human intercession alongside the monitoring and support. Pretty much every water system framework can be robotized with the assistance of clocks, sensors, and PCs and additionally automated apparatuses which can be available in the framework as implanted gadgets [1].

A recent paper of 2020 proposed by S. IYER has stated that a highly modern system based on cloud networks can be used. This paper also proposed the use of an app they have made in order to control and check the status remotely. The proposed system can be cost-effective. However, there is an issue with this system as it uses too much nodes and networks, means if there is a simple issue within the system sensors then whole framework will suffer [2].

R. Suresh in 2014 proposed a rain gun irrigation system that is highly reliable for irrigation when there is urgent need of water. This system saves a huge volume of water needed for agriculture, these frameworks carry a change to the executives of farm asset, and they have built a stack software and used Android and Java language for its key application and system design. They have used GPRS system for the irrigation which is not an cost-effective method for smart irrigation [3].

Ms. Rane in 2014 proposed a system of mechanized water system framework that depends on microcontroller, and Sun-based power was utilized uniquely for wellspring of intensity gracefully. Different sensor was set on farm land. Sensors detect amount of water in soil ceaselessly and give the data to the system through GSM. System administrator can remotely control the framework utilizing wireless technology without visiting the farm, and if the soil is evenly irrigated, then system will automatically get shut off automatically [4].

B. Ravi in 2020 proposed about a system that uses WeMosD1 controller. The ESP8266 is already built-in available in Wemos D1 with this NodeMCU, or Arduino both can work and can be programmed. They also used SIM900A GSM module for end-to-end connection as this module is Indian phone friendly that operates at 900 MHz. This system is best of its own kind, but it uses static IP address for its Web access, and this makes their system risky and unsafe. Another issue is the GSM module that is prone to easy illicit access [5].

J. Karpagam in 2020 presented a very simple smart irrigation framework. They did an experiment on automated irrigation system, and the outcome of that was the soil moisture sensor and is the most crucial sensor in any IoT-based irrigation frameworks. Apart from soil moisture sensor, type of soil and type of crop also play an important role in smart irrigation system. A unique thing with their system was the use of water level sensors of tank as if water is depleted in water tank, system will send a message to farmer [6].

Anamika Chauhan and team in 2020 presented a paper in which major focus was on the hardware and software selections and management of major hardware equipment that are not only less costly but on the same way cordial to the farmers. They used Arduino Uno microcontroller with GSM to communicate with the microcontroller to turn on and off the motor from a distance and LCD screen for monitoring the crops or any threat for the crops. In their another paper of 2020, they have surveyed the various possible hardwares and microcontrollers with various communiqué systems present in the field of IoT-based irrigation systems [7, 8].

In another paper of 2020, Himanshu and others have surveyed the various machine learning algorithms and data-mining approaches for the smart irrigation systems like fuzzy logic, naïve Bayes algorithm, linear regression, and neural network and tested the different hardwares used in the systems with these algorithms for an optimal working machine for smart irrigation.

63.3 Basic Designs

The viable strategy for controlling and observing the water system framework depends on IoT and climate data access. There are different kinds of technology

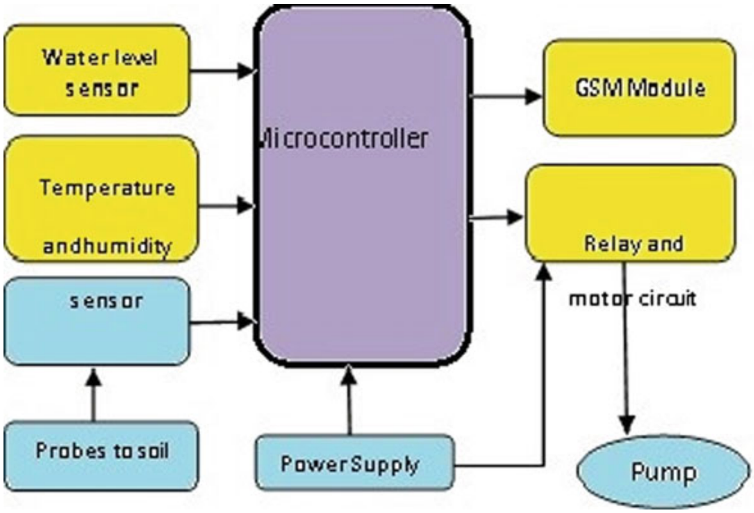


Fig. 63.1 Diagram for mechanized irrigation framework

for checking and controlling framework accessible with different highlights dependent on their plan and establishment properties for various purposes. The essential object of the cutting-edge remote agribusiness framework is as per the following.

Remote sensing—From the far distant point, we can collect messages, live images or video, significant data like temperature, pressure, humidity, soil dampness, and rainfall from the system apart from that irrigation framework, and status of agricultural land, grazing animal or insect infestation can be observed remotely.

Remote access—Various sensors or system can be remotely accessed like water pump framework/system. Alarm system for grazing animals can be switched on.

Simple access to Data-Climatic condition, runtime-data utilitarian information to far off area can be accessed easily. Actual information of the field can be accessed. Information transference to various other equipment or persons can be done. Preprogrammed integration of environmental information into the decision process and making system and agriculture models.

Remote Communiqué—Data can be seen graphically, audio, video, message, or demo call (Fig. 63.1).

63.4 Microcontrollers

An microcontroller is a chip-based small mircroprocessor/microcomputer in its own, and it has everything that requires a normal computer to function like RAM, ROM, input/output ports CPU, USB port, etc., depending on the model. It is utilized for different simple tasks having fluctuation, comparability and is cordial to the breadboard for the automated watering. The programmed brilliant watering framework is utilized to check and control the different sensors utilizing a microcontroller (Arduino or Raspberry Pi). Arduino [9] is a costless, modifiable and redistributable platform to easily work on hardware and software). Arduino is a software for programming that plans and makes single board miniature computers for building devices. It is fit for perusing commands like for turning sensors on or off, components ON/OFF, and go around into a yield like communicating something specific. It contains a bunch of guidelines for the microcontroller. Arduino programming language and its programming can be utilized to determine directions to the microcontroller for its different parts and sensors. Raspberry Pi [10] is cost-effective and small controller, and it is most easily and doable IoT platform. Raspberry Pi has different models, and due to this variability, it is not as cost-effective as Arduino although it has all properties of Arduino like 256 MB RAM greater CPU processor, but its operation is bit tedious in accordance with Arduino IDE.

63.5 Sensors

Major sensors that are used in major irrigation system are given below as follows:

1. Soil dampness sensor
2. Temperature sensor
3. Hygro/humidity sensors.

63.5.1 Soil Sensor

Soil sensor [11, 12] is used to measure the dampness present in the soil, and after the temperature sensor, this is the major sensor used in automated irrigation systems. This sensing device works in two simple types, on water tension and other on difference in di-electric constant. In water tension type, a sensor named Tensiometer is cost-effective but has difficulty in prolongation, another is granular matrix, a highly cost-effective sensor but also highly sensitive to salt present in soil, and hence, results are inconsistent and faulty. The sensors based on difference in dielectric are less sensitive to properties of soil like TDR soil sensor, and this is highly accurate and fast sensor, but its cost is very high. Similar to this, FDR sensor system has a plus point that is the appraisal of not at all like length for an indistinguishable position is doable but sensitive to soil type. A VH400 sensor is independent to the salt present in soil as salt present in soil does not affect its accuracy, but it uses a costly keylogger system. The easy-to-use cost-efficient and durable soil dampness sensor is LM358 [13]. It has two-prong for sensing, that should be established in the land, which capacities as a cooperator and a function up to change the affectability of the sensors. The estimation of soil dampness sensor can change from 0 to 1023. 0 means most damp

Fig. 63.2 Gifkun capacitive soil sensor

condition, and 1023 means least damp condition. An another high standard sensor is [14] Gifkun (Fig. 63.2). This is dielectric-type capacitive sensor, it has voltage regulator in its own, and it is cheap and can be inserted deep into the soil.

63.5.2 Temperature Sensor

Temperature [15] is the most crucial parameter for any crop; hence, choosing a correct temperature sensor is a crucial task. Temperature is measured using temperature sensors. Accurate and exact measurement is required for a good yield of the crop, not just in the field of agriculture; students, scientists, doctors, engineers, etc., every field has to use temperature sensors for their purpose. These sensors sense the temp by measuring any difference in physical properties. Temperature sensors can be touchless or with contact depending on the type. The majorly used type of temperature sensor is a digital thermometer (Fig. 63.3) that is majorly used to measure temperature worldwide. Digital thermometer is simple but less correct, easy to use but less durable can measure temperatue from + 25 to + 45 °C means it has a limited range but no external source is required for same.

Infrared thermometer [16] senses the infrared radiation of the object and accurately tells the precise temperature reading. In the COVID-19 pandemic, it has become a boon to the world as this is a non-touch-type temperature sensor. A thermocouple sensor [17] is cost-effective, self-powered sensor. As it do not heat itself and comes with a range of − 200 to + 2150 °C, the downside of this sensor is accuracy, less sturdiness, cold junction, and sensitivity. An RTD sensor in case is more stable,

Fig. 63.3 Digital thermometer

accurate, and linear, but it has resistance error, high cost, little slow in terms of measurement and is bulky.

A thermistor sensor [15] is highly sensitive, but nonlinearity is there. It is more accurate and rugged, but current source is needed for working, and it is fragile with self-heating issue with a temperature range of − 100 to + 260 °C.

Semiconductor-type sensor [17] has accurate output. It is cost-effective and gives nonexponential sensing, but it needs source, a little slow in measurement, heats the machine, and has a limited sensor range of − 50 to + 160 °C. An accurate and precise sensor, that is [18] DS1820 temperature sensor, can measure a temperature between − 15 to 125 °C. DS1820 temperature sensor is cheap, highly accurate, and easily available, and the best thing about this sensor is fully waterproof and can be placed deeply into the soil for most accurate readings (Fig. 63.4).

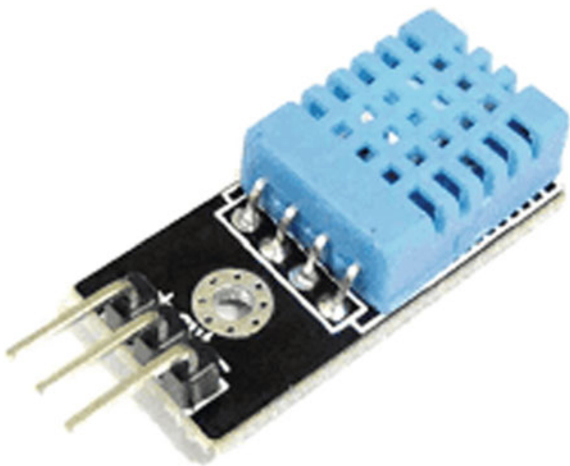
63.5.3 Hygrometers

Hygrometers/humidity sensors [19] sense the humidity levels of the air. Humidity simply means the percentage of H_2O molecules present in the air. In agriculture, humidity plays an important key role for a good crop, not just in the field of agriculture. It has uses in domestic, industrial, medical, electronics, and semiconductors also. There are majorly three types of hygrometers, namely in accordance with their principal capacitive, resistive, and thermally conductive. Capacitive RH hygrometers are linear, has a wide range of sensing, and stable, but its downside is limited

Fig. 63.4 DS1820 temperature sensor

signaling. The [19] resistive sensor is cost-efficient but easily moderable by pollutants. It has high range of sensing, but outcome differs with hydophilic products. The thermal conductive hygrometer is sturdy and has a good resolution but sensitive to thermal gases. A best humidity sensor is named DHT 11. It has in-built temperature and humidity sensor, but its temperature sensing part is not that accurate. So, it can be used for only sensing humidity level (Fig. 63.5).

Fig. 63.5 DHT-11 sensor



63.6 Communication Systems

Communication tools for smart and remote irrigation can be classified into three major parts named as below:

1. GSM.
2. Bluetooth.
3. Wireless network/Internet.

63.6.1 Global System for Phones/Mobiles (GSM)

GSM [20–22] is used to transfer voice and data from one phone to another phone. The bell laboratories in 1970 implemented this concept via a phone application radio arrangement. GSM is an open-source, digitated phone application telecommunication used to transfer voice/data. The frequency at which GSM today operates is 850, 900, 1800, 1900 MHz. GSM is a highly cost-effective phone application technology. It is exceptionally proficient and joining innovation through a versatile organization. GSM has improved rangeability, low estimated phone application set, and raised standard discourse. GSM aids the decrease of time pull out a human mistake in controlling the dirt dampness level and overstate their net creation. Some attributes of the GSM are world roaming, IDN, FDN, and majorly wide acceptance over the earth.

SIM900A [23] is a popular module among GSM modules and works on 900 megahertz that is Indian phone-friendly frequency. It is majorly used as GSM module, and it has six pins that are Vcc and double Rx and Tx pins. It uses serial communication for its operation.

SIMENS TC35 module [24] is a GSM-based module that uses ARM7 controller and can run on both computer and phone that use C language. Nokia FBUS [25] is an GSM-based module by 8051 family microcontroller, works on Kiel, LinuxOS, and uses both C and Java language for its operation but can only be worked on computer system. SIMENS TC35 [26] module with Msp430f149 controller is an GSM-based works on C430 IDE and uses both C language for its operation but can only be worked on computer system.

63.6.2 Bluetooth

Bluetooth [27] is a remote innovation that is designed to interface gadgets and different versatile gear. It has a low consumption of electricity for its work and is pretty quick. Bluetooth is operated only for a short distance. Bluetooth is also determined as a less energy-consuming correspondence. It has a scope of approximately 8–10 m (25–28 ft.). It underpins the recurrence band of 2.45 GHz. It is used in pan networks or within a personal network.

CBOEMSP [28] a312 is a Bluetooth-based module that uses Atmega 32 microcontroller and works on mobile with AVR studios and Eclipse frameworks on C or Java language. Similarly, Tdkblue2i [29] is an Bluetooth-based module that uses Atmega 64 controller, works on computer, and uses AVR studio with C programming language.

Blue-giga Wt11 [30] is another Bluetooth-based module with Atmega 168 controller, and it uses AVR studio and Symbian framework with C or Python language. Blur-giga can be used only on computer.

63.6.3 Wireless Network/Internet

Wireless network/Internet [31, 32] Commonly, WI-FI is a remote system administration innovation that permits gadgets, for example, cell phones and other gear (printers and camcorders), to interface with the Web. It permits these gadgets—and some more—to trade data with each other, making an organization. Web network happens through a remote switch. Recently, by the introduction of 5G technologies, Internet has gone up to the speed of several gigabit-per-second (Gbps), billions of people are connected with this at the point when you access WI-FI or Internet, and you are associating with a remote that permits your WI-FI-viable gadgets to interface with the Web and Internet. This [33] remote monitoring and accessing framework based on net uses 8051 microcontroller with Keil IDE framework. A computer system is required for this with programming language Java or Interactive C for its operation.

A PIC 16 microcontroller [34] is used in this remote monitoring and accessing framework. Based on Internet system, it uses MPLAB IDE as tool and works on UML and C programming languages.

63.7 Outcome and Benefits

1. These mechanized watering arrangements grounded on the Internet of Things are highly cost-effective, efficient and easy for the farmer, both software and hardware uses.
2. It reduces the water overuse and uses less water to irrigate whole farmland.
3. It not only saves the water requirement but also helps the farmer in case of drought as less water is required.
4. It makes works less tedious and increases the productivity.
5. It causes ranchers to spare water, increment yields, and to expand the nature of the harvest. The sensor assists with conveying results right away. As the GSM module help in Remote transmission is conceivable. Soil sensor help in less soil aggravation i.e., keeps up the fruitfulness of the dirt.
6. No consistent upkeep is needed for the framework. This is an energy and normal assets sparing cycle. Inhumane toward typical saltiness.

7. It forestalls over-water system. Information produced through the sensors is exceptionally exact and momentary. It gives ongoing information and aides in water preservation with low operational expenses.
8. It gives exact homestead and field assessments. It improves animal cultivating—reduced ecological impression, and helps in far off checking of the field.

63.8 Conclusion

It is concluded from the above automated models that these systems propose an automated irrigation system that monitors and maintains the desired soil moisture content via an automatic irrigation system. This system not only saves the water requirement but also helps the farmer in case of drought as less water is required in this system. This system can be very useful in the severe drought-affected area of India like Rajasthan and Maharashtra. It makes the work of the farmers less tedious using different sensors like moisture sensors, temperature sensors, humidity sensors, and efficient WI-FI networks. It helps in the high productivity of crops and hence a stable income to farmers.

63.9 Future Scope

There is a very high scope of these technologies as future aspects. This setup will help farmers with less tedious work with efficient and highly productive crops. If these technologies get implemented on the ground level, it will help the farmers to save water, save money on fertilizers and chemicals needed for agriculture.

The setup of various sensors and microcontrollers will save farmers from overhead work and provide them a better version of traditional irrigation. It contributes and comes up with new technology for farmers or countrymen to upgrade the quality of the irrigation system and helps in production. It will become an aid for the farmers where water crisis or drought situation occurs or does not have an adequate amount of water.

References

1. Iyer, S., et al.: IoT based cost-effective centralised smart irrigation system using LoRa. Available at SSRN 3565531 (2020)
2. http://www.ensembletech.in/lora-frequency-bands-india
3. Rane, D.: Review paper based on automatic irrigation system based on rf module. IJAICT **1**(9) (2015)
4. Prakash, B.R., Kulkarni, S.S.: Super smart irrigation system using internet of things. In: 2020 7th International Conference on Smart Structures and Systems (ICSSS), Chennai, India, pp. 1–5 (2020). https://doi.org/10.1109/ICSSS49621.2020.9202361

5. Karpagam, J., Merlin, I.I., Bavithra, P., Kousalya, J.: Smart irrigation system using IoT. In: 2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS), Coimbatore, India, pp. 1292–1295 (2020). https://doi.org/10.1109/ICACCS48705.2020.9074201
6. Meena, H., Nandanwar, H., Pahl, D., Chauhan, A.: IoT based perceptive monitoring and controlling an automated irrigation system. In: 2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT), pp. 1–6. IEEE (2020)
7. Chauhan, A., Nandanwar, H., Pahl, D., Meena, H.: Smart IoT based monitoring irrigation system-present Indian perspective. Int. J. Adv. Sci. Technol. **29**(7), 2307–2314. Retrieved from http://sersc.org/journals/index.php/IJAST/article/view/1796
8. Nandanwar, H., Chauhan, A., Pahl, D., Meena, H.: A survey of application of ML and data mining techniques for smart irrigation system. In: 2020 Second International Conference on Inventive Research in Computing Applications (ICIRCA), pp. 205–212. IEEE (2020)
9. https://www.arduino.cc/en/guide/introduction
10. https://www.raspberrypi.org
11. https://wiki.eprolabs.com/index.php?title=Moisture_Sensorn
12. https://www.elprocus.com/soil-moisture-sensor-working-and-applications/
13. https://www.vernier.com/product/soil-moisture-sensor/
14. Vulpe, C.L., Mihai, C.: Research on infrared body temperature measurement—virus spreading prevention. In: 2020 12th International Conference on Electronics, Computers and Artificial Intelligence (ECAI), Bucharest, Romania, pp. 1–4 (2020). https://doi.org/10.1109/ECAI50035.2020.9223195
15. https://www.electronics-tutorials.ws/io/io_3.html
16. https://www.researchgate.net/publication/262199118_REVIEW_OF_TEMPERATURE_MEASUREMENT_AND_CONTROL
17. https://www.omega.co.uk/temperature/z/thermocouple-rtd.html
18. Nasution, T.H., Muchtar, M.A., Siregar, I., Andayani, U., Christian, E., Sinulingga, E.P.: Electrical appliances control prototype by using GSM module and Arduino (2017)
19. https://www.electronicshub.org/humidity-sensor-types-working-principle/#Humidity_Sensors_Classification_and_Working_Principle
20. https://www.academia.edu/9587271/GSM_based_Automatic_Irrigation_System?show_app_store_popup=true
21. https://www.tutorialspoint.com/gsm/gsm_overview.html
22. https://www.elprocus.com/gsm-technology-architecture-its-applications/
23. Ding, J., Zhao, J., Ma, B.: Remote monitoring system of temperature and- humidity based on GSM. In: 2nd International Conference on Image and Signal Processing, pp. 1–4 (2009)
24. Ahmed, V., Ladhake, S.A.: Design of ultra low cost cell phone based embedded system for irrigation. IEEE Trans. Consum. Electron. **55**(2), 718–721 (2010)
25. Huang, W., Wang, G., Lu, J., Gao, F., Chen, J.: Re-search of wireless sensor networks for an intelligent measurement system based on A.R.M. In: International Conference on Mechatronics and Automation Conference on, pp. 1074–1079 (2011)
26. Ahmed, V., Ladhake, S.A.: Novel ultra low cost remote monitoring system for home automation using cell phone. In: International Conference on Computational Intelligence and Communication Systems, pp. 569–573 (2011)
27. https://www.elprocus.com/how-does-bluetooth-work/
28. Wijetunge, S.P., Wijetunge, U.S., Peiris G.R.V., Aluthgedara C.S., Samarasinghe, A.T.L.K.: Design and implementation of a bluetooth based general purpose controlling module. In: 4th International Conference on Information and Automation for Sustainability, pp. 206–211 (2008)
29. Piyare, R., Tazil, M.: Bluetooth based home automation system using cell phone. In: IEEE 15th International Symposium on Consumer Electronics, pp. 192–195 (2011)
30. Al-Ali, R., Al-Rousan, M.: Java-based home automation system. IEEE Trans. Consum. Electron. **50**(2), 498–504 (2004)
31. https://www.home-network-help.com/wireless-network.html

32. https://www.cisco.com/c/en_in/products/wireless/what-is-wifi.html
33. Alkar, Z., Buhur, U.: An internet based wireless home automation system for multifunctional devices. IEEE Trans. Consum. Electron. **51**(4), 1169–1174 (2005)
34. Suresh, R., Gopinath, S., Govindaraju, K., Devika, T., Suthanthira Vanitha, N.: GSM based automated irrigation control using Raingun irrigation system. Int. J. Adv. Res. Comput. Commun. Eng. **3**(2) (2014)

Chapter 64
Compressed Sensing MRI Reconstruction Using Convolutional Dictionary Learning and Laplacian Prior

Mrinmoy Sandilya and S. R. Nirmala

Abstract In compressed sensing magnetic resonance imaging (CS-MRI), applications of dictionary learning techniques have craved a decade long way with the development of methods like K-SVD, matching pursuits, etc. Dictionary learning methods are particularly useful in context of input signal adaptability. The data acquisition process of MRI is noisy in nature with various types of noise associated, like Rician, Gaussian, Rayleigh noise, motion artefacts like breathing artefacts, etc. In this context, training a dictionary directly with the noisy training samples may lead to an inefficient dictionary. Moreover, complexity and size of the constructed dictionary may be very big. This paper proposes a Laplacian sparse dictionary (LSD) technique for obtaining a concise and more representative dictionary which utilizes the concepts of manifold learning and double sparsity for MR image. This can be utilized to reconstruct an MR image using any of the existing compressed sensing methodology. The method along with online convolutional dictionary learning (CDL) has been demonstrated in this manuscript. Keeping in mind increased efficiency and reduced reconstruction time, the proposed method attempts to tackle the problem of MR image reconstruction. The results obtained from the proposed method have been compared with traditional CS-MRI methods using metrics—PSNR and SSIM.

Sparse reconstruction techniques like compressed sensing (CS) have gained momentum and popularity since last decade in the field of image and video processing. In magnetic resonance imaging (MRI), sparse reconstruction techniques have been used. Traditional MRI is ill-reputed to have greater exposure time for patients due to slow data acquisition system, motion artefacts and breathing artefacts in the reconstructed image, in spite of all the advantages it poses. In this context, sparse representation techniques did some pioneering work towards popularizing the idea of saving time in the expanse of obtaining under-sampled data and trying to reconstruct the image [1]. The basic requirements of traditional sparse reconstruction methods

M. Sandilya (✉)
Department of Electronics and Communication Engineering, Gauhati University, Guwahati, India

S. R. Nirmala
School of Electronics and Communication Engineering, KLE Tech. University, Hubballi, India

T. Senjyu et al. (eds.), *IOT with Smart Systems*, Smart Innovation, Systems and Technologies 251, https://doi.org/10.1007/978-981-16-3945-6_65

are; a sparsifying transform, an under-sampling scheme, and a reconstruction strategy [2]. It has been an established fact that better the ’sparsifying’ ability of the transform used, better or more concise is the sparse representation [3–5]. Numerous research work has been dedicated towards building an efficient sparsifying transform both fixed basis functions and adaptive basis using learning. In [6], a sub-sampled contourlet transform has been used as the sparsifying transform. Hao et al. [7] proposed the use of the contourlet as a sparsifying transform, combined with the fast iterative shrinkage/threshold algorithm (FISTA) for compressed sensing magnetic resonance imaging reconstruction. However, fixed transform bases have their limitations in representing variations in data. Sometimes, the same transform may represent a particular set of data better or perform worse in another.

In this context, data-adaptive, dictionary learning-based approaches like [8, 9] play important role in adapting to the input, which can further explore sparsity in input data.

In [8], the authors proposed a method called dictionary learning for MRI (DLMRI), which has been considered as a pioneering work in this field. The traditional DL-MRI alternating reconstruction algorithm is a two-step process—in first step, it learns the sparsifying dictionary and subsequently restores and fills in the k-space data in the other step. DLMRI depends upon K-SVD algorithm [10] to reconstruct. In [11], a beta-Bernoulli process as a Bayesian prior for dictionary learning which adaptively infers the dictionary size, the sparsity of each patch and the noise parameters was used. In [12], dictionary learning scheme is used at two resolution levels, i.e. a low-resolution dictionary is used for sparse coding, and a high-resolution dictionary is used for image updating.

The original information in MR representation usually has redundant, noisy information latched data which may add redundancy or effect the training of the dictionary. The data representation may be distorted by the atoms representing artefacts in methods like DL-MRI which is also computation costly. Also, over-complete dictionaries may not be optimal in all cases. Hence, prior preprocessing step is required to send the actual essence of the data to train the model. In this context, Li et al.[13] proposed a Laplacian sparse dictionary for natural image classification based on sparse representation, which implements manifold embedding and double sparsity dictionary learning for the issue in hand. The method maps a Laplacian weighted graph to the original sparse representation to obtain a truthful dictionary representation. Also, an l_1 norm sparsity was imposed on the dictionary itself to obtain a more compact dictionary so that computation complexity decreases as compared to computing a big redundant dictionary. This paper explores a compact dictionary for reconstruction along with Laplacian prior further as to find potential application in MRI reconstruction to obtain better results.

64.1 Methodology

64.1.1 Problem Formulation

Let us consider a signal $x \in \mathbb{C}^N$, and $y \in \mathbb{C}^M$ is the measurements available, such that $A_u x = y$, where $A_u \in \mathbb{C}^{M\times N}$ is the under-sampling matrix. Along with noise N, the noisy model can be formulated as

$$y = A_u x + N \tag{64.1}$$

In the similar limelight to Eq. 64.1, given $F_u \in \mathbb{C}^{M\times N}$ is the under-sampled Fourier transform matrix, compressed sensing MRI poses a problem formulation where we need to obtain this unknown, original signal x from some known representations of the signal y, by solving an under-determined system of linear equations -

$$\min_x \|\Psi x\|_0 \text{ s.t. } F_u x = y \tag{64.2}$$

where $\Psi \in \mathbb{C}^{k\times N}$ is a sparsifying transform like wavelet, curvelet, contourlet, etc., and $\| \|_0$ is the l_0 norm. Here, in Eq. 64.2, the signal X may not be a generic sparse signal for a Fourier transform, for which we introduce $D_u \in \mathbb{C}^{M\times N}$, such that we can represent X as $X = D\alpha$, where $\alpha = [\alpha_1, \alpha_2 \ldots \alpha_M] \in \mathbb{C}^M$ is the set of coefficients. Along with the dictionary learning model for data-adaptive transforms, the equation can be formulated as

$$\min_y \sum_j \left\|\alpha_j\right\|_0, \text{ s.t. } \left\|D_j\alpha_j - X_j\right\|_F^2 \le \eta \quad \text{where } j = 1, \ldots N \tag{64.3}$$

Here, η is a error threshold, $\|.\|_F$ denotes the Frobenius norm, and j is the number of patches. This l_0 sparse coding problem in Eq. 64.3 has already been established as an NP-hard problem [4]. Some Greedy algorithms have also been proposed for the same. Also, the l_0 norm in the formulation can be relaxed to l_1 norm, and the problem can be attempted to be solved using a very large number of methods like K-SVD [10], DLMRI [8], and it has been shown that these methods can actually reconstruct an MR image from its under-sampled k-space counterparts. The l_1 version of Basis pursuit denoising (BPDN) is formulated as follows:

$$\min_{\alpha_j, D} \frac{1}{2} \left\|D\alpha_j - x_j\right\|_2^2 + \lambda\left\|\alpha_j\right\|_1 \text{ subject to } \Sigma D \le c \tag{64.4}$$

Here, λ is a trade-off constant, and c is an constant threshold. Equation 64.4 is similar to solving two least squares optimizations iteratively, one for finding the sparse coefficients α while keeping the dictionary fixed, and the other is updating the dictionary D while keeping coefficients fixed.

An alternate online dictionary learning formulation [14, 15] for the same given in Eq. 64.4 can be done in a streaming fashion with the availability of the data with two steps as follows:

1. The sparse coefficent is optimized first. Here, t stands for a time instance.

$$x_t = \min_{\alpha_j} \left\| D_{t-1}\alpha_{j_t} - x \right\|_2^2 + \lambda \left\| \alpha_{j_t} \right\|_1 \tag{64.5}$$

2. The dictionary D_t is updated with the given set of information

$$\left\{ \{D_\tau\}_{\tau=0}^{\tau=t-1}, \{\alpha_\tau\}_{\tau=1}^{\tau=t}, \{x_\tau\}_{\tau=1}^{\tau=t} \right\}.$$

The proposed methodology is based on convolutional form of the BPDN (CBPDN). Given dictionary atom set $\{d_m\}_{m=1}^{M}$ of D and x, the convolutional dictionary learning (CDL) is given as

$$\min_{\{d_m\},\left\{\alpha_{j_{k,m}}\right\}} \frac{1}{2} \sum_{k=1}^{N} \left\| \sum_{m=1}^{M} d_m * \alpha_{j_{k,m}} - x_{j_k} \right\|_2^2 + \lambda \sum_{k=1}^{N} \sum_{m=1}^{M} \left\| \alpha_{j_{k,m}} \right\|_1 \tag{64.6}$$
$$\text{subject to } \|d_m\| \leq 1, \ \forall m \in \{1, \cdots M\}$$

Here, the norm constraint $\|d_m\| \leq 1$ eliminates any scaling mismatch between d_m and $\alpha_{k,m}$ by keeping dictionary atom values under 1. The coefficients set $\alpha_{k,m}$ represents the signal x_k.

64.1.2 Proposed Method

In order to obtain clean patches from noisy ones, the proposed method has introduced a Laplace prior designed to operate on a basis comprising set of the eigenvectors of the Laplacian graph of the image. In [16], a Laplacian prior system algorithm called eigenvectors of the graph Laplacian (EGL) is designed such that it preserves the intra-image geometrical features like edges and curves. Similar to [16], the mathematical formulation for the proposed Laplacian prior is given based on following rules:

1. Nearest neighbour graph G: G is formulated with a set of feature vectors f_i and f_j such that if f_j is a nearest neighbour of f_i then f_i and f_j are connected in G.

2. Weight matrix: The weight matrix W with entries w_{ij} is formulated as

$$w_{ij} = \begin{Bmatrix} e^{-\frac{\|y_i - y_j\|^2}{k}}, & \text{if } y_i, y_j \text{ are connected} \\ 0, & \text{otherwise} \end{Bmatrix} \tag{64.7}$$

Here, k is a scaling constant. The normal Laplacian matrix L is given as

$$L = I - B^{-1/2}\, W B^{-1/2} \tag{64.8}$$

Here, I is an identity matrix, B is a diagonal matrix, and its diagonal entries are the row sums of W. If we suppose the eigenvectors of L are $\{u_i\}_{i=1}^{N}$ and $U = \{u_1, u_2 \ldots u_M\}$ is the set of first M eigenvectors, and then, the denoising estimation prior can be given as

$$\tilde{Y} = (YU)\, U^T \tag{64.9}$$

Here, $\tilde{Y}$ is the estimation measurement matrix. If we take $C = UU^T$, the equation in 64.5 can be expressed as

$$\widehat{x}_t = \min_{\alpha_j} \left\|\alpha_{j_t}\right\|_1 \text{s.t. } \left\|D_{t-1}\alpha_{j_t} - \rho E_j g_{ij}\right\|_2^2 \leq \varepsilon' \tag{64.10}$$

Here, $\widehat{x}_t$ is a denoised estimate x_t, $E = [e_1 e_2 \ldots e_N]$ is the residue matrix, $e_k = y_k - \sum_{j \neq i} D x_j c_{j,l}$ for a particular patch e_k, g_i is the ith row vector of C, and ρ is a normalizing constant. The proposed method of incorporating Laplacian prior has been compared with traditional BPDN with OMP and online CDL.

64.1.3 Alternating Direction Method of Multipliers (ADMM)

The optimization algorithm chosen here for BPDN reconstruction and proposed method is alternating direction method of multipliers (ADMM). The ADMM algorithm is a very versatile algorithm with basic Lagrangian formulation for a equation of the form $Ax = b$ as follows :

$$L(x, y) = f(x) + y^T(Ax - b) \tag{64.11}$$

Here, x, ye are the primal and dual variables, respectively. Then the primal and dual variable feasibility conditions are given as

$$0 = \nabla L(x^*, \cdot) \Rightarrow Ax^* - b = 0 \;\&\; 0 \in \partial L(\cdot, y^*) \Rightarrow 0 \in \partial f(x^*) + A^T y^* \tag{64.12}$$

Here, ∂ is the sub-differential operator [17]. The residuals of primal and dual variables can be used to define stopping criteria and convergence of the optimization [18]. Here, Eq. 64.10 has been formulated as an ADMM consensus problem which then can be solved iteratively.

64.2 Experimental Set-Up

The experimental set-up comprises a CPU having specifications (20 core, 64 GB RAM, 2080 TI—11 GB graphics memory). The coding part has been done in Python 3.6–64 bit, library used is SPORCO [19], DLMRI [8], PyCurvelab and default python libraries. The data-set used for training the model has been obtained from "FastMRI" data-set [20]. The training data-set used here comprises 50 fully sampled knee MRI obtained under magnetic field at 3 or 1.5 T.

64.3 Results

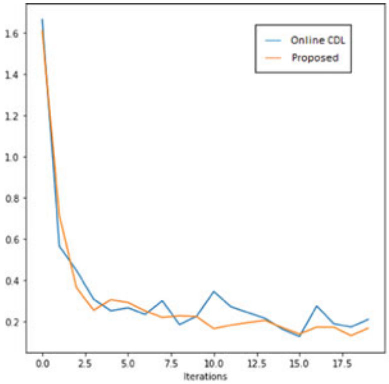
Dictionary $D \in 64 \times 128$ is learned in a set of image with 8×8 image patches over 50 iterations. The reconstruction examples for BPDN and proposed method have been shown in Fig. 64.1. The knee MRI original image shown in Fig. 64.1a is under-sampled by an under-sampling mask. The BPDN reconstruction in Fig. 64.1b is smoothed out, and proposed reconstruction shown in Fig. 64.1c retains the edges and curves of the image. The reconstruction comparisons have been made with PSNR and SSIM and computation time (in seconds) required for the algorithm to run. The comparisons have been shown in Table 64.1. It has been observed that BPDN has the lowest PSNR and SSIM among the three methods although it has the least

(a) Original+noise (b) BPDN (c) Proposed

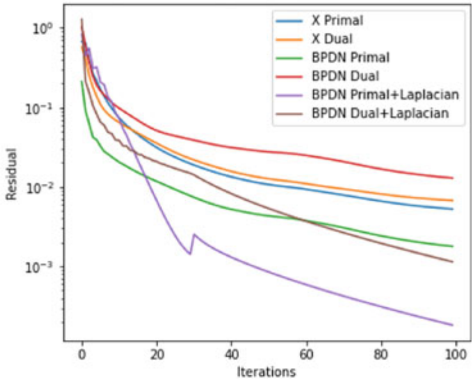
Fig. 64.1 Reconstruction examples—original, BPDN, proposed

Table 64.1 Performance comparison using BPDN, online CDL and proposed

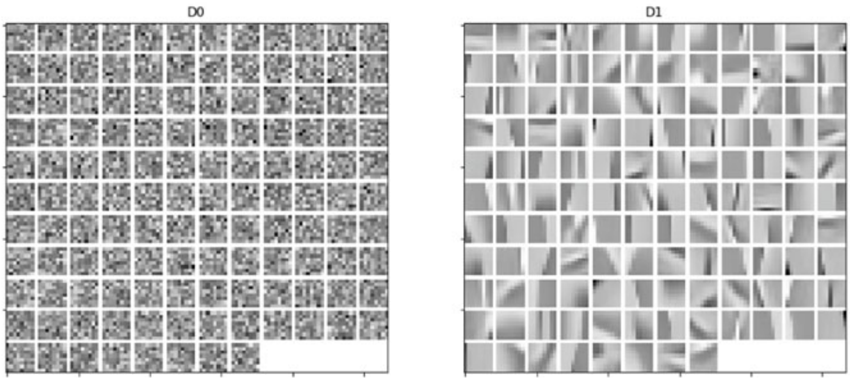
Method	Computation time (s)	PSNR (dB)	SSIM
BPDN	11.01	30.17	0.76437
Conventional online CDL	113.45	42.53	0.8912
Proposed method	154.23	42.62	0.92893



(a) Cost function versus iterations



(b) Residue versus iteration



(c) Dictionary learned : D0: initial dictionary D1: Final dictionary for Online CDL

Fig. 64.2 **a** Cost function versus iterations for online CDL and proposed, **b** residue versus iteration shown for X primal, X Dual, BPDN primal, BPDN Dual, BPDN+Laplacian primal, BPDN+Laplacian Dual **c** Dictionary learned

computation time required. Proposed method outperforms BPDN and online CDL in terms of PSNR and SSIM although it has the highest computation time.

Figure 64.2a shows plot between cost function minimization over time in terms of number of iterations. The convergence rate of minimization for both online CDL and proposed method is almost similar. Figure 64.2b shows the residue reduction for both primal and dual formulation of the cost function. The comparisons have been shown for original X primal-Dual, BPDN primal-Dual and BPDN along with Laplacian prior primal- Dual. As can be observed from the graph, the Laplacian BPDN formulation has better residual convergence rate as compared to its counterpart. Figure 64.2c shows the initial dictionary and final dictionary learned using the proposed method. As seen in figure, the initial dictionary D0 is almost pure noise and very random, whereas the final dictionary D1 learns the high level features of the image like edges and curves. The final dictionary represents the fine features for the smooth reconstruction.

64.4 Conclusion

Compressed sensing dictionary learning has seen tremendous growth in recent times, along with commercial MRI machines integrated with the technology already rolling out in the market. Keeping in mind the importance of the technology, a Laplacian prior has been proposed here for online CDL reconstruction method for MR imaging. The Laplacian prior is shown to be effective for online CDL as well as BPDN reconstruction of MR imagesusing performance parameters—PSNR and SSIM. The reconstruction algorithm used here is ADMM, and the versatility and popularity of which is well established. However, ADMM is more suitable for only convex optimization and sometimes takes a long time to converge to an optimum quality of reconstruction. There are some robust improvements in accelerating ADMM which shall be considered as a future scope for the work in this manuscript. Also, a dedicated perceptual error prior has been planned for the future, which can be used in dictionary learning reconstruction.

References

1. Sandilya, M., Nirmala, S.R.: Compressed Sensing trends in magnetic resonance imaging. Eng. Sci. Technol. Int. J. **20**(4), 1342–1352 (2017)
2. Candès, E.J., Romberg, J., Tao, T.: Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. IEEE Trans. Inform. Theory **52**(2), 489–509 (2006)
3. Donoho, D.L.: Compressed sensing. IEEE Trans. Inform. theory **52**(4), 1289–1306 (2006)
4. Lustig, M., Donoho, D., Pauly, J.M.: Sparse MRI: the application of compressed sensing for rapid MR imaging. Magnetic Reson. Medi. Off. J. Int. Soc. Magnetic Reson. Med. **58**(6), 1182–1195 (2007)
5. Lustig, M., Donoho, D.L., Santos, J.M., Pauly, J.M.: Compressed sensing MRI. IEEE Signal Process. Mag. **25**(2), 72–82 (2008)
6. Qiaoling, W., Lin, N., Delong, H.: Compressed sensing based on nonsubsampled contourlet transform for image reconstruction. J. Univ. Sci. Technol. China **2** (2012)
7. Hao, W., Li, J., Qu, X., Dong, Z.: Fast iterative contourlet thresholding for compressed sensing MRI. Electron. Lett. **49**(19), 1206–1208 (2013)
8. Ravishankar, S., Bresler, Y.: MR image reconstruction from highly undersampled k-space data by dictionary learning. IEEE Trans. Med. Imaging **30**(5), 1028–1041 (2010)
9. Huang, Y., Paisley, J., Lin, Q., Ding, X., Fu, X., Zhang, X.-P.: Bayesian nonparametric dictionary learning for compressed sensing MRI. IEEE Trans. Image Process. **23**(12), 5007–5019 (2014)
10. Aharon, M., Elad, M., Bruckstein, A.: K-SVD: an algorithm for designing overcomplete dictionaries for sparse representation. IEEE Trans. signal Process. **54**(11), 4311–4322 (2006)

11. Ding, X., Paisley, J., Huang, Y., Chen, X., Huang, F., Zhang, X.-P.: Compressed sensing MRi with Bayesian dictionary learning. In: 2013 IEEE International Conference on Image Processing, pp. 2319–2323. IEEE (2013)
12. Song, Y., Zhu, Z., Lu, Y., Liu, Q., Zhao, J.: Reconstruction of magnetic resonance imaging by three-dimensional dual-dictionary learning. Magnetic Reson. Med. **71**(3), 1285–1298 (2014)
13. Li, F., Sheng, J., Zhang, S.-Y.: Laplacian sparse dictionary learning for image classification based on sparse representation. Front. Inform. Technol. Electron. Eng. **18**(11), 1795–1805 (2017)
14. Liu, J., Garcia-Cardona, C., Wohlberg, B., Yin, W.: Online convolutional dictionary learning. In: 2017 IEEE International Conference on Image Processing (ICIP), pp. 1707–1711. IEEE (2017)
15. Liu, J., Garcia-Cardona, C., Wohlberg, B., Yin, W.: First-and second-order methods for online convolutional dictionary learning. SIAM J. Imaging Sci. **11**(2), 1589–1628 (2018)
16. Tang, Y., Chen, Y., Xu, N., Jiang, A., Zhou, L.: Image denoising via sparse coding using eigenvectors of graph laplacian. Digital Signal Process. **50**, 114–122 (2016)
17. Hiriart-Urruty, J.-B., Lemaréchal, C.: Fundamentals of Convex Analysis. Springer Science & Business Media, Berlin (2004)
18. Wohlberg, B.: ADMM penalty parameter selection by residual balancing (2017). arXiv preprint arXiv:1704.06209
19. Wohlberg, B.: SPORCO: a python package for standard and convolutional sparse representations. In: Huff, K., Lippa, D., Niederhut, D., Pacer, M. (eds.) Proceedings of the 16th Python in Science Conference, pp. 1–8 (2017)
20. Knoll, F., Zbontar, J., Sriram, A., Muckley, M.J., Bruno, M., Defazio, A., Parente, M., Geras, K.J., Katsnelson, J., Chandarana, H., et al.: FastMRI: a publicly available raw k-space and dicom dataset of knee images for accelerated MR image reconstruction using machine learning. Radiol. Artif. Intell. **2**(1), e190007 (2020)

Chapter 65
Internet of Things (IoT)-Based Distributed Denial of Service (DDoS) Attack Using COOJA Network Simulator

Harshil Joshi and Dushyantsinh Rathod

Abstract The revolution in Internet of Things (IoT) technology has carried the astonishing ability to interconnect various devices that are traditionally known as "dumb" devices. On the contrary side, it has also welcomed hazard of billions of unprotected and easily hackable devices. These unexpected flooding of newly added vulnerable devices has welcomed unsought threats like distributed denial of service (DDoS) attacks. In this paper, a brief introduction of the IoT has been given, and broader categories of DDoS attacks have been explained. This paper is mainly focused on various types of DDoS attacks which are done in the recent era and chance to make unwanted damage in nearby future. At the end of paper, DDoS attack is implemented using Contiki operating system-based Cooja network simulator. Care must be taken till a strong intrusion detection system or any other technique is developed for the IoT-based network.

65.1 Introduction

The Internet of Things (IoT) in the twenty-first century stands unquestionably the most important development that can connect devices beyond personal computer or phone. IoT devices play a noteworthy responsibility in a world that are dominated by digital technology (i.e., automation). From a tiny device to a large driverless car, everything has become achievable because of the implementation of the IoT. The Internet of Things and its application can turn any object into a smart object, and these objects can be remotely available. Contrarily, at present time, the enterprise or the firms are in the race to quickly develop the IoT devices. Due to this rush in the development of the new IoT devices, the security of the devices is poorly designed [1, 2]. It can be claimed that this rush and competition in the revolution in the Internet

H. Joshi (✉)
Gujarat Technological University (GTU) and Devang Patel Institute of Advance Technology and Research (DEPSTAR), Charotar University of Science and Technology (CHARUSAT), Anand, Gujarat, India

D. Rathod
Sakalchand Patel Institute of Technology, Visnagar, India

T. Senjyu et al. (eds.), *IOT with Smart Systems*, Smart Innovation, Systems and Technologies 251, https://doi.org/10.1007/978-981-16-3945-6_66

of Things could lead to a potential disaster [3]. An increase in the number of linked or unsecured devices in the market can increase the potentials of hackers to access critical data of IoT devices by targeting them and as a result oversee our devices and control our lives [4–6]. This shift in the Internet of Things toward insecurity could take us back to distributed denial of service (DDoS) attacks powerful and more complex than earlier which could even not be identified or traced. Because of this increase in insecured Internet of Things devices, the number of those attacks has grown to a large extent, and this can lead to more and more criminal activities [7, 8].

The year 2016 was remembered as the year of Mirai as it was the critical point in which the combination of the DDoS and the insecured Internet of Things devices was piled up to create the largest DDoS attack ever noted. Thousands of connected devices were infected, and the major DDoS attack was seen which reached the nasty capability to about 1.2 terabytes (TB) per second [9, 10].

Interestingly, the impressive thing about the Mirai attack is not only the power of attack but was how the worm was able to infect those large ranges of devices or units. Those devices were affected by a very simple dictionary attack which was based on only sixty entries approximately. The attack was made only a very simple fact that the infected devices never changed the default login credentials, and some devices were not able to change because of technical reasons.

Recently in January 2019, an undisclosed client of Imperva experienced an outsized DDoS attack that received 500 million packets per second on their network. In the January 2019 attack, each packet was around 800–900 bytes in size. This means 500 packets of 850 bytes resulting in approximately 3.4 trillion bits of data targeting the Web site and resulting in an unresponsive one. This could affect the thousands of Internet of Things devices that were connected to it. All the above things lead to the unquestionable need to face the Internet of Things security problems.

65.2 How Does a DDoS Attack Work?

The Internet itself makes the DDoS attacks possible and more powerful. When the aim is to improve the functionality of the Internet and not the security, the Internet becomes inherently vulnerable to various security issues. Due to these security issues, the DDoS attacks are possible [11, 12].

To carry out a DDoS attack, an attacker needs to gain control of a network of online machines. These machines (Internet of Things devices) are infected with malware, and thus, each machine turns into a bot. This group of bots forms a "botnet", and the attacker has complete remote control over this botnet.

Once the attacker has complete remote control over this botnet, the attacker can send the updated instructions to each bot. Each bot from the botnet will send a request to the IP address of the victim and hence causing the network to overflow capacity.

As these bots are also legitimate Internet devices, they cannot be separated from the traffic and it results in the denial of service to the normal traffic. The DDoS needs to go throughout the following states to be struck [11, 12].

1. **Recruitment**: The attacker scans for the bots that can later be used to penetrate the attack to the victim machine. Previously, the attackers manually recruit the bots, but nowadays, there are many scanning tools available that can be used for this purpose.
2. **Forming a botnet**: Malicious code is injected into the vulnerable machines, and a botnet is created. At present, this stage is also been automated (the bots are recruited by self-propagating tools).
3. **Command and Control**: The attacker communicates with the bots via command-and-control architecture. In this way, the attacker gets to know about the active bots, upgrades the bots, and schedules an attack.
4. **Attack**: The attacker sends the command to the botnet, and the bots start sending malicious packets to the target. The attacker adds the parameters such as victim, number of packets, and duration of the attack. The attackers hide the IP address of the agent machines (bots); hence, machines, during the attack, will remain undiscoverable.

65.3 Distributed Denial of Service (DDoS) Attacks

There are many different types of DDoS attacks that are noted in today's life, and a very extensive range of classification is suggested in literature today. Elements that are seen in DDoS attack are as below.

1. The attacker is the main person behind the attack and on which the whole attack is based.
2. The organizer which is negotiable hosts a program running on them that can control many operators.
3. The person or the online work which is attacked by the attackers.

65.3.1 Types of Distributed Denial of Service (DDoS) Attacks

The DDoS attack can be categorized into three types, which are as below.

65.3.1.1 Volume-Based Attack

The volume-based attacks are also called volumetric DDoS attacks. In this attack, the attackers consistently surge the sufferer with the immense volume of vast connecting equipment, networks, high-frequency resources, and servers. This is called the most common DDoS attack. This attack includes various spoofed-packet floods like UDP floods and ICMP floods. The motive of the attacker is to consume the bandwidth of the victim's site. The degree of the attack is deliberate in bits per second (bps).

65.3.1.2 Protocol-Based Attacks

Protocol-based attacks mainly target manipulating a deficiency in Layer 3 or Layer 4 of the OSI layer. The motive of the attacker is to consume resources of the actual server. The degree of the attack is deliberate in packets per second.

65.3.1.3 Application Layer Attack

The application layer attack is also called the Layer 7 attacks. This attack refers to the type of malignant act which is designed to target the application layer of the OSI model where the common Internet requests such as HTTP POST and HTTP GET occur. The motive of the attacker is to target the Apache, Windows, or open BSD vulnerabilities, and more.

65.3.2 Examples of DDoS Attack

In today's era, we can find various DDoS attacks; few of them are described here.

65.3.2.1 UDP Flood

This is a subtype of volume-based attack. UDP is an abbreviation for User Datagram Protocol. UDP is a type of attack in which the attacker crushes random ports on the targeted host. UDP runs with lower overhead as it does not require a three-way handshake like TCP.

65.3.2.2 ICMP Flood

This is also a subtype of volume-based attack. ICMP flood is an abbreviation for Internet Control Message Protocol. It is also called the ping flood. In ping flood, an attacker takes control of the a victim's computer by sending ICMP echo requests.

65.3.2.3 TCP SYN Flood

This flood attack, a subtype of protocol-based attack, which exploits part of the normal TCP three-way handshake to consume resources on the targeted server and revive it passively. In a SYN flood attack, the targeted server is flooded with the repeated SYN packets sent by the attacker using a fake IP address.

65.3.2.4 HTTP Flood

This is a subtype of the application layer attack. During this, a web server or application is flooded via HTTP GET or POST requests sent by the attacker. HTTP flood attacks are volumetric attacks, habitually employing a botnet "zombie army."

65.3.2.5 Hydra

An open-source prototype malware that appeared in 2008 was the router that worked as an open-source tool by their own. The main purpose of the malware was to access the routers by using brute-force methods for performing DDoS attacks. The access to the router was possible by either with dictionary attack or with the use of the D-Link authentication bypasses method [13].

65.3.3 Implementation of DDoS Attack

In this paper, Contiki operating system-based Cooja network simulator is used to implement DDoS attack in the IoT network. Cooja network simulator provides environment that is nearer to real-time IoT network.

As shown in Fig. 65.1, we have created IoT network scenario with 25 nodes. In this network, Node 1 acts as server and Node 2 to Node 25 act as client. Server node is indicated by green color, and client node is indicated by yellow color.

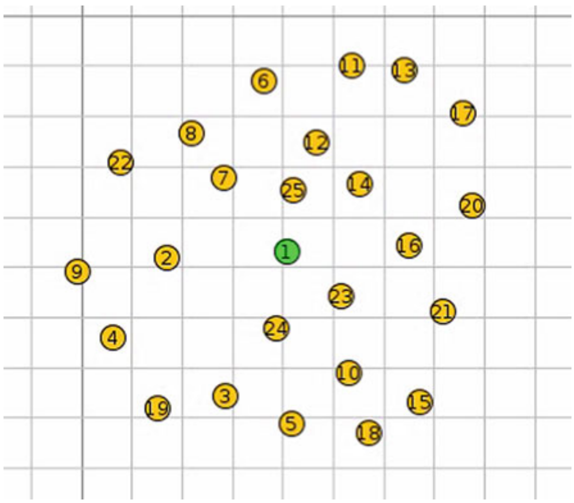


Fig. 65.1 IoT network scenario when simulation is not initialized

As shown in Fig. 65.2, each node starts communication with each other via sending HELLO packet. When any attacker node is connected, it may start to send multiple packets to each node (as shown in Fig. 65.3).

As we all know, IoT devices have limited amount of computational power which is not sufficient to handle large amount of packet simultaneously. Processing of such a huge amount of data will affect IoT devices badly. We must provide some lightweight solution which can work for the IoT network.

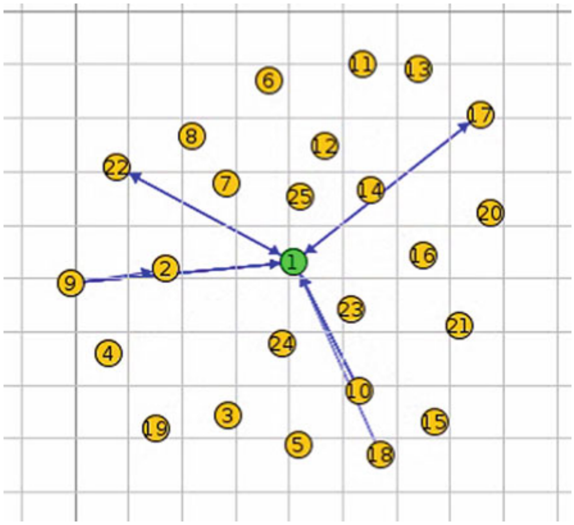


Fig. 65.2 IoT network scenario when simulation is initialized

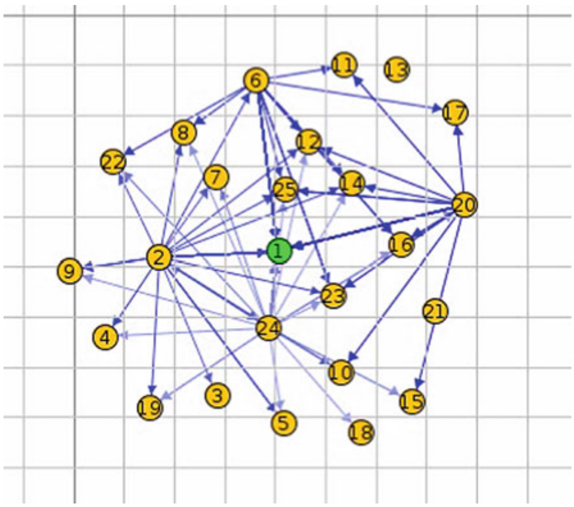


Fig. 65.3 IoT network is flooded with multiple packets sent via attacker node

65.4 Conclusion

In the recent era, we have seen insignificant growth in the IoT devices which are poorly designed and with limited security features. Due to lack of attention, fundamental urgency has been emerged in the market to redesign security parameter that can work with IoT devices with limited computational power and heterogeneously designed IoT network. Inspired by the intensified situation, various taxonomy of DDoS attack has been indentified and represented in the paper. Moreover, small scenario has been implemented to provide nearly real-time scenario of IoT-based network during DDoS attack. However, some techniques need to develop such as lightweight instution detection system for IoT-based network.

References

1. Granjal, J., Monteiro, E., Sa Silva, J.: Security for the internet of things: a survey of existing protocols and open researchissues. IEEE Commun. Surv. Tutorials **17**(3), 1294–1312 (2015)
2. Arias, O., Wurm, J., Hoang, K., Jin, Y.: Privacy and security ininternet of things and wearable devices. IEEE Trans. Multi-Scale Comput. Syst. **1**(2), 99–109 (2015)
3. Dragoni, N., Giaretta, A., Mazzara, M.: The internet of hack able things. In: Ciancarini, P., Litvinov, S., Messina, A., Sillitti, A., Succi, G. (eds.) Proceedings of the 5th International Conferencein Software Engineering for Defense Applications (SEDA16), Advances in Intelligent Systems and Computing. Springer, Berlin (2017)
4. Hughes, D.: Silent risk: new incarnations of longstanding threats. Netw. Secur. **2016**(8), 17–20 (2016)
5. Shukla, S.K.: Editorial: cyber security, IoT, block chains—risks and opportunities. ACM Trans. Embedded Comput. Syst. (TECS) **16**(3, article 62), 1–2 (2017)
6. Goyal, R., Dragoni, N., Spognardi, A.: Mind the tracker youwear: a security analysis of wearable health trackers. In: Proceedingsof the 31st Annual ACM Symposium on Applied Computing (SAC'16), pp. 131–136, Pisa, Italy (2016)
7. Bertino, E., Islam, N.: Botnets and internet of things security. IEEE Comput. **50**(2), 76–79 (2017)
8. Kolias, C., Kambourakis, G., Stavrou, A., Voas, J.: DDoS in the IoT: mirai and other botnets. IEEE Comput. **50**(7), 80–84 (2017)
9. York, K.: Dyn statement on 10/21/2016 DDoS attack, Dyn Blog (2016). http://dyn.com/blog/dyn-statement-on-10212016-ddosattack/
10. Hilton, S.: Dyn analysis summary of Friday October 21 attack, Dyn Blog (2016). http://dyn.com/blog/dyn-analysis-summary-offriday-october-21-attack/
11. Mirkovic, J., Reiher, P.: A taxonomy of ddos attack and ddos defense mechanisms. Comput. Commun. Rev. **34**(2), 39–53 (2004)
12. Douligeris, C., Mitrokotsa, A.: DDoS attacks and defensemechanisms: classification and state-of-the-art. Comput. Netw. **44**(5), 643–666 (2004)
13. Janus, M.: Heads of the hydra. Malware for network devices. Tech. Rep. Securelist (48) (2011)

Chapter 66
Designing an Email Security Awareness Program for State-Owned Enterprises in Namibia

Mamoqenelo Priscilla Morolong, Fungai Bhunu Shava, and Victor Goodson Shilongo

Abstract Emails offer easy and timely communication enabling businesses to connect to their stakeholders in real time. State-owned enterprises are no exception to this practice. Being a global enabler of communication, emails are a preferred medium for delivering malicious payload for hackers. Emails have well-known vulnerabilities such as malware, phishing, and spamming that can cripple and devastate an organization. Most attacks depend on the user performing some action such as clicking a link or sharing sensitive information making email security a human issue. Through their insecure behavior, employees can open doors for hackers into their organizational infrastructure. Change in behavior needs more than availing information on threats, risks, and unethical behaviors, and employees must be able to know or understand what is at risk and how to mitigate the email security risks. End-users/IT professionals must be motivated and willing to apply security practices/policies, and the latter requires a change of attitude and intentions. This qualitative exploratory research study presents results of an email security awareness study conducted with one state-owned enterprise in Namibia. Findings will be used to design an email security awareness program that will enhance the email security within the state-owned enterprise organizations and improve ICT awareness.

66.1 Introduction

State-owned enterprises (SOE) in Namibia make use of information and communication technology (ICT), and as a result, ICT security in SOE is of utmost importance. Namibia as one of the developing countries has seen a rise in Internet and email usage [1]; hence, it is imperative that researchers investigate the security level and

M. P. Morolong (✉) · F. B. Shava · V. G. Shilongo
Namibia University of Science and Technology, 13 Jackson Kaujeua Street, Windhoek, Namibia

F. B. Shava
e-mail: fbshava@nust.na

V. G. Shilongo
e-mail: shilogov@telecom.na

T. Senjyu et al. (eds.), *IOT with Smart Systems*, Smart Innovation, Systems and Technologies 251, https://doi.org/10.1007/978-981-16-3945-6_67

the degree of awareness regarding email security. In business environments, information has over the years been of great significance and is regarded as a very critical asset; it is often exposed to a variety of threats and vulnerabilities that are usually mitigated by a combination of both technical and procedural controls.

Electronic mail (email) enables business interactions and information exchange among customers, partners, and internal staff over the Internet or internal networks. These extensive capabilities have caused email to be widely adopted as the official communication method in organizations and for personal use, through a diverse number of compatible software clients and via web browsers [2]. With email being the professional global means of communication, companies should enforce comprehensive email security measures to mitigate spam infiltration targeting their employees.

Reference [3] reports that there are 3.7 billion email accounts around the world with around 269 billion emails sent per day. However, [4] coins that employees lack knowledge to protect themselves against email threats, thus, creating room for leakage of sensitive information such as identity numbers and new business prototype designs. In today's era, organizations heavily depend on their digital information, hence, if that information is compromised; business institutions suffer severe consequences such as loss of income, loss of customer trust, damage of organizational reputation, and sometimes legal actions [5].

In multiple cases of email data breaches, emails were used as an attack vector as reported by [5] where through emails: About 55% computer devices were infected with malware in 2018, stolen user credentials through phishing was 51%, senior executives credentials stolen was 51%, and accounts take over-based attacks were 42%. These statistics reveal that even though companies have embarked on investing in general security, more effort is still needed to secure email communication. Report by [6] justifies the number of spam emails received daily globally, and it also reveals the number of spambots received. Results from MacAfee require SOE to invest in email security awareness to protect themselves against email attacks/threats. The goal of email security awareness should encourage the adoption of secure online and offline behaviors in a working environment. The next section will present the research objectives of this study.

66.1.1 Research Objectives

The objective of this study is to design an email security awareness program for SOE in Namibia. The sub-objectives of the study are: examine threats to emails; investigate the causes of email data leakages in SOE; describe the methods of email data leakage in SOE; propose email security awareness model for SOE; how to help mitigate email security issues in SOE?

66.2 Literature Review

Electronic mail has become a primary target and transmitter of attacks because of its global use in the corporate and enterprise worlds; hence, it must be protected on both ends: sending and receiving [2]. This brings us to the discussion on threats to emails, causes of email data leakage, and methods of email data leakage.

66.2.1 Threats to Emails

Since their invention by Ray Tomlison in the 1960s, emails have been handled using different email protocols such as Post Office Protocol (POP), POP3, and Internet Message Access Protocol (IMAP). Like any other network-based platform, attackers have taken advantage of the weaknesses in the email protocols and are exploiting them for different purposes such as monetary gains, blackmails, and social destruction [7]. The common email threats are discussed in the following sections:

Malware. According to [7, 8], emails malware transmission has increased during the past three years to date. Reference [8] reports that email is twice likely to transmit malware to users than through exploit kits. Additionally, [9] alludes that malware can steal credentials, install additional threats, and spread itself further via email. The email malware statistics detailed by [8] reflects how emails spread malware to businesses.

Spam. Spam is said to include adware and phishing emails; there is evidence that spam escalate botnets as sending spam emails is a common strategy for monetizing botnets. Reference [10]. Even though Bill Gates declared in 2004 that all spam emails will be solved by introducing a spam box, there still exist myriad reports of spam cases [11].

User email behavior. Employees are the last line of defense, and most breaches are a result of human actions. At least spear-phishing attacks recorded 91% of data breaches, while a 1-billion-dollar criminal business ransomware was recorded in 2016. About $5.3 billion is recorded as CEO fraud or business email compromise (BEC) and continues to grow [12]. Regardless of the many email security solutions presented in the literature, report by [13] demonstrates that in 2019, there have been 94% of reported phishing email attacks. The articles show that about 54.9% of users click on a phishing link in less than 60 min, 12.38% in 60–120 min, 7.23% in 120–180 min, 5.18% in 180–240 min while 20.29% up to 17 h [14]; thus, user behavior exacerbates the email attacks in SOE. Having presented all these challenges, the next section focuses on existing solutions to address the threats.

66.2.2 Email Security Solutions

According to the report published by [5], decision-makers should focus on developing security policies to guide employees on the best practices on email usage. The report further indicates the importance of email security awareness and training to all employees. As opined by [15], companies tend to utilize security warning as a means of warning end-users against threats such as downloading malicious email attachments, installing malicious applications, or accessing suspicious emails. However, for this method to work, end-users need to understand what the security warning means; hence, some users were reported to ignore warnings due to complicated jargon. Employees need to undergo continuous training on what the security warnings mean and how to behave when they receive such warnings, in line with [5]'s claim of educating employees on email security as well as analyzing user behavior.

To address email security challenges, [16] introduced a personal-centered security awareness for information of human-based security challenges which proved to be a good way of providing security awareness to the employees.

66.3 Methodology

Purposive sampling as discussed by [17] was applied in selecting the study participants. The SOEs participants selected by the researcher included five end-users from different departments except IT department and two experienced IT personnel, adding up to a sum of seven participants which according to [17, 20] is an appropriate number for a qualitative study. To gather data from the participants, interviews and questionnaire were the main data collection tools used as they are part of qualitative data collection tools [18]. The tools were structured following this structure:

- Section 1: Demographic information of the participants to collect age, gender, department, work experience, and SOE.
- Section 2: Email security awareness gathered information on password management, spam email knowledge, behavior toward spam emails, and virus and its prevention knowledge.

The qualitative data analysis for interview responses was coding, and interview data were grouped into themes and finally coded following [19], while descriptive statistics was used to analyze questionnaire data.

66.4 Findings and Results

This section presents the findings per tool used.

66.4.1 Questionnaire Results

- **Age Distribution**: Participants are mature ranging from 30 to 50 years of age. This age group is generally tech-savvy as per the research done by the Smart Energy Consumer Collaborative.
- **Qualification Distribution**: Most participants have either a degree or a post-graduate degree (57%), and the remaining 43% is allocated to others (diploma or less).

The research questionnaire was devoted to capture five themes (i) attitude toward spam emails, (ii) what actions were taken toward spam emails, (iii) knowledge of being affected by the virus, (iv) reasons for opening spam emails, and (v) action taken when a laptop is being suspected of virus infection.

Of the participants, 14% opened spam compared to 86% who had not opened the spam mail, meaning participants are careful when it comes to suspicious emails. The actions with suspicious email flood their inbox reflects that most participants archive emails (43%) and delete and block senders' address (43%). These are the best practices of dealing with spam emails and 14% of the participants do nothing. This demonstrates secure email behavior by most participants; however, none of them reported to the IT department, and hence, technical controls could not be improved based on their experiences.

Some participants indicated that they open suspicious email out of ignorance or lack of knowledge on what to do, due to lack of awareness. Knowledge of being affected will help users know what to do, in this study; even though SOE does not have security strategies in place, participants responded positively, six out of the seven participants knew when their devices were infected by the virus. Furthermore, findings indicate that 34% of the participants report suspected viruses on work unit to the PC technicians for further investigation or repair, while 11% report their suspected work unit/laptop to email providers, and 11% took no actions. About 22% preferred repairing the work unit themselves because of their IT background, and the other 22% preferred to run anti-virus software to scan for virus on their work unit. In addition to the questionnaire data, the interview results are presented next.

66.4.2 Participants Interviews Data Themes

Support. Interview data revealed a lack of support from top management on email security which is influenced by lack of knowledge in the subject matter. Findings indicate that when higher positions are occupied by well-educated and security-driven personals and when genuine management activities are executed, guidance toward email security in SOE will follow. There is need for management's commitment and leadership, process and policies, skills and training, and a proper information security structure related to email security.

Professional Choices. There is a lack of information security which reflects the general lack of IT specialists in the SOE supported by Kron [13]. Developing countries fall victims' to hackers due to the limited funds and information security experts [19], and Namibia is not an exception.

Knowledge. This goes along with support, employees lack knowledge about email threats. They are not aware, nor well-trained about email threats. This is due to that the top management lacks advisers on this matter, and they also do not have enough knowledge about any security issues. The development of an email security culture within SOE is dependent on all employees' levels of awareness. Hence, it is crucial to educate and inform employees on potential email threats organizations might be faced with. Information security awareness also plays a critical role in the utilization of information system procedures, policies, and technology by end-users [19]. The current study identified that lack of email security awareness program in SOE to help improve the email security culture in organizations as a big challenge encountered by SOEs in Namibia.

Behavior. The interviews demonstrated positive behavior toward email security. This is also demonstrated by their level of knowledge regarding malware and virus. These results show the individual responsibility of each employee as they do not have any email security awareness program within their organization which confirms questionnaire results.

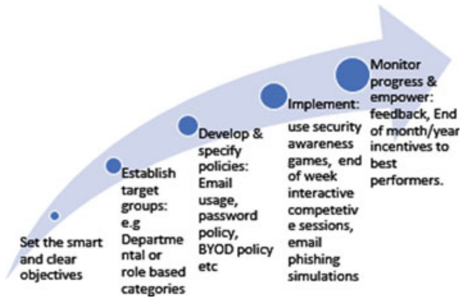
Bias/Discrimination. Interviews showed that employees are motivated by incentives such as professional opportunities which may include free training with certificates offered by their organizations. Incentives such as the best employee of the month who practiced maximum security where email security program is availed and monitored. These results are coupled with the questionnaire results presented in Section 66.4 assisted with the development of the proposed email awareness security model for SOE that is presented in the next section.

66.5 Proposed Email Security Awareness Model

An effective email security awareness program implementation for SOE requires proper planning, maintenance, and evaluation. In this section, the research study outlines the step to be taken to implement/design the email awareness program. Figure 66.1 presents the proposed email security awareness model for SOE.

Objective Setting: The program scope, goals, and objectives are core considering the research findings in the previous section which reflect the following elements: no email security awareness programs in place; lack of IT security specialists in SOE; malware/viruses are on the rise both globally and in SOE. From these set scopes, the researcher then identified the need for the SOE to set smart and clear objectives for the email security awareness tailored toward the organization's context. The objectives need to be focused on addressing the three elements.

Fig. 66.1 Proposed email security awareness model



Target Group: When designing an email security awareness program, the SOEs should consider grouping the targets, e.g., end-users, IT staff, senior management, human resource, and finance personnel. The program should distinguish groups of people as well as their roles and present the relevant information for a specific target group to harvest the best results. The following presents the suggested grouping strategy.

- Segment based on departments, job category, task, or function. Users need to be separated into groups according to their job descriptions or duty stations.
- Segment based on their level of email security awareness. Users will be grouped according to the level of their email security awareness knowledge. Easier to determine what information and method of communication to use. Phishing email simulations, surveys, questionnaires, and online security games can be used to measure employees' level of awareness.
- Segment based on the type of technology devices or IT system used. Users will require security training on the specific application, environment, operations, and devices they use.

Policy development: The SOE needs to implement an email security awareness and training program for all employees as such, the following steps should be considered: developing policies to deal with employees who fail to comply with related SOE email security policies; Ensuring that data systems activities records are revisited regularly; periodic auditing of system records, this can be effectively done quarterly; enforcement of policies; availing email policies to the employees through monitored and automated deployment on their email applications. This can be achieved by setting automatic reminders or banners when performing certain email processes.

Implementation: To change employee behaviors, SOE security team must use marketing strategies to sell their product to their audience as done by marketers. The audience is the targeted groups, e.g., end-users and IT staff. Custodians need to utilize different tools to deliver email security awareness to their targets. This can be achieved through email phishing simulations; security interactive games and/or end of week interactive awareness sessions or competitions as to measure the program's success.

Progress Monitoring: Metrics are measures of the email security awareness program, including human behaviors and service effectiveness, thus are important for measuring email awareness performances, planning program items, measuring productiveness of the program, etc. The SOE email security awareness program design will have two types of metrics: compliance metrics which measures the deployment of the email security awareness program and policy enforcement and impact metrics which measures the impact of the email security awareness program, e.g., running evaluation test on employees, and the results should answer the following questions: Is the program changing employees' behaviors?, Is email security risk being managed?, and Are email security risks being reduced?

The evaluation and review period can be monthly or quarterly depending on the target. Additionally, empowerment should be considered to motivate employees demonstrating a degree of change in behavior and compliance with email policy. Answering the above questions will determine the effectiveness of the designed awareness program.

66.6 Recommendations

SOE should invest in a consistent email security awareness training seasons that are a great tool to prepare employees on how to manage all kind of information security risks emphasizing on the types of threats facing employees and SOEs; that will aid employees to identify spam emails and study how to avoid clicking/opening wrong attachments and make use of strong passwords; strong email password that expires in 30 days, and employees must make sure passwords change regularly, which is the key to ensuring robust protection against cyber threats. In situations where employees find passwords too complex to memorize, password manager systems that include two-factor authentication (2FA) may be used. Email and information security policies based on the needs of an organization need to be developed and enforced to strengthen the awareness program.

66.7 Future Work

A highly possible future research study will consider investigating other SOEs and the investigations of the identified matters; with the findings, a refined email awareness security program for all SOEs will be proposed.

66.8 Conclusion

The study was aimed at proposing a model for an email security awareness program for SOE in Namibia. The study results revealed that the sample population of the study was aware of the email attacks; however, their organization lacked email security awareness program to consistently empower systematically. The study identified some email threats and how data are leaked through emails and thus found out lack of email security awareness program is the main contributing factor; hence, we proposed an email security model that SOE in Namibia can adopt and utilize to minimize email data leakage.

References

1. Sumaili, A., Dlodlo, N., Osakwe, J.: Adopting dynamic capabilities of mobile information and communication technology in Namibian small and medium enterprises. In: 2018 Open Innovations Conference OI 2018, pp. 213–222 (2018). https://doi.org/10.1109/OI.2018.8535941
2. CSA: SecaaS Category 4 Email Security Implementation Guidance. *Secaas Implement. Guid.* (2012)
3. Morolong, M., Gamundani, A., Bhunu Shava, F.: Review of sensitive data leakage through android applications in a bring your own device (BYOD) workplace. In: 2019 IST-Africa Week Conf. IST-Africa 2019, pp. 1–8 (2019). https://doi.org/10.23919/ISTAFRICA.2019.8764833
4. 'Examination of Insider Threats : A Growing Concern by Cecil L. Hartline Jr. A Capstone Project Submitted to the Faculty of Utica College December 2017 in Partial Fulfillment of the Requirements for the Degree of Master of Science in Cybersecurity' (2017)
5. Venter, I.M., Blignaut, R.J., Renaud, K., Venter, M.A.: Cyber security education is as essential as "the three R's". Heliyon**5**(12), 0–7 (2019). https://doi.org/10.1016/j.heliyon.2019.e02855
6. Osterman, R.: Best Practices for Protecting Against Phishing, Ransomware and Email Fraud (2018) [Online]. Available: http://www.lightsondata.com/white-paper-on-best-practices-for-powerful-dashboards/
7. 'McAfee Labs Threats Report: December 2018. Comput Fraud Secur (1), 4 (2019). https://doi.org/10.1016/s1361-3723(19)30004-1
8. Clark, J., van Oorschot, P.C., Ruoti, S., Seamons, K., Zappala, D.: Securing Email (2018) [Online]. Available: http://arxiv.org/abs/1804.07706
9. Malwarebytes LABS: 2019 State of Malware. Tech. Rep., p. 33 (2019) [Online]. Available: https://resources.malwarebytes.com/resource/2019-state-malware-malwarebytes-labs-report/
10. Vayansky, I., Kumar, S.: Phishing—challenges and solutions. Comput. Fraud Secur. **2018**(1), 15–20 (2018). https://doi.org/10.1016/S1361-3723(18)30007-1
11. Zhuang, L., et al.: Characterizing botnets from email spam records. In: LEET 2008—1st USENIX Work. Large-Scale Exploit. Emergent Threat. Botnets, Spyware, Worms, More (2008)
12. Symantec: Email Threats 2017 An ISTR Special Report Analyst: Ben Nahorney Internet Security Threat Report (2017) [Online]. Available: https://www.symantec.com/content/dam/symantec/docs/security-center/white-papers/istr-email-threats-2017-en.pdf
13. Kron, E.: Email Safety and Security: The Human Firewall About Erich Kron (2017) [Online]. Available: http://www.fgfoa.org/docs/default-source/past-presentations/annual-conference/2017-Annual-Conferenece/christopher-ramos-and-erich-kron---safety-and-security---the-human-firewall.pdf?sfvrsn=9176dfd5_0

14. Mimecast: The State of Email Security Report 2019. Mimecast (2019) [Online]. Available: https://www.mimecast.com/resources/ebooks/dates/2019/8/the-state-of-email-security-report-2019/
15. Bada, M., Sasse, A.M., Nurse, J.R.C.: Cyber Security Awareness Campaigns: Why do they fail to change behaviour?, pp. 118–131 (2019) [Online]. Available: http://arxiv.org/abs/1901.02672
16. Amran, A., Zaaba, Z.F., Singh, M.M., Marashdih, A.W.: Usable security: revealing end-users comprehensions on security warnings. Procedia Comput. Sci. **124**, 624–631 (2017). https://doi.org/10.1016/j.procs.2017.12.198
17. Ki-Aries, D., Faily, S.: Persona-centred information security awareness. Comput. Secur. **70**, 663–674 (2017). https://doi.org/10.1016/j.cose.2017.08.001
18. Creswell, J.W.: Research Design: Qualitative, Quantitative, and Mixed Methods Approaches (2013)
19. Ma, U., Ma, U.: The analysis. Do More with Less, pp. 74–100 (2018). https://doi.org/10.4324/9781315441641-7
20. Gray, D.E.: Doing Research in the Real World. [Electronic resource]., THIRD edit. Sage (2014)

Chapter 67
An Approach for Credit Card Churn Prediction Using Gradient Descent

P. M. Saanchay and K. T. Thomas

Abstract A very important asset for any company in the business sectors such as banking, marketing, etc. are its customers. For them to stay in the game, they have to satisfy their customers. Customer retention plays an important role in attracting and retaining the customers. Customer retention means to keep the customer satisfied so that they do not stop using their service/product in the domain of banking; the banks provide various kinds of services to the customers especially in the electronic banking sector. For this study, we have selected the service of credit card. For a bank to give a loan or amount on credit basis, the e-bank should make sure if its customers are eligible and can repay their money. The purpose of this project is to implement a neural network model to classify the churners and non-churners.

67.1 Introduction

Deep learning algorithm is used for implementation as the objective of this project. With the amount of increase in data, the machine learning algorithms does not work well. The functionality of neural networks is that by increasing data points the depth of the neural network can be increased for better learning. Artificial neural network (ANN) is implemented in this project with feature engineering and optimized for better classification of churners and non-churners. In the banking sector especially in the credit card service, churn rate has been increasing in the recent years. There is a lot of research going on this area on how to improve the churn rate by using various classification algorithms such as applying machine learning algorithms or deep learning algorithm in helping them obtain a model which gives more accuracy. The scope of the customer churn classification can be further extended to different domains like retail business, banking sector, motor insurance company. In insurance

P. M. Saanchay (✉) · K. T. Thomas
Christ (Deemed To Be University), Lavasa, India
e-mail: saanchay.pm@science.christuniversity.in

K. T. Thomas
e-mail: thomas.kt@christuniversity.in

T. Senjyu et al. (eds.), *IOT with Smart Systems*, Smart Innovation, Systems and Technologies 251, https://doi.org/10.1007/978-981-16-3945-6_68

company the customers may move easily from one to another company easily after their policy ends and there is need to retain customers to avoid incur losses. In this research, an artificial neural network model is proposed using stochastic gradient descent to classify the churners and non-churners.

The section organization of the paper starts with the introduction which gives a glimpse into Churn prediction models, Sect. 67.2 is comprised of the literature review, Sect. 67.3 mentions the data pre-processing, Sect. 67.4 describes the methodology used in the research, Sect. 67.5 contains the results and interpretation of the proposed model, and Sect. 67.6 consists of the conclusion and future work followed by references.

67.2 Related Work

Altinisik et al. [1] in this paper have compared generic machine learning algorithms and deep learning algorithms to predict the customers of commercial banks who may be in risk of cancelling the credit card subscription by following three months after a year or less. The models that are compared here are logistic regression, support vector machine, multilayer perceptron and deep neural network. The deep neural network algorithm has given the most accuracy.

Raja Mohamed et al. [2] in this paper, they have proposed a hybrid data mining model. Here, they have pre-processed and normalized the data. After pre-processing the data, they have used unsupervised methods like K-means and rough K-means algorithms to cluster the data. They have used various algorithms, but support vector machine combined with rough k-means clustering algorithm has given them better accuracy.

Keramati et al. [3] in this paper have talked about how customer retention by banks can increase profits by 85% by improving retain rate up to by 5%. The study shows long-term customers bring more profit than acquiring new customers. Data mining is used to find hidden patterns in the data. Here, a classification model decision tree is used, and a confusion matrix is used to find the accuracy. Then, the model is cross validated using k-fold cross validation.

Wongchinsri et al. [4] in this paper have discussed about various data mining techniques used in the application of credit card process. They have studied research works which were published between 2007 and the first quarter of 2015. In their analysis, they have found that there are three important factors to make decision models more accurate. They include a systematic framework, an appropriate selection of data set and a suitable time period of data set.

Ganesh Sundarkumar et al. [5] in this paper, they have used decision tree (DT), support vector machine (SVM), logistic regression (LR), probabilistic neural network (PNN) and group method of data handling (GMDH) for classification. They have used two data sets for their analysis, i.e. automobile insurance fraud data set and credit card customer churn data set.

Farquad et al. [6] in this paper have used support vector machine algorithm and for tuning, naïve Bayes tree is used. The data set used by the researches is an unbalanced data set where 93.24% are loyal and 6.76% are churned customers. They have used balancing techniques such under-sampling, over-sampling, combinations of under-sampling and over-sampling are used to balance the data.

Nie et al. [7] in this paper have talked about using data mining and classification models. Logistic regression is used as a classification model. The data used is by a Chinese commercial bank; here, they have new measure criterion called misclassification which takes in to account the economic cost for the evaluation. They have used decision tree algorithm to compare with logistic regression; they have observed logistic regression has higher accuracy than decision tree.

Kim et al. [8] in this paper, they have used support vector machine algorithm to find the hidden patterns in the credit card customer churn analysis. To check the accuracy of the model, they have used back propagation neural network as a benchmark in their analysis; they have found that SVM has a higher accuracy than back propagation neural network algorithm.

To obtain better accuracy in machine learning models, feature selection has to be done which is missing in the discussed literature. Activation functions play a major role in deep neural networks. The need to use efficient activation functions in the hidden layers to maximize accuracy has not been considered in the review of the literature. Some data sets has more non churners and it may lead to under sampling of data and some of the works did not perform oversampling techniques like SMOTE algorithm. The main objective is to predict the customer churn in the banking sector in the credit card domain using various machine learning and deep learning algorithms. The relevance of this research is that in the banking industry, the churning rate is very high which is leading to merger of banks or them ending up bankrupted. So in the present scenario, it is costly to invest in new customers so retaining the existing customers is important.

In this research, we aim to do churn prediction on a bank's credit card data set using an ANN model using a logistic function for classification and stochastic gradient descent (SGD) as an optimizer along with feature engineering to achieve better performance/accuracy.

In Fig. 67.1, the steps of the model are shown. For the evaluation of the performance, metrics like AUC curve, confusion matrix, accuracy, etc. are used.

67.3 Data Pre-processing

Bank customer churn data set is obtained from the Kaggle website. The data set contains 10,000 rows and 14 attributes. The data set contains two classes, i.e. churners and non-churners. It is created for the purpose of research, analysis, etc. to classify the churners and non-churners in the banking industry (credit card). The data set is of size 668 kbytes. Here, each customer is identified using a unique customer id.

In Fig. 67.2, there are three types of data, i.e. integer, float and object.

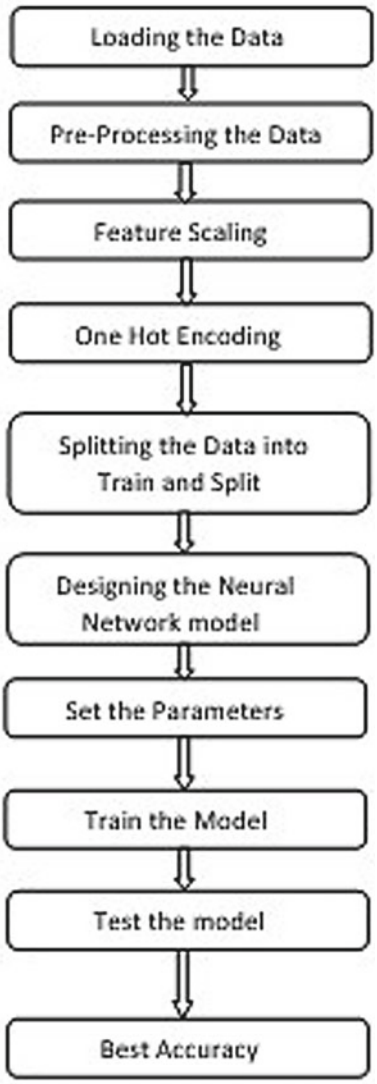


Fig. 67.1 Model process

Data pre-processing is done to remove noise, standardize the data, remove null values so that the model can learn and perform better.

Data is converted into dummies using functions available in pandas library. All categorical values are converted into zeroes and ones.

There are many columns with different categorical values represented in different formats. These values are converted into a standard format by replacing the values. The attributes such as “gender” and “geography” are of type object. The attribute “gender” has values “male”, “female” which converted to 0 and 1, and the other attribute “Geography” has three classes which are converted to 0 and 1.

Fig. 67.2 Data types

```
Out[7]: CreditScore          int64
        Geography           object
        Gender              object
        Age                  int64
        Tenure               int64
        Balance            float64
        NumOfProducts        int64
        HasCrCard            int64
        IsActiveMember       int64
        EstimatedSalary    float64
        Exited               int64
        dtype: object
```

The data set is imported into the python environment using the Numpy library to get the data into an array. The library Sklearn is used to standardize the data using the minmax function. The data is standardized to make the data internally consistent which means the data type of the data should be in the correct format and should not change. Standardizing the data makes it easy to track the data.

The concept of feature engineering is very important to be understood in order to successfully manage data. For this analysis, as there are only 12 features/columns. We aim to add a few more columns that are likely to have an impact on the probability of churning. Here, a few columns are added to the original data set like TenreByAge, BalanceSalaryRatio, etc.

The bank customer churn data set's "Exited" column has imbalanced classes of churners and non- churners. So the data is balanced using synthetic minority over sampling technique (SMOTE) algorithm. Here, the minority class is over sampled. If the model is not balanced, the model will be biased as it will learn more from the majority class, i.e. is non-churners. The library used to handle the imbalanced data is Imblearn over_sampling. Using the SMOTE technique which is used to oversample the minority class.

The features are selected using mutual_info_classif function from scikit learn library. The function will take inputs as features and target vector and will return the score as to which predictor variable explains the target variable better.

In the Fig. 67.3, the attributes "age", "TenureByAge" are highly correlated with the target variable "exited" and the attributes "HasCrCard" is the least correlated with the target variable.

The data is spilt into train and test using train_test_split function available in sklearn library. The split ratio is chosen as 80:20, which means that 80% of the data is chosen for training and remaining 20% is chosen for testing the new observations and determining the class.

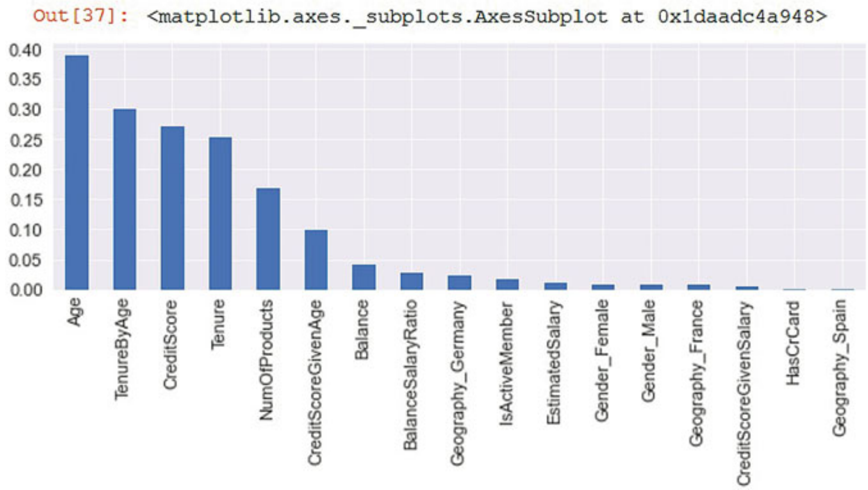


Fig. 67.3 Feature selection

67.4 Methodology

In the model building phase, hyper parameters such as number of hidden layers, batch size and learning rate are tuned to reach optimal values for the model. The optimal values are found using random search and grid search from scikit learn library.

67.4.1 Model

The functionality of neural network is that with increase in data points, the depth of neural network can be increased for better learning. The artificial neural network (ANN) works well on huge data and model complex patterns better than machine learning algorithms. So ANN is used for classification for this problem.

In the model, the first layer has 40 neurons with uniform as kernel initializer and relu as an activation function. In the second layer, batch normalization is done to normalize the values along with the relu activation. For the output layer, sigmoid activation is used, as the target column is of type binary classification. Thus, the sigmoid function acts like a logistic function. Model is compiled with a binary_cross_entropy loss function and stochastic gradient optimizer with optimal value of 0.01 learning rate is used. The model is trained on training data starting from 50 epochs. The epochs are selected when the model stops learning. Model stops learning when the epochs reach 450. After 400 epochs, the test accuracy remains at 84%, and only the training accuracy improves. The value of epochs is 400 in this model.

Table 67.1 Confusion matrix for the model

	Actual (0)	Actual (1)	
Predicted (0)	TN = 1400	FN = 193	1593
Predicted (1)	FP = 148	TP = 259	407
	1548	452	2000

67.5 Results and Interpretation

The model is trained on the training data for 400 epochs and is validated on the test data to ensure goodness of fit, accuracy, etc.

The various performance metrics used for the evaluation of the model are as follows.

67.5.1 Confusion Matrix

In Table 67.1, the confusion matrix has classified 1400(TN) non-churners and 259(TP) churned. The confusion matrix has incorrectly classified 193(FN) non-churners as churners and 148(FP) churned customers as non-churners.

67.5.2 Evaluation Metrics

In Table 67.2, we can see that the precision, recall and $F1$ score for the class 0 is higher than that of class 1.

Table 67.2 Evaluation matrix for the model

Model name	Class	Precision	Recall	$F1$ score	AUC score
Model 1	0	0.90	0.88	0.89	0.75
	1	0.57	0.64	0.60	

Table 67.3 Accuracy

S. No.	Model name	Training accuracy	Testing accuracy
1	Model 1	77.1	84.4

67.5.3 Training and Testing Accuracy

In Table 67.3, for the model, the training accuracy obtained is 77.1, and the testing accuracy obtained is 84.4.

67.6 Conclusion and Future Work

In the banking industry, customer churn is a major problem, and it might lead to bankruptcy of the industry. Retaining existing customers for business growth is imperative because the cost of acquiring new customers is much more than maintaining existing ones. In this project, artificial neural network-based models are built on customers into churners and non-churners. Before building model, missing values are handled and values are converted into required format, and data is split into training and test set, and balancing of data is carried out. After pre-processing, the model is trained and tested on unseen data and evaluated using performance metrics. The model uses relu activation function in the input and hidden layer, and for the out layer, sigmoid activation function is used. The input layer consists of 15 neurons, the hidden neuron consists of 10 neurons. All the layers have uniform kernel initializer. The models training and testing accuracy is 77 and 84%, respectively. Without feature engineering, the training accuracy and testing accuracy decreases by 1%. While finding optimal values of model, the grid search cv takes a lot of computation of time. So in place of gridsearch cv random search cv is used. The limitation of the proposed system is that various other optimizers and feature engineering can be done to improve accuracy, also ensemble technique can be used to compare different models and pick the model with the highest accuracy. With the amount of increase in data, the machine learning algorithms does not work well. The functionality of neural network is that by increasing data points, the depth of neural network can be increased for better l learning. The artificial neural network (ANN) works well on huge data and model complex patterns better than machine learning algorithms. The limitation for deep learning algorithms is data. If you have sufficient data, the algorithm works well, and it will learn data points very well. The model can be trained with huge amount of data and model architectures can be built with different layers, activation functions, epochs and regularization techniques for achieving better performance.

References

1. Altinisik, F.: Predicting Customers Intending to Cancel Credit Card Subscriptions Using Machine Learning Algorithms: A Case Study. IEEE (2020)
2. Raja Mohamed, R.: Improved Credit Card Churn Prediction Based on Rough Clustering and Supervised Learning Techniques. Springer, Berlin (2017)

3. Keramati, A.: Developing a Prediction Model for Customer Churn from Electronic Banking. Springer, Berlin (2016)
4. Wongchinsri, P.: A Survey—Data Mining Frameworks in Credit Card Processing. IEEE (2016)
5. Ganesh Sundarkumar, G.: One-Class Support Vector Machine Based Under Sampling: Application to Churn Prediction and Insurance Fraud Detection. IEEE (2015)
6. Farquad, M.A.H.: Churn prediction using comprehensible support vector machine: an analytical CRM application. Appl. Soft Comput. J. (2014)
7. Nie, G.: Credit Card Churn Forecasting by Logistic Regression and Decision Tree. Elsevier (2011)
8. Kim, S.: An Application of Support Vector Machines for Customer Churn Analysis: Credit Card Case, Korea

Chapter 68
Secure SDLC Using Security Patterns 2.0

E. R. Aruna, A. Rama Mohan Reddy, and K. V. N. Sunitha

Abstract Bug-free software application is expected by all the users. Trustworthiness of the application is value to software. The cyber-attacks are increasing day by day as the software-intensive systems are gradually increasing and pervading our everyday lives and at the same pace software vulnerabilities are also increasing due to implementation flaws during software development. Users, hackers and developers open many vulnerabilities. Hence, secure software development is essential and urgent need to mitigate all the known vulnerabilities during all stages of software development life cycle. Hence, we proposed "**Secure SDLC using Security Patterns 2.0 (SSDLC using SPs2.0)**", and this framework enhances security by minimizing the known vulnerability. Identifying the security requirements using security discoverer process, selection of security pattern for identified security requirements, design security requirements using security building blocks, creating test templates to support pattern implementation during development stage, vulnerability scanning and secure configurations are key functionalities in our SSDLC using SPs 2.0 framework. The proposed framework integrates security concerns from initial to disposal stage, and hence, software security vulnerabilities are found and mitigated at SDLC initial stages and save huge amount of reengineering cost for post-implementation vulnerabilities.

68.1 Introduction

In SDLC all phases, security concerns must be implemented to deliver secure software to end user. Several existing secure software engineering areas are Microsoft SDL (Security Development Life Cycle), Software Assurance Maturity Model

E. R. Aruna (✉)
Vardhaman College of Engineering, Hyderabad, India

K. V. N. Sunitha
SVU College of Engineering, S.V University, Tirupati, India

A. Rama Mohan Reddy
BVRIT HYDERABAD, Hyderabad, India

T. Senjyu et al. (eds.), *IOT with Smart Systems*, Smart Innovation,
Systems and Technologies 251, https://doi.org/10.1007/978-981-16-3945-6_69

(SAMM), Software Security Framework (SSF), Open Web Application Security Project (OWASP), McGraw's Touchpoints, Comprehensive Lightweight Application Security Process (CLASP), etc.

The state of the art of the secure software development in the context of global software development (GSD) is limited and in the SDLC context is significant which was examined from different digital libraries [1]. This is because secure SDLC needs more cost, time and resources to implement it.

Security patterns are the packages of reusable security solutions. Security patterns receive significant research attention over the last two decades. In the last ten years,

Software practitioners, researchers have led more research projects using security patterns to realize the considerable support of security patterns in secure software development systems [2]. Security patterns have security expert's knowledge, and they provide guidelines to improve security characteristics. For non-security experts incorporate security using security patterns as a promoting tool in their designs effectively with optimum time utilization.

68.2 Related Work

In secure software engineering, the dedicated methodologies are MS security development life cycle, OWASP comprehensive lightweight application security process and McGraws Touchpoints. These three methodologies provide a collection of activities for secure software development life cycle [3]. These three processes compared and identified their pros and cons likewise SDL is well-defined process, provides good guidance, and supports security as quality. SDL is considered as heavyweight process and is suitable for few organizations.

OWASP CLASP provides a list of resources to perform code review as checklist in the application. CLASP provides little support to map with SDLC phases and software engineering methods. CLASP is supported by limited organizations. The security practices gathered by industry people with their experience over several years are McGraw's Touchpoints. This practices can be tailored with secure software development process. These Touchpoints are less support to address security issues and detailed design.

Khari et al. [4] confirmed that SDLC must provide high-level security in addition to functionality. They embedded traditional SDLC with security properties. The objective of their work is to merge traditional SDLC with security requirements. Risk mitigation, management, assessment and evolution are the security requirements in all the stages of SDLC. The team compared their eSSDLC with MS SDL, open-SAMM [3]. Few of the compared features are not detailing by eSDLC. The authors initiated security factors within SDLC likewise security requirements, its design, secure coding principles, security testing tools and security training and monitoring.

Shehab et al. [5] proposed a methodology with security descriptions at each phase of SDLC to enhance security and minimize the vulnerabilities. The authors also proposed measures to assess the security at each level. They also developed a real-world software application with product requirements stability, improved process, software product readiness and stability.

The proposed work is measured in all the efforts likewise schedule variance, stability, quality measure, software quality, work and progress, time to deliver variance rate and work product quality. More emphasized on Security Assessment in every SDLC process without focus on Security Concerns on every phase of SDLC.

Assal et al. [6] have interviewed two groups of industry software practitioners as security adopters and security inattentive. The authors stated that security practices in literature considerably vary from real security practices. The observations are the security adopters treat security as very essential, priority and implementation must be formal where as Security Inattentive are Security considerations are adhoc, limited feature-driven, misuse perceptions of existing security frameworks, security consideration is minimal. Based on these observations, the authors have stated and listed the best practices to be followed during the development of secure software.

Fernandes et al. [7] designed the IoT-based health monitor by following the secure SDLC stages. The authors developed the secure SDLC by customizing the security requirements and its evolution. Identification of security requirements, threat analysis, static security analysis, verification and validation of security requirements are the essential activities.

From the literature review, software practitioners must be expertise in security to implement security requirements in each phase of SDLC. Integrating security effectively with optimized resources is also significant to meet the looming deadlines.

Secure software can be developed securely by thinking security from the beginning to disposal stage. A well-defined secure software development framework with efficient and flexible tools needs to be designed. Empirical secure software development from inception to disposal to reduce cost for post-implementation security vulnerability by mitigating critical vulnerabilities during early stages is a challenge in this study.

Hence, we proposed "SSDLC using Security Patterns 2.0" to enhance security by reducing the minimizing the critical security vulnerabilities.

Proposed methodology is explained in Sect. 68.3, implementation is detailed in Sect. 68.4, results are discussed in Sect. 68.5, conclusion and future scope are discussed in Sect. 68.6.

68.3 The Proposed Methodology

We tailored SSDLC with Security Patterns 2.0 with security discoverer process [8], text categorization process [9], Security Patterns 2.0 building blocks [10], security patterns validation [11] and security testing approaches.

The overview of **SSDLC using SPs2.0** is shown Fig. 68.1.

The steps in our framework are as follows:

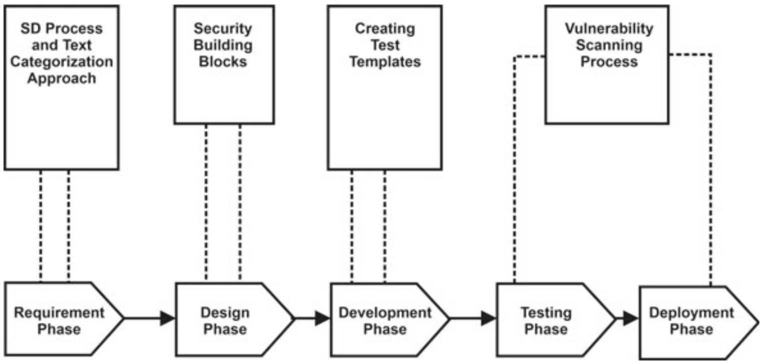


Fig. 68.1 SSDLC using SPs2.0 methodology

1. Using security discoverer (SD) process identification of implied security requirements from functional requirements.
2. Using text categorization selection of suitable security patterns for all the security requirements identified in step 1.
3. Modeling of security requirements using reusable security building blocks of security patterns.
4. Implementation and validation of security patterns mechanisms as per the design in step 3 during the development of functional code.
5. Testing the security features using combination of static and dynamic vulnerability scanning tools.
6. Deploying the software product in secure configured environment.

68.4 Implementation

To justify our proposed framework, we have implemented IoT-based health monitor application. The design of IoT-based health monitor application as shown in Fig. 68.2 is used to monitor remote patient by the doctor. Patient bio-data is captured through sensors. The patient data is transferred to database through Arduino. The patient records are disposed to authenticated doctor/remote caretaker. The remotely located doctor/caretaker requests the specific patient records through Web application.

The functional requirements of the IoT-based health monitor is

FR0: Patient data must be captured through sensors and then sent to database.
FR1: Data must be received and stored in the database.
FR2: Patient records must be present to doctor/guardian whenever they required.

IoT-based health monitor application design using SSDLC with security patterns framework is as follows:

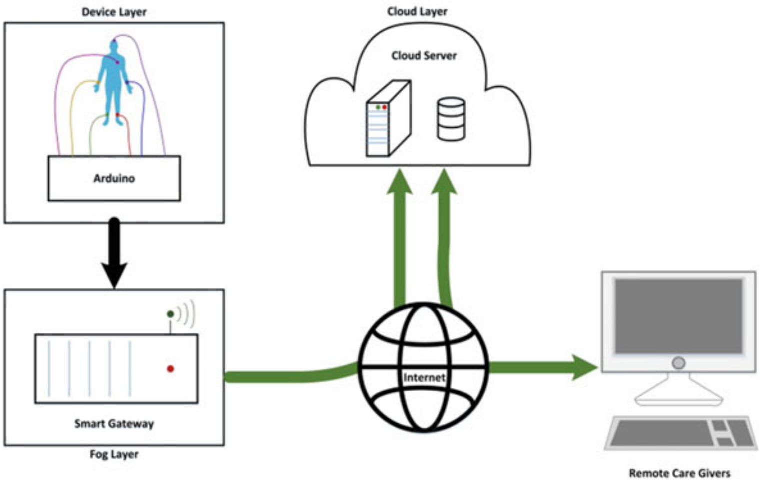


Fig. 68.2 IoT health monitor system design

Step 1 The implied security requirements are identified using SD process. The sample implied security requirements of first functional requirement (FR0) are shown in Table 68.1.

Step 2 Using text categorization approach with cosine similarity, the selected most suitable patterns for security requirements identified in step 1 are shown in Table 68.2. To select the security patterns, we have used the security patterns repositories [12–15].

Step 3 Designe each security requirement using reusable security building blocks likewise decision points, active entities, data storage, enforcement points, data creation, cryptographic primitives, etc., of Security Patterns 2.0 version. This building block will guide the developers about the security concerns while developing.

Sample design of security requirements remote caretaker/doctor authentication shown in Fig. 68.3. Security building blocks such as enforcement points will specify when the security concern is mandatorily required to be implemented. Decision points will be made whether privileges will be granted or not to active entities based on their true identities.

Cryptographic keys generation, data creation like creation of session IDs, encrypted storage of sensitive data such as session IDs, public and private keys are also modeled in the design of IoT health monitor in the form of security building blocks.

Step 4 The selected patterns mechanisms such as security building blocks are implemented during development stage and tested the behavior of each pattern mechanism using JUnit test cases. These test cases will analyze the

Table 68.1 Implied security requirements for FR0

Security objective	Suggested security requirements templates	Instantiated security requirements
Identification and Authentication (IA)	<subject> = Arduino, Sensors, channel and Database <resource> = Patient data <action> = Receive/Process/Write/Encrypt/Transit The Arduiono <subject> shall receive <action> Patient Data <resource> from sensors Arduino <subject> shall process <action> received data <resource> **Security Requirement**: Arduino <subject> shall authorize <action> before writing the data to the channel (IAI) IAI: Arduino Authorization	Before writing sensitive data to channel, the Arduino has to be authorized
Confidentiality	Arduino <subject> shall write <action> processed data to channel <subject> **Security Requirement**: Channel <subject> shall encrypt & transit <action> data <resource> to database (C3) C3: confidentiality during transmission	The sensitive patient records need to be encrypted and transferred to the database

Table 68.2 Selected security patterns

Security requirements identified	Suitable patterns
Arduino must be authorized before writing the sensitive data to channel (IA1)	Secure Adapter Pattern
Encrypt and transfer the sensitive patient data to the database (C3)	Secure Channel

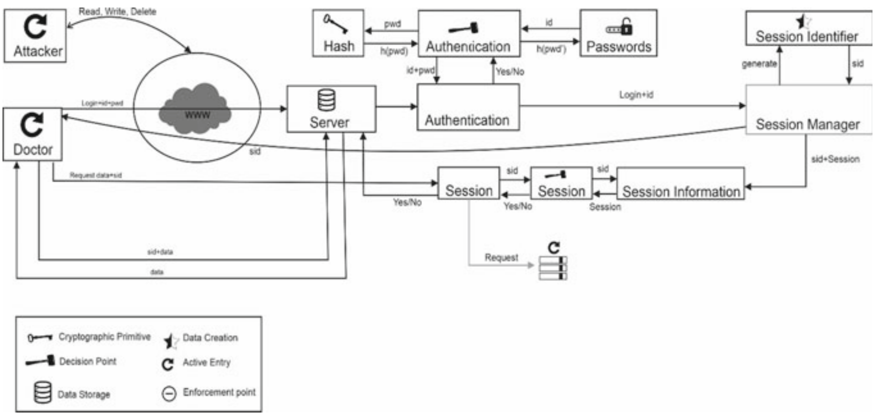


Fig. 68.3 Design of remote doctor authentication

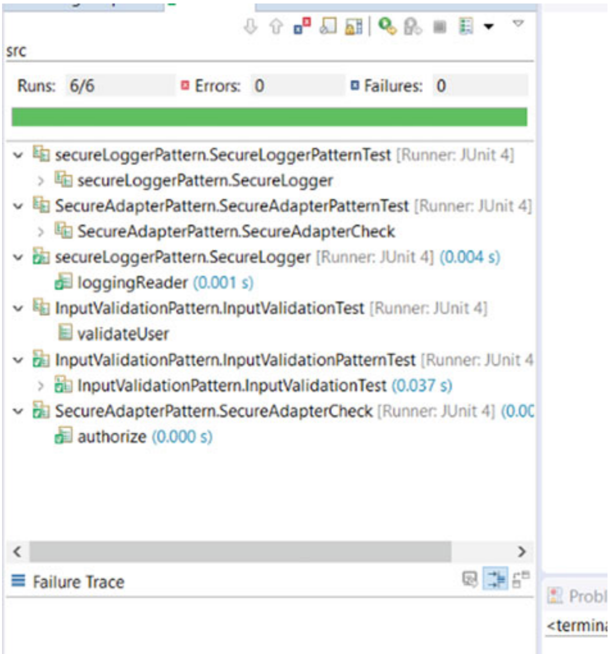


Fig. 68.4 Correct behavior of applied patterns

correct and incorrect behavior of the security concerns within the application. Sample test case for the multiple combinations of security aspects of the security patterns is shown in Figs. 68.4 and 68.5, respectively.

Step 5 Testing is done using “Test-a-User” approach using ZAP tool of OWASP for vulnerability scanning is done.

Step 6 The application is deployed in secure environment and retested. Regular updates of secure configuration, sending the sensitive data with Httponly flag, autocomplete options must be set to off, removing hard-coded comments in the code are some of secure configuration strategies that are suggested.

68.5 Results and Discussion

Using our SSDLC using Security Patterns 2.0 approach, the critical bugs are mitigated during early stages of SDLC. The results are shown in Figs. 68.6 and 68.7, respectively.

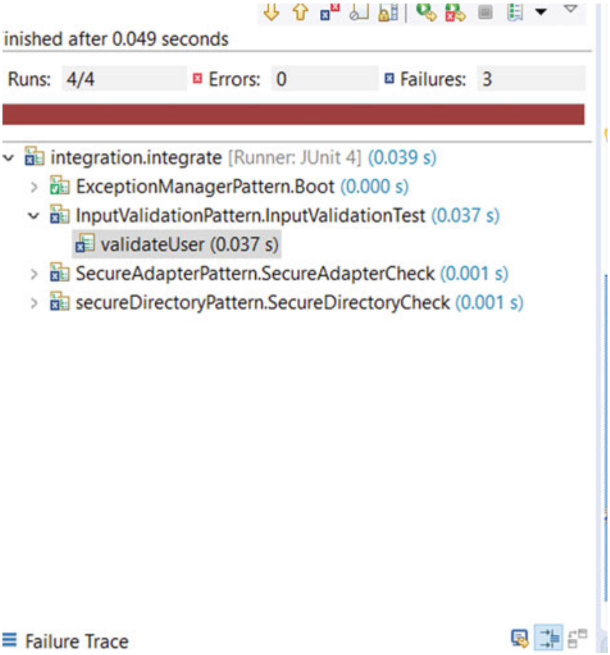


Fig. 68.5 Incorrect behavior of applied patterns

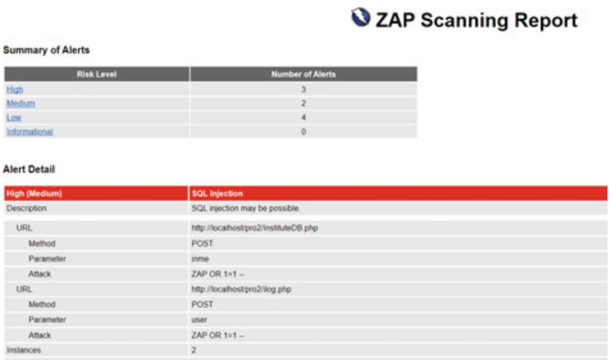


Fig. 68.6 Bugs severity after before SSDLC framework

68.6 Conclusion

The IoT-based health monitor application with improved security is designed with our proposed framework and successfully deployed in the user environment. Our SSDLC framework becomes a flexible guiding tool for non-security software development experts. Our framework provides time-bound proven solutions. The security

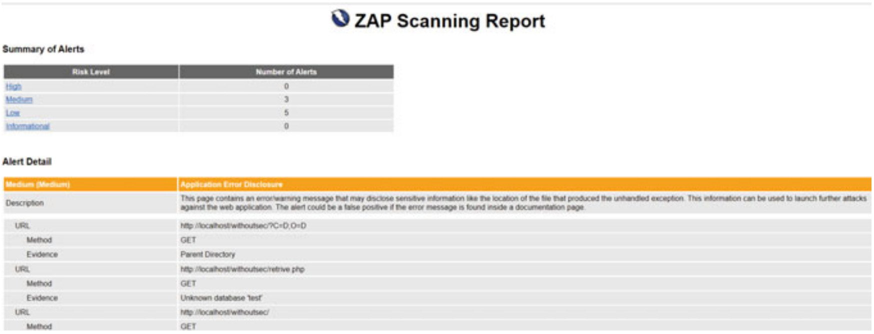


Fig. 68.7 Bugs severity after applying SSDLC framework

patterns become popular tools in the entire SSDLC. In future work, proposed SSDLC efficiency and performance need to be measured by comparing with other popular software development methodologies.

References

1. Khan, R.A., Khan, S.U.: A preliminary structure of software security assurance model. In: 13th International Conference on Global Software Engineering, (ICGSE), pp. 137–140. ACM/IEEE
2. Washizaki, H.: Security patterns: research direction, metamodel, application and verification. In: International Workshop on Big Data and Information Security (IWBIS). IEEE (2017)
3. Win, B.D., Scandariato, R., Buyens, K., Gregoire, J., Joosen, W.: On the secure software development process: CLASP, SDL and touchpoints compared. J. Inf. Softw. Technol. **51**(7), 1152–1171 (2009)
4. Khari, M., Vaishali, Kumar, P.: Embedding security in software development life cycle (SDLC). In: 3rd International Conference on Computing for Sustainable Global Development (INDIACom). IEEE (2016)
5. Farhan, S.A.R., Mostafa, M.G.M.: A methodology for enhancing software security during development processes. In: 21st Saudi Computer Society National Computer Conference (NCC). IEEE (2018)
6. Assal, H., Chiasson, S.: Security in the software development lifecycle. In: Proceedings of the Fourteenth USENIX Symposium on Usable Privacy and Security (2018). ISBN 978-1-939133-10-6
7. Fernandes, A.M., Pai, A.: Secure SDLC for IoT based health monitor. In: Proceedings of the 2nd International conference on Electronics, Communication and Aerospace Technology. IEEE (2018)
8. Riaz, M., King, J., Slankas, J., Williams, L., Massacci, F., Quesada-López, C., Jenkins, M.: Identifying the implied: findings from three differentiated replications on the use of security requirements templates. Int. J. Empirical Softw. Eng. **22**, 2127–2178 (2017)
9. Ali, I., Asif, M., Shahbaz, M., Khalid, A., Rehman, M., Guergachi, A.: Text categorization approach for secure design pattern selection using software requirement specification. IEEE Open Access J. **6**, 73928–73939 (2018). https://doi.org/10.1109/ACCESS.2018.2883077
10. van den Berghe, A., Yskout, K., Joosen, W.: Security patterns 2.0, towards security patterns based on security building blocks. In: ACM/IEEE 1st International Workshop on Security Awareness from Design to Deployment (2018)

11. Yoshizawa, M., Kobashi, T., Washizaki, H., Fukazawa, Y., Okubo, T., Kaiya, H., Yoshioka, N.: Verification of implementing security design patterns using a test template. In: Proceedings of 9th International Conference on Availability, Reliability and Security (ARES), pp. 178–183, Fribourg, Switzerland (2014)
12. . Schumacher, M., Fernandez-Buglioni, E., Hybertson, D., Buschmann, F., Sommerlad, P.: Security Patterns-Integrating Security and Systems Engineering. Wiley Ltd. (2006). ISBN: 978-0-470-85884-4
13. Lee, W.-T., Law, P.-J.: A case study in applying security design patterns for IoT software system. In: Proceedings of the 2017 IEEE International Conference on Applied System Innovation, IEEE-ICASI (2017)
14. Kienzle, D.M., Elder, M.C., Tyree, D., Edwards-Hewitt, J.: Security Patterns Repository Version 1.0. DARPA, Arlington,VA, USA, Tech. Rep. (2002). http://www.securitypatterns.com
15. Dougherty, C., Sayre, K., Seacord, R.C., Togashi, K.: Secure Design Patterns (2009). http://www.cert.org/

Chapter 69
Eklavya—Shaping the Leaders …

Mannat Amit Doultani, Smith Gajjar, Hrithik Malvani, Jai Soneji, and Sonia Thakur

Abstract Eklavya is a virtual classroom platform for educational institutes. Currently, Google Forms are widely used to create online tests. In the process of generating examination through Google Forms, there is a lot of manual work done by faculties. This is a very tedious and time-consuming job. Eklavya enables faculties to conduct and automate this online examination and with different options to generate it. Our portal Eklavya is a one stop solution for all the examinations that are being conducted online nowadays. Eklavya provides a proctored mode for conducting online tests, a virtual classroom feature where students and faculties can collaborate and manage different assignments, announcements, and it can also be used for generating reports.

69.1 Introduction

As we all know that we are moving into the digital era where almost everything takes place online, for example, banking, learning, etc. Nowadays, most of the educational activities like assignment submission, quizzes, aptitude tests, interviews etc. are also conducted online using Google Classroom, Google Forms, mettl and various other e-platforms. These platforms have many positive points, but they are clustered into specific requirements. Users need to switch between different platforms to get the

M. A. Doultani (✉) · S. Gajjar · H. Malvani · J. Soneji · S. Thakur
VESIT, Mumbai, India
e-mail: pooja.nagdev@ves.ac.in

S. Gajjar
e-mail: 2017.smith.gajjar@ves.ac.in

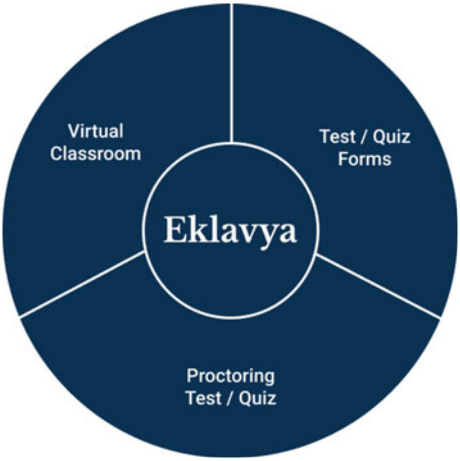
H. Malvani
e-mail: 2017.hrithik.malvani@ves.ac.in

J. Soneji
e-mail: 2017.jai.soneji@ves.ac.in

S. Thakur
e-mail: 2017.sonia.thakur@ves.ac.in

T. Senjyu et al. (eds.), *IOT with Smart Systems*, Smart Innovation, Systems and Technologies 251, https://doi.org/10.1007/978-981-16-3945-6_70

Fig. 69.1 Eklavya platform



work done. Using different applications for various works is a tedious job for any user.

Eklavya will provide users with a unified application which basically adds up all the features along with the existing ones to make it better to use. Eklavya will have an option for importing questions from different platforms or sites, scan documents to import questions or choose questions from existing repositories. Eklavya will also provide users with proctored features like having a tab lock while ongoing quiz and tests. Monitoring of users for the exam invigilator. Eklavya will also feature a best in class virtual classroom for faculties and students.

The paper is structured as follows: below, Sect. 69.2 discusses different papers which tackle similar problems and the motivation of the current work, while Sect. 69.3 illustrates the proposed strategies. In Sect. 69.4, the techniques or methodology proposed by our system is discussed. Section 69.5 covers the limitations of our proposed system, and Sect. 69.6 describes the expected result and discussions on the same. The paper ends with Sect. 69.7 which discusses conclusions and future work (Fig. 69.1).

69.2 Background and Related Work

As discussed above, there are various platforms to carry out online examinations and other academic activities. Some of the lacuna in the existing systems are discussed (Table 69.1).

In [1], the author is proposing a smart education concept, the development of different new technologies that enable learners to learn more successfully, efficiently, pliability, and comfortably. Learners can make use of digital devices. They described "Smart education" as a concept of learning in the digital age that has gained increased

Table 69.1 Lacuna in the existing systems

Name	Points	Lacuna in the existing systems	Solution provided by our system
Google forms	Protected mode:	During online examinations that are conducted via the Google Forms are not protected, hence allowing students to search from other tabs	Eklavya has a tab lock feature and video proctoring of students
	Manual quiz formation:	Faculties use Google Forms to generate quizzes. While creating quizzes, they manually need to type the questions and answers, or copy paste them. This is very tedious and time consuming	Eklavya has different options for faculties like fetching questions from prebuilt repository, data scraping, scanning or predicting documents
Google classroom	Hierarchical structure:	In Google Classroom when teachers upload a document it just gets dumped into it, there is no hierarchical structure. This is quite unorganized and is also difficult while searching for any documents	Eklavya directly integrates the Google Classroom feature and builds a customized feature on top of it, which allows one to create nodes with the subject name and creating sub nodes with the topic name like: lessons, labs, etc.
	User statistics:	In Google Classrooms, the documents or experiments that are submitted by the students need to be graded by the teacher. After grading it, the teacher needs to individually go through each experiment of the student to generate a statistics/ performance of the student	Eklavya automatically generates reports of the student not only from the assignments but also from the quiz or test conducted on our platform directly into the Google Classroom without any intervention of the faculties
Mettl	No mobile support:	Mettl is one the best softwares for conducting examinations online, but it does not support mobile phones	Eklavya allows students to take their test directly from our Web application, which can be installed in smart devices

(continued)

Table 69.1 (continued)

Name	Points	Lacuna in the existing systems	Solution provided by our system
	No communication channel support:	Most of the companies are using mettl for college and off-campus placements, but there is no communication channel directly integrated into mettle	Eklavya allows organizations to establish communication channels directly through creating classrooms with students or candidates without having to create any other channel

Fig. 69.2 Four-tier architecture of smart pedagogies [1]

attention. This paper discusses smart education and presents a four-tier framework of smart pedagogies (Fig. 69.2).

The classroom should be differentiated and responsive to various learners' readiness levels, interests. Whether learning happens in the physical classroom or any other online method, who have various performances are required to learn altogether within a particular group to achieve a goal. In this process, the students can have very competitive abilities also including solving problems with critical thinking.

In [2], the author describes the importance of tagging questions so as to generate questions which are according to the course objectives or according to the specifications of teacher(s). The same is proposed in our system which helps mentors to generate questions at very high granularity (Fig. 69.3).

In [3], the author describes the practical applications that have been gradually emphasized during the last five years, particularly in medical and remote sensing areas. In the decision-theoretic approach, they are still looking for effective and efficient feature extraction and selection techniques, particularly in nonparametric and small sample situations. Image models should also be used more extensively in the

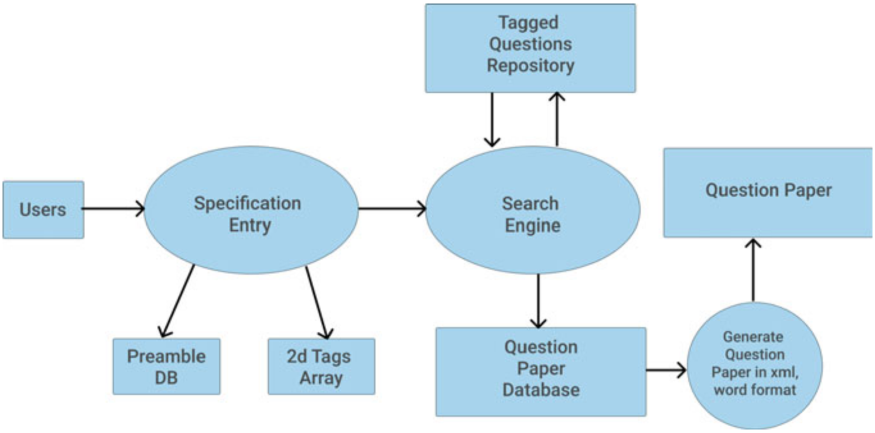


Fig. 69.3 System architecture [2]

design of optimal image segmentation and feature extraction procedures. When the goal is not objectively specifiable, but rather involves general-purpose man machine dialog about images, the computer will also need to understand the visual capabilities and limitations of its human partner. Thus, image and visual models need further development in both image processing and recognition.

In [4], the authors have used natural language processing techniques to generate MCQs out of given text. They have used Wikipedia data as test data. For now, the system is only capable of generating MCQs for a particular topic that is Physics. The MCQs generated are single choice questions thus limiting the number of different questions that can be generated (multiple choice questions).

69.3 One Unified Platform

Eklavya is a unified platform for schools and colleges for managing virtual classes and tests. The platform is being developed natively that it supports all the major devices available. One major drawback during this COVID-19 that we are trying to solve is to provide a solution for conducting online examinations in proctored mode.

Eklavya tries to automate the entire process of creating tests for the faculties and to provide them with different options. Faculties can manage their previously conducted tests, view individual student reports, and reuse the previous test templates.

Online proctored mode provides faculties with monitoring student activities during the quiz and reports. In the proctored mode, once the session is started, students doing any mischievous activity, i.e. changing tabs, browsers, or going into multi window will be detected and reported to faculties. Faculties and students will be able to manage all the previous sessions conducted, and notifications will be given for future sessions. This platform provides a virtual classroom experience

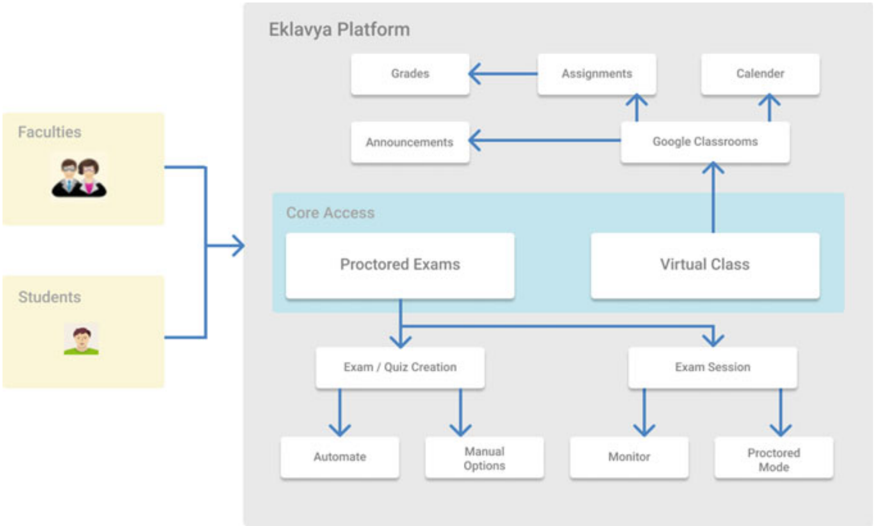


Fig. 69.4 This system architecture of Eklavya platform with different modules, submodules, users, and services it uses

where faculties can create assignments, or they can announce some important notice (Fig. 69.4).

69.4 Methodology

Eklavya is divided into four major modules across different services for faculties and students. Also, some of the services are only meant for developers, i.e. API Service. Let's discuss all modules in details (Fig. 69.5).

69.4.1 Quiz Generation

This module enables faculties to create exams or quiz forms using four different question techniques.

Fetch from repository: Inbuilt ever-growing repository database consisting different questions based on subjects, topics or keywords.

Scrape from websites: Scraping MCQ-based questions from different websites using html-based parser.

OCR: Scan different multimedia files such as pdf documents or images using Tesseract SDK.

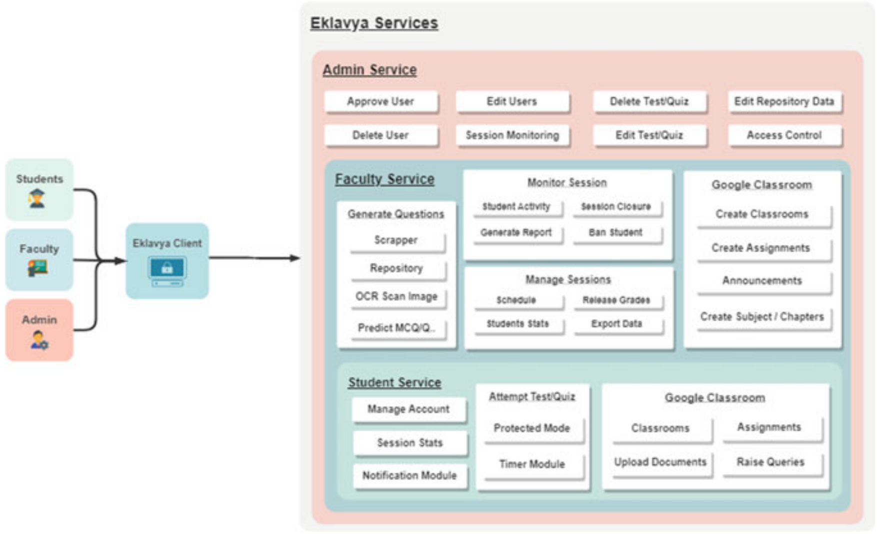


Fig. 69.5 Eklavya modular diagram

Prediction of question: Predictions of different types of questions such as MCQ-based, true or false-based using Questgen SDK.

69.4.2 Proctored Exams

This module allows students to attempt exams in a proctored environment. Some of the features of this module is:

Tab lock: Restricted and three-stage warning system preventing students from switching tab using native browser window API to detect tab change.

Video analysis: Video proctoring service for analysing student behaviour using Google's TensorFlow SDK.

Live dashboard: Dashboard for live statistics of the users appearing for exams and will enable them to view any warnings generated for the users using our own backend services.

69.4.3 Classroom Integration

This module integrates Google Classroom features using Google API's. All the Google Classroom functionalities such as creating assignments, announcements,

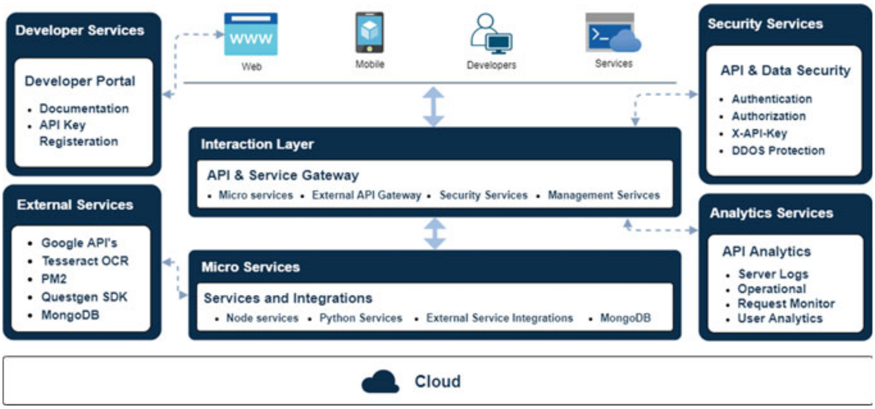


Fig. 69.6 Eklavya API architecture containing different services, micro services and gateways which are hosted on cloud

grading, etc. along with some custom features unavailable in Google Classroom. Apart from all the above features, some under the hood features are available such as user reports, forms management, user management, etc. using our backend service as a wrapper on Google API.

69.4.4 API Architecture with Diagram

This module enables developers to interact with our backend API microservice using subscription-based service and pay as per use model. In today's world, having to provide developers with an API service can greatly increase productivity and build time rather than developing those modules in house. Having to provide developers an API service, the platform needs to be secured, robust, monitoring service, and a documentation for the ease of developers. Developers need to first register themselves to get API keys, these keys are unique for all developers and can be generated one per registration; once the registration is complete, can proceed further by making requests to the endpoints as mentioned in the documentation along with the API key provided (Fig. 69.6).

69.5 Limitations

Eklavya consists of three major modules for automating or fetching questions from different sources like Web scraping, repository, or ocr, while it may be noted that there are many different options available yet to be discovered. One of them being having to provide faculties with an option for uploading an image of a paragraph and then generating different types of questions, i.e. MCQ, short answers, true or false questions. The scrapping features is being developed using map to schematics method, i.e. every website needs to be explicitly configured as per the structure exists, this can be improved by having to build a generalized model which would try to predict the structure and map quiz data into a json format. Google Classroom integration is subject to institution weather they have educational workspace or availability of some features might not be accessible.

69.6 Results/Discussion

Eklavya platform would provide students and faculties with a set of features or modules to conduct online exams and virtual classroom experience; however, it should be noted that the module for generating quiz or exam questions from a paragraph or image consisting of paragraphs would require a high computing server to process all the required data available. During the exams, live dashboards would be made available for faculties to view any kind of mischief being done by students and to take necessary actions. The end result would be an all in one progressive Web application for any institution looking forward to setting up virtual classroom and online examination. This platform might be optimized for organizational purposes like conducting virtual hiring processes.

69.7 Conclusion

Eklavya will enable schools and colleges to conduct any online examinations quickly by using different options available like MCQ generation model; in a secure environment. The system will ensure that it is being conducted in protected mode to reduce the chances of any malpractice and alternatively provide developers with out-of-the box API service.

References

1. Zhu, Z.T., Yu, M.H., Riezebos, P.: A research framework of smart education. Smart Learn. Environ. **3**, 4 (2016). https://doi.org/10.1186/s40561-016-0026-2
2. Nalawade, G., Ramesh, R.: Automatic generation of question paper from user entered specifications using a semantically tagged question repository. In: 2016 IEEE Eighth International Conference on Technology for Education (T4E), Mumbai, India, pp. 148–151 (2016). https://doi.org/10.1109/T4E.2016.038
3. Fu, K-S., Rosenfeld: Pattern recognition and image processing. IEEE Trans. Comput. **C-25**(12), 1336–1346 (1976). https://doi.org/10.1109/TC.1976.1674602

4. Kumar, S., Prasad, V., Santhanavijayan, A., Balasundaram, S.R., Narayanan, S.: Automatic generation of multiple choice questions for e-assessment. Int. J. Signal Imag. Syst. Eng. **10**, 54 (2017). https://doi.org/10.1504/IJSISE.2017.10005435

Chapter 70
An Overview on Mobile Edge Cloud System

Sunanda Dixit, Sheela Kathavate, and S. K. Gautham

Abstract A model that allows the clients or users to access a collection of resources used for computing such as storage, services provided by peer devices, server and applications which is available to all the mobile users as it presents at the edge of the wireless network is called as mobile edge cloud system. It provides wireless network information and local context awareness along with low latency and bandwidth conservation, so it is much better than traditional central cloud. This paper represents the characteristics, application, model services, architecture and open challenges of mobile edge cloud system; the application is also classified into different criteria. It also gives an idea of MEC for future research direction.

70.1 Introduction

The knowledge of combining both information technologies and communication helps in creating a network structure that helps to improve the user requirements and develop a large number of options for selecting new internet application. So, one among them is cloud computing. But the model of cloud computing which is based on centralized data centers comes with a number of limitations like:

- For real-time applications, it is very much difficult and also costly to move all the data from different edge nodes to the center cloud.
- It difficult to take care of the latency requirements of time-critical application.
- It difficult to adapt in traffic changes for applications that works on local network and user content using far away data centers.

S. Dixit (✉) · S. Kathavate · S. K. Gautham
BMS Institute of Technology and Management, Bengaluru, India
e-mail: sunandadixit_cse@bmsit.in

S. Kathavate
e-mail: sheela@bmsit.in

S. K. Gautham
e-mail: 1by19scs02@bmsit.in

T. Senjyu et al. (eds.), *IOT with Smart Systems*, Smart Innovation,
Systems and Technologies 251, https://doi.org/10.1007/978-981-16-3945-6_71

These limitations as lead to the development of an integrated model called mobile edge cloud system also known as edge cloud. These limitations are often addressed employing a new integrated model called mobile edge clouds (MEC) also called edge cloud. The system consists of base stations, access points, routers, switches and a collection of computation and storage resources which is used by the clients or users in mobile devices for performing a task [1–4]. Fog computing [5] provides a virtualized platform concept which is placed at the sting of wireless network that provides resources for computing, storage and networking services between the datacenter and different edge nodes. Mobile edge computing [6–9] is analogous to cloud computing that provides IT and cloud-computing services that is in the range of the radio access network which can be referred as RAN at a close range of mobile users. The mobile edge computing server can be placed at the base stations, radio network controller or at sites with different types of MultiTechology Aps. [10–12] proposed various segmentation techniques.

70.2 Characteristics and Applications of MEC

MEC has the subsequent characteristics [5, 6] (Fig. 70.1).

(1) *Status awareness for local network*: To access channel information and real-time network, MEC is established at the string.
(2) *Context awareness for local user*: The user location and contents are locally available to the application.
(3) *Distributed*: The MEC resources, service and application are distributed in nature placed at different locations, and the wireless access networks are

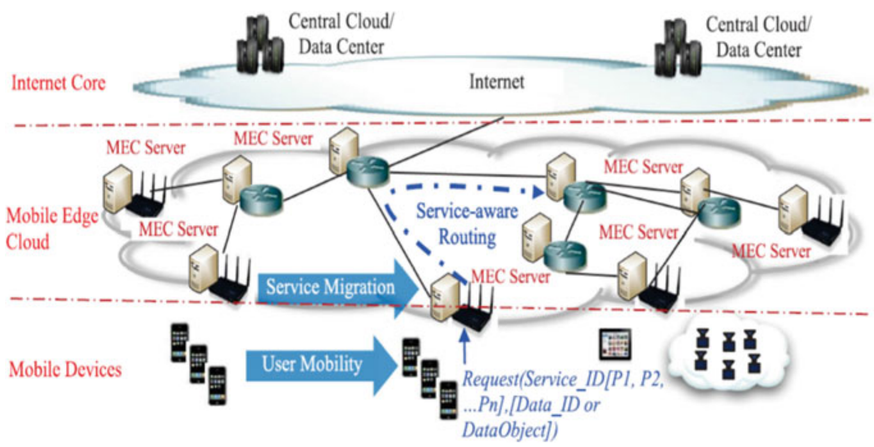


Fig. 70.1 The overview of mobile edge cloud system

arranged in a hierarchy order which are base stations, access points, switches, gateways and also end-point devices.

(4) *Multiformity*: The MEC nodes are heterogeneous in nature with different processing and storage capabilities with changing network bandwidth and connectivity.

(5) *Mobility and undependable access*: MEC services needs to be hosted in the mobile devices itself because the services might be accessed through unexpended wireless network or the point of connection to the network might change so mobility support is must.

(6) *Ultra-low latency*: To support tangible Internet applications which are robotics, virtual, AR and real-time application MEC must be able to perform in low-latency so that the application can overcome delay and end-end latency.

(7) *Interaction with mesial cloud*: MECs provide local processing of data along with low latency and content awareness, but there are clouds which provides global centralization which as advanced storage and processing capabilities, and it also provide support to the traditional central cloud.

Along with the properties, the applications of MEC are as follows (Fig. 70.2).

(1) Real time application like live gaming can reduce the experience of delay if the process is executed at the sting so that the data does not need to move between the end devices and other data centers.

(2) The application needs to adjust with the local network because there will be situation where the network degenerate or load in the traffic increases. One of

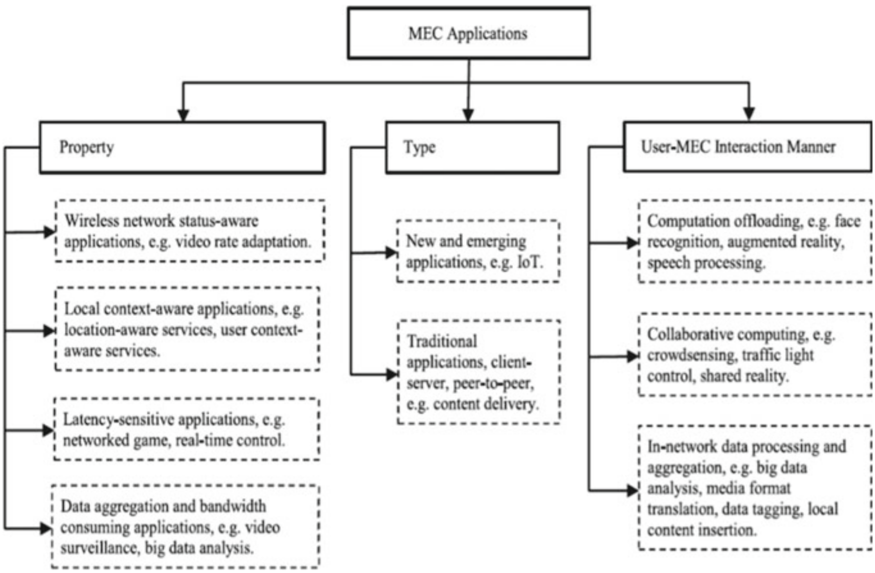


Fig. 70.2 The different methods of MEC applications

the solutions is reducing the video data rate which is suitable for the required bandwidth.

(3) The application which are local in nature process and compute the information on smaller dataset locally and maintains location privacy of the user by not sharing with any kinds of network. The application involves insertion, sharing, crowd sensing of local data or content.

(4) Huge amount of data like large data analysis, surveillance monitoring is processed and aggregated at the sting in order to reduce the bandwidth and cloud storage of the mesial cloud.

MEC can provide three levels of services to its customers:

- Infrastructure as a service (IaaS)
- Software as a service (SaaS)
- Platform as a service (PaaS).

The characteristics difference between traditional cloud and MEC is listed in Table 70.1.

For MEC SaaS, mobile users do not need to think about the problems that involves system configuration like application that is within the data center, take on cloud computing. MEC application services can be of any kind of service such as requested or transparent service and these are provided by the service providers, when the data from the client passes through mobile network, i.e., content optimization of radio-aware shown in Fig. 70.3.

Several use cases of MEC platform services are:

(1) To serve applications to mobile end-users, third-party service provider charter a share of the platform. (2) Similar to crowd sensing, MECs allow the mobile end user to release their own application in the platform. (3) In normal cloud architecture, wireless access network contains different levels which contains resources without

Table 70.1 Characteristics difference between traditional cloud and MEC

	Traditional cloud	MEC
Service location and resource	The data are sent to dedicated data centers over the Internet	Present at base stations (BS), access point (AP), gateways, switches and mobile devices
Awareness of service content	Application report provides user content and location	Aware of user content and status of local radio network
Heterogeneity of resource	The devices and network are well-planned in the machine room	Distribution and heterogeneity depend on variation in process and storage resources and network connectivity
Latency	High latency	Low latency
Resource capabilities	Limited resources	More powerful compared to traditional cloud

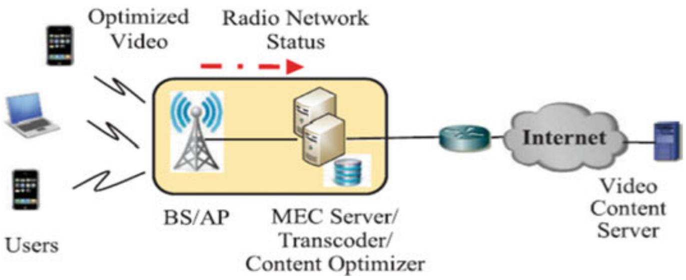


Fig. 70.3 Content optimization of radio-aware

any planned connection, but in MEC infrastructure, the user can send the code for execution in the platform service automatically.

70.3 Architecture of MEC

Fig. 70.2 illustrates the components like access point (AP), base station (BS), switches, router and aggregation for cell which makes up the MEC server [6] which provides the client to access network data and radio data, resources for computing and storage.

70.3.1 Server

Fig. 70.4 shows the architecture and interfaces of MEC server. The server consists of two parts: one is hosting infrastructure and other is virtualized application platform (VAP). The VAP provides a virtualization manger which provides a sharing environment that is used for deploying and running application that ranges from network operator to third party application and also services to host the different applications such as platforms.

Middleware services like communication, registry and traffic offload function (TOF) are provided by the application platform to the application. The registry stores information about the services that is present in the server that helps to get services to the application at different end points. One more middleware service is radio network information service (RNIS) which provides application to know about its services and also to tell other applications.

The TOF contains policy that routes the client's data between different applications and also set priority to the traffic.

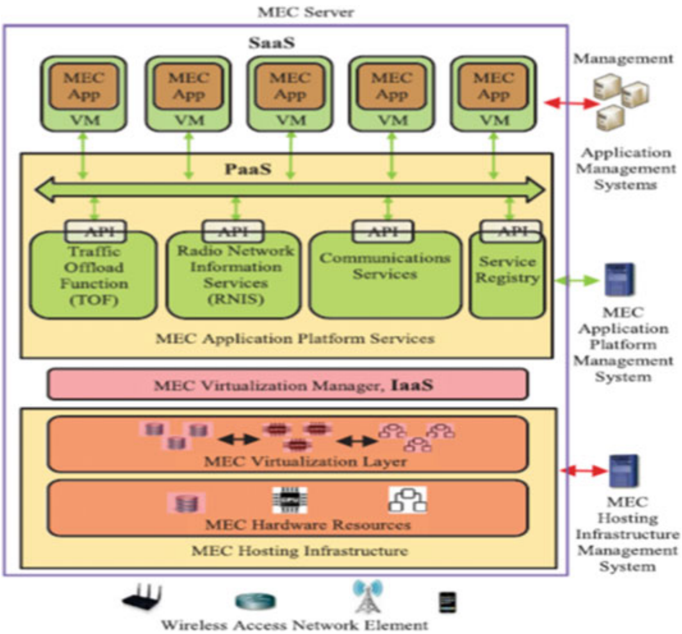


Fig. 70.4 Overview of server platform

70.3.2 *Specification of Server*

The wireless access network is arranged in a hierarchical level and are placed at different locations and each server host different type of application.

Services such as Web services, TCP/IP protocol and other services which contains structure of traditional IP network can be provided by the MEC server [9, 13]. Three of the programming frameworks, i.e., cloudlet, MAUI, and clone cloud and comparison between them is shown in Table 70.2.

Table 70.2 Programming frameworks of offloading

	Approaches of offloading	Remote execution of unit
Cloudlet	To create a user-defined environment for execution, the dynamic VM synthesis by the base VM and VM overlay from mobile devices	Virtual machines
MAUI	Partitions the application program and offloads certain methods to the cloud	Method
Clone Cloud	Partitions the application program and certain threads are migrated to cloud using application-level VM	Thread

70.4 Research Challenges of MEC

The major challenges faced in MEC are (Table 70.3).

70.4.1 Proper Platform with APIs

Network algorithm and control protocol are importantly needed to build a good MEC system, or else it leads to problems in planning. The main problems include: (1) to route quickly and efficiently the traffic to the simplest service case within the network. (2) An adaptive allocation for moving users for allocating function and the required data which will handle user mobility and the required services. (3) To maintain session state while migrating a service from one virtual network to another virtual network without causing any interruption in service.

70.4.2 Management of Resource and Computation Balance

Networks usually contains highly dynamic and complete resources of varies dimension which are placed at geo-spatial points and follows multiple levels of hierarchal network. Further, in unreliable wireless network the changing of mobile device points brings problems in allocation of resources and scheduling algorithm. So, strategies need to be developed to overcome this problem. So, for that, new mechanisms are required to virtualize the complex network and provide resources which are required to the distributed environment.

Table 70.3 Overview of MEC system architecture

MEC architecture	Service model	Targeted layer	Service naming	Service discovery and routing
Application based on TCP/IP	SaaS	Application layer	URI	Based on IP routing and dynamic DNS
Mobility first	SaaS	Transport and network layer	Based on flat service ID	Distributed name resolution servers are used to resolve service ID to network address
Cloud edge	SaaS	Application, network and transport layer	Based on flat service ID	The data plan is programmed with central controller
NFN	SaaS	Network layer	Based on λ calculus expression	The service name prefix diffusion

70.4.3 Interaction with Mobile Devices and Traditional Cloud

The execution of application requires the task to jump from one end mobile node to another, i.e., MEC network nodes and distant central clouds. And it may also involve several stages of computation especially in pipelined fashion where data flow from one stage to another. So, it is important that innovative runtime management schemes needs to be developed for effectively executing the distributed jobs and tasks which as real-time constraints.

70.5 Conclusion

The paper describes about the different MEC systems, their architecture, challenges and models. At first, the paper briefs the characteristics and also explored the applications of MEC based on criteria. Service models are also discussed. Different types of architecture are explained and also the challenges faced by them. As future work, problems faced by MEC in security can be explored and studied.

References

1. Drolia, U., et al.: The case for mobile edge-clouds. In: 2013 IEEE 10th International Conference on Ubiquitous Intelligence and Computing and 2013 IEEE 10th International Conference on Autonomic and Trusted Computing, Vietri sul Mere, pp. 209–215 (2013). https://doi.org/10.1109/UIC-ATC.2013.94
2. Liu, H., Smith, K.: Improving the expected quality of experience in cloud-enabled wireless access networks. In: 2015 IEEE 12th International Conference on Mobile Ad Hoc and Sensor Systems, Dallas, TX, pp. 519–524 (2015). https://doi.org/10.1109/MASS.2015.53
3. Satyanarayanan, M., Bahl, P., Caceres, R., Davies, N.: The case for VM-based cloudlets in mobile computing. IEEE Pervasive Comput. **8**(4), 14–23 (2009). https://doi.org/10.1109/MPRV.2009.82
4. Chen, Y,. Chen, Y., Cao, Q., Yang, X.: PacketCloud: a cloudlet-based open platform for in-network services. IEEE Trans. Parallel Distrib. Syst. **27**(4), 1146–1159 (2016). https://doi.org/10.1109/TPDS.2015.2424222
5. Fog computing aims to reduce processing burden of cloud systems. [Online]. Available: http://www.eweek.com/
6. Liu, H., Eldarrat, F., Alqahtani, H., Reznik, A., de Foy, X., Zhang, Y.: Mobile edge cloud system: architectures, challenges, and approaches. IEEE Syst. J. **12**(3), 2495–2508 (2018). https://doi.org/10.1109/JSYST.2017.2654119
7. Nunna, S., et al.: Enabling real-time context-aware collaboration through 5G and mobile edge computing. In: 2015 12th International Conference on Information Technology—New Generations, Las Vegas, NV, pp. 601–605 (2015). https://doi.org/10.1109/ITNG.2015.155
8. Orsini, G., Bade, D., Lamersdorf, W.: Computing at the mobile edge: designing elastic android applications for computation offloading. In: 2015 8th IFIP Wireless and Mobile Networking Conference (WMNC), Munich, pp. 112–119 (2015). https://doi.org/10.1109/WMNC.2015.10

9. Takahashi, N., Tanaka, H., Kawamura, R.: Analysis of process assignment in multi-tier mobile cloud computing and application to edge accelerated web browsing. In: 2015 3rd IEEE International Conference on Mobile Cloud Computing, Services, and Engineering, San Francisco, CA, pp. 233–234 (2015). https://doi.org/10.1109/MobileCloud.2015.23
10. Dixit, S., Pai, V.G., Rodrigues, V.C., Agnani, K., Priyan, S.R.V.: 3D reconstruction of 2D X-ray images. In: 2019 4th International Conference on Computational Systems and Information Technology for Sustainable Solution (CSITSS), Bengaluru, India, pp. 1–5 (2019). https://doi.org/10.1109/CSITSS47250.2019.9031045
11. Dixit, S., Ranjitha, S., Suresh, H.N.: Segmentation of handwritten Kannada text document through computation of standard error and weighted bucket algorithm. Int. J. Adv. Comput. Technol. **3**(2), 55–62 (2014). ISSN: 2319-7900
12. Dixit, S., Narayan, S.H., Belur, M.: Kannada text line extraction based on energy minimization and skew correction. In: 2014 IEEE International Advance Computing Conference (IACC), Gurgaon, pp. 62–67 (2014). https://doi.org/10.1109/IAdCC.2014.6779295
13. Mannweiler, C., Schneider, J., Schotten, H., Klein, A.: Access schemes for mobile cloud computing. In: 2013 IEEE 14th International Conference on Mobile Data Management, Kanas City, Missouri, USA, pp. 387–392 (2010). https://doi.org/10.1109/MDM.2010.79. https://doi.org/10.1109/MDM.2010.79

Chapter 71
Effective Use of E-tutoring System: Social WhatsApp Messenger on Social Identity Development

Mthobeli Nogubha and Siphe Mhlana

Abstract The present research study outlines the importance of WhatsApp application in the tutoring environment. It analyses utilities related with informative WhatsApp groups. The study additionally investigates how the presence of the tutor in the educational WhatsApp group can have impact on the tutees behaviour and the performance of the tutees. In the study, virtual interviews were convened with computer science first year students. The findings of the study revealed that the significant capacities these WhatsApp educational groups serve are generally for educational purposes. Nevertheless, apart from academic purposes, tutors do utilize WhatsApp groups for sharing career-related vacancies, as well as for entertainment purposes. In addition, the outcomes show that the presence of the tutor has a positive impact on the educational WhatsApp group discussions. Despite the fact that tutees alludes that sometimes these educational WhatsApp groups cause some challenges and can take a great deal of time, but they also accept that educational groups are unavoidable. This research used a qualitative method for data collection and an interpretive phenomenological analysis (IPA) was used to analyse data recorded in the Experience Journal. The results also identify themes that emerged during data analysis.

71.1 Introduction

Nowadays, mobile innovation is commonly used in online teaching in colleges, universities, primary, and high schools around the world [1]. It offers pupils decisions and prospects with regards to online teaching. Online courses with mobile technologies are turning into a more successive factor in learning institutions, and the quantity of mobile courses has expanded [2]. Online tutors and tutees are using mobile innovation technologies in learning institutions around the globe. According

M. Nogubha (✉) · S. Mhlana
University of South Africa, Florida Park, Pretoria, South Africa
e-mail: 49458175@mylife.unisa.ac.za

S. Mhlana
e-mail: mhlans2@unisa.ac.za

T. Senjyu et al. (eds.), *IOT with Smart Systems*, Smart Innovation, Systems and Technologies 251, https://doi.org/10.1007/978-981-16-3945-6_72

to Preston et al. [3], tutees learn equally well from online tutors compared to face-to-face classes. The main aim of this study is to gauge the viability of the blended modes regarding improving tutees' psychological and attitudinal levels with regards to an introduction to programming module. First year university students learn the course entitled C++ or a similar object-orientated language *if–then-else* construct in university. Tutor and e-tutee used WhatsApp as a social media to facilitate the learning of the programming module. Social media are Web 2.0 innovations that encourage social connection and coordinated effort and cultivate a sense of community [4]. Possibly, we should think about another venture—one that will improve tutoring and learning in all content areas utilizing creative strategies in innovation. WhatsApp is a free messenger application that works over various platforms and is generally used among tutees to send multimedia messages like photographs, videos, sounds with instant messages. High invasion of smartphones has started developing utilization of WhatsApp for groups of tutors and their tutees to help the learning procedure by permitting direct access to heaps of online assets.

The rest of the paper is structured as follows: Sect. 71.2 provides the literature review of the paper, Sect. 71.3 discusses the methodology used in this study, Sect. 71.4 discusses data collection techniques, Sect. 71.5 provides data analysis, Sect. 71.6 gives the results of the paper, and Sect. 71.7 presents the conclusion of the paper.

71.2 Literature Review

Over the past years, the high intrusion of smartphones into the market has started increasing utilization of WhatsApp Messenger as a correspondence stage for different tutees educational groups, and lately for groups of tutors and their tutees also. Tutors can make an educational WhatsApp group for their tutees that comprise a kind of simple social network for the class [5]. Nowadays, WhatsApp Messenger has points of interest over other technical devices utilized by the training institutions, for example, low cost, simplicity, openness, effectiveness, and normal language [6]. WhatsApp Messenger underpins information sharing among tutors and tutees. hatsApp Messenger enables studying outside of the classroom, and the tutors' significant availability to tutees, queries, and feedback might improve the learning process. WhatsApp Messenger empowers a simple and fast transaction of connections to contemplate materials [7].

71.2.1 Social Media Environment for Academic Purposes

Learning is a synergistic exertion spent by resolving issues identified with the real structure, appointing in autonomous, unrepeated assignments during the time spent critical thinking, and making a social conditions by forming groups as expressed in the constructivist approach. The strategy in which data is found is significant, so advancement of study tutoring and thinking procedures is vital. Social networking sites, which are known as the main kind of social media by numerous individuals, furnish clients with association, correspondence and cooperation between tutors and tutees by profile pages with individual data. Tutees can make companions or schoolmates; they can chat with each other through instant texting. The profile page can include different media types, for example, photographs, video, sound documents, and web journals. WhatsApp Messenger is the most broadly utilized social networking platform. Social networking sites are utilized by the youthful explicitly, which excite another issue called "WhatsApp addiction", which makes them victims of wasting their precious time on their smartphones, and this habit can be transformed into an advantage by making social networking tutoring groups.

Understanding the intensity of Information and Communication Technology (ICT) and forth Industrial Revolution in the current educational system, pupils need to focus on learning strategies for new times, like social learning where the tutor's duty and direction is vital. Working on social media online networking sites appropriately, they may positively influence the three key learning results; information, aptitudes, and mentality [8].

71.2.2 Users

E-tutoring is slowly gaining recognition with the growth of educational innovations especially in open distance learning environments. E-tutoring is a model of e-learning which aims to bridge the geographical gap between learners and their teachers as well as learners and their peers [9]. E-tutoring involves all of the activities that are aimed at supporting the students in their process of learning [10]. The key users of the e-tutoring system are the e-tutor and e-tutee or e-students. An e-tutor is an individual who facilitates effective knowledge building via supportive interactions with learners in collaborative virtual learning environments [11]. The key responsibility of an e-tutor is to effectively manage the virtual learning environment by making sure that it provides a conducive learning environment where learning can occur [12]. As part of some of the activities or tasks, the e-tutor plan activities that will be used to improve the learning experience, but also prepare for each session by familiarizing with the study material, and other resources delivered to the students.

71.3 Methodology

The nature of this study is the exploration of the social media application to accomplish e-tutoring systems that assist tutor and tutee in mastering the introduction of programming languages in an open distance learning (ODL). Qualitative research approach was selected to help in the exploration of the research objectives and for data collection. Qualitative method is appropriate in situations where the nature of the problem is explorative in nature and researcher would like to gain the deeper understanding of a problem at hand [13]. The method characterized the way information or data is gathered [14]. The research adopted the qualitative method because of the explorative nature and allowed the research collect enough data for this study.

71.4 Data Collection

Data collection techniques include interviews, study groups, journaling, and observation. Interviews can be utilised to discover the opinions, skills, beliefs, and inspirations of specific participants. The research adopted journaling to record the communication between e-tutors and e-tutee when they interact using WhatsApp. A journal is described as both a diary and a log as it mixes an individual's reflections, accounts of actions, and descriptions of knowledge. The main aim of journaling is to record and reflect on experiences in relation to how an individual is thinking, learning, and understanding. According to Lutz and Paretti [15], journals have been utilised in disciplines outside engineering education to promote reflection and deep learning. Reflective journaling means explaining a new experience and unpacking noticeable features, example individuals, resources, events that impacted learning, and doing so in a continuous method over time. Alsaleem [16] describes journaling as the art of written conversation where a learner and educator converse frequently, this can be daily, weekly, or monthly depending on the educational setting. The advantages of journaling include, being an effective instrument to support learning prior to (assisting people explain their expectations), during (responding correctly to the present condition), and after (understanding their skills) educational activities. The question that the students will ask on WhatsApp, either during lessons or after lessons will help us keep a journal because we have something on record. The platform will allow the students to journal their thoughts and feelings in the platform after each interaction.

71.5 Data Analysis

This study adopted an interpretive phenomenological analysis (IPA) to analyse the qualitative data collected. This method of analysis was recommended for this study because it may pay much attention on individual's lived experiences and the manner in which the participants make meaning from these experiences [17]. IPA is a qualitative research method aimed at examining how people make sense of major life experiences, implemented with a relatively small sample size with an aim to reveal some of the experience of the subject under research [18]. IPA approach is also best suited in research studies where the data collection method is done through journal. These data collection methods enable participants provide information that is rich and detailed account of their experiences [18]. The study is aimed at exploring the users' experiences in the context of social media. This was achieved by presentation of lessons to the e-tutee through the use of a social media application tool while observing the participants feelings and reactions in the process. The observed experiences were recorded in an experience journal at the end of each session.

71.6 Results

Participants of the research were the tutee also referred to as the student who is the 1st year student who is registered for a module called Introduction to Programming in the first year module. The students were taught if-then-else statements in C++ through WhatsApp Application. The following table summary of the themes that were identified during the meeting of the tutor and tutee.

Situation	Thoughts
Day 1 Sunday 25 October 2020 First get to know each other session I introduced myself my background in regard to studying. I explained the aim of the class, that it is for a research and that I will try to assist as much as possible. We had the cameras on so that we can see each other. Video was a bit dark, however, we could see each other. I created a Room, which is what enabled us to do the video call on the group, the room was scheduled to start at 21:00, and she acknowledged seeing the agenda for the meeting and the room Student introduced herself as well, I shared the planned lessons and topics for the week with the student, and I also posted it on the wall plan below: Lesson 30 min per day, Sunday, 21:00—Get to know each other session. Monday, 20:00—Simple If then statement examples using algorithm and flowchart, start with nested If then statement. Tuesday, 20:00—Nested if–then statement algorithm and flowcharts and start with coding the simple if then statements on C++. Wednesday, 10:00—C++ coding and revision, questions from student Thursday, 20:00—Student to write an assessment.	Tutor–Feeling a bit anxious about the first lesson, about meeting the student, as I did not understand her level of experience with the topic. I had learnt C++ as part of the research, and I am not a programmer by profession, so I doubted my skill since I did not see myself as an expert. Tutee—was a bit uncomfortable, nervous when I suggested the video call, she commented that she did not prepare her looks, and her hair looks like a mess. Tutee—feeling excited about the tutorial lessons, the module is one of the modules she has been struggling with, and she finds the module very challenging and would appreciate the help. Looking forward to the classes and looking forward to the assistance, Tutee did not mind the data costs as she has been spending a lot of time on YouTube, she has been listening to videos to help her with the module. Tutee—seemed happy to see the plan for that evening as it was discussed on the group wall. Tutee—feeling delighted the exercises done that evening were also going to be posted on the group. Tutor—Worried that the student might have missed the time since the session was not setup, tutee joined about 5 min late. Tutor felt relieved and relaxed as soon as the tutee joined

(continued)

(continued)

Situation	Thoughts
Day 2, 26th October 2020 Session was scheduled for 20:00 as both tutor and tutee were comfortable with the time. I only setup the room invite before the session started. We used video to try and make it close to the natural environment. We picked up that when sharing. The video automatically switches off anyway, and I did not know that that is how it works. We started with easy examples, one example was using if–then-statement to login, and the other one was to get a student grade, converting the examples to a flowchart. I asked the student to bring her examples from her textbook the next day. I shared the link of the drawing tool, and I also shared the questions and solution of the exercises we did in the group. Student was able to ask some questions on the chat during the lesson, I encouraged her to do this, keep a record so that she does not forget	Tutor—disappointed as the quality of the video was not good, I ended up switching off the camera option while projecting. Tutee—mentioned that the examples were very easy and boring. Tutee—was very cheerful to find out about my career, what I currently do, and how it started. Tutor—delighted that I can share my journey with someone. Tutee—showed interest in the drawing tool I used, it is something that could come handy with her other assignments, was excited to see that it is so easy to use

(continued)

(continued)

Situation	Thoughts
Day 3, 27th October 2020 Session was scheduled for 20:00; however, student was running late, she had mentioned previously that she will be coming from another appointment. She sent a message by the messenger at 19:45, asking that we meet at 20:30. I missed the notification, saw it around 20:20 as all my social media notifications are off. I also did not pay attention to the messenger alert when I was in front of the PC. We started by looking at the nested loop flowchart and algorithm, I shared the screen since tutee did not have the drawing tool. The session had lots of network issues and there were times we could not hear each other. We then used the chat function and the messages were not coming in immediately. We troubleshoot and restarted the student PC, and there was some improvement. This evening was filled with a lot of challenges. When the student wanted to share, we picked up that her PC only allowed her to share the browser that Facebook was running on, she could not share the code building tool. I had to lead the writing of the code as had wanted the student to write the code herself so that I can understand her logic. Students provided examples, a combination of a for loop with an if then statement, we agreed we would look at it the following day. Lessons took longer than the planned 30 min due to the technical difficulties	Tutor—feeling puzzled on where the student might be, before I saw the message on the chat function. Tutor—felt a relief when I saw the message. Tutee feeling a bit sad when she found out about the misunderstanding. Tutor feeling confused when the student was responding due to network glitches, later the feeling changed to frustration. Tutee—feeling frustrated at not being understood and heard. Tutors—felt disappointed when the environment did not work as expected as the lesson plan had to change. Tutor and Tutee were not bothered by the lesson ending late as there was not travelling involved
Day 4: 28th October 2020 Session was scheduled to start at 10:00, tutor sent a message by messenger requesting that we start at 10:20 as she was held up somewhere. Students acknowledge saw the message and the meeting started at 10:20. There was no time for me to look at the example before we met as we finished late and we started early this day. The exercise was a bit complex; however I tried as much as possible to assist the student	Tutor—I felt excited when I saw the acknowledgement response from the student about the time change. I was using the message feature from my phone as I was not in front of the laptop Tutor—a bit nervous to tackle the exercise since I did not have time to work on it prior the lesson. Tutor felt disappointed as the solution did not work as expected; there was something we were missing in our code on the For loop. Tutor feeling relieved that the student felt like I had assisted tutee—feeling excited that some of the items that were not clear to her were made clear while tackling the exercise. Tutee-feeling grateful on the lessons learnt and appreciated the lessons

71.7 Conclusion

The main objectives of this study were to introduce an environment where e-tutoring takes place and to explore how the use of social media as e-tutoring platform addresses the academic needs of learners and their user experience. The study efforts were much focused on gaining a deeper understanding on how the social media environment can be used as a platform, and how it makes sense in fulfilling the academic needs of learners, but also examining their user experience associated with such learning spaces. The major themes identified in this study are a contribution to the already existing literature in terms of the multiple benefits that the online learning platform offers. The major focus of this study was limited to the social media context, but the observed benefits of implementing social media as a teaching and learning platform are not far from those of traditional online learning strategies such as time flexibility, convenience, accessibility, and cost effectiveness [19–21]. However, this study exposed extra advantages that are specific to a social media context, such rapid feedback from e-tutors, participants' freedom to probe for more details on areas that proved challenging to their learning process. While the social media platform to some extent is regarded as an informal environment for transacting academic information between the researcher and the participants, this is the amount of academic benefit that was disclosed in this study that matters most. This study, however, did have some limitations. The major challenge was that the study sample was too small to provide a good representation participants' observations, therefore, affecting results of the findings to some extent.

Acknowledgements The results presented in this research forms part of a larger project in the School of Computing. I would like to acknowledge Mrs. Petra le Roux, the project champion, for supporting me in this research.

References

1. Jimoyiannis, A., Tsiotakis, P., Roussinos, D., Siorenta, A.: Preparing teachers to integrate web 2.0 in school practice: Toward a framework for pedagogy 2.0. Austr. J. Educ. Technol. **29**, 248–267 (2013). https://doi.org/10.14742/ajet.157
2. Inan, F.A., Flores, R., Grant, M.M.: Perspectives on the design and evaluation of adaptive web-based learning environments. Contemp. Educ. Technol. **1**, 148–159 (2020). https://doi.org/10.30935/cedtech/5971
3. Preston, G., Phillips, R., Gosper, M., et al.: Web-based lecture technologies: highlighting the changing nature of teaching and learning. Austr. J. Educ. Technol. **26**, 717–728 (2010). https://doi.org/10.14742/ajet.1038
4. Bingham, T., Conner, M.: Social Learning: A Guide to Transforming Organizations Through Social Media, pp. 1–22 (2011)
5. Fischer, Y.: The Facebook is Dead—Long Live WhatsApp (2013)
6. Church, K., De Oliveira, R.: What's up with WhatsApp? Comparing mobile instant messaging behavior with traditional SMS. In: MobileHCI 2013—Proceedings of 15th International

Conference on Human-Computer Interact with Mobile Devices Service, pp. 352–361 (2013). https://doi.org/10.1145/2493190.2493225
7. Bouhnik, D., Deshen, M.: WhatsApp goes to school: mobile instant messaging between teachers and students. J. Inf. Technol. Educ. Res. **13**, 217–231 (2014). https://doi.org/10.28945/2051
8. Orús, C., Barlés, M.J., Belanche, D., et al.: The effects of learner-generated videos on YouTube on learning outcomes and satisfaction. Comput. Educ. **95**, 254–269 (2016). https://doi.org/10.1016/j.compedu.2016.01.007
9. Matoane, M., Mashile, E.O.: Key considerations for successful e-tutoring: lessons learnt from an institution of higher learning in South Africa. In E-Learn: World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education. Assoc. Adv. Comput. Educ. 863–871 (2013)
10. McPherson, M., Baptista Nunes, M.: Organisational issues for e-learning: critical success factors as identified by HE practitioners. Int. J. Educ. Manag. **20**, 542–558 (2006). https://doi.org/10.1108/09513540610704645
11. Kopp, B., Matteucci, M.C., Tomasetto, C.: E-tutorial support for collaborative online learning: an explorative study on experienced and inexperienced e-tutors. Comput. Educ. **58**, 12–20 (2012). https://doi.org/10.1016/j.compedu.2011.08.019
12. Mashile, E.O., Matoane, M.C.: E-Learning Development: the case of UNISA. In E-Learn: World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education. Assoc. Adv. Comput. Educ. 1152–1157 (2012)
13. Creswell, J.: Design: Choosing Among Five Approaches (2012)
14. Oates, B.J.: New frontiers for information systems research: computer art as an information system. Eur. J. Inf. Syst. **15**, 617–626 (2006). https://doi.org/10.1057/palgrave.ejis.3000649
15. Lutz, B.D., Paretti, M.C.: Development and implementation of a reactive journaling method for qualitative research. In: Proceedings of ASEE Annual Conference on Expo, pp. 9461–9475 (2019)
16. Alsaleem, B.: The effect of "Whatsapp" electronic dialogue journaling on improving writing vocabulary, word choice and voice of EFL undergraduate Saudi Students. Arab World English J **4**, 213–225 (2013)
17. Braun, V., Clarke, V.: Successful Qualitative Research: A Practical Guide for Beginners, 400 (2013)
18. Smith, J.A., Shinebourne, P.: Interpretative phenomenological analysis. Am. Psychol. Assoc. (2012)
19. Graham, C.: Blended learning systems. In: The Handbook of Blended Learning: Global Perspectives, Local Design, vol. 1, pp. 3–21 (2006)
20. Allen, I.E., Seaman, J.: Going the distance: online eduation in the US, 2011. Sloan Consortium **44**, 1–44 (2011)
21. Roy, A., Raymond, L.: e-Learning in support of SMEs: Pipe dream or reality. In: Proceedings of 5th European Conference on e-Learning, pp. 383–388 (2005)

Chapter 72
Recognizing Abnormal Activity Using MultiClass SVM Classification Approach in Tele-health Care

Aniruddha Prakash Kshirsagar and L. Shakkeera

Abstract Anomalies are cases of an active class that vary in their performance from the standard or planned sequence of events. Early identification of body sensor networks is an essential and difficult role and has many possible benefits. Both these tasks are very important for the monitoring of suspicious behavior to capture human activities and evaluate these results. Data collection is important if the true evidence of human behavior is to be discovered. Since it uses the technique of state table transformation to find suspicious behaviors, identify them and track them from a distant location for elderly or children. This article offers an automatic activity surveillance system that identifies common activities that usually occur in a human routine. SVM is used here E-MultiClass (Support Vector Machine). It is used to rate two data groups with a binary score. This paper describes our task of data acquisition, recording them in the field of imagery processing and validating our algorithms on data obtained from remote sensors. It is also used to detect improvements in human routine as well as everyday life across these capabilities. It describes our activities in the field of image processing.

72.1 Introduction

The climate has changed significantly with the advent of new technology in the last century. Smart devices were distributed around the globe over the last decade. Various emerging innovations increase urban productivity and well-being by maximizing the usage of diverse services, such as power, water, governance, healthcare, and other [1–5]. In the field of electricity delivery, smart meters are among the key facets of technology upgrades all the time. Many countries around the world use intelligent power meters to monitor trends of energy consumption. These device meters are key components in intelligent digital infrastructure which report more detailed information on electricity consumption than classic electric meter [6]. Any house with many smart meters communicates its readings in the smart society to the

A. P. Kshirsagar (✉) · L. Shakkeera
VIT Bhopal University, Bhopal, Madhya Pradesh, India

T. Senjyu et al. (eds.), *IOT with Smart Systems*, Smart Innovation, Systems and Technologies 251, https://doi.org/10.1007/978-981-16-3945-6_73

main portal which sequentially transfer data to the community center in safe and privacy-friendly ways via the community network. The business center pays for the use of services in each household and stores detailed statistics. The data is regularly transmitted to the call center which transmits it for computer billing to the utility provider. Because there is no direct participation of the individual in the facilities maintenance process, the administrative expense of the service provider is minimized. The privacy and protection of the household have been enhanced on the customer side because there has been no measuring staff approaching their residences. Smart meters operate across the home area network as a two-way interface for consumer side applications.

72.1.1 Motivation

The most important field for future study in video monitoring is the examination and identification of human activities [7]. This is justified by the need to track and support the everyday lives of aged or disabled persons [8]. This need, along with the relative costs of cameras and sensor technology [8], has encouraged researchers in diverse situations or settings for tracking purposes a broad ranging of ways and methods of identifying various forms of human activities. The identification of suspicious events, by humans or computers, in the event of possibility or need is one achievement of recent activity in this field [9]. This axis is not yet rich in particular tools and dedicated libraries, but it uses the approaches used for the detection of human behaviors by applying them to unusual practices. Advent in this axis is often inspired by the degree of monitoring that is followed for learning different human behaviors (supervised, unmonitored or semi-supervised), the methods used to remove noise and thus to ensure the robustness of interest characteristics and the related steps for the comparison of descriptors and classifiers as defined in [4]. In addition, the question emerged as to the abstraction and representation, as also the essential characteristics that allow the choice of characteristics to be drawn, of the actions of individuals in a given scene.

Sensor technologies and smart can be used to track human movement in smart houses. The following is the reasons why we performed this study:

(i) The everyday life tasks (ADL) of sensors, portable strips and CCTV cameras are currently known. The sensors are correlated with a high price, which does not allow sensors to be used for large-scale activity detection.

(ii) Wearable bands and the protection of CCTV impact consumers and many people refuse to allow them in their daily lives. In such a situation, it tends to be a decent scheme to use such passive methods such as smart meter/abnormal behavior monitoring with a certain activity pattern for reconciliation.

(iii) Anomaly detection methods have not been conducted in the detection of irregular patterns so far.

Fig. 72.1 Abnormal activities in ATM

Abnormal Activities

An irregular activity is characterized as any rare and uncommon activity which could lead people in some contexts [1] to hazards. The writers of [2] describe "abnormal behaviors" as "connected and not planned activities." The word can vary depending on the area of research or context, for example, when examining crowd actions on the streets where crime or robbery habits may be anomalous. Anomalous acts are of various kinds, analogous to regular activities: movements, simple actions, events, interactions, habits, and group actions [3]. The writers of [4] identify irregular behaviors into: community activities which may include evacuation (rapid panic dispersion), single-direction herds and crowd forming, local dispersion and separation, and the activities of individuals as dropping [6], non-mobile as [10], odd pace [11], and misdirection (Fig. 72.1).

72.2 Related Work

1. In terms of instruction, the bulk of studies in this field used supervised learning approaches (e.g., decision-making, climbing, biking, etc.), with the use of discriminatory classificators and generative mechanisms.
2. Semi-supervised and transition learning: The previous study has integrated semi-supervised education into behavior or background recognition structures to minimize the dependence on branded education data and to maximize the advantages of plentiful, unlabeled data [10]. Semi-monitored methods of study may boost the precision of the identification by refining the judgment limit based on unlabeled data distribution or allocating highly confident approximate labels for unlabeled data.
3. The accuracy of recognition of human behavior was increased with active learning. The previous work is strengthened by incorporating active learning into the behavior detection process, so that the machine can recognize undefined behaviors.

4. Another study path is unattended learning. Unattended learning is linked. Clustering or discovery trend rather than grouping is the subject of unregulated learning [7]. The perception of human behavior is categorized into the identification of activity and activity patterns. The first category focuses on the detailed identification of human behavior on the basis of a predefined or pretrained model of activity, while the second category detects unknown patterns directly from low sensor results. These methods are generated by a variety of unidentified clusters that cannot be used for classification or identification purposes. Labels are still needed to link the discovered patterns to the existing groups in order to perform identification.
5. Approach based on laws: There are also several rules regarding the identification of events. In [9], Store et al. recommended to use rules and manual configurations in extensible markup language (XML) format a multi-agent based structure. The authors have used fluorescent logic for example. Detection of the operation "meal preparation" covers a variety of situations and guidelines including the mixture of a microwave, a fridge for use, a kitchen counter, etc. Approaches based on guidelines can be difficult to enforce without any domain experience, or where there is no clear law and so data must be studied.

72.3 Proposed System

Context and human activities comprehension are a central factor that promotes all aspects of context perception and makes it possible. The study proposed developed a fully supervised learning approach that correctly detects irregular behavior with a limited number of field true mark applications. And where there is no preparation material for a specific case type, consider human behavior. This approach will generalize information that has been studied before and improve the capacity to understand emerging business groups. Like the elderly, they tend to live independent lives. At the age of the infant, people become vulnerable to multiple injuries. A smart home monitoring system [1] offers an insight into the faulty logic system used by physioprotection sensors, microphones, infrasound sensors as well as debit sensors and status change sensors to identify human everyday life. The system is incredibly useful and important to develop smart home monitoring systems. In order to acknowledge behavior conduct and to forecast irregular conditions, real-time data processing is therefore necessary.

SVM uses a transformation table to minimize computational time, helping to eliminate unattainable classification states. The method suggested is based on the following observations: Multiple human behaviors and viewpoints have the same simple semantic parameters:

Sit, run, walk, slip, stand, etc., parameters are normal occurrences which can be used in various situations such as lunch, desk, telephone chat, etc. The mathematical parameter model can then be moved from one operation to another. Incorporating human intelligence can transcend the limitations of supervised learning:

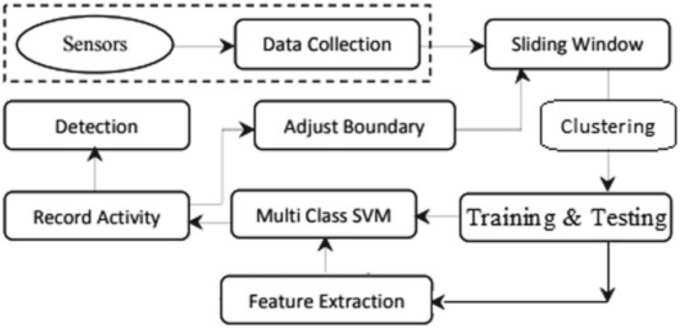


Fig. 72.2 E-MultiClass SVM working

Sensor data and labels for all contexts allow people to identify a context category with nameable parameters without the selection process of sensor data. For example, the action "Office work," "Hands On the Table," and sound-related attributes like "Printer sound," "Keyboard sound" and "Conversations" can easily be connected together with motion-related attribute like sitting (Fig. 72.2).

SVM E-Class Framework Working

It demonstrates here how anomalous activity identification is effectively carried out by remote sensors in the concept of smart home on the given data.

The first step is to define the daily and repayable behavior that is observable by sensors or cameras comprising our intelligent home concept. Once the operation and the related events have been discovered and categorized, we retain the subsequent entries in the database. It will then create a model for identifying the activity, beginning to evaluate current and new events. But how can you find a recent or run-time operation not in the database? It uses an active directory definition that senses a non-differentiated concept of learning in a run-time manner. Numerous functions are given here, for example, filtering, normalization, E-MultiClass SVM, K-means and Random forest algorithm for some input data. Systems obstruct our system's capabilities.

(a) **Input data**: Sensing data is gathered and then sent continuously to a hospital program or other remote site, to aid in conditions known as receiver for any suspicious operation. The collected data was tracked constantly and compared to various metrics, attributes and patterns.

The input requested from a Proposed approach is the following:

```
INPUTS
1. TRAINING SET { xi, yi, i=1..l}
2. WHEIGHTS qi, i=1..l
3. BIAS b
4. TRAINING SET PARTITION INTO
SUPPOTSET(S) , ERRORSET(E)
   AND REMAINING SET(R)
5. PARAMS: e , C , KERNELTYPE AND
KERNEL PARAMs
6. R MATRIX
7. NEW SAMPLE C = (xc, yc)
        Where, x and y are the position of each
value with i^th number of sets.
```

At the beginning of a training, the training set, coefficients, support set, error set and remaining set is set to 0.

(b) Tracking, learning, and reconnaissance system are not based on sensor types and computer types such that sensor data sources are any sort of information: data collection (data tracking and discovery). Data collection. To improve recognition accuracy, it is necessary to choose the correct set of parameters or attributes. Assume that individuals with the desire to do fitness tasks, such as warm up, carrying, resting, pitching, walking and running, are carried out with various activities. The thin grains of members, joints, and muscles will then further break into each subactivity and it has been discovered correctly on this basis. The irregular behavior is seen using the transformation table which includes all possible conditions. The machine is qualified to classify and monitor anomalies in the behaviors of the people. The system distinguishes nine separate individual operations using E-MultiClass SVM.

(c) **E-MultiClass SVM**: Support vector machines (SVMs), which analyzes data and identifies patterns used for classification and regression analyzes, are supervised learning models with related learning algorithms. SVM allows categorizing because it is used to identify images and their details. It provides substantial precision in the search [12].

```
IF (c € REMAINING SET)
REMOVE   SAMPLE   FROM   REMAINING
SET
REMOVE SAMPLE FROM TRAINING SET
EXIT
IF (c € SUPPORTSET)
        REMOVE SAMPLE FROM SUPPORT
SET
IF (c €  ERROR SET)
REMOVE SAMPLE FROM ERROR SET
```

Initially, it is easy to delete if it is in the remaining package.

If the values of qc are equal to 0, they can otherwise be excluded from their sample and begin a loop that ends only.

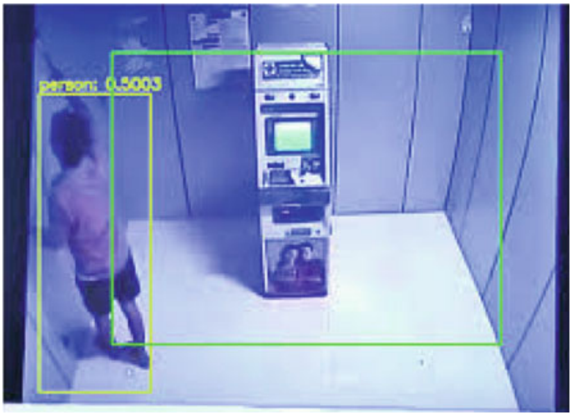


Fig. 72.3 Abnormal activity detection using E-MultiClass SVM method

(d) **Active Activity Detection**: Active Directory events not identified but regular activity and the unidentified ones that cannot be tagged as an irregular activity can be detected by categorizing the defined activities. In order to describe regular operations, the transformation table used to pick the E-MultiClass SVM can be used. The transformation table describes all possible states which a person can execute. If there are events outside the transformation table spectrum, the incident is marked as irregular.

OUTPUTS

1. NEW TRAININGSET { xi, yi, i=1..l+1}
2. NEW COEFFICIENTS qi, i=1..l+1
3. NEW BIAS b
4. NEW TRAININGSET PARTITION
5. NEW R MATRIX

The output comprises all the values that are changed in the input. The action identification as seen below (Fig. 72.3).

72.4 Results and Discussions

Figure 72.4 gives the details of proposed model where we had trained 80% dataset after training from sample dataset, we are using 20% dataset for testing and observed the result with some stander assumption those we have considered and predict the model working. From that it shows proposed model give good predication in less time period. Details predication is shown in Fig. 72.4.

The above sample dataset contains the near about 400 to 500 data entry which contain the different parameter for evaluation propose from that we are able to predict the output.

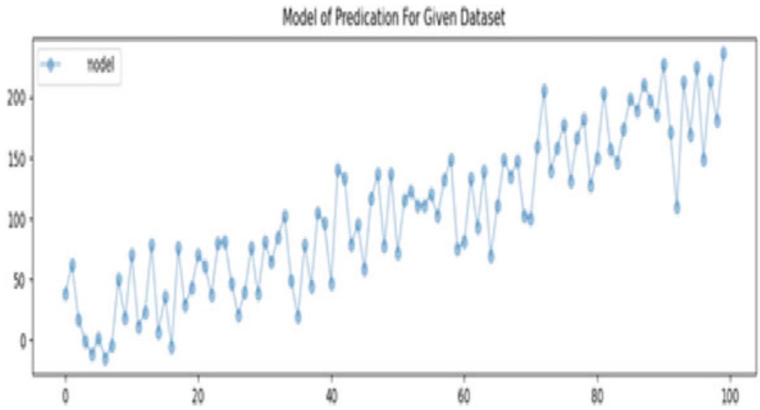


Fig. 72.4 Model of predication of given dataset by training the datasets

In case of mapping, results of both existing as well as proposed model were compared with that of statistical analysis.

By considering the sample dataset of near about 400 sample we have performed the prediction using SVM and proposed E-MultiClass SVM model. Figure 72.5 gives the details of SVM model where we had trained 80% dataset after training from sample dataset, we are using 20% dataset for testing and observed the result with some stander assumption. From observation, its shows that proposed model gives good accuracy which is good as compare to conventional SVM.

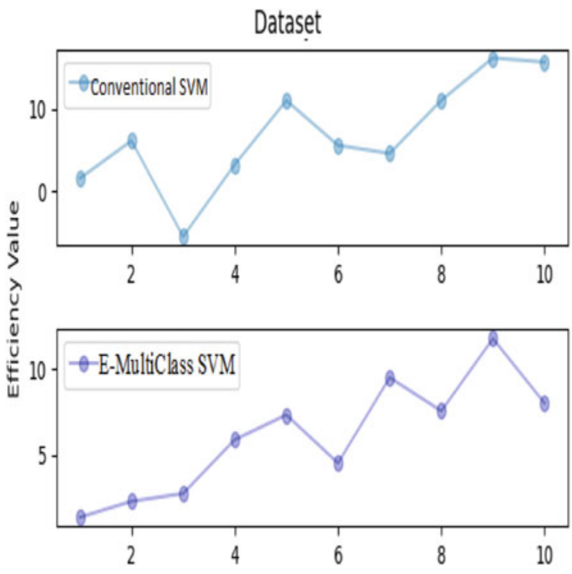


Fig. 72.5 Efficiency of using conventional SVM and E-MultiClass SVM approach

From above result shows that proposed model give better result as compare to conventional SVM. So it's give better accuracy and efficiency with less time stamp.

Efficiency: From above model SVM with E-MultiClass SVM model, we have calculated following efficiency using efficacy formula for that we had consider the give dataset value with time period on that basis we are able to find the accurate efficiency of model. It is observed in Fig. 72.5 that Conventional SVM model give the varying efficiency as compare to Proposed E-MultiClass SVM approach. Proposed work with AI give the simple and good efficiency as compare to other show in Fig. 72.5.

$$c(x) = \arg\max_{c \in C} \sum_{i=1}^{k} \delta(c, c(y_i)),$$

where $c(y_i)$ is the class of y_i, and δ is a function that $\delta(u, v) = 1$ if $u = v$.

Accuracy

From above model SVM with E-MultiClass SVM approach, we have calculated following accuracy for that we had consider the give dataset value with time period on that basis we are able to find the accurate accuracy of given dataset. It is observed in Fig. 72.6 that SVM model gives the varying accuracy as compare to proposed approach. Classification obtained results of accuracy for the use of AI classifier method 80.7%, whereas SVM gives 79.30%. E-MultiClass SVM approach gives the good accuracy as compare to other show in Fig. 72.6.

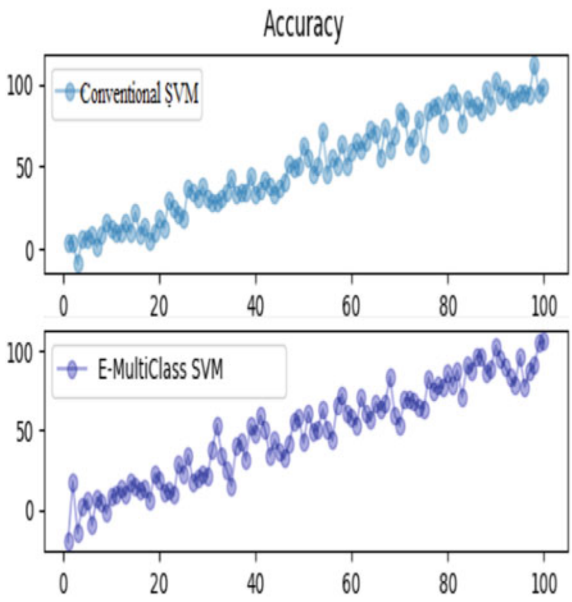


Fig. 72.6 Accuracy of given dataset as compare to Conventional SVM and E-MultiClass SVM approach

Perception

From above model SVM with E-MultiClass SVM approach, we have calculated following accuracy for that we had consider the give dataset value with time period on that basis we are able to find the accurate perception of given dataset. It is observed in Fig. 72.7 that SVM model gives the varying accuracy as compare to E-MultiClass SVM approach.

We track the suspicious behavior with the use of the E-MultiClass SVM system and observation of brain activity with numerous clinical approaches.

Besides the estimation time, the classification's performance metrics are also compared, and the results are assessed and performance validated after classification. It is evident from Table 72.1 that the classic SVM classification has increased performance. It needs more computing power to have high efficiency.

At the rate of high computing times, the proposed classifier has obtained high results. The method suggested uses less computing power, but achieves better efficiency than the standard classifier.

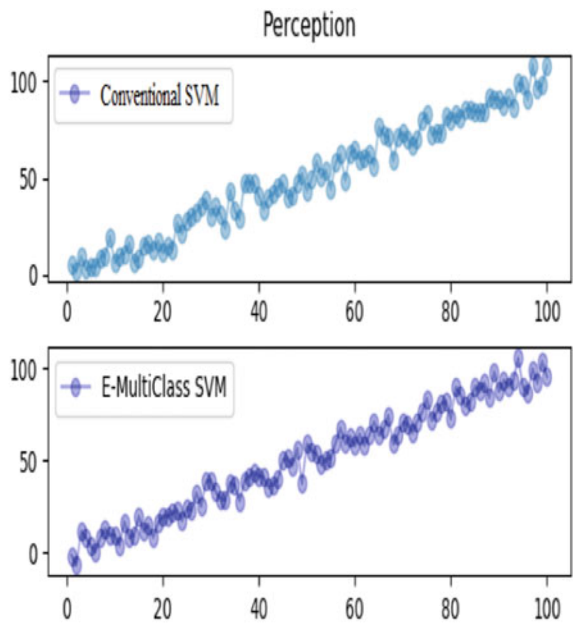


Fig. 72.7 Perception of given dataset as compare to Conventional SVM and E-MultiClass SVM Approach

Table 72.1 Performance comparison

	Conventional SVM (%)	E-MultiClass SVM (%)
Accuracy	95.8	96.4
Precision	96.1	96.7
Efficiency	94.2	95.8

72.5 Conclusion

The framework suggested in this paper for irregular behavior detection is functioning effectively in real time. Availability is defined, and abnormality is observed with high precision in the method suggested. A new SVM classification method for E-MultiClass is adopted, searching for and avoiding states which cannot meet the current state. The transformation table can be used to pick the classifiers that assign the data into one of the operations. Without losing the precision, it decreases measurement time. Improved performance measures, such as accuracy, precision, and sensitivity of the proposed unit, make it efficient for real-world applications. Future studies may use data from environmental and physiological sensors. More sensors can be used to consider the background facts and health conditions of the patient.

References

1. Beretta, I.: Te social effects of eco-innovations in Italian smart cities. Cities **72**, 115–121 (2018)
2. Kim, B.: A distributed coexistence mitigation scheme for IoT based smart medical systems. J. Inform. Process. Syst. **13**(6), 1602–1612 (2017)
3. Sharma, P.K., Moon, S.Y., Park, J.H.: Block-VN: a distributed blockchain based vehicular network architecture in smart city. J. Inform. Process. Syst. **13**(1), 184–195 (2017)
4. Jo, H., Yoon, Y.I.: Intelligent smart home energy efficiency model using artificial TensorFlow engine. HCIS **8**(1), 1–18 (2018)
5. Sharma, P.K., Ryu, J.H., Park, K.Y.J.H., Park, J.H.: Li-Fi based on security cloud framework for future IT environment. HCIS **8**(1), 23–36 (2018)
6. Ibrahim, Z., Rahim, F.A., Mokhtar, S.: Digital forensics issues in advanced metering infrastructure. J. Fundamental Appl. Sci. **10**(6S), 2714–2726 (2018)
7. Wu, X., Chu, Z., Yang, P., Xiang, C., Zheng, X., Huang, W.: TW-See: human activity recognition through the wall with commodity Wi-Fi devices. IEEE Trans. Veh. Technol. **68**(1), 306–319 (2019)
8. Passerini, F., Tonello, A.M.: Smart grid monitoring using power line modems: anomaly detection and localization. IEEE Trans. Smart Grid **7**(1), 27302–27312 (2019)
9. Wang, J., Chen, Y., Hao, S., Peng, X., Hu, L.: Deep learning for sensor-based activity recognition: a survey. Pattern Recogn. Lett. **119**(1), 3–11 (2019)
10. Siano, P.: Demand response and smart grids—a survey. Renew. Sustain. Energy Rev. **30**(5), 461–478 (2014)
11. Kelly, J., Knottenbelt, W.: Te UK-DALE dataset, domestic appliance-level electricity demand and whole-house demand from fve UK homes. Sci. Data **2**(1), 150007–150021 (2015)
12. Malasinghe, L.P., Ramzan, N., Dahal, K.: Remote patient monitoring: a comprehensive study. J. Ambient. Intell. Humaniz. Comput. **10**(1), 57–76 (2019)
13. Nweke, H.F., Teh, Y.W., Mujtaba, G., Al-Garadi, M.A.: Data fusion and multiple classifer systems for human activity detection and health monitoring: review and open research directions. Inform. Fusion **46**(1), 147–170 (2019)
14. Hossain, M.S.: Patient state recognition system for healthcare using speech and facial expressions. J. Med. Syst. **40**(12), 272–280 (2016)
15. De, P., Chatterjee, A., Rakshit, A.: Recognition of human behavior for assisted living using dictionary learning approach. IEEE Sens. J. **18**(6), 2434–2441 (2018)

16. Ghayvat, H., Mukhopadhyay, S., Shenjie, B., Chouhan, A., Chen, W.: Smart home based ambient assisted living: recognition of anomaly in the activity of daily living for an elderly living alone. In: Proceedings of the 2018 IEEE International Instrumentation and Measurement Technology Conference (I2MTC 2018), pp. 1–5. MTC, May 2018
17. Ye, J., Stevenson, G., Dobson, S.: Detecting abnormal events on binary sensors in smart home environments. Pervasive Mob. Comput. **33**, 32–49 (2016)
18. Skocir, P., Krivic, P., Tomeljak, M., Kusek, M., Jezic, G.: Activity detection in smart home environment. Procedia Comput. Sci. **96**(1), 672–681 (2016)
19. Lu, L., Qing-ling, C., Yi-Ju, Z.: Activity recognition in smart homes. Multimedia Tools Appl. **76**(22), 24203–24220 (2017)
20. Yao, L., Sheng, Q.Z., Benatallah, B., et al.: WITS: an IoTendowed computational framework for activity recognition in personalized smart homes. Computing **100**(4), 369–385 (2018)
21. Hassan, M.M., Uddin, M.Z., Mohamed, A., Almogren, A.: A robust human activity recognition system using smartphone sensors and deep learning. Futur. Gener. Comput. Syst. **81**(1), 307–313 (2018)
22. Sendra, S., Parra, L., Lloret, J., Tomas, J.: Smart system for ´ children's chronic illness monitoring. Inform. Fusion **40**(1), 76–86 (2018)
23. Pham, M., Mengistu, Y., Do, H., Sheng, W.: Delivering home healthcare through a cloud-based smart home environment (CoSHE). Futur. Gener. Comput. Syst. **81**(1), 129–140 (2018)

Chapter 73
Ontology Based Food Recommendation

Rohit Chivukula, T. Jaya Lakshmi, Saleti Sumalatha, and Kandula Lohith Ranganadha Reddy

Abstract Eating right is the most crucial aspect of healthy living. A nutritious, balanced diet keeps our bodies support fight off diseases. Many lifestyle related diseases such as diabetes and thyroid can often be avoided by active living and better nutrition. Having diet related knowledge is essential for all. With this motivation, an ontology related to food domain is discussed and developed in this work. The aim of this work is to create on ontology model in the food domain to help people in getting right recommendation about the food, based on their health conditions if any.

73.1 Introduction

There is enormous information available in the internet for related to the domain of food science. Currently, information retrieval systems are centred on text mining [1]. Often, these techniques retrieve irrelevant information if the user query is not clear. This may lead to adverse effects on user's food requirements. A solution to the above chaos is ontology. Ontologies are an interesting avenue to connect different information systems via the Semantic Web. Ontology is a data model that describes concepts in a specific domain and the relationships between them.

In this work, a food ontology is presented for aiding user's needs. Some key concepts related to food domain comprising the types of food, flavours and textures, and different kinds of food courses and meals are considered. In addition to them, information regarding some recipes, ingredients used in them along with their nutritional facts are incorporated. Finally, user with the details of physical attributes such as age, weight with health history is also maintained. Food ontology designed in this work can be used by people with culinary interests, nutritionists, restaurants and chefs. Following are some of the situations where the food ontology discussed in this work could be useful.

R. Chivukula (✉)
University of Huddersfield, Huddersfield, UK

T. J. Lakshmi · S. Sumalatha · K. L. R. Reddy
SRM University, Amaravati, Andhra Pradesh, India
e-mail: sumalatha.s@srmap.edu.in

T. Senjyu et al. (eds.), *IOT with Smart Systems*, Smart Innovation, Systems and Technologies 251, https://doi.org/10.1007/978-981-16-3945-6_74

- Personalized nutrition recommendation for a user.
- To recommend best matched substitutions for ingredients where appropriate in order to meet the users taste preferences or food allergy related needs.
- To suggest the users best recipes for the specified meal types (e.g. lunch) or course types (e.g. starter).
- To suggest food for a user based on taste and interest. For example, if a user requests not to have recipes that contain high Glycaemic Index (GI) ingredients, the system will avoid offering these high glycaemic recipes in the recommendations.
- To suggest appropriate recipes to a user having disease like diabetes.

The main elements of ontology are classes, individuals and properties/ relationships. A class is a collection of objects that describe concepts in the domain. There can be subclass to a class. In food domain, example classes can be ingredient, nutrient and recipe. Individuals are the instances in ontology. For example, almond is an instance of an ingredient class, and banana_bread is an instance of a recipe class. Relations describe the way in which classes relate to each other. Relations can normally be expressed directly between individuals (such as banana_bread has ingredient almond) or between Classes (such as ingredient has nutrient). Properties describe the attributes of individuals or relations. There are two types of properties. First one is the datatype used to assign a value to a property (such as banana has glycaemic index 55), and second is between two objects (such as recipe A has ingredient b).

73.2 Modelling Food Ontology

There are several food ontologies available in the literature for various purposes. Section 73.4 specifies potential works published online. The food ontology modelling done here is inspired from the work of Haussmann et al. [2] and Bailoni et al. [1]. The key terms of food ontology considered in this work are Food, Chemical Food, Course, Meal, Ingredient, Recipe, Nutrient, User and Disease. Some meaningful relations exist between a pair of classes and a few classes have some attributes. Figure 73.1 describes the overall ontology of food domain that is considered in this work.

73.2.1 Classes

The food ontology is modelled with 10 classes initially. Some of these classes have sub classes. Table 73.1 describes the purpose of each class.

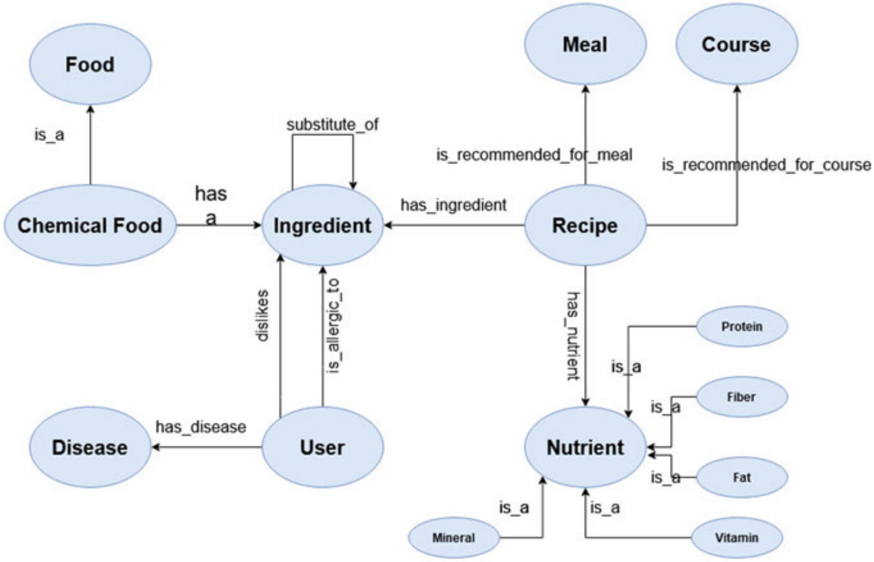


Fig. 73.1 Proposed food ontology

Table 73.1 Classes and their descriptions

Class	Description
Food	A Food is something that can be eaten. A recipe uses foods as part of ingredients and also produces foods to be eaten. Chemical Food is a subclass of Food
Characteristic	Characteristic is a differentiating feature of food. For instance, different food items can have different flavors and textures. Subclasses include Flavour and Texture
Course	A specific set of food items that are served together during a meal. Starter, Main Course and Dessert are subclasses of Course
Meal	Represents the time of taking food. Breakfast, lunch, dinner are a few examples of meal
Recipe	A combination of ingredients and a method of preparing. Subclasses include Breakfast Recipe, Dinner Recipe, High Glycaemic Recipe, Quick Recipe, Lunch Recipe and Side Recipe
Measurement	This class is considered to represent the information of time taken to prepare or temperature at which it is to be prepared. For example, a sandwich may take 15 min to prepare, a cake needs to be baked at 350 °F etc. Temperature Measurement and Time Measurement are subclasses of Measurement
Ingredient	A food substance used to prepare a food
Nutrient	Nutrients are compounds in foods that are essential to life and health and providing us energy. Examples include protein, vitamin, fibre, etc. Subclasses of Nutrient are vitamin, protein, mineral, fibre and fat
User	A user that gets benefited from this ontology
Disease	Disease a user may have like diabetes, hypertension etc

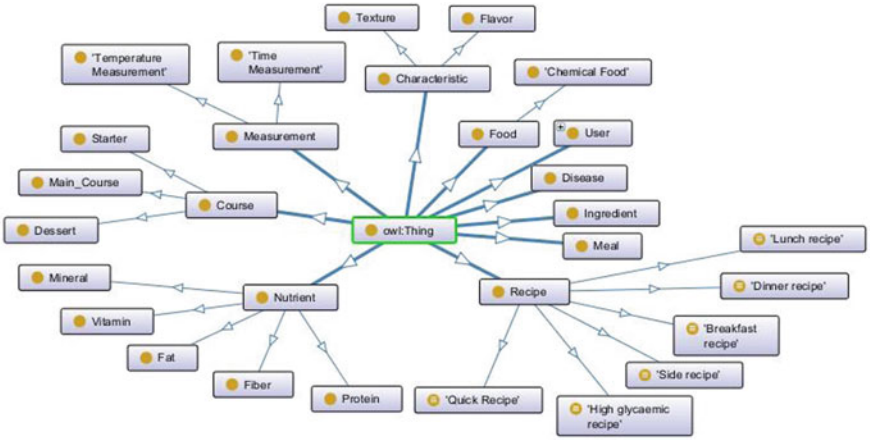


Fig. 73.2 A snapshot of classes of food ontology created using protege

These classes are created using Protégé [3]. Ontology structure is depicted in Fig. 73.2. There are two kinds of attributes in an ontology: Object properties and Data properties.

- **Object properties** connect pair of objects. The object properties considered for modelling are given below.
  - $substitute\_of$: Information regarding the substitute of a particular ingredient in case of dislike, not suitable for a particular user or unavailability of that ingredient.
  - $has\_ingredient$: Connects a Recipe to Ingredient, indicating the ingredient used in a Recipe.
  - $has\_disease$: Connects a user instance to a disease instance, depicting the information of a particular user having a disease.
  - $has\_nutrient$: Maintains the nutrient information of an ingredient.
  - has_flavor: Stores the flavour contained in a food item.
  - $dislikes$: Connects a user instance to ingredient instance, reflecting the information of a user disliking an ingredient.
  - $is\_allergic\_to$: Maintains the data of a user allergic to a particular ingredient.
  - $is\_recommended\_for\_course$: This property is used to recommend a course to a particular user.
  - $is\_recommended\_for\_meal$ A particular recipe is recommended for a meal.
- **Data properties** connect objects with literals. These are the functional data properties of objects, also called as attributes. Some of the selected data properties taken in this model are given below.

- *hasBMI*: Attribute of a user instance that indicates the Body Mass Index of a user.
- *hasLowActivityLevel* Indicates true, if a user has low-physical activity level otherwise false.
- *hasAge* Indicates the age of a user.
- *hasWeight* Indicates the weight of a user
- *hasCookTime* Property of a recipe instance that indicates the time needed to cook that recipe.
- *hasGluten* Property of ingredient instance. Indicates whether an ingredient contains gluten.
- *hasGlycaemicIndex* Ingredient Integer Indicates a glycaemic index value that measures carbs to determine how they affect blood sugar levels after meals.

In protégé, the objects/instances are called as individuals. In this ontology, all together 75 instances of ingredient, 6 instances of Course, 5 instances of Meal, 22 instances of recipe, 2 instances of disease, 5 instances of vitamin, 2 instances of protein, 5 instances of mineral, 2 instances of fibre, 4 instances of Fat, 4 instances of Chemical Food, 7 flavours, 7 textures, 2 temperature measurements, 2 time measurements and three users are created. The individuals of selected classes are given below.

- Class: **Course**
  Instances: Appetizer, dessert, entree, salad, Side and soup.
  Object Properties: Nil
  Data Properties: Nil.
- Class: **Recipe**
  Instance: Almond biscotti
  Object Properties: *has_ingredients*: chiken_egg, butter, almond, olive oil, all_purpose_floor, white sugar, salt and baking_powder.
  is_recommended_for_course: dessert; is_recommended_for_meal: snack
  Data properties: *hasCookTime*: 120; *hasCookingTemperature*: 350
- Class: **User**
  Instance:Raj
  Object Properties: *has_disease*: hypertension; *dislikes*: beef; *is_allergic_to*: cheese
  Data Properties: *hasAge*:35; *hasLowActivityLevel*: true; *hasWeight*:90; *has BMI*: 32
- Class: **Ingredient**
  Instance: Almond
  Object Properties: *has_nutrient*: plant_protein, saturated_fats; *substitute_of*: walnut, pumpkin_seed, pecan; *has_texture*: crunchy
  Data Properties: *hasGlycaemicIndex*:15; *hasGluten*:false

73.3 Recommendation Using Food Ontology

The usefulness of any ontology is proved by the queries it answers. The following are a few queries that can be used to recommend food related information to the users of food ontology using SPARQL.

- Which recipes use Salt, Chiken and Onion?
  PREFIX food: <http://example.org/food/>
  PREFIX ingredient: <http://example.org/food/>
  SELECT DISTINCT ?recipe
  WHERE {
  ?recipe food:hasIngredient ingredient:Salt .
  ?recipe food:hasIngredient ingredient:Chicken .
  ?recipe food:hasIngredient ingredient:Onion .}
- What are the foods that has banana, gluten and do not contain walnut and.
  PREFIX food: <http://example.org/food/>
  PREFIX ingredient: <http://example.org/food/>
  SELECT DISTINCT ?recipe
  WHERE {
  ?recipe food:hasIngredient ingredient:Banana .
  FILTER NOT EXISTS
  ?recipe food:hasIngredient ingredient:Walnut .
  ?recipe food:hasGluten false. }
- What chiken recipes with low sugar content?
  This kind of nutritional-based recommendations may be ideal for a diabetic user. In the ontology, information about sugar is represented using glycaemic index.
  PREFIX food: <http://example.org/food/>
  PREFIX ingredient: <http://example.org/food/>
  SELECT DISTINCT ?recipe
  WHERE {
  ?recipe food:hasIngredient ingredient:Chicken .
  FILTER NOT EXISTS
  ?recipe food:hasIngredient ?ingredient .
  ?ingredient food:hasGlycemicIndex ?GI .
  FILTER (?GI ≥ 50)}
- What are the foods containing basil and takes less than 20 min of time?
  PREFIX food: <http://example.org/food/>
  PREFIX ingredient: <http://example.org/food/>
  SELECT DISTINCT ?dinner ?CT
  WHERE {
  ?dinner food:hasIngredient ingredient:Basil .
  ?ingredient food:hasCookTime ?CT .
  FILTER (?CT <= 20) }

73.4 Other Food Ontologies in Literature

There are a few works in the literature on food ontology. Many of the researchers have published their ontology on web. Table 73.2 shows the food ontologies available in the literature and their contributions.

Table 73.2 Literature summary

Citation	Contribution
Esnaola et al. [4]	PFEEPSA ontology (Poultry Farm Energy Efficiency Prediction Semantic Assistant)
Haussmann et al. [5]	Developed a knowledge graph for food domain and provided recommendation to users
Zulaika et al. [6]	Person, Diet, Ingredient. Instances: Cheese, ice cream, yoghourt
Griffiths et al. [7]	20 food ontologies (wine, cheese, ricotta, oil, meat, fish, honey, essential oil, etc.)
BioPortal Food Ontology [8]	Semantics for food safety, food security, agricultural and animal husbandry practices related to food production, culinary, nutritional and chemical ingredients and processes
BBC Food [9]	Published data about recipes, foods they are made from, diets, menus, seasons, courses and occasions they may be suitable for
Gyrard et al. [10]	Concepts: Food, Emotion, Disease, Nutrients, Recipe, Color, Season, Weather, Diets, Allergies
Pizzuti and Mirabelli [11] and Pizzuti et al. [12]	An ontology oriented to the domain of food traceability. Concepts covered: Food, Beverage, Food additive, Cultivator, Farmer
Tumnark et al. [13]	Concepts: Athlete (height, weight, age), Food, Process type (baked, roasted, smoked), food menu (dish, dessert snacks or beverages), type of meals (dinner, lunch, breakfast, training session), Nutrition, Sport
Espin et al. [14, 15]	Concepts: Fish, Fruit, vegetable, animal fat, vegetable fat, beverage, canned food, nut, snack, per calories, proteins, daily recommended, alcoholic content isRichIn, proteinLvel, caloricLevel, recommendedRationsAWeek
Shakir et al. [16]	Ontology for suggesting information of nutrition food for a traveller. Concepts: Dishes (Breakfast, Lunch, Snack-Time, Dinner), Person, MedicalCondition (Diabetes, Healthy, HeartDisease, Obesity, ScurvyPatient = vitamin c deficiency), hasNutrition, hasCalories, hasCarbohydrates, hasFats, hasProteins, vitamins
Chi et al. [17]	A chronic kidney disease dietary consultation system based on semantic information

73.5 Conclusion

In this work, a food related ontology is modelled. The developed ontology contains a total of 29 classes, 10 object properties and 9 data properties with more than 100 individuals. The efficacy of the ontology is tested using a few SPARQL queries. The food ontology model could guide the users with useful recommendations. In the future, we would like to enhance the quality of recommendation using advanced semantic similarity techniques.

References

1. Bailoni, T., et al.: Healthy lifestyle support: the PerKApp ontology. In: OWL: Experiences and Directions—Reasoner Evaluation, pp. 15–23. Springer, Cham (2016)
2. Haussmann, S., Seneviratne, O., Chen, Y., Ne'eman, Y., Codella, J., Chen, C. H., Zaki, M. J.: FoodKG: a semantics-driven knowledge graph for food recommendation. In: International Semantic Web Conference, pp. 146–162. Springer, Cham (2019)
3. https://protege.stanford.edu/
4. Esnaola-Gonzalez, I., Gómez-Omella, M., Ferreiro, S., Fernandez, I., Lázaro, I., García, E.: An IoT platform towards the enhancement of poultry production chains. Sensors **20**(6), 1549 (2020)
5. Esnaola-Gonzalez, I., Fernandez, I., García, E., Ferreiro, S., Gomez, M., Lázaro, I., Garca, A.: Towards animal welfare in poultry farms through semantic technologies. In: Proceedings of the IoT Connected World Semantic Interoperability Workshop (IoT-CWSI), pp. 22–25. Bilbao, Spain (2019)
6. Zulaika, U., Gutiérrez, A., López-de-Ipiña, D.: Enhancing profile and context aware relevant food search through knowledge graphs. Multidiscip. Digit. Publ. Inst. Proc. **2**(19), 1228 (2018)
7. Griffiths, E.J., Dooley, D.M., Buttigieg, P.L., Hoehndorf, R., Brinkman, F.S., Hsiao, W.W.: FoodON: a global farm-to-fork food ontology. In: ICBO/BioCreative, pp. 1–2 (2016)
8. http://bioportal.bioontology.org/ontologies/FOODON
9. https://www.bbc.co.uk/ontologies/fo
10. Gyrard, A., Bonnet, C., Boudaoud, K.: Enrich machine-to-machine data with semantic web technologies for cross-domain applications. In: 2014 IEEE World Forum on Internet of Things (WF-IoT). IEEE (2014)
11. Pizzuti, T., Mirabelli, G.: FTTO: an example of food ontology for traceability purpose. In: 2013 IEEE 7th International Conference on Intelligent Data Acquisition and Advanced Computing Systems (IDAACS), vol. 1. IEEE (2013)
12. Pizzuti, T., et al.: Food track & trace ontology for helping the food traceability control. J. Food Eng. **120**, 17–30 (2014)
13. Tumnark, P., et al.: Ontology-based personalized dietary recommendation for weightlifting (2013)
14. Espín, V., et al.: Semantic-based recommendation of nutrition diets for the elderly from agroalimentary thesauri. In: International Conference on Flexible Query Answering Systems. Springer, Berlin, Heidelberg (2013)
15. Espín, V., Hurtado, M.V., Noguera, M.: Nutrition for Elder care: a nutritional semantic recommender system for the elderly. Exp. Syst. **33**(2), 201–210 (2016)
16. Karim, S., Shaikh, U.U., Rajput, Q.: Ontology-based personalized dietary recommendation for travelers. In: Southern Association for Information Systems Conference (SAIS) (2015)

17. Chi, Y.-L., Chen, T.-Y., Tsai, W.-T.: A chronic disease dietary consultation system using OWL-based ontologies and semantic rules. J. Biomed. Inform., 208–219 (2015)

Chapter 74
A Survey of Single Image De-raining in 2020

Hasal Fernando, Mohamed Ayoob, and Guhanathan Poravi

Abstract Impact of rain streaks is often considered a disturbance which causes performance reductions in many computer vision systems. To get rid of this problem, various rain removal systems have been introduced. The process of removing rain is called "de-raining". There are two main categories of de-raining. Those are single image de-raining and video de-raining. Performing single image de-raining is more challenging than video de-raining because it has no temporal information to be used in the process. This paper undertakes a comparative analysis of existing novel and credible single image de-raining works, and it is followed by an empirical study defining research gaps and possible future works.

Keywords Single image de-raining · Video de-raining · Rain removal · Computer vision · Image synthesis · Image-to-image translation

74.1 Introduction

Many works focused on de-raining have taken place during the past decades as it is a challenging task to remove rain streaks due to them being non-uniform and having diverse shapes and directions [1]. Chen [2] defines three types of rain streaks. They are occluding, veiling and accumulated rain streaks. Emergence of rain in images and videos differs due to the distance from the capturing device as well [3]. That describes how different the characteristics of rain streaks could be due to various reasons. Rain removal has two aspects to it [4], namely video de-raining [5–7] and single image de-raining. Single image de-raining is greatly challenging than video

H. Fernando (✉) · M. Ayoob
University of Westminster, London, UK
e-mail: w1697758@my.westminster.ac.uk

M. Ayoob
e-mail: w1654551@my.westminster.ac.uk

G. Poravi
Informatics Institute of Technology, Colombo, Sri Lanka
e-mail: guhanathan.p@iit.ac.lk

T. Senjyu et al. (eds.), *IOT with Smart Systems*, Smart Innovation,
Systems and Technologies 251, https://doi.org/10.1007/978-981-16-3945-6_75

de-raining [8] as it has no way of recovering the lost pixels due to rain streaks, using interframe information [4]. It is possible to use single image de-raining techniques in video de-raining as well to remove rain streaks in videos, frame by frame. That makes single image de-raining more challenging and useful out of the two.

Barnum [5] is treating the problem of video de-raining by detecting rain streaks through analysing the spatio-temporal frequencies. Liu [6] treats the video de-raining problem using deep recurrent neural networks based on special texture appearances while Wang [7] uses a tensor optimization model and group sparsity to address the problem. Ren [9] modelled different types of rain streaks in videos separately, using a matrix decomposition framework.

In the case of single image de-raining, traditional techniques such as dictionary learning [10], discriminative sparse coding [11] and layer priors [4] were based on mathematical models in the past.

With the advent of deep learning, deep neural network models were experimented for single image de-raining. Fu [12] introduced a deep detail network based on deep residual learning [13] which simplified the process. It focused on high-frequency details of the image which were mapped from input to output, in order to remove rain streaks. Subsequently, there have been many other deep learning-based approaches for single image de-raining.

As technology moves forward, computer vision applications are considerably employed in both the industry and academia. Rain is often regarded as a form of interference which causes degraded visibility. According to [8, 12, 14], rain streaks are known to affect the vision and performance of many computer vision applications such as self-driving cars and drones, image/video coding, human action recognition/detection, object detection/tracking, real-time data analytics and person re-identification to name a few. Most of these systems are trained to perform well under controlled clear weather conditions [15]. A successful de-raining system can make these applications much useful in rain weather conditions.

This paper gives out a comprehensive comparison of existing single image de-raining models. Section 74.2 of this paper brings out a comparison between existing important single image de-raining systems. Section 74.3 of the paper leads up to the existing research gaps while Sect. 74.4 brings up the limitations of the survey. Section 74.5 describes possible future works, while Sect. 74.6 concludes the complete survey.

74.2 Single Image De-raining

74.2.1 Single Image De-raining Variants

The following is a broad analysis of popular single image de-raining variants. Each model is partitioned individually to begin with, and later, a comparison table is built up to compare techniques, improvements and future works of all traditional and deep learning models. The models were chosen based on their credibility, recency and performance. The performance was measured based on the peak signal-to-noise ratio (PSNR) and structural similarity index measure (SSIM) of the models for popular data sets such as Rain100L and Rain100H. Both PSNR and SSIM are image quality metrics which are used to measure the difference of pixel representations between the de-rained image and the ground truth image.

Rain Streak Removal Using Layer Priors. Li [4] follows a traditional approach based on the assumption that de-raining can be worked out as a layer decomposition problem. It uses simple patch-based priors on both rain streak and background layers, which are formed on patches learned using Gaussian mixture models (GMMs) which proved to be more effective than methods which were based on low-rank constraints and dictionary learning. To model background patch priors, GMMs trained on natural image patches were used with an additional gradient sparsity constraint employed to regularize the background. Same approach was used to model rain patch priors as well. This had better results with reduced over de-rained effects compared to other traditional approaches.

Removing Rain from Single Images via a Deep Detail Network. Fu [12] introduced a deep CNN architecture hinged on residual learning [13]. To make the learning procedure much easier, this reduced the mapping array from input image detail layer to output by focusing on high-frequency details, which set the focus on high-frequency rain streaks while it managed to effectively avoid background intrusion. It used a lossless "negative residual mapping" to recognize the difference between rain and rain-free images. This work contributed a paired synthetic data set (Rain1400) consisting of 14,000 images.

Deep Joint Rain Detection and Removal from a Single Image. Yang [16] introduced a novel multitask deep learning architecture which added a binary map consisting of rain streak and background layers, to an existing region-dependent de-rain model. A contextualized dilated network was introduced to extract and refine features of the image, observing the contextual information which improved the accuracy of rain detection. The network was capable of handling both rain streak accumulation (occurrence optically close to fog or mist) and differing shapes and directions of overlapping rain streaks.

Density-Aware Single Image De-raining using a Multi-stream Dense Network. Noticing the limitations of [12] and [16] in using a single network to perform both rain detection and removal, Zhang and Patel [8] introduced a densely connected CNN-based algorithm. It was based on the assumption of using residual as an improved feature representation to better characterize rain density information. This is aware and capable of automatically determining the rain density using the novel residual-aware classifier which was built separately. With the separate classifier, it was effective in getting rid of over and under de-rained effects in outputs. They also put out a rain image data set (Rain1200) containing 12,000 training images with rain streak feature labels.

Image De-raining Using a Conditional GAN. Zhang [15] followed a new approach, exploring the effectiveness of conditional GAN [17] for image de-raining. Architectural originalities with a novel loss function bound for reducing artefacts in generated outputs were introduced in this approach. This network consisted of a densely connected symmetric generator sub-network and a multi-scale discriminator network influenced by both local and global information. This method not only concentrated on the characterization of rain streaks but also focused on maintaining the visual quality in the optimization function. Hence, it got rid of any additional image post-processing to enhance the image as well.

Single Image De-raining via Decorrelating the Rain Streaks and Background Scene in Gradient Domain. Based on the assumptions that (1) there is a rain-free gradient in which the image is least affected by rain, (2) rain streaks in an image or a particular block of an image consist of similar patterns and (3) each pixel of a rainy image is representing either rain or background, guiding to believing a low possibility of a correlation between rain and background, Du and Liu [18] built a decomposition framework. Combining total variation and low-rank constraint, another L1 norm decorrelation term was utilized in illustrating the proper decorrelation between the rain and background gradients.

Multi-Scale Progressive Fusion Network for Single Image De-raining. To uncover correlations of rain streaks, Jiang [3] introduced a multi-scale progressive fusion network which used a pyramid representation. To characterize rain streaks of multiple scales, the network was devised to three main modules. A coarse fusion module (CFM) and a fine fusion module (FFM) were used to successfully extract multi-scale information of rain streaks with varying scales in a cooperative approach while a reconstruction module (RM) was used to construct an image identical to the rain-free image. Information from all three modules is collaboratively used to characterize rain streak distributions. Due to that, predicted rain streak distributions from the network were proven to be more correct, and it performed well in experimental real-world scenarios as well.

Conditional Variational Image De-raining. Du and Xu [19] constructed an image de-raining network using the generative ability of the conditional variational auto-encoder (CVAE), which consisted of a spatial density estimation module which created a rain density map for each rainy input and used a channel wise de-raining scheme to analyse varied rain densities among different colour channels. This network was built to produce multiple corresponding de-rained output forecasts to obtain a better performance.

Syn2Real Transfer Learning for Image De-raining using Gaussian Processes. Yasarla [20] introduced a semi-supervised transfer learning approach to do single image de-raining with the use of Gaussian processes to model transitional latent spaces. That created a pseudo-ground truth image to unlabelled inputs. This used a nonparametric method to use unlabelled data for training. It used "l1" and "perceptual" supervised loss functions when training on labelled data. With the help of that, generated pseudo-ground truth was used to supervise unlabelled samples.

AI-GAN for Single Image Rain Removal. Jin et al. [1] brought in feature-wise disentanglement, a novel and distinct perspective to single image de-raining, using a general, contemporary mathematical model. This disentangled the rain image into two latent spaces (rainy latent space and background latent space) and introduced learnable parameters to describe them. This was built on interactions and constraints between the two latent spaces using a GAN framework which follows a two-branch architecture. It consisted of two generators asynchronously communicating with each other while generating the de-rained image. This follows a synthesis strategy in two phases, communicating feed forward information and feedback gradients (Table 74.1).

An Overview of Single Image De-raining Variants. Existing novel literature about the domain of single image de-raining describes fewer traditional approaches, while there are many deep learning approaches addressing different issues. It is common to see deep learning approaches perform better than traditional approaches due to their better learning ability. All the works are using paired images of rain and background data for training, and none of those have the capability to use unpaired images (rain images and background images with no links to each other) for training.

74.3 Exploring Research Gaps

Considering the traditional systems, systems based on mathematical models such as the decomposition model [18] introduced based on a low-rank constraint, a decorrelation term and a total variation used less computational power, but it had problems such as residual rain and blocking artefacts in outputs. Instead of using dictionary learning and low-rank constraints, using patch-based modelling [4] proved to be better for the purpose of de-raining, but this model giving less attention to the density of rain streaks and to different shapes of rain streaks resulted in some rain streaks being left out in the de-rained outputs.

When it comes to deep learning systems, Fu [12] introduced a de-raining system using a deep network architecture, which was based on residual learning. It reduced the mapping from an input to output making the training process much easier. This too suffered the problem of residual rain streaks in outputs. Yang [16] used a binary map to detect rain streaks, and it was used within a multitask deep learning framework to perform the de-raining process. Having a rain model which relies on regions, it was effective in removing both light and heavy rain, but using just a single network to fulfil both purposes of rain detection and removal came at a cost of over de-rained outputs and a loss of important background details. Recognizing the common limitation of both those works [12, 16], which was using a single network to both detect and remove rain streaks, Zhang [8] came up with a multi-stream dense network which used a separate classifier to classify the high, medium and light densities of rain streaks. This had much success than [12] and [16] in getting rid of over and under de-rained effects, but it came at a cost of requiring more paired images for training.

Table 74.1 Summary of single image de-raining variants

Variant	Technique	Improvements	Future works
Li [4] (2016)	Patch-based priors formed on GMM	Patch-based priors formed on GMM	1. Best possible size and placement of patches to approximate the GMMs 2. Handling of overloaded rain pixels 3. Handling the residual artefacts in background
Fu [12] (2017)	CNN architecture hinged on residual learning	1. Uses a deeper network structure which proves to be effective in image de-raining 2. Generalizes well to many other image processing tasks such as image denoising and artefact reduction 3. Contributes with 14,000 rainy and clean paired images as a synthetic data set	1. Not specified
Yang [16] (2017)	Multitask deep learning architecture which uses a binary map, contextualized dilated network and a recurrent rain detection and removal network	1. First method to use a contextualized dilated network in rain removal 2. First method to tackle heavy rain, using a recurrent detection and removal network	1. Not specified
Zhang and Patel [8] (2018)	CNN-based density-aware multi-stream dense network with a rain density classifier network	1. Contributes the first rain data set with rain density labels which consists of 12,000 training images and 1200 test images. 2. Gets rid of over and under de-rained effects quite well, with a separate rain density classifier network	1. Not specified
Du and Liu [18] (2018)	A decomposition framework which uses a L1 norm decorrelation term, combined with low-rank constraint and total variation	1. Effective in dealing with rain streaks of different sizes and patterns 2. Using Poisson's equation to reduce rain removal to reconstructing a de-rained image free of rain gradients	Improvements in, 1. Searching for a direction along the Y-axis, for least rain-affected gradient 2. Assessing X axis direction for a rain-free gradient
Zhang [15] (2019)	Generative adversarial network consisting of a densely connected symmetric generator sub-network and a multi-scale discriminator network	1. Improved stability in training 2. Reducing artefacts generated by GANs in de-rained outputs	1. Not specified

(continued)

Table 74.1 (continued)

Variant	Technique	Improvements	Future works
Jiang [3] (2020)	Fusion network with a CFM, FFM and a RM	1. First method which is applied in vision-related tasks to assess the de-rain performance 2. Predicted rain streak distribution generated through collaborative representation of rain streaks of various scales is high in accuracy, compared to other models	1. Not specified
Du and Xu [19] (2020)	Conditional variational auto-encoder (CVAE) framework with a SDE module and a CW scheme	1. First work to do rain removal using the generative ability of the CVAE framework 2. Increasing the performance by outputting multiple predictions	1. Not specified
Yasarla [20] (2020)	Gaussian process-based semi-supervised transfer learning framework	1. Successfully uses unlabelled data in training the network 2. Creates an intermediate latent space within the network using Gaussian processes to create a pseudo-ground truth image to unlabelled inputs. 3. Accomplishes on-par performance with unlabelled limited training data, to fully trained networks with labelled data.	1. Not specified
Jin [1] (2020)	Generative adversarial network using a two-branch architecture	1. Brings in feature-wise disentanglement to single image de-raining, with a novel de-rain model. 2. Has a strong generalization capability through several multimedia applications	1. Optimizing the performance of the framework. 2. Extending the framework to unsupervised learning, to enable the use of unpaired training data. 3. Simplifying and modularizing the structure to improve sturdiness of the network

This requirement for more paired image data is evident in GAN-based works [1] and [15] which produce high-quality outputs compared to other works.

Throughout all the works, residual rain streaks left in outputs (under de-rained), blurry outputs (over de-rained), residual artefacts and loss of background information are consistent problems when applied in real-world scenarios. Hence, those could be considered as gaps which exist within the domain. One of the main reasons for the problem of residual rain streaks in de-rained outputs is synthetic training data which are used to train models not containing real-life rain streak features. Training models on real-life paired training data are almost impossible due to the inability of collecting real paired data for the same scenario with and without rain, with the same lighting conditions and background details. Also, such paired data being expensive to collect is a renowned problem in the world of data science [21]. Due to this, extending single image de-raining to be done with the use of unpaired training data is considerably a significant gap left in the domain, which is clearly mentioned in the conclusion chapter of AI-GAN [1].

74.4 Limitations of the Survey

There are a few limitations to this survey which need to be addressed. In total, there are ten single image de-raining models which were surveyed. These works were chosen based on their novelty, performance, popularity and credibility.

Works on video rain removal using different approaches and techniques also have depth to it. Different types of rain streaks in scenes are the only one type of a disturbance caused in computer vision systems. There are many other disturbances such as rain drops on camera lens [22] and disturbances from snow [5] as well. There have been many other traditional approaches for the single image de-raining problem in the past.

Above-mentioned areas were not taken into high consideration throughout this survey since (1) the scope of the research is to survey about highly challenging single image de-raining problem and (2) deep learning-based solutions proving to be highly effective in solving the problem compared to traditional methods.

74.5 Future Works

The domain of single image de-raining is a very active domain, and it has been researched a lot to achieve the results it has had so far, but there is a lot more to be researched and improved, mostly in terms of utilizing natural rain data to learn natural rain streak features, which could lead to much better results in the domain. All works are using paired image data to train and test models. Manually synthesized data being used in training the models is a reason why most systems struggle in real-world scenarios, due to them being unable to detect some natural rain streak features.

Focus of the paper is to provide future researchers with a rapid understanding about novel and credible single image de-raining works and possible future work paths along which the works on domain can be extended.

A summary of future works that can be done on the domain would be (1) to improve available algorithms in order to get rid of over de-rained and under de-rained effects and residual artefacts, (2) to simplify and modularize structures of networks to improve efficiency in performance and (3) to extend single image de-raining to use unpaired training data which will improve natural rain streak detection in real-world scenarios.

74.6 Conclusion

Rain is a disturbance considered to be decreasing the effectiveness of many computer vision systems. Many researches have been carried out in the past in the domain of single image de-raining. They fall under two main categories which are video de-raining and single image de-raining. Video de-raining solutions tend to use temporal and contextual information in the process, which single image de-raining solutions are unable to use. Hence, single image de-raining is considered more challenging compared to video de-raining. There have been many traditional approaches tried for single image de-raining, and most of them are treating it as a signal removal problem. But, when deep learning approaches were tried out for image de-raining, those had more success compared to traditional approaches.

This paper thoroughly surveys ten single image de-raining variants that is followed by a comprehensive tabular analysis of techniques, limitations and future works mentioned for each work, which will be extremely useful to many future researchers of the domain.

Acknowledgements I am deeply grateful to my supervisor Mr. Guhanathan Poravi and co-supervisor Mr. Mohamed Ayoob, who were always willing to guide me in the right path, providing the much needed knowledge. Without their guidance, this research would never have been possible. Also, I would like to thank every person who supported me even by a word in making this research a success.

References

1. Jin, X., Chen, Z., Li, W.: AI-GAN: Asynchronous interactive generative adversarial network for single image rain removal. Patt. Recogn. **100**, 107143 (2020). https://doi.org/10.1016/j.patcog.2019.107143
2. Chen, J., Tan, C.-H., Hou, J., Chau, L.-P., Li, H.: Robust video content alignment and compensation for clear vision through the rain. arXiv:1804.09555 [cs], Apr 2018, Accessed 16 Oct 2020. [Online]. Available: http://arxiv.org/abs/1804.09555

3. Jiang, K., et al.: Multi-scale progressive fusion network for single image deraining, pp. 8346–8355, Accessed: 22 Oct 2020. [Online]. Available: https://openaccess.thecvf.com/content_CVPR_2020/html/Jiang_Multi-Scale_Progressive_Fusion_Network_for_Single_Image_Deraining_CVPR_2020_paper.html
4. Li, Y., Tan, R.T., Guo, X., Lu, J., Brown, M.S.: Rain streak removal using layer priors, pp. 2736–2744. Accessed: 25 Sept 2020. [Online]. Available: https://www.cv-foundation.org/openaccess/content_cvpr_2016/html/Li_Rain_Streak_Removal_CVPR_2016_paper.html
5. Barnum, P.C., Narasimhan, S., Kanade, T.: Analysis of rain and snow in frequency space. Int. J. Comput. Vis. **86**(2), 256 (2009). https://doi.org/10.1007/s11263-008-0200-2
6. Liu, J., Yang, W., Yang, S., Guo, Z.: D3R-Net: dynamic routing residue recurrent network for video rain removal. IEEE Trans. Image Process. **28**(2), 699–712 (2019). https://doi.org/10.1109/TIP.2018.2869722
7. Wang, Y.-T., et al.: A total variation and group sparsity based tensor optimization model for video rain streak removal. Signal Process. Image Commun. **73**, 96–108 (2019). https://doi.org/10.1016/j.image.2018.11.008
8. Zhang, H., Patel, V.M.: Density-aware single image de-raining using a multi-stream dense network, pp. 695–704 (2018). Accessed: 27 Sept 2020. [Online]. Available: https://openaccess.thecvf.com/content_cvpr_2018/html/Zhang_Density-Aware_Single_Image_CVPR_2018_paper.html
9. Ren, W., Tian, J., Han, Z., Chan, A., Tang, Y.: Video desnowing and deraining based on matrix decomposition, pp. 4210–4219 (2017). Accessed: 16 Oct 2020. [Online]. Available: https://openaccess.thecvf.com/content_cvpr_2017/html/Ren_Video_Desnowing_and_CVPR_2017_paper.html
10. Kang, L.-W., Lin, C.-W., Fu, Y.-H.: Automatic single-image-based rain streaks removal via image decomposition. IEEE Trans. Image Proces. **21**(4), 1742–1755 (2012). https://doi.org/10.1109/TIP.2011.2179057
11. Luo, Y., Xu, Y., Ji, H.: Removing rain from a single image via discriminative sparse coding. In: 2015 IEEE International Conference on Computer Vision (ICCV), Santiago, Chile, pp 3397–3405, Dec 2015. https://doi.org/10.1109/ICCV.2015.388
12. Fu, X., Huang, J., Zeng, D., Huang, Y., Ding, X., Paisley, J.: Removing rain from single images via a deep detail network. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, pp. 1715–1723, July 2017. https://doi.org/10.1109/CVPR.2017.186
13. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. arXiv:1512.03385 [cs], Dec 2015. Accessed: 25 Sep 2020. [Online]. Available: http://arxiv.org/abs/1512.03385
14. Wei, W., Yi, L., Xie, Q., Zhao, Q., Meng, D., Xu, Z.: Should we encode rain streaks in video as deterministic or stochastic?, pp. 2516–2525 (2017). Accessed: 20 Nov 2020. [Online]. Available: https://openaccess.thecvf.com/content_iccv_2017/html/Wei_Should_We_Encode_ICCV_2017_paper.html
15. Zhang, H., Sindagi, V., Patel, V.M.: Image de-raining using a conditional generative adversarial Network," arXiv:1701.05957 [cs], Jun. 2019, Accessed: Oct. 11, 2020. [Online]. Available: http://arxiv.org/abs/1701.05957
16. Yang, W., Tan, R.T., Feng, J., Liu, J., Guo, Z., Yan, S.: Deep joint rain detection and removal from a single image, pp. 1357–1366 (2017). Accessed: 27 Sept 2020. [Online]. Available: https://openaccess.thecvf.com/content_cvpr_2017/html/Yang_Deep_Joint_Rain_CVPR_2017_paper.html
17. Isola, P., Zhu, J.-Y., Zhou, T., Efros, A.A.: Image-to-image translation with conditional adversarial networks. arXiv:1611.07004 [cs], Nov 2018. Accessed: 23 Sept 2020. [Online]. Available: http://arxiv.org/abs/1611.07004
18. Du, S., Liu, Y., Ye, M., Xu, Z., Li, J., Liu, J.: Single image deraining via decorrelating the rain streaks and background scene in gradient domain. Patt. Recogn. **79**, 303–317 (2018). https://doi.org/10.1016/j.patcog.2018.02.016

19. Du, Y.-J., Xu, J., Zhen, X.-T., Cheng, M.-M., Shao, L.: Conditional variational image deraining. IEEE Trans. Image Process. **29**, 6288–6301 (2020). https://doi.org/10.1109/TIP.2020.2990606
20. Yasarla, R., Sindagi, V.A., Patel, V.M.: Syn2Real transfer learning for image deraining using Gaussian processes. arXiv:2006.05580 [cs], June 2020. Accessed: 22 Oct 2020. [Online]. Available: http://arxiv.org/abs/2006.05580
21. Choi, Y., Choi, M., Kim, M., Ha, J.-W., Kim, S., Choo, J.: StarGAN: unified generative adversarial networks for multi-domain image-to-image translation. arXiv:1711.09020 [cs], Sept 2018, Accessed: 23 Sept 2020. [Online]. Available: http://arxiv.org/abs/1711.09020
22. Qian, R., Tan, R.T., Yang, W., Su, J., Liu, J.: Attentive generative adversarial network for raindrop removal from a single image. arXiv:1711.10098 [cs], May 2018. Accessed: 02 Dec 2020. [Online]. Available: http://arxiv.org/abs/1711.10098

Chapter 75
Wi-Fi and LTE Coexistence in Unlicensed Spectrum

V. Kiran, Prajwal S. Telkar, Deekshith Nayak, and D. N. Rahul Raj

Abstract Comprehensive growth of the internet and supporting technologies has increased the competition in the technological market to provide highly reliable services. With improved services with exceptional features in new applications, it's driving the need for high bandwidth in the licensed spectrum. But because of the increase in need, licensed spectrum has been exhausted by the usage and bandwidth for service is decreasing. To mitigate this problem and to upgrade the current cellular infrastructure to support high bandwidth needs and to serve reliable services, researchers are looking at usage of unlicensed spectrum. Utilization of unlicensed spectrum in the past has been very minimal compared to the total bandwidth available. Hence, using unlicensed spectrum will be a promising option to support ever increasing bandwidth needs. Deploying the data from licensed band into unlicensed band requires careful observation of unlicensed band to make sure it supports all of its native protocols such as Wi-Fi. LTE-U and LTE-LAA have emerged as the technology of choice to deal with the need. Both the technologies follow different approaches to serve the licensed spectrum data through the unlicensed spectrum without harming its normal functionality. The problem associated with coexistence of Wi-Fi with LTE, and their deployments in the 5 GHz band has been illustrated. The paper is concluded by depicting the flow of the LBT design structure depicting functionality in different layers to achieve coexistence of Wi-Fi and LTE in unlicensed band.

75.1 Introduction

The opportunity to communicate remotely outside the licensed band pulled in a wide assortment of expected clients. In 2017, the greater part of the world's IP traffic is extended Wi-Fi, utilizing in excess of nine billion Wi-Fi gadgets, with in excess of three billion new gadgets sold every year. Wi-Fi at first was centered around interfacing workstations to the Internet. It is presently utilized for everything

V. Kiran (✉) · P. S. Telkar · D. Nayak · D. N. R. Raj
R V College of Engineering, Bengaluru, India
e-mail: kiranv@rvce.edu.in

T. Senjyu et al. (eds.), *IOT with Smart Systems*, Smart Innovation, Systems and Technologies 251, https://doi.org/10.1007/978-981-16-3945-6_76

from modern machines to PCs to telephones to remote media administrations. One potential barrier for licensed spectrum users is that they need to get consent from FCC to begin their new services and to perform testing for their deployment. But no such mandate exists for unlicensed band users.

The increased throughput, system capacity and efficiency have made it possible for the LTE to gather a wide range of data subscribers all over the globe since it emerged. Due to the exponential growth in the number of data users over the past years, there is a need to accommodate all these users while still maintaining the required capacity and efficiency. The existing licensed spectrum (1.4, 3, 5, 10, 15, 20 MHz) is not able to serve the increasing users as it is limited. So the use of unlicensed spectrum and aggregating unlicensed carriers with licensed carriers makes it possible to accommodate a wide range of users. But the unlicensed spectrum already holds up several users of wireless technologies such as Wi-Fi, Bluetooth, radar networks. Unlicensed spectrum has become a medium for connecting remote/rural users at a considerably least cost. And it also serves as a home for innovation, for testing and implementing new technologies thus acting as a cost-effective platform. This deployment creates several coexistence problems with the already prevalent Wi-Fi users in the 5 GHz band. LTE without any protocol modifications can transmit without worrying about the conditions of surrounding Wi-Fi users. Hence, there exists a need to deploy suitable protocol changes to LTE in order to ensure fair and peaceful coexistence with Wi-Fi users.

The fundamental idea of using unlicensed band is to increase spectral efficiency and not to dismount the Wi-Fi users from that band. Therefore, several variants of LTE have been proposed as the solution to the above problem.

75.2 Background

LTE technology has a centrally scheduled architecture in which channel access is scheduled by base station in licensed bands. The industry consortium LTE-U forum proposed the LTE-U technology that particularly focused on the duty cycled transmission. Adjustments on duty cycle by control elements of LTE-MAC or by utilizing dynamic duty-cycle mechanism can further improve the performance. Third 3GPP Release 13 proposed the novel LTE-LAA technique which makes use of the concept of carrier aggregation where the controlling and certain level of data would be transmitted through primary licensed carrier, while remaining data would be through secondary unlicensed carrier. Here, the primary carrier would always be ON, while secondary would be deployed depending on the availability needs.

75.2.1 Related Work

The use of unlicensed spectrum by LTE in order to satisfy the demand of hike in number of users and to increase the system capacity further aims to enhance the user experience. The authors in [1] implemented LTE-U channel selection and CSAT on a prototype platform developed by Qualcomm. In LTE-U, three mechanisms were proposed to ensure peaceful coexistence. [2] shows significant improvement in throughput for different deployment scenarios. Duty-cycle mechanism for LTE-U was proposed in [3, 4], in which a suitable transmission duration and probability are chosen to access the channel thereby ensuring proportional fairness among Wi-Fi and LTE-U entities. Here, all the network nodes are provided with equal channel times and tested for a number of successful transmissions and collisions, if any. Performance evaluation of the aforementioned coexistence situation is done by NS3 simulator in [5]. Wi-Fi stations and LTE nodes are modeled in a typical indoor scenario with the Additive White Gaussian Noise (AWGN) error modeling. In [6, 7], a mathematical approach was followed for efficient evaluation of performance and to ensure transparent comparisons of different deployment techniques. A stochastic geometrical [8] based framework to evaluate the coexistence scenarios unlike the conventional simulation techniques which were time consuming due to tangled dynamic nature of the overlaid Wi-Fi and LTE networks are proposed. The operation of LTE-LAA is explored with emphasis on CCA in [10, 11]. This leaves the system with improved coverage and ensures mobility management. Coexistence of two LTE-U networks utilizing the same unlicensed resources [13] is described to exist in presence of an interface that coordinates and ensures fair competition among the nodes thus utilizing the spectrum efficiently. As proposed by European Telecommunication Standards Institute (ETSI), two mechanisms of LBT, frame based, load based equipment are discussed. In frame based equipment (FBE) based LBT mechanism, in-device coexistence within the same user equipment and solution for the same is discussed, i.e., by appropriately managing TDM or notifying the LTE (eNB) by discontinuous reception cycles or by completely detaching one over the other. In paper [14], NS3 is used to demonstrate different scenarios. NS3 provides a provision to model Wi-Fi net devices and LTE net devices with realistic error and delay models. The LTE module is aimed to evaluate performance parameters such as throughput, latency and different radio resource management, and scheduling algorithms are studied. If the communication environment is chosen such that not all Wi-Fi access-points/ stations are within the operable service area of the eNB station, then such nodes are unhindered by LTE transmission. The authors in paper [15] have considered hidden terminal scenario. Non-victim users which are in the outer coverage region attempt to transmit packets to the victim users in the LTE-U ON period, but victim users cannot participate in communication as they should stand still during half the duty cycle in accordance with LTE-U protocol. Therefore, the non-victim users have to re-transmit the packets which leads to channel resource wastage. Conventional ways of dealing with hidden terminal problems by employing RTS/CTS frames does not hold good in this case as Wi-Fi and LTE are two entirely different technologies. It has been found that the average throughput of non-victim users is high compared to the victim users for all values of duty cycle because of disproportionate access to the medium and unfair treatment. Improvement of 35% (at = 0.5) was observed with the proposed scheme.

75.2.2 *Mechanisms for Wi-Fi and LTE Coexistence in 5 GHz Spectrum*

In some countries, LBT is not mandatory. Thus, different mechanisms could be adapted for coexistence of LTE and Wi-Fi. The LTE-U forum focused on three different approach to implement Wi-Fi and LTE in unlicensed band.

They are (1) Channel selection mechanisms (CHS). (2) Carrier sensing adaptive transmission (CSAT). (3) Opportunistic supplementary downlink (OSDL).

The flowchart in Fig. 75.1 depicts the flow of how the channel is selected based on the availability and the algorithms adopted to select the channel. The initial step is the channel selection that facilitates the small cells to adopt a clean channel stationed on the Wi-Fi and LTE analysis. Due to this interference between small cells and adjacent, Wi-Fi devices are restrained.

Thus, the channel selection algorithm establishes the quality of operating channel and selects most stable one. After this is a clean channel is not available then CSAT algorithm adopts a mechanism of TDM transmission which involves small cells of LTE-U, which is based on the analysis of perceiving the channel for long period co-channel actions of Wi-Fi. If the cell is loaded with lot of traffic, OSDL transmission is utilized such that secondary carrier can be toggled accordingly so that the transportation overhead like cell specific reference signals can be avoided. And

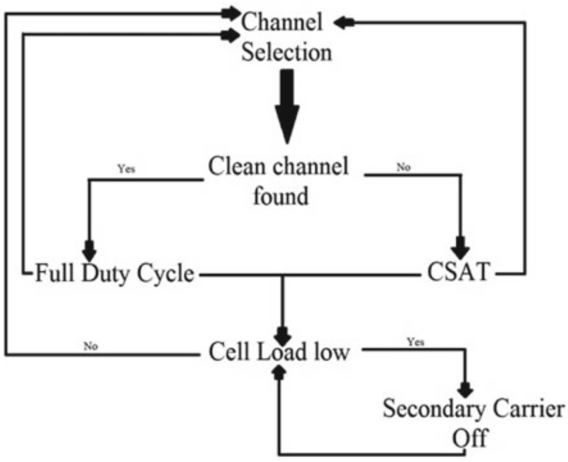


Fig. 75.1 Flowchart depicting LTE-U mechanism

thus reduces further interference to neighboring Wi-Fi access point (AP)s. This is achievable as the primary carrier always operates in a licensed band.

1. Channel selection mechanism [CHS]: Channel selection mechanism in LTE-U is used for continuous auditing of the supplement downlink transportation. This auditing when leads to a clean channel, secondary cell takes over the clean channel. If the interference is observed in this clean channel operation, a corresponding clean channel is searched, and the supplement downlink transportation is carried in the new available channel. If nil clean channels were available, then CSAT algorithm is adopted. Channel selection mechanism can be used where the traffic density is low such that clean channels are available.

2. Carrier Sensing Adaptive Transmission [CSAT]: The algorithm makes use of the concept of TDM, in which LTE-U can coexist the common channel with Wi-Fi, or different LTE-U. In the structure, the small cell always hears the channel for more period around 10–200 ms which would be greater than CSMA and LBT technique and stationed on the channel observation and analysis of the algorithm allows ("gating off") for transmission of LTE. The time cycle is a period in which small cell transfer the data in part of cycle and the remaining time span for transmission of LTE. The duty cycle of transmission small cell and LTE signals is done according to the sensing of channel activity of other technical knowledge. The LTE-U, on secondary cells, is turned-on and turned-off regularly employing LTE MAC control elements. At the time when the LTE-U is in turned-off period, the channel is free to access, and thus, the adjacent Wi-Fi can regain the regular Wi-Fi transportations. The small cell will analyze the Wi-Fi channel based on activities when LTE-U is turned-off and based on these results, it adjusts the On/Off duty cycle appropriately. Based on these activities, TDM cycle is set.

3. Opportunistic SDL (OSDL): The mechanism is used on the opportunistic base of the SDL carrier available in unlicensed band. When the downlink traffic in a small cell exceeds threshold, the SDL carrier is turned-on to off load. If the small cell has moderate traffic, then the secondary component is turned-off to reduce CRSs. This algorithm is suitable in areas where dense deployments, where no clean channel is available since it reduces the impact on co-channel communications.

75.3 Coexistence of Wi-Fi and LTE-LAA in 5 GHz Spectrum

In some countries, LBT is mandatory. In such situations, we use a method of licensed-assisted access. This method adopts the mechanisms of carrier aggregation. It combines the carriers from both licensed and unlicensed spectrum such that PCell on licensed bearer and SCell on unlicensed bearer. This increases the performance parameters. Licensed-assisted access is a technology such that the primary carrier is always turned-on and the secondary carrier can be turned-on or turned-off based according to the channel opportunity. Hence, some data and control signals are carried by both primary and secondary carriers.

Channel Access Mechanism: The Wi-Fi channel is accessed through DCF based on CSMA/CA access method. In this mechanism, the channel is sensed before each transmission, and announces busy if the sensed energy in the channel is above threshold. If it is less than threshold then the Wi-Fi device is permitted to transmit. But there is no such mechanism in LTE to identify collision and no LBT. Thus, due to difference in MAC layers of both Wi-Fi and LTE, coexistence becomes difficult, and thus, LTE to coexist in unlicensed band leads to performance degradation of Wi-Fi.

75.4 Assisted Access for LTE (LAA)

LAA is a protocol based technology used in LTE that makes use of unlicensed 5 GHz band in succession with licensed band to bring up better end user experience. This technology adds LTE in licensed spectrum and makes use of carrier aggregation techniques to bind LTE in unlicensed band to deliver higher speed in data transmission and enhanced end user experience.

Since Wi-Fi operates in the unlicensed 5 GHz band, where LAA deploys LTE, it must be able to exist side-by-side by avoiding channels which are being used for Wi-Fi. Sharing of channels between Wi-Fi and LTE can happen in friendly or fair manner, in order to achieve this, LAA uses a technique called LBT (listen before talk) also called listen before transmit. Bringing together Wi-Fi and LTE is a key breakthrough toward 5G. LAA boosts the performance for end users and it optimizes resources available in wireless networks for all users.

Listen before talk is an approach used in radio communications that sense the radio environment to analyze a clear channel before it starts transmitting. If there is no pleasant channel accessible, LAA will share a channel fairly with others. Radio devices use LBT to find either a network it is allowed to operate on or to search a clear channel. But the difficulty in the latter is the minimum signal strength it has to listen. Countries, such as Japan and Europe, have a regulation that LBT mechanisms have to be used necessarily. This scheme basically employs a sensing scheme similar to CSMA as shown in Fig. 75.2. With the LBT algorithm (Figs. 75.3 and 75.4).

75.5 LBT Algorithm

Step-1: The LBT algorithm has an initial CCA period and an extended CCA period. CCA refers to clear channel assessment which is the process of checking for a clear or idle channel available for transmission in the spectrum. In the initial CCA period, if the request for transmission exists then it checks for a clear channel in the spectrum. Step-2: If it found a channel which was idle for the initial CCA period (*T*

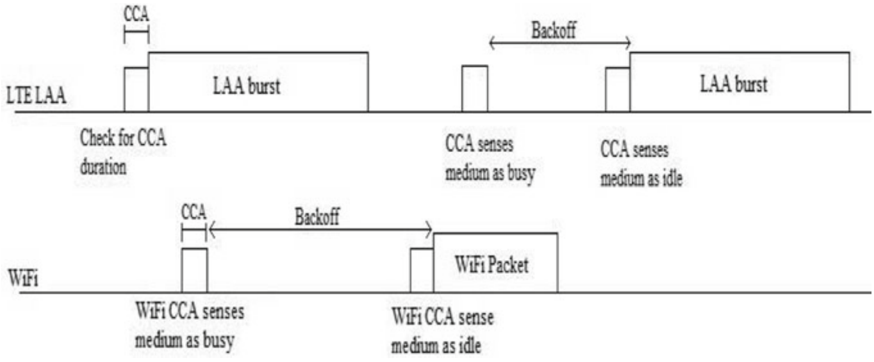


Fig. 75.2 LTE-LAA mechanism with channel assessment

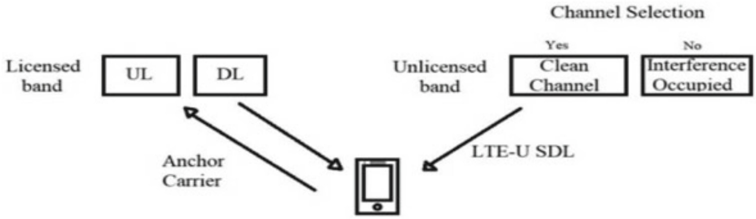


Fig. 75.3 Channel selection in LTE-U

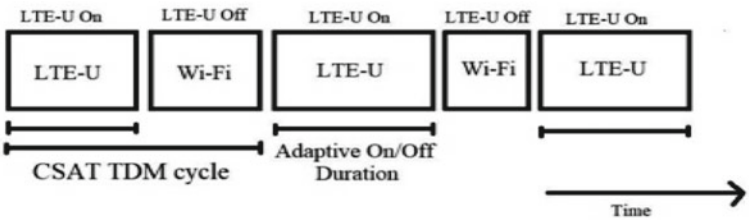


Fig. 75.4 Wi-Fi and LTE-U enabled by CSAT

init) then it declares that channel as clear and accepts the request for transmission. Algorithm goes back to step-1 once the transmission is completed. Step-3: If no clear channel was found or if the transmission could not be completed in the available idle channel then the algorithm runs into extended CAA period, which is step-4. If the transmission is completed, then the channel goes back to its idle state and the process continues from step-1. Step-4: In the extended CAA period, the algorithm selects a random value for N between 0 and $q - 1$. Value of q (contention window size) will be decided by the ACK (acknowledgment) received from HARQ (hybrid automatic repeat request) mechanism. Step-5: It waits for eCCA (extended CCA defer period) D to check if any channel is in idle state for that duration. Step-6: Once it finds a channel idle for D duration, then it starts decreasing the value of N and the algorithm again runs into step-5. This loop continues till the value of N equals 0. Step-7: If the channel goes out from idle state before N reaches 0 then the algorithm again encounters extended CCA period, which is step-3. Step-8: Transmitter and channel both enter into idle state once the transmission is completed.

75.5.1 LBT Design

LTE-RLC: There are 3 modes of operation for radio link control layer of LTE transparent mode, unacknowledged mode, and acknowledged mode. RLC layer is responsible for upper layer PDUs. It handles error correction in acknowledged mode. Its functions in unacknowledged and acknowledged mode include concatenation, segmentation and reassembly of RLU SDUs.

LTE-MAC: MAC stands for medium access control its responsible for mapping between logical and transport channels, multiplexing and demultiplexing of MAC SDUs, error correction through HARQ, implementing priority scheduling between different UEs and logical channels of the same UE. Error correction technique used in LTE is HARQ, which is the implementation of ARQ using high-rate forward error correction. In normal ARQ, the message will be padded with redundant bits using error detecting codes, and a new message will be transmitted on request of the receiver whenever it detects erroneous code. But in hybrid technique, the parity bits will be sent either along with the original message data or on detection of error by the receiver. These FEC codes are capable of error detection and correction, but this adds an expense of significantly lower throughput even in good signal SNR.

LTE-PHY: Physical layer carries physical signals into the spectrum shared by LTE and Wi-Fi. This layer takes care of adaptation, cell search, power control and other measurements required for transmission.

LBT Channel access manager: Access manager handles the access to channels in unlicensed band. It overhears the transmission ongoing in the common spectrum channel shared by both LTE and Wi-Fi. In LTE traffic, RLC layer interacts with MAC layer. The MAC layer is responsible for collision detection and avoidance. Hence, it interacts with LBT access manager to check for possible collision in the channel. Physical layer of LTE is responsible for transmitting the physical signal in the spectrum. Hence, it collaborates with LBT access manager to check whether any free channel is available for transmission (Figs. 75.5 and 75.6).

AP Station manager: AP Access manager handles all the APs (Access points) of the Wi-Fi consortium. It interacts with the MAC and physical layer to transmit physical signals in the spectrum from AP. It does not have any special collision detection algorithm as the MAC layer implements its own mechanism to avoid any

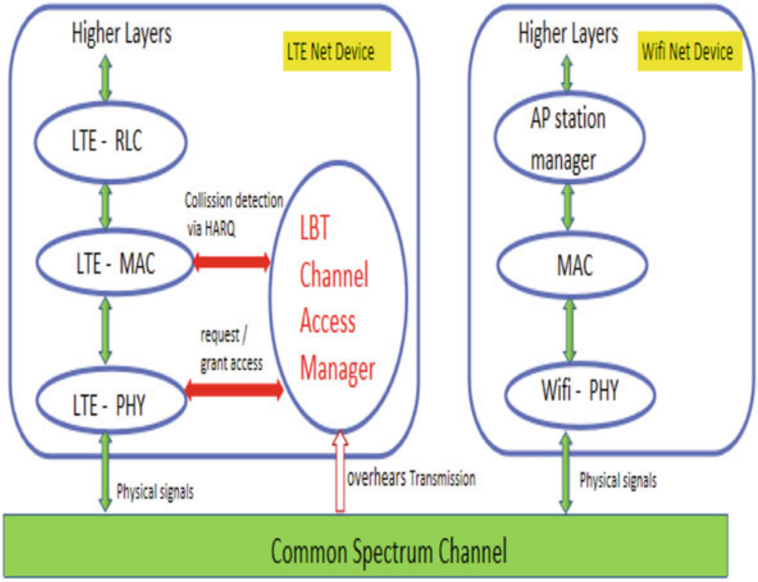


Fig. 75.5 Structure depicting functionality in different layers

interference from different APs. Wi-Fi has only an unlicensed band to operate upon, and hence it will be given higher priority over LTE traffic.

Wi-Fi-MAC: MAC layer coordinates access to the shared radio environment so that access point and Wi-Fi station in range can communicate. It handles multiple access by carrier sensing, channel contention and random back off. It uses DCF and PCF for collision avoidance.

Wi-Fi-PHY: This is the layer where data gets transmitted in terms of bits 1 and 0 s. This layer contains physical equipment of Wi-Fi. It transmits physical signals into the common spectrum shared by Wi-Fi and LTE.

75.6 Conclusion

Analysis for various scenarios for coexistence of Wi-Fi with LTE in unlicensed bands was explained. Acceleration in technology has facilitated more potent usage of spectrum as they have allowed for the concurrent use by multiple equipment and technologies without conflict or need of license. A detailed discussion was brought up on the rules that govern the deployment of LTE in unlicensed band and for different coexistence scenarios. Later, the paper concentrated on the existing mechanism for coexistence of LTE and Wi-Fi. Despite the fact that performance of Wi-Fi gets degraded considerably when LTE technology transmits data continuously without

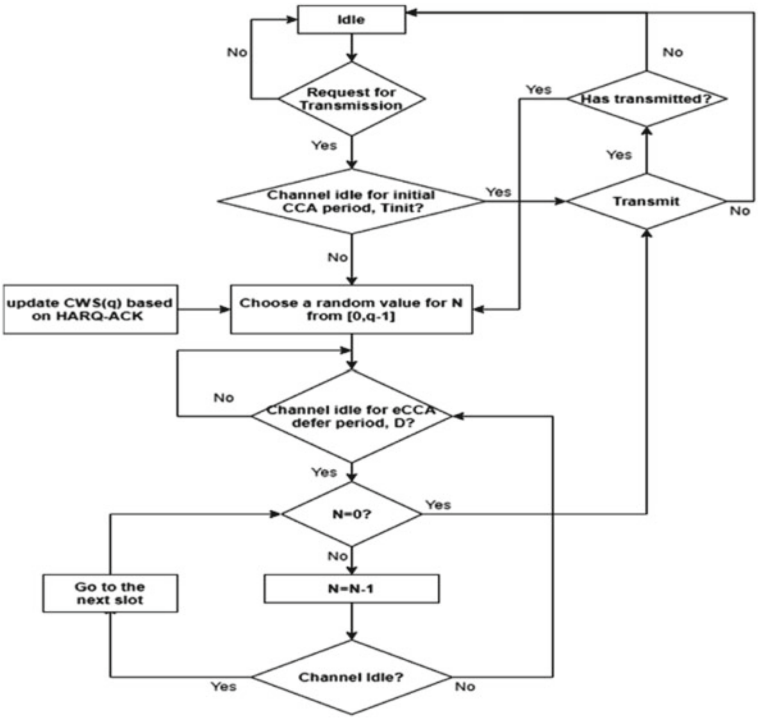


Fig. 75.6 Flowchart for LBT Algorithm

any changes to the protocol, the survey indicated that LTE can be made to be a good neighbor to Wi-Fi in unlicensed band by altering duty cycle for transmission, energy threshold for sensing, or priority for channel access of LTE. Different algorithms and approaches involved combating performance degradation and issues, while maintaining the throughput of users of both the technologies are explored. Spectrum de-licensing is a malleable way to spectrum management that encourages innovation and market evolution.

References

1. Qualcomm: LTE in Unlicensed Spectrum: Harmonious Coexistence with Wi-Fi. Qualcomm Technologies, Inc., Morehouse Drive, San Diego, CA, USA (2014)
2. Pei, E., Jiang, J.: Performance analysis of licensed-assisted access to unlicensed spectrum in LTE release 13. IEEE Trans. Veh. Technol. **68**(2), 1446–1458 (2019). https://doi.org/10.1109/TVT.2018.2886332
3. Cano, C., Leith, D.J.: Coexistence of WiFi and LTE in unlicensed bands: a proportional fair allocation scheme. In: 2015 IEEE International Conference on Communication Workshop (ICCW), 2015, pp. 2288–2293. https://doi.org/10.1109/ICCW.2015.7247522

4. Mukherjee, A., Cheng, J., Falahati, S., Falconetti, L., Furuskär, A., Godana, B., Kang, D.H., Koorapaty, H., Larsson, D., Yang, Y.: System architecture and coexistence evaluation of licensed-assisted access LTE with IEEE 802.11. In: 2015 IEEE International Conference on Communication Workshop (ICCW), 2015, pp. 2350–2355. https://doi.org/10.1109/ICCW.2015.7247532
5. Giupponi, L., Henderson, T., Bojovic, B., Miozzo, M.: Simulating LTE and Wi-Fi coexistence in unlicensed spectrum with ns-3 (2016) arXiv:1604.06826 [cs.NI]
6. Dalili Shoaei, A., Derakhshani, M., Le-Ngoc, T.: Efficient LTE/Wi-Fi coexistence in unlicensed spectrum using virtual network entity: Optimization and performance analysis. IEEE Trans. Commun. **66**(6), 2617–2629 (2018). https://doi.org/10.1109/TCOMM.2018.2801789
7. Yuan, X., Qin, X., Tian, F., Hou, Y.T., Lou, W., Midkiff, S.F., Reed, J.H.: Coexistence between Wi-Fi and LTE on unlicensed spectrum: a human-centric approach. IEEE J. Sel. Areas Commun. **35**(4), 964–977 (2017). https://doi.org/10.1109/JSAC.2017.2680819
8. Li, Y., Baccelli, F., Andrews, J.G., Novlan, T.D., Zhang, J.C.: Modeling and analyzing the coexistence of Wi-Fi and LTE in unlicensed spectrum. IEEE Trans. Wireless Commun. **15**(9), 6310–6326 (2016). https://doi.org/10.1109/TWC.2016.2582866
9. Xu, S., Li, Y., Gao, Y., Liu, Y., Ga˘canin, H.: Opportunistic coexistence of LTE and WiFi for future 5G system: experimental performance evaluation and analysis. IEEE Access **6**, 8725–8741 (2018). https://doi.org/10.1109/ACCESS.2017.2787783
10. Wang, X., Mao, S., Gong, M.X.: A survey of LTE Wi-Fi coexistence in unlicensed bands. GetMobile: Mob. Comput. Commun. **20**(3), 17–23 (2017). ISSN: 2375-0529. https://doi.org/10.1145/3036699.3036705
11. Cui, Q., Gu, Y., Ni, W., Liu, R.P.: Effective capacity of licensed-assisted access in unlicensed spectrum for 5G: from theory to application. IEEE J. Sel. Areas Commun. **35**(8), 1754–1767 (2017). https://doi.org/10.1109/JSAC.2017.2710023
12. Yi, J., Sun, W., Park, S., Choi, S.: Performance analysis of LTE-LAA network. IEEE Commun. Lett. **22**(6), 1236–1239 (2018). https://doi.org/10.1109/LCOMM.2017.2779752
13. Huang, Y., Chen, Y., Hou, Y.T., Lou, W., Reed, J.H.: Recent advances of LTE/WiFi coexistence in unlicensed spectrum. IEEE Netw. **32**(2), 107–113 (2018). https://doi.org/10.1109/MNET.2017.1700124
14. Alhulayil, M., Lopez-Benitez, M.: Coexistence mechanisms for LTE and Wi-Fi networks over unlicensed frequency bands. In: 2018 11th International Symposium on Communication Systems, Networks Digital Signal Processing (CSNDSP), 2018, pp. 1–6. https://doi.org/10.1109/CSNDSP.2018.8471826
15. Baswade, A.M., Atif, T.A., Tamma, B.R., Franklin, A.: A novel coexistence scheme for IEEE 802.11 for user fairness and efficient spectrum utilization in the presence of LTE-U. Comput. Netw. **139**, 1–18 (2018). ISSN: 1389-1286. https://doi.org/10.1016/j.comnet.2018.04.002

Chapter 76
Comparison of Machine Learning Algorithms for Vehicle Routing Problems

V. S. Vamsi Krishna Munjuluri, Yashwanth Reddy Telukuntla, Parimi Sanath Kumar, Aravind Mohan, and Georg Gutjahr

Abstract Vehicle routing is a classical problem in combinatorial optimization. A large number of exact and heuristic solution methods have been developed in the past. In the last few years, machine learning algorithms have been applied to such problems with some success. This paper investigates three recent machine learning algorithms: reinforcement learning, the dynamic attention model and neural large neighborhood search. These algorithms are compared on a variety of benchmark problems from the literature. It is found that the neural large neighborhood approach gave the best quality solutions. The dynamic attention model was found to require the largest amount of memory and was not able to be trained for larger instances. Reinforcement learning provided a good compromise between runtime and solution quality.

V. S. Vamsi Krishna Munjuluri (✉) · P. Sanath Kumar
Department of Computer Science and Engineering, Amrita Vishwa Vidyapeetham, Amritapuri, India
e-mail: munjulurivkrishna@am.students.amrita.edu

P. Sanath Kumar
e-mail: psanathkumar@am.students.amrita.edu

Y. R. Telukuntla
Department of Electronics and Communication Engineering, Amrita Vishwa Vidyapeetham, Amritapuri, India
e-mail: yashwanthreddy@am.students.amrita.edu

A. Mohan · G. Gutjahr
Department of Mathematics, Amrita Vishwa Vidyapeetham, Amritapuri, India
e-mail: aravindmohan@am.amrita.edu

G. Gutjahr
e-mail: georgcg@am.amrita.edu

G. Gutjahr
Center for Research in Analytics & Technologies for Education (CREATE), Amrita Vishwa Vidyapeetham, Amritapuri, India

T. Senjyu et al. (eds.), *IOT with Smart Systems*, Smart Innovation, Systems and Technologies 251, https://doi.org/10.1007/978-981-16-3945-6_77

76.1 Introduction

The vehicle routing problem (VRP) is a combinatorial optimization and integer programming problem and can be seen as an extension of the traveling salesman problem. We have a set of customer locations that needs to be serviced. The services may include collection or distribution of certain commodities and even both. In some cases, services can be simple visits. The customer locations are visited using either a single or a fleet of vehicles. The objective of the problem is to find a set of routes visiting all the customer locations with minimum total cost. Additionally, there can be some requirements (constraints) that should be satisfied when finding the routes.

Like the traveling salesman problem (TSP), VRP instances can also be modeled as a graph in a similar way. The set of customer locations will form the vertex set, and the path connecting these locations will form the edge set. In a road network, there can be multiple paths connecting two locations. The edges can be directed or undirected according to the restrictions given in the problem. Also, different costs can be given in each direction of the same edge. The path with the least cost between two customer locations is considered as the edge between their corresponding vertices. Usually, the cost of the path will be the travel distance or travel time. So, the path with the least travel distance (travel time) is taken to be the edge, and the cost of this shortest path will be the weight of the corresponding edge. These can easily be found by solving the shortest path problem between each of the customer locations. Thus, we get a complete graph from the original sparse graph.

VRP has many important applications in the field of logistics [5], transportation [1], distribution [18], disaster relief [12] and humanitarian logistics [11]. In many industrial fields, transportation has become an important component of production cost and optimizing transportation costs necessary for the overall cost effectiveness of the industry. Depending upon the constraints that we use, we have different variations of VRP, for example the capacitated VRP. More information on VRPs can be found in the book by Golden et al. [10].

Exact solution methods are techniques that extensively search the solution space and try to return a global optimum. Exact methods for solving VRPs include branch and price, branch and cut and dynamic programming methods [16]. But vehicle routing problems are generally computationally hard, and only smaller instances can be solved exactly. So we usually rely on heuristics and metaheuristic techniques that try to give a near optimum or good quality solutions. Details on metaheuristics for solving VRPs can be found in the Handbook of Metaheuristics [9].

Solution techniques that are based on machine learning can also be used to solve VRPs. Different learning approaches can be used to achieve this. In some cases, a set of targets is provided, and a function that fits more closely to the target for every input is found. This learning technique is known as supervised learning. On the contrary, in unsupervised learning no such targets are used but try to take some properties from the joint distribution of the observed random variables. Another learning process is the reinforcement learning, where an agent is used with the aim of maximizing the expected sum of values or rewards. In each step the agent chooses an action to receive a reward from the existing environment and move to the next state. The environment for the agent is typically set up in the form of a Markov decision process, i.e., the future states are solely dependent on previous states via the current state. An overview for machine learning-based methods can be found in Bengio et al. [4].

The paper compares three recent machine learning algorithms for solving vehicle routing problems, namely the reinforcement learning algorithm by Nazari et al. [19], a dynamic attention model by Peng et al. [20] and a neural large neighborhood search algorithm by Hottung and Tierney [13].

The paper is organized as follows. Section 76.2.1 describes the three machine learning algorithms in more detail. Section 76.2.2 describes the methodology for the comparison. The results of the comparison are given in Sect. 76.3. Conclusion and other remarks are given in Sect. 76.4.

76.2 Methodology

76.2.1 Description of the Algorithms

Reinforcement Learning (RL) An efficient reinforcement learning algorithm for the VRP was proposed by Nazari et al. [19]. The main motivation when designing this approach was to present a framework which requires the least amount of hand-tuning and one which performs well with both static and dynamic settings of the VRP. The Markov decision process (MDP) formulation for solving the VRP was used which optimizes the reward function based on some stochastic sequence of decisions. The policy model in the framework does well even when the instance is newly given to the algorithm which has been trained on the same number of customer nodes and demand distributions.

The pointer network (PN) [26] to decode the solution restricts the problem by assuming that the system stays static over time. This greatly reduces the potential of the model, and moreover, the slightest of changes in the customer locations and the demands can really degrade the performance of the models using instance-specific policies which is the case with classic heuristics for the VRP. This explains why the policy model was implemented using recurrent neural networks (RNNs) coupled with attention mechanisms. By using RNNs, the system makes better decisions when the instance is dynamic in nature, and also, it adds an additional benefit that it allows multiple vehicles to service any particular node (split delivery) which is considered a desirable feature.

The model mainly has two components. The first being the embedding layer that is used to encode the customer locations and the corresponding demands [22] and the second being the decoder network which, at every decoding step, points to an input [3, 6, 23]. It is to be observed that unlike many of the models in the literature [3, 14, 17], the model presented has no encoder RNN due to the fact that it does not add any meaningful information to the system and hence is not included. Up until this point, there have been components that only can handle static inputs but to handle dynamic settings of the CVRP, the attention layer is employed.

The attention mechanism addresses different parts of the input and is a differentiable structure. Right at the decoder step, context-based attention mechanism is used [25] which specifies the relevance of every input in the next decode step. It is also pointed out that when pointer networks are used for combinatorial optimization problems, performance also depends on relative ordering of the inputs but since the embeddings and the attention mechanisms are invariant to the input order, policy model in the article shows no reduction in the performance due to such factors.

Dynamic Attention Model (DAM) In contrast to the previous algorithm, the idea of the dynamic attention model (DAM) by Peng et al. [20] is to have a dynamic encoder–decoder architecture. Dynamic encoder–decoder architecture is very similar to encoder–decoder architecture in [6, 15], one of the key differences being the embedding of each node which will immediately be recomputed when the vehicle returns to the depot in DAM, whereas the embedding of each node is fixed in the normal attention model in encoder–decoder architecture [17] which can only represent the initial state of the input instance. But in the proposed model, the encoder and decoder are used alternately to recode the embedding of each node and to know which node to be selected to visit. And there are two strategies proposed: one is a sample rollout that selects a node using sampling which is a stochastic one, and the other is a greedy rollout that selects the node with maximum probability which is a deterministic policy. We generally prefer the stochastic one.

Firstly, the input instances are extracted by the encoder, the structural features are embedded to the encoder and passed onto the decoder to construct a partial solution at each step, and finally, the solution is constructed incrementally from each step and is summed up to give us the final solution at the end. A key difference between this encoder–decoder architecture and the vanilla architecture is that every time a vehicle returns to the depot, the embeddings of each node are recomputed.

When the models trained for 20 customer nodes are used to infer for instances with 50 and 100 customer nodes and vice versa, it is mentioned that they perform quite well, i.e., the models generalize well.

Neural Large Neighborhood Search (NLNS) The method by Hottung and Tierney [13] employs a form of large neighborhood search (LNS) which is a metaheuristic that iteratively searches through the entire solution space using destroy (removes a part of the solution) and repair (fixes the incomplete or the destroyed solution) operators. It is argued that making use of such improved search techniques reduces the performance gap between the state-of-the-art optimizations and ones powered by machine learning. This explains why the authors proposed a model which integrates learned heuristics into a metaheuristic that is not instance-specific and is hence guaranteed to have good generalization capabilities which is a highly desirable property for any learning model.

But, unlike LNS, NLNS uses multiple destroy and repair operators much like the adaptive large neighborhood search [21], and the performance greatly depends on the quality of these operators used. The formulated model uses two simple destroy operators, more specifically, a point-based destroy operator which removes, from all the tours of a solution, a subset of customers that are closest to a randomly selected point and a tour-based destroy operator which removes, from a solution, all the tours closest to a randomly selected point in combination with learned repair operators. The repairing process is modeled as a reinforcement learning problem where an agent interacts with an environment. These interactions happen over multiple time steps, and in each time step, an incomplete tour gets connected with another incomplete tour or with the depot.

Each action corresponds to connecting tours, and this process is repeated until there are no incomplete tours. A masking procedure is put in place to prevent the model from creating infeasible solutions. The model is trained to repair the solutions that the destroy operator destroys via reinforcement learning using the approach proposed by [27]. The embedded inputs are fed to the attention mechanism [26] to compute a context vector with all the relevant inputs (determined by another embedding). The objective is to minimize the expected loss while repairing a particular solution destroyed using a particular destroy operator.

76.2.2 Description of the Comparison

The comparison uses benchmark datasets for VRPs that have been previously published in the literature. The datasets are as follows: Augerat et al. (Set P, Set B), [2] Christofides and Ellion (Set E) [7], Christofides, Mingozzi and Toth (Set M, CMT) [8] and Uchoa et al. [24]. Uchoa created a public library for all these CVRP datasets[1], and all abovementioned datasets were taken from this library.

All simulations were done on Google Colab's standard cloud-based runtime. Training data were generated on the fly for all three algorithms. Each model was trained until the Colab resource restriction was reached, which are about 10 h of runtime. After training, each algorithm was run 5 times on each dataset on each scenario. We reported mean objective values and standard deviations from these runs.

76.3 Results

The RL model is trained for problem instances of sizes 20, 21, 51 and 100–20,000, 18,500, 12,000 and 2000 train steps, respectively. The model takes about 30–90 s, depending on the size of the problem instance, to build the agent. After the agent is built, the inference times of small, medium and large scenarios are 6, 14 and 30 s, respectively.

[1] http://vrp.atd-lab.inf.puc-rio.br/index.php/en.

NLNS was applied with a pre-trained model given by the original authors of algorithm.[2] The pre-trained model was trained to 250,000 train steps on each operator. The testing times were fixed to be 180, 500 and 1000 s for small, medium and large scenarios, respectively.

DAM model has been trained to 7000 epochs and 6000 epochs for small and medium scenarios, respectively. Larger number of epochs can potentially produce better quality solutions. Due to time and memory constraints on our setup, we had to settle for the mentioned number of epochs. DAM model produced results in about 2 s, for both small and medium scenarios.

Table 76.1 shows the results, on how close to the optimal solution, the models could give after training to the specified number of epochs. While the RL and the DAM models were very quick to generate solutions, the NLNS model outperformed them in terms of quality of the results. NLNS, being part metaheuristic and part reinforcement, offers a trade-off between its runtime and the quality of its solutions. For small scenarios, 180 s seemed promising, but for larger scenarios a greater runtime might potentially produce a better quality solution. DAM model is the next best to NLNS on both small and medium scenarios; on the other hand for large scenarios, we were not able to train the DAM model on our setup due to memory overflow. We also point out that the reason for zero standard deviation in the case of DAM pertains to the fact that the model uses a deterministic policy to assess which nodes to visit next. Note that RL model's performance was very close to DAM model on both small and medium scenarios. RL model has two simple types of searches, greedy and beam search. While greedy has been the quicker to produce solutions, beam search proved to be the closest to the optimal in most cases. The test results reported in the table are the better among the two searches RL model is capable of. Both RL model and DAM model were much faster than NLNS in terms of time taken to produce a solution.

76.4 Discussion and Conclusion

This paper considered the application of machine learning-based methods to solve vehicle routing problems. Three machine learning-based algorithms, namely reinforcement learning, dynamic attention model and neural large neighborhood search, were trained, and their performances on some benchmark datasets were compared. The neural large neighborhood search came out to be better in terms of solution quality. But, the reinforcement learning and dynamic attention methods were faster to produce solutions. While dynamic attention model performed reasonably on smaller instances, it was unable to be trained for larger instances as it run out of memory.

A restriction of the study was the maximum training time of 10 h for the models. Both the RL and the DAM approaches might benefit from additional training to increase the quality of the solutions.

[2] https://github.com/ahottung/NLNS/tree/master/trained_models/cvrp/XE1.

Table 76.1 Simulation results

Scenario	Size	Optimal	NLNS		RL		DAM
			Mean	SD	Mean	SD	Mean
Augerat Set P (P-n21-k8)	20	211	224.8	6.3	496.5	162.6	280.0
Chirstofides Set E (E-n22-k4)	21	375	375.2	0.4	623.2	155.8	567.1
Augerat Set P (P-n22-k2)	21	215	239.8	7.1	472.6	99.9	283.0
Augerat Set P (P-n22-k8)	21	603	603.8	16.8	698.4	44.7	713.2
Augerat Set P (P-n51-k10)	50	741	1028.2	20.9	1496.0	88.7	1053
Christofides Set E (E-n51-k5)	50	521	618.0	13.2	1344.1	142.3	867.0
Augerat Set B (B-n51-k7)	50	1032	1032.8	5.4	2449.2	758.3	1362.0
Chirstofides et al. (CMT 1)	50	524	628.8	13.2	1429.3	210.1	867.5
Chirstofides et al. (CMT 6)	50	555	640.0	12.6	1507.5	215.3	867.5
Uchoa et al. (X-n101-k25)	100	27,591	27880.7	34.3	52263.4	9653.8	–
Chirstofides et al. (M-n101-k10)	100	820	1052.0	38.8	3286.0	386.4	–
Christofides et al. (CMT8)	100	866	1190.0	22.6	2895.0	457.9	–
Christofides Set E (E-n101-k14)	100	1071	1421.8	27.7	2826.7	290.9	–
Augerat Set P (P-n101-k4)	100	681	997.8	38.5	2680.7	305.7	–

Future works include providing more training to the methods to make them even better. Also, using a longer runtime would result in even better solutions. Training of dynamic attention model for larger instances can also be done by using a setup with more memory capacity. More machine learning methods can be applied for vehicle routing problems, and their performances can be compared with other solution techniques. Machine learning approaches could also be tried to be implemented for some real-world problems as well. They can also be applied for other variants of vehicle routing problems such as pickup and delivery problem or the dial-a-ride problem.

References

1. Ajayan, S., Dileep, A., Mohan, A., Gutjahr, G., Sreeni, K., Nedungadi, P.: Vehicle routing and facility-location for sustainable lemongrass cultivation. In: 2019 9th International Symposium on Embedded Computing and System Design (ISED), pp. 1–6. IEEE (2019)
2. Augerat, P.: Approche polyèdrale du problème de tournées de véhicules. Ph.D. thesis, Institut National Polytechnique de Grenoble-INPG (1995)
3. Bahdanau, D., Cho, K., Bengio, Y.: Neural machine translation by jointly learning to align and translate. arXiv preprint arXiv:1409.0473 (2014)
4. Bengio, Y., Lodi, A., Prouvost, A.: Machine learning for combinatorial optimization: a methodological tour d'horizon. Eur. J. Oper. Res. (2020)
5. Buhrkal, K., Larsen, A., Ropke, S.: The waste collection vehicle routing problem with time windows in a city logistics context. Procedia—Soc. Behav. Sci. **39**, 241–254 (2012)
6. Cho, K., Van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., Bengio, Y.: Learning phrase representations using rnn encoder-decoder for statistical machine translation. arXiv preprint arXiv:1406.1078 (2014)
7. Christofides, N., Eilon, S.: An algorithm for the vehicle-dispatching problem. J. Oper. Res. Soc. **20**(3), 309–318 (1969)
8. Christofides, N., Mingozzi, A., Toth, P.: Exact algorithms for the vehicle routing problem, based on spanning tree and shortest path relaxations. Math. Progr. **20**(1), 255–282 (1981)
9. Gendreau, M., Potvin, J.Y., et al.: Handbook of Metaheuristics, vol. 2. Springer (2010)
10. Golden, B.L., Raghavan, S., Wasil, E.A.: The Vehicle Routing Problem: Latest Advances and New Challenges, vol. 43. Springer (2008)
11. Gutjahr, G., Krishna, L.C., Nedungadi, P.: Optimal tour planning for measles and rubella vaccination in Kochi, South India. In: 2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI), pp. 1366–1370. IEEE (2018)
12. Gutjahr, G., Viswanath, H.: A genetic algorithm for post-flood relief distribution in Kerala, South India. In: ICT Systems and Sustainability, pp. 125–132. Springer (2020)
13. Hottung, A., Tierney, K.: Neural large neighborhood search for the capacitated vehicle routing problem. arXiv preprint arXiv:1911.09539 (2019)
14. Jean, S., Cho, K., Memisevic, R., Bengio, Y.: On using very large target vocabulary for neural machine translation. arXiv preprint arXiv:1412.2007 (2014)
15. Kool, W., Van Hoof, H., Welling, M.: Attention, learn to solve routing problems! arXiv preprint arXiv:1803.08475 (2018)
16. Laporte, G., Nobert, Y.: Exact algorithms for the vehicle routing problem. In: North-Holland Mathematics Studies, vol. 132, pp. 147–184. Elsevier (1987)
17. Luong, M.T., Pham, H., Manning, C.D.: Effective approaches to attention-based neural machine translation. arXiv preprint arXiv:1508.04025 (2015)
18. Mohan, A., Dileep, A., Ajayan, S., Gutjahr, G., Nedungadi, P.: Comparison of metaheuristics for a vehicle routing problem in a farming community. In: Symposium on Machine Learning and Metaheuristics Algorithms, and Applications, pp. 49–63. Springer (2019)
19. Nazari, M., Oroojlooy, A., Snyder, L.V., Takáč, M.: Reinforcement learning for solving the vehicle routing problem. arXiv preprint arXiv:1802.04240 (2018)
20. Peng, B., Wang, J., Zhang, Z.: A deep reinforcement learning algorithm using dynamic attention model for vehicle routing problems. In: International Symposium on Intelligence Computation and Applications, pp. 636–650. Springer (2019)
21. Ropke, S., Pisinger, D.: An adaptive large neighborhood search heuristic for the pickup and delivery problem with time windows. Transp. Sci. **40**(4), 455–472 (2006)
22. Scarselli, F., Gori, M., Tsoi, A.C., Hagenbuchner, M., Monfardini, G.: Computational capabilities of graph neural networks. IEEE Trans. Neural Netw. **20**(1), 81–102 (2008)
23. Sutskever, I., Vinyals, O., Le, Q.V.: Sequence to sequence learning with neural networks. arXiv preprint arXiv:1409.3215 (2014)

24. Uchoa, E., Pecin, D., Pessoa, A., Poggi, M., Vidal, T., Subramanian, A.: New benchmark instances for the capacitated vehicle routing problem. Eur. J. Oper. Res. **257**(3), 845–858 (2017)
25. Vinyals, O., Bengio, S., Kudlur, M.: Order matters: Sequence to sequence for sets. arXiv preprint arXiv:1511.06391 (2015)
26. Vinyals, O., Fortunato, M., Jaitly, N.: Pointer networks. arXiv preprint arXiv:1506.03134 (2015)
27. Williams, R.J.: Simple statistical gradient-following algorithms for connectionist reinforcement learning. Mach. Learn. **8**(3–4), 229–256 (1992)

Chapter 77
Prediction and Comparative Analysis Using Ensemble Classifier Model on Leafy Vegetable Growth Rates in DWC and NFT Smart Hydroponic System

P. Srivani, C. R. Yamuna Devi, and S. H. Manjula

Abstract A comparison of leafy green spinach species growth rates in two different hydroponic systems was performed in a controlled environment. The integration of several sensors to monitor the parameters of plant growth has been deployed using the Internet of things (IoT) technology. Intelligent models to predict the plant growth in the hydroponic system are necessary for better decision making in controlling the parameter during plant growth. This research compares the plant growth dynamics in deep water culture (DWC) and nutrient film technique (NFT) systems. The results demonstrate efficient plant growth in the NFT system compared to DWC in terms of height and number of leaves. The study also discusses the observations during the growth time to analyze the most suitable hydroponic structure for spinach growth. The growth prediction is implemented using an ensemble classifier model, which gives an accuracy rate above 79% on DWC and 64% on the NFT dataset based on binary classification.

77.1 Introduction

The traditional agricultural practices use heavy chemical fertilizers in order to provide efficient crop growing conditions in the soil. The Urban Controlled Environment Agriculture (CEA) is more advantageous as the production rate is faster and utilizes significantly less water resource by pairing with hydroponic technology. The recent advances in modernized farming for urban cultivation have promoted technology and

P. Srivani (✉)
BMS Institute of Technology and Management, Bengaluru, India
e-mail: srivanicse@bmsit.in

C. R. Yamuna Devi
Dr. Ambedkar Institute of Technology, Bengaluru, India
e-mail: yamunadevicr.et@drait.edu.in

S. H. Manjula
University of Visvesvaraya College of Engineering, Bengaluru, India

T. Senjyu et al. (eds.), *IOT with Smart Systems*, Smart Innovation, Systems and Technologies 251, https://doi.org/10.1007/978-981-16-3945-6_78

intelligent solutions for efficient plant production. Hydroponic, a soil-less culture, is based on supplying nutrients non-stop to roots. The nutrient-rich water constitutes macronutrients like nitrogen, phosphorus, and calcium and micronutrients like iron, magnesium, chloride, zinc, etc. Hydroponic farming has got well established in recent days for crop cultivation in urban cities. Among different categories of hydroponic systems, the most suitable indoor-based hydroponics are NFT and DWC techniques. The DWC system has a naive structure where the plant's roots are directly submerged in an oxygenated and nutrient-rich reservoir tank, and plants are supported through a mesh basket. NFT system uses the recirculating nutrient solution, where plants get oxygenated with a simple piped structure [1], and the contact of roots with the solution is partial.

Leafy green vegetables have short growth cycles than any flowering or fruiting crops, which are the best suitable crops for indoor production. These leafy veggies have greater planting density which consumes lesser energy, cost, and power utilization. This can be produced in larger quantities in commercial-based hydroponics and can also aid as model crops to analyze the plant growth dynamics. They are well adjusted to grow in Aquaponics, DWC, and NFT systems with low-level nutrients to medium-level nutrients. The rate of flow of nutrient solution in NFT systems [2] is very significant for optimal plant growth, which aids in better absorption of nutrients. Indoor hydroponic cultivation is an approach to guard the crops against adverse climate, pest and insect attacks, and soil-borne diseases. The photosynthesis process happens through LED lights, which supplements the source of sunlight, and environmental parameters [3] directly influences the health of the plant.

Internet of things delivers enormous real-time data on crop allied parameters fed to machine learning models for better crop productivity and optimized resource utilization. Data analytics is vital to an agricultural industry which is a ground for modern farming methods. The yield can be increased through customized current technology approaches like the Internet of things, machine learning, and intelligent models that use complex data. This data helps achieve the benefits of crop management, choosing the best crop genetics, setting suitable conditions and weather that can best suit the plant growth dynamics.

NFT systems [4] are simple in their design structure and assembly, suitable for indoor cultivation within a small space. The number of plants in the NFT system can be well defined based on the number of pipes, shelves, and holes in each pipe which aids in deciding the right amount of solution to be supplied. The primary motivation is to analyze the optimized and better plant growth in a hydroponic system for leafy vegetables. The rest of the article is systematized as follows: Sect. 77.2 discusses the state of the art of literature, Sect. 77.3 articulates materials and methods used, Sect. 77.4 deliberates about the results, and Sect. 77.5 concludes the research work.

77.2 Literature Review

Environmental parameters and light intensity also influence plant leaf growth. The growth rate of spinach in hydroponics was analyzed [5] when treated with red and blue LED lights in terms of leaf height, area, and number. The growth rates and photosynthesis process were improved in higher intensity ratios. In a deep flow technique-based spinach growth, the uptake of nutrients at various stages of plant growth has a significant impact on yield. The effect of macronutrients such as potassium, nitrogen, and phosphorus during the mid-growing days [6] showed increased plant shoot growth rate. The concentration of microlevel nutrients did not affect much, which helped control and manage the proportions of other nutrient elements required.

A comparative study of NFT, DWC, and "Shallow Aggregate Ebb-and-Flood (SAEF)" system in a controlled greenhouse environment on basil plant was experimented to analyze the yield of crops. In all three methods, pH and electrical conductivity (EC) were well maintained between 5.5–6.5 and 1.5–2.0 μS/cm, respectively. The rate of growth in the SAEF system [7] was higher compared to the other two methods, and DWC was better than the NFT system. A similar comparative study of growth rates between NFT hydroponic and NFT Aquaponics [8] was experimented on lettuce crop and different herb species. The authors have strived to improve the Aquaponics yield and operating design compared to the standard NFT system. A multiple linear regression model [9] was implemented to automate the control mechanism in adjusting pH values. Similarly, a fuzzy logic control methodology was implemented [10] to control the level of nutrients at different stages of plant growth automatically (Table 77.1).

Prediction and analysis of plant growth systems play a prominent role in an agriculture industry that guides controlling and rapid decision making. Several artificial intelligence (AI) models [14] can predict the best crop yield based on seasonality and trend. Several studies have implemented statistical analysis models [15] to analyze the linear relationship between the yield and growth parameters and act upon it. An

Table 77.1 An existing system's comparative study

References	Comparative study	Crop	Outcome
[11]	Automated NFT and soil-based	Spinach, lettuce	Around 87% better crop yield than soil-based
[12]	Semi-automated NFT on different crops	Red spinach, green spinach, water spinach	Red and green had 2% higher relative growth in an automated system compared to the conventional approach
[13]	NFT, DWC, and soil cultivation	Lettuce	64% water savings from NFT, the yield was high in hydroponics compared to soil-based cultivation

optimized deep neural network (DNN) model [16], together with spatial statistical approaches, showed the best results. An analysis of the variance model [17] was implemented to test the growth comparability between soil-based and soilless agriculture systems on cucumber plants. However, this soilless system proved efficient and faster plant growth. This research aims to compare the crop growth rate of spinach in two different hydroponic cultivation systems. The NFT and DWC systems yield and overall plant vitality are determined using a machine learning model.

77.3 Materials and Methods

The trials experimented in an indoor environment in DWC culture, where roots are directly submerged in the nutrient solution with 25 L of water. Similarly, the plant growth section in NFT was built using standard polycarbonate pipes (9 m wide × 24 m long), where the nutrient solution is circulated continuously from the nutrient tank using a submersible pump. The plant growing system is monitored in real time using different sensors by integrating the Internet of things. The photosynthesis process for plant growth is done through grow white LED lights. The water temperature of both the systems was maintained at 25 °C.

77.3.1 Monitoring System Using IoT

EC, pH, environmental parameters (temperature and humidity), and nutrient solution temperature are recorded by the sensors that are implemented using IoT technology. The grow light was positioned above the plant growth systems (DWC and NFT) to provide the light intensity with around 26–28 W for around 10 h of light and rest hours of darkness. The LED light and the pumps are automated to work for 10 h and 12 h, respectively. All the required data parameters are automatically collected through a cloud application and stored in Google Spreadsheet. The dataset referred contains five columns and one output column that determines the growth in terms of height. In data preprocessing output, the column is encoded into binary class value keeping a threshold of plant growth parameter as class 1 if the size is more significant than 0.1 cm and 0 if lesser than or equal to 0.1 cm. There were neither null values nor missing values.

77.3.2 Methodology

In statistical analysis, multivariate classification models are thoroughly associated with predictive modeling techniques. The ensemble learning approach is inferred from training data and classifies the instances of data samples based on the class labels which are predefined. Further, the model is tested on the rest data to make predictions and through which the performance of the model is evaluated. In the proposed methodology, an ensemble of machine learning classifier models is Naïve Bayes and support vector machine.

The Naïve Bayes model is a supervised classification approach, which can make predictions built on prior knowledge. One of the efficient Naïve Bayes models works well with small and medium datasets, and a supervised model based on Bayes theorem is shown in the equation below Eq. 77.1.

$$P(X|Y) = \frac{P(Y/X) * P(X)}{P(Y)} \tag{77.1}$$

The above equation shows the conditional probability of an event called posterior probability. Later, the association or hypothesis to a response variable is formed to construct a model that learns and predicts new data variables. For a single input parameter, the Naïve Bayes classifier can predict the growth similar to the above equation, which is shown in Eq. 77.2.

$$P(\text{GrowthClass}|\text{inputfeature}) = \frac{P(\text{inputfeature}|\text{GrowthClass}) * P(\text{GrowthClass})}{P(\text{inputfeature})} \tag{77.2}$$

The Naïve Bayes model assumes that the input features are independent and classifies based on the below Eq. 77.3 for multiple input features.

$$P(C_i, x_1, x_2, x_3 \ldots, x_k)\alpha P(x_1|C_i)\alpha P(x_2|C_i)\alpha \ldots P(x_k|C_k = \prod_{i=1}^{n}(Px_i|C_i) * p(C_i) \tag{77.3}$$

where $x_1, x_{2,} \ldots, x_n$ are multiple features, and the probability that each row belongs to a binary class C_i is determined. When there is a continuous data attribute in the dataset, then a Gaussian Naïve Bayes model, which features a normal distribution, is best for classification problems. The Gaussian probability density function for classifying is calculated by considering the input parameter to predict the probability of a new variable for that class is shown in Eq. 77.4.

$$\text{Guassian(input)} = \frac{1}{\sqrt{(2 * 3.14) * \text{sd}))}} \mathrm{e}^{\frac{-(x-\text{mean})^2)}{2*(\text{sd})^2}} \tag{77.4}$$

Support vector machine is a supervised approach to learning that can be constructed in an N-dimensional space. A predictive classification analysis algorithm is more suitable for binary classification. This model uses a nonlinear kernel function to predict the target value. The most commonly used prediction classifier better than decision trees is logistic and Naïve Bayes classifier models. It evaluates based on the closest point to the hyperplane [18] and computes the distance to find the optimal hyperplane for class prediction. It transforms the input feature into a high-dimensional space using the radial basis function (RBF) using Eq. 77.5.

$$k(x_j, x_k) = \left(-\gamma \left\| x_j - x_k \right\|^2\right) \tag{77.5}$$

RBF function is capable of identifying short-term prediction in the classification of data and finds the angular distance between the two points (x_j, x_k).

Ensemble learning approach aids in improving the efficiency of the predictive models by implementing multiple methodologies on a specific dataset. This [19] ensemble learning generally combines an individual set of models to improvise on the strength and predictive effect of the model. The ensemble model is implemented and evaluated by splitting the dataset into 70% of the training set and 30% of the testing dataset and further applying tenfold cross-validation. In this research experiment, an ensemble learning method to predict plant growth based on hydroponic and environmental parameters is used. Also to improve the accuracy, a voting classifier is used with this ensemble model, which wraps above-mentioned models and aggregates the model predictions. The growth prediction is the categorical value converted into binary data that determines growth. The ensemble learning classification model was fitted using cross-validation. In K-fold cross-validation, the data samples are randomly partitioned into equal-sized subsamples. A single subsample is randomly picked for testing and remaining for training.

77.4 Results and Discussions

A comparative study of growth rates of NFT and DWC systems was analyzed in this experiment on the spinach plant. However, few differences are observed, which are shown in Table 77.2.

The plant growth comparison of spinach plant for day 30 is shown in Fig. 77.1. The first part is a visualization of the spinach in the DWC system shown in Fig. 77.1a, and the second part shows the NFT system as shown in Fig. 77.1b. The parameters monitored in both the system are pH, EC, root zone temperature, and environment parameters like humidity and air temperature. The light played an essential role in the growth of the plant and also with respect to the variation of pH value of the solution.

The pH variation in two different systems is visualized in the above Fig. 77.2. Although pH varied around 5–10% out of the specific control range (5.5–6.5) in NFT, the growth rate had no effect on the plants. The plant growth in terms of the number of leaves and height increased even with varied pH values. The nutrient solution of both systems was maintained well within the specific range as per the nutrient requirement. In the DWC system, when the pH value was not controlled, there was a decline in plant growth rate since roots were directly submerged in the nutrient solution tank.

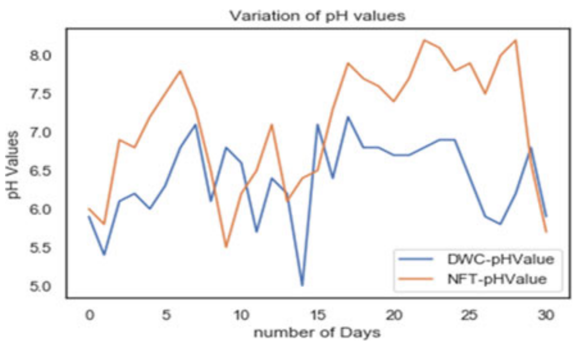
Table 77.2 Observations on growth of spinach in two different hydroponic systems

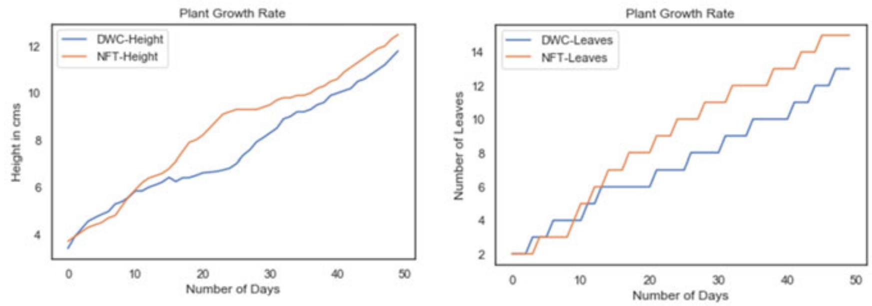
S. No.	DWC	NFT
1	Roots are directly submerged in the nutrient solution tank	Roots are partially submerged in recirculating polyvinyl chloride (PVC) pipes
2	The nutrient solution pH values are the same all over the tank	The recirculating water in pipes has different pH values compared to the nutrient solution tank
3	The temperature would be the same all over the tank	The solution tank has a different temperature compared to PVC pipe temperature
4	There is a need for a separate device to oxygenate the nutrient solution tank	As the water is recirculated continuously, the system is self-oxygenated

(a) Plant growth in DWC (b) Plant growth in NFT

Fig. 77.1 Plant growth in two different hydroponic on 25th day after transplantation

Fig. 77.2 Comparison of pH variations in NFT and DWC





(a) Height (b) Number of leaves

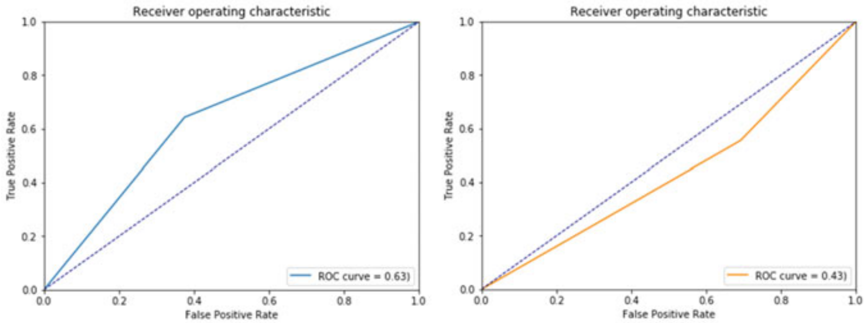
Fig. 77.3 Plant growth rate comparison in terms of height in cm

The growth rate between the NFT system and DWC hydroponic system is compared in terms of the height of plants and the number of leaves, as shown in Fig. 77.3. As observed in Fig. 77.3a, the growth rate concerning height is faster with the NFT system for leafy vegetables. Initially, in NFT for the first eight days, the plant was not exposed to any light. During this situation, the growth was prolonged.

The number of leaves changed day to day is compared in both the hydroponic system. As shown in Fig. 77.3b, the number of leaves in the initial eight days was less in the NFT system. This was due to no exposure to artificial LED light. This reduced the photosynthesis process in plants. However, very little research has been conducted on the prediction of hydroponic plant growth using intelligent machine learning models. The research work was implemented using ensemble classifier models to predict plant growth with binary classification. The model's performance was evaluated using 10 cross-validation on the dataset collected from NFT and DWC systems. The accuracy rate of the classifier ensemble model for the DWC and NFT datasets was 64% and 41%. As observed, the prediction rate is very low because of consideration of the smaller dataset. The leafy vegetables can be grown with a shorter life cycle.

The performance of the ensemble learning model is measured by considering a mean area under receiver operating characteristics (ROC) curve for each independent attribute over the probability predictions. ROC obtained on the DWC dataset is shown in the below figure, which is used in binary classification. The ROC curve features and interprets the probabilistic forecast for binary classification in terms of true positives and false positives on the graph.

The interpretation of Fig. 77.4 determines that for the DWC dataset, there are lesser false positives and greater true negatives with a ROC value of 0.63, which is shown in the above Fig. 78.4a. Similarly, for the NFT dataset, there are greater true positives and lesser false negatives with a ROC value of 0.43, shown in Fig. 78.4b.



a) DWC System b) NFT System

Fig. 77.4 ROC curve of hydroponic systems

77.5 Conclusion

CEA-based indoor hydroponic has been advancing as a precision farming technique in places with uncertain climate, inadequate land space, etc. Spinach plant growth showed better and faster crop yield in the NFT system compared to DWC hydroponics. The growth rate can also be assessed based on fresh and dry weight of plants. The ensemble learning classifier model predicted the growth by considering a threshold value of plant height based on binary classification. However, the dataset used in this model was small to get accurate prediction. Also, further the model can be implemented on different leafy veggies to analyze the performance of the deployed model. This sort of predictions would help in maximizing the productivity and increase profits.

References

1. Puno, J.C.V., Haban, J.J.I., Alejandrino, J.D., Bandala, A.A., Dadios, E.P.: Design of a nutrient film technique hydroponics system with fuzzy logic control. In: 2020 IEEE Region 10 Conference (TENCON), pp. 403–408. IEEE (2020)
2. Al-Tawaha, A.R., Al-Karaki, G., Al-Tawaha, A.R., Sirajuddin, S.N., Makhadmeh, I., Wahab, P.E.M., Youssef, R.A., Al Sultan, W., Massadeh, A.: Effect of water flow rate on quantity and quality of lettuce (*Lactuca sativa* L.) in nutrient film technique (NFT) under hydroponics conditions. Bul. J. Agric. Sci. **24**(5), 791–798 (2018)
3. Lakhiar, I.A., Jianmin, G., Syed, T.N., Chandio, F.A., Buttar, N.A., Qureshi, W.A.: Monitoring and control systems in agriculture using intelligent sensor techniques: a review of the aeroponic system. J. Sens. **2018** (2018)
4. Chowdhury, M.E.H., Khandakar, A., Ahmed, S., Al-Khuzaei, F., Hamdalla, J., Haque, F., Reaz, M.B.I., Al Shafei, A., Al-Emadi, N.: Design, construction and testing of Iot based automated indoor vertical hydroponics farming test-bed in Qatar. Sensors **20**(19), 5637 (2020)
5. Nguyen, T.P.D., Tran, T.T.H., Nguyen, Q.T.: Effects of light intensity on the growth, photosynthesis and leaf microstructure of hydroponic cultivated spinach (*Spinacia oleracea* L.) under a combination of red and blue LEDs in house. Int. J. Agric. Tech. **15**(1), 75–90 (2019)

6. Maneejantra, N., Tsukagoshi, S., Lu, N., Supoaibulwatana, K., Takagaki, M., Yamori, W.: A quantitative analysis of nutrient requirements for hydroponic spinach (*Spinacia oleracea* L.) production under artificial light in a plant factory. J. Fertilizers Pesticides **7**(12), 2–5 (2016)
7. Doty, S., Dickson, R.W., Evans, M.: Evaluation of a novel shallow aggregate Ebb-and-flood culture system and transplant size effects on hydroponic basil yield. HortTechnology **1**(aop), 1–8 (2020)
8. Lennard, W., Ward, J.: A comparison of plant growth rates between an NFT hydroponic system and an NFT aquaponic system. Horticulturae **5**(2), 27 (2019)
9. Helmy, H., Janah, D.A.M., Nursyahid, A., Mara, M.N., Setyawan, T.A., Nugroho, A.S.: Nutrient solution acidity control system on NFT-based hydroponic plants using multiple linear regression method. In: 2020 7th International Conference on Information Technology, Computer, and Electrical Engineering (ICITACEE), pp. 272–276. IEEE (2020)
10. Mashumah, S., Rivai, M., Irfansyah, A.N.: Nutrient film technique based hydroponic system using fuzzy logic control. In: 2018 International Seminar on Intelligent Technology and Its Applications (ISITIA), pp. 387–390. IEEE (2018)
11. Mehta, M., Chawla, P., Jot, G.: Farming of spinach and lettuce in hydroponics system with IoT. Int. J. Adv. Sci. Technol. **29**(10s), 1743–1762 (2020)
12. Yue, S.J., Hairu, C., Hanafi, M., Shafie, S.M., Salim, N.A.: IoT based automatic water level and electrical conductivity monitoring system. In: 2020 IEEE 8th Conference on Systems, Process and Control (ICSPC), Melaka, Malaysia, pp. 95–100 (2020), http://doi.org/10.1109/ICSPC50992.2020.9305768
13. Majid, M., Khan, J.N., Shah, Q.M.A., Masoodi, K.Z., Afroza, B., Parvaze., S.: Evaluation of hydroponic systems for the cultivation of lettuce (*Lactuca sativa* L., var. Longifolia) and comparison with protected soil-based cultivation. Agric. Water Manag. **245**, 106572 (2021)
14. Yasrab, R., Zhang, J., Smyth, P., Pound, M.P.: Predicting plant growth from time-series data using deep learning. Remote Sens. **13**(3), 331 (2021)
15. Khudoyberdiev, A., Ahmad, S., Ullah, I., Kim, D.H.: An optimization scheme based on fuzzy logic control for efficient energy consumption in hydroponics environment. Energies **13**(2), 289 (2020)
16. Kim, N., Ha, K.-J., Park, N.-W., Cho, J., Hong, S., Lee, Y.-W.: A comparison between major artificial intelligence models for crop yield prediction: case study of the midwestern United States, 2006–2015. ISPRS Int. J. Geo Inf. **8**(5), 240 (2019)
17. Gashgari, R., Alharbi, K., Mughrbil, K., Jan, A., Glolam, A.: Comparison between growing plants in hydroponic system and soil based system. In: Proceedings of the 4th World Congress on Mechanical, Chemical, and Material Engineering, pp. 1–7. ICMIE, Madrid, Spain (2018)
18. Rodríguez-Pérez, R., Vogt, M., Bajorath, J.: Support vector machine classification and regression prioritize different structural features for binary compound activity and potency value prediction. ACS Omega **2**(10), 6371–6379 (2017)
19. Ilyas, Q.M., Ahmad, M.: An enhanced ensemble diagnosis of cervical cancer: a pursuit of machine intelligence towards sustainable health. IEEE Access **9**, 12374–12388 (2021). https://doi.org/10.1109/ACCESS.2021.3049165

Chapter 78
Twitter Sentiment Analysis of the 2019 Indian Election

Kalpdrum Passi and Jaydeep Motisariya

Abstract In 2019, the Indian Loksabha Election saw around 360 million tweets on Twitter, giving their opinions and showing their sentiment toward the political leaders and their parties. The political parties have used this technique to run their campaigns and understand the opinions of the public; this also enables them to modify their campaigns accordingly. We performed text mining on approximately 2 million tweets collected over four months that referenced four national political parties in India during the campaigning period for the Loksabha election in 2019. We have identified the sentiment of Twitter users toward each of the considered Indian political parties (Congress, BJP, AAP and BSP) using Valence Aware Dictionary and sentiment Reasoner (VADER).

78.1 Introduction

India is the world's largest democracy, during the year 2019, India witnessed 17th Loksabha elections. The general elections are the most significant event every five years to form the government at the center and elect the Prime Minister of India. The elections were held for 543 parliament seats [1] all over India. A party requires to win 272 seats for the majority to be elected at the center. As per the Election Commission of India, 900 million [1] people were qualified to vote with 84.3 million voters added since the last elections in 2014. The 2019 general elections were the largest-ever election in the world. The Election Commission of India conducted the elections in a total of seven phases over seven weeks starting from April 11, 2019 with first phase and seventh phase commencing on May 19, 2019 with results to be declared on May 23, 2019. A total of 600 million voters polled their votes in 2019, which was the highest recorded turnout in the history of Indian general elections. In this

K. Passi (✉) · J. Motisariya
Laurentian University, Sudbury, ON P32C6, Canada
e-mail: kpassi@laurentian.ca

J. Motisariya
e-mail: jmotisariya@laurentian.ca

T. Senjyu et al. (eds.), *IOT with Smart Systems*, Smart Innovation, Systems and Technologies 251, https://doi.org/10.1007/978-981-16-3945-6_79

Fig. 78.1 Flow diagram for Twitter sentiment analysis

paper, we used Twitter archive sources [2] to collect the tweets related to the Indian Loksabha elections from February to May 23, 2019, as this was the period where most of the election campaigns were conducted. This research tried to relate the public opinion and emotions with the parties campaigning timelines and find out if there were any hidden patterns. The research aims to use Valence Aware Dictionary and Sentiment Reasoner, commonly known as VADER [3], an advanced lexicon rule-based sentiment analyzer which is specialized in analyzing social media texts. The proposed flow for Twitter sentiment analysis is shown in Fig. 78.1.

78.2 Related Work

Analysts have obtained a rule-based model for sentiment analysis called VADER [3]. They discovered that with a sentiment lexicon and several syntax rules, their model could exceed both individual human raters and machine learning methods. The ideas that people express in social media are also related to having an impact on people's choices of political parties [4]. Some investigations have proved to predict political election results with sentiment analysis of tweets using the lexicon method for the Swedish elections [5]. In the study concerning the Brazilian elections, Oliveira et al. [6] examined whether sentiment analysis of data from Twitter could appraise the citizen's political decisions, as public sentiment polls do. Their results were positive. Lexicon-based approach uses lexicon and then a chi-square analysis to classify new tweets. The method presented by Barbosa and Feng [7] classifies in subjective and objective terms, then abandons the actual tweets and organizes the first group as positive or negative. Tumasjan et al. [8] examined the influence of Twitter and tweet sentiment in the 2009 German national parliament election. Chatterjee [9] performed sentiment analysis on two parties BJP and Congress, using crawling Twitter data through APIs. They have tried to predict different moods on parties using moods, applying standard machine learning algorithms and deep learning to do multiclass mood classification for two prominent parties in the election.

78.3 Data Extraction and Data Preprocessing

78.3.1 Data Collection

Twitter provides an API for developers to integrate it with their existing applications. Developers use Twitter API to extract the data of user tweets to generate insights. Twitter has both the REST API and Streaming API for data extraction. However, there are limitations to it; the data cannot be fetched above 90 days from the current date, which means historical data of last year cannot be fetched with Twitter API. In this paper, we used one such Twitter archiver [2] source to extract the historical data for the relevant keywords from February 1, 2019 to May 23, 2019. This is the time wherein most of the events took place for the 2019 Loksabha Elections. The data was downloaded as a compressed ".tar" file for each day of the month. The dataset downloaded from the source was huge since it contains the user tweets from over the world for the mentioned period; the average size of data per month was approximately equal to 50 GB. Also, there were challenges to extract data from the JSON file, which is a semistructured data file format, also popularly known as data interchange file format. After filtering the user tweets based on the keywords for each party and its leader, the raw dataset contains approximately 2 million tweets and is stored as a CSV file for further preprocessing.

78.3.2 Data Preprocessing

Data preprocessing is the fundamental step in any text mining analysis. It is the required step before performing natural language processing (NLP); without data preprocessing, it can be difficult to calculate the polarity of any text data and we may end up getting incorrect results. The dataset extracted is a linguistic data from Twitter, which means it can contain multiple dialects in a single user tweet as well as other noise like repeated words, slangs, URLs, images, extra whitespaces, etc. All this noise in the user tweets makes it difficult in performing sentiment analysis. Therefore, it is necessary to remove noise and irregular data, so that the true meaning of a user text can be understood. Moreover, data preprocessing also reduces the overall size of the dataset as the noise is removed from the data which in turn reduces the overall execution time as well as improve the accuracy of calculating the sentiment of the tweet. Table 78.1 describes the algorithm implemented in Python for cleaning and removing noise from each user tweet.

Table 78.1 Data preprocessing algorithm

Input—User tweets **Output**—Processed and cleaned user tweets **For** each user tweet in **dataset**:
1. Remove all https:// or URL using regular expression methods 2. Replace all '@username' with the word 'username' 3. Filter All #Hashtags and RT from the tweets 4. Look for repetitions of two or more characters and replace with the character itself 5. Filter all additional special characters (: \; { } - | [] + ( ) ? ! @ # < > % * ,) from the tweets. This also includes striping the extra white spaces 6. Removing stop words like I, am, each, the, and, etc. 7. Replace wrong contractions with custom Python dictionary containing mapping of correct English contractions 8.Replacing keywords related to political parties
Return processed tweet

78.4 Sentiment Analysis

Humans are termed as social beings, people live across different countries in the world, and it is a diverse environment. Twitter itself has more than 300 million active users monthly. People use these platforms to communicate their opinions and emotions freely. They use common languages like English, Hindi, French, Spanish, Japanese, etc. Communication by means of language is referred to as linguistic communication [10].

78.4.1 Natural Language Processing Toolkit

Natural Language Processing Toolkit (NLTK) in a Python programming language is a framework which allows developers, researchers, and analysts to perform linguistic data processing. NLTK is a set of modules and corpora published under the GPL open-source license, which authorizes users to learn and to conduct research in NLP [11]. It provides an interface to various lexical resources like WordNet, SentiWordNet, and even advanced corpuses like VADER, which is specifically used for processing and analyzing social media texts. In combination with corpuses and lexical resources, it provides functionality like word tokenization, stemming, tagging, parsing, stop words filtering, etc.

78.4.2 Sentiment Lexicons

A sentiment lexicon is a record of lexical features, e.g., words commonly labeled according to either positive or negative semantic orientation [12]. Lexicons are the core part of any sentiment analysis procedure; the discussion will be presented for majorly two types of sentiment lexicons, one polarity-based and other valence-based. Furthermore, the discussion will be concluded by how VADER uses both the approaches in combination to calculate the polarity of social texts efficiently and accurately, which prior two approaches are not good at handling.

78.4.2.1 Sentiment Orientation (Polarity-Based) Lexicons

Linguistic Inquiry and Word Count (LIWC) is a computer program that was built to analyze the various components in a sentence such as emotional, cognitive, structural, and process the one that is present in the text samples. LIWC uses its own dictionary (76 categories and around 4500 words). In [13], the authors pointed out, assessing sentiment in a social media text, LIWC does not include an application for sentiment lexical analysis parts such as "acronyms, emoticons, initialisms, or slang" which are known to be essential for sentiment analysis of the social text [14].

78.4.2.2 Sentiment Intensity (Valence-Based) Lexicons

Valence-based lexicons allow us to find the sentiment intensity in text. It is very crucial and helpful to understand how the sentiment intensity for a product has changed over a period of time. The extension of the WordNet database is Senti-WordNet, which provides the valence strength of the text. Each word in the sentence is attached with a numerical score, which ranges between 0 and 1. These scores are calculated using highly complex semi-supervised algorithms. But often, SentiWordNet fails to determine the sentiment intensity for social media text correctly.

78.4.2.3 Vader Aware Dictionary for Sentiment Reasoning (VADER)

The major short coming of the two approaches discussed was that they are unable to handle social media texts. VADER promises to leverage the benefits of rule-based modeling for calculating the sentiments of text by adding the support of social media style texts, yet able to efficiently handle textual data from other domains too. In [13], the authors point out that the VADER lexicon performs very well in the social media domain. The correlation coefficient determines that VADER ($r = 0.881$) performs as well as singular human raters ($r = 0.888$) at rivaling ground truth (aggregated group mean from 20 human raters for sentiment intensity of each tweet). Even though the corpus used in VADER has more than 9000 lexical features which is higher than other gold standard lexical resources, it is still fast enough to be used online with streaming data and does not have any performance lag. VADER sentiment lexicon administers both polarity and strength of sentiments shown in social media contents. It provides a "normalized, weighted composite score" also called as compound score which is the sum of valence scores of each word in the lexicon, and then it is normalized to a range of $(-1, 1)$ where -1 is most negative and $+1$ is most positive.

In this paper, VADER has been used to calculate the sentiment polarity and intensity of every user tweet. VADER provides the compound scores; the user tweets are then classified into positive, negative, and neutral categories based on the threshold value. In this analysis, the value of 0.1 is used. The rules for labeling each user tweet into sentiment category based on the threshold are shown below:

- Positive sentiment: compound score ≥ 0.1
- Negative sentiment: compound score ≤ -0.1
- Neutral sentiment: $-0.1 <$ compound score ≤ 0.1.

78.5 Results and Discussion

78.5.1 Results

The election campaigns picked up from the month of March 2019 to May 2019, but to understand the transition and analyze campaign growth, the data was extracted from user tweets from February 1, 2019 to May 23, 2019. The dataset contains approximately 200,000 tweets; this is after filtering the tweets based on the keywords listed in the data collection section. After conducting data preprocessing by cleaning the raw tweets and mapping the tweets to the party and its candidates, Tables 78.2 and 78.3 show the distribution of tweets by party and candidate.

The distribution of user tweets shows that during the election campaign, BJP and Congress were the two major parties for whom the public was showing more interest, followed by AAP and BSP.

Table 78.2 Tweets by party

Party	Number of tweets
BJP	154,667
Congress	33,260
AAP	8497
BSP	185

Table 78.3 Tweets by candidate

Candidates/party	Number of tweets
Narendra Modi/BJP	135,822
Rahul Gandhi/Congress	27,868
Arvind Kejriwal/AAP	20,222
Mayawati/BSP	161

78.5.1.1 Analyzing Change in Sentiment

Figure 78.2 displays the average sentiment value (its compound sentiment value or sentiment strength provided by VADER) for each month during the campaign period. Figure 78.3 shows that the overall sentiment of user tweets for the candidate Narendra Modi has doubled since election campaigns started. The candidate appointed by Congress, Rahul Gandhi, has a significant drop in the overall sentiment; this is the

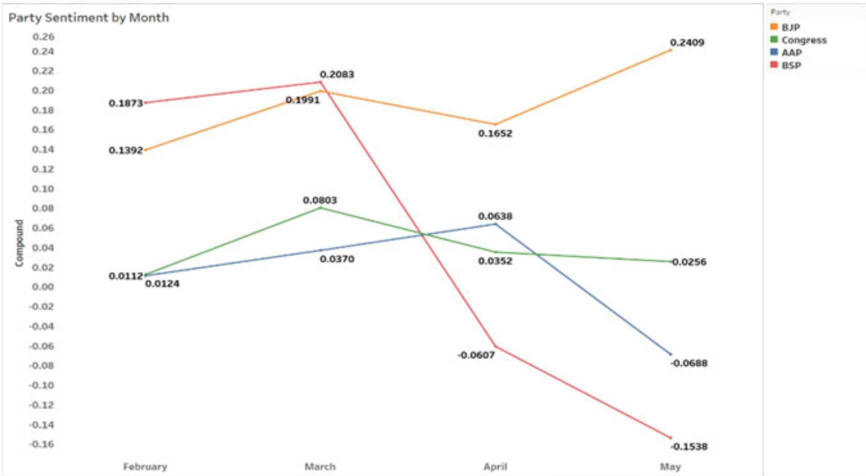


Fig. 78.2 Change in party sentiment for each month

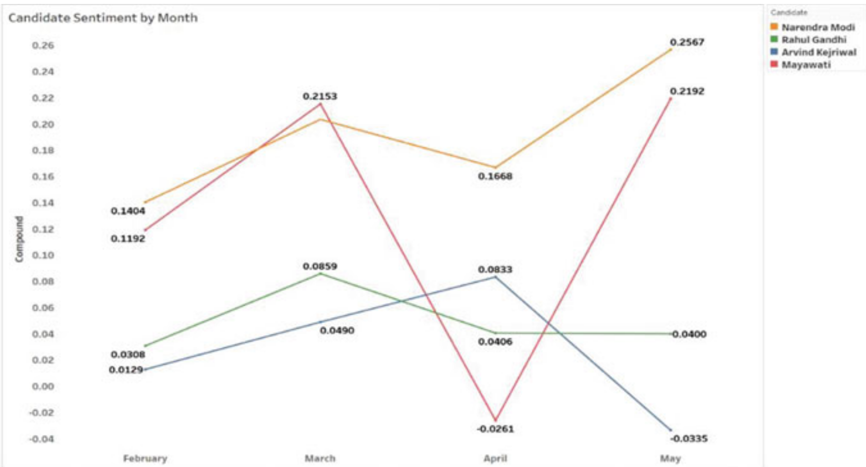


Fig. 78.3 Change in candidate sentiment for each month

same for Arvind Kejriwal, AAP, for whom the overall sentiment was negative for the month of May.

Mayawati, BSP, although sees a sharp rise in positive sentiment from April to May, again, the overall number of user tweets for Mayawati is much lower than other candidates; therefore, those numbers are not significant concerning other candidates.

78.5.2 Discussion

The sentiment analysis was conducted using VADER for calculating the polarity and sentiment intensity of text. Tables 78.4 and 78.5 show the distribution of sentiment score for the party and its candidates in percentage since the number of tweets extracted for each party and candidate is different, so it is incorrect to compare the absolute count of tweets for each of positive, negative, and neutral polarity, respectively.

In the overall results shown in Tables 78.4 and 78.5, the positive sentiment for BSP party is 22.94%, whereas its candidate Mayawati has a positive sentiment of 38.29%. There is an approximately 15% gap between positive sentiment between BSP Party and its leader Mayawati as the number of tweets distribution for BSP and Mayawati is very low compared to all other parties.

It is difficult to compare party and its leader in this case as people show more negative sentiment for other BSP candidates except for Mayawati, as a result the overall party sentiment is more negative than its leader. That is why BSP sentiment is showing 49.54% negative and only 22.94% positive. But as people think positive for Mayawati, so candidatewise, Mayawati gets a positive sentiment of 38.29%. The tweets distribution for other parties and its leaders is more than BSP and its leader. BJP, Congress, and AAP and its leaders Narendra Modi, Rahul Gandhi, and Arvind Kejriwal have consistent results with the tweet distribution. It can be clearly

Table 78.4 Distribution of tweet sentiment by party

Party	Positive (%)	Negative (%)	Neutral (%)
BJP	50.02	18.94	31.03
Congress	35.78	27.80	36.42
AAP	36.98	38.51	24.51
BSP	22.94	49.54	27.52

Table 78.5 Distribution of tweet sentiment by candidate

Party	Positive (%)	Negative (%)	Neutral (%)
Narendra Modi	50.51	18.54	30.96
Rahul Gandhi	36.80	27.39	35.82
Arvind Kejriwal	36.65	34.16	29.19
Mayawati	38.29	36.57	25.14

inferred from Tables 78.4 and 78.5 that both BJP and its leader Narendra Modi have higher positive sentiment. Also, the number of tweets extracted based on keywords is more than the other parties and their leaders, which shows Narendra Modi had more popularity and positive opinion among the people. Due to higher popularity of Narendra Modi among people and since he was the existing Prime Minister during the 2019 elections, BJP won the election with a whopping 303 seats out of contested 542 seats. Rahul Gandhi, the leader of Congress party, failed to reserve the majority and in contrast, Congress was able to win merely 52 seats.

78.6 Conclusion and Future Work

78.6.1 Conclusion

In this paper, our target was to conduct Twitter sentiment analysis for the 2019 Indian Loksabha Election for major parties and their candidates. User tweets were extracted from Twitter archiver between February 1, 2019 and May 23, 2019 based on keywords and hashtags related to the trend observed during the election period. A model was proposed to perform sentiment analysis on a linguistic dataset, highlighting the use of the VADER sentiment analyzer specifically attuned to calculate the polarity of text expressed in social media rather than using a traditional dictionary-based approach that maintains WordNet for positive, negative, and neutral polarity keywords. It was seen that people had higher positive sentiment and emotions for Narendra Modi and its party BJP than Congress, AAP, and BSP. The results of sentiment analysis follow the actual election results which resulted in BJP winning the elections by a majority vote.

78.6.2 Future Work

The future scope of this work can include more keywords for the data extraction process and increase Tweet distribution for the party and its candidates and more parties and candidates can be added in the tweet extraction. Additionally, changes in data extraction techniques can be made by using advanced big data frameworks like Apache Spark, as such sophisticated event processing engines can process vast amounts of data. The data size of social media texts is growing manifolds and using Apache Spark or Spark Streaming will help us perform the extraction, transformation, and loading in near real time. In this study, the focus was on Twitter, but more social networks could be added. Additionally, it would be interesting to create datasets with a more balanced distribution of tweets from different classes with a balancing algorithm.

References

1. 2019 Indian general election. https://en.wikipedia.org/wiki/2019_Indian_general_election
2. Twitter Archive: https://archive.org/details/twitterstream
3. VADER Sentiment Analyzer. https://github.com/cjhutto/vaderSentiment
4. Eirinaki, M., Pisal, S., Singh, J.: Feature-based opinion mining and ranking. J. Comput. Syst. Sci. **78**(4), 1175–1184 (2012). https://doi.org/10.1016/j.jcss.2011.10.007
5. Liu, B.: Sentiment Analysis and Opinion Mining. Morgan & Claypool Publishers (2012)
6. Oliveira, D.J.S., de Souza Bermejo, P.H., dos Santos, P.A.: Can social media reveal the preferences of voters? A comparison between sentiment analysis and traditional opinion polls. J. Inform. Tech. Polit. **14**(1), 34–45 (2017). https://doi.org/10.1080/19331681.2016.1214094
7. Barbosa, L., Feng, J.: Robust sentiment detection on twitter from biased and noisy data. In: Proceedings of the International Conference on Computational Linguistics (COLING), pp. 36–44, Beijing, China, Aug 2010
8. Tumasjan, A., Sprenger, T.O., Sandner, P.G., Welpe, I.: Predicting elections with Twitter: what 140 characters reveal about political sentiment. In: Proceedings of the Fourth International Conference on Weblogs and Social Media, ICWSM 2010, Washington, DC, USA, May 23–26, 2010, The AAAI Press (2010)
9. Chatterjee, S.: https://hackernoon.com/twitter-sentiment-analysis-for-the-2019-lok-sabha-elections-b43o320e
10. Bird, S., Klein, E., Loper, E.: Natural Language Processing with Python. O'Reilly Media, Inc, USA (2009)
11. Bird, S.: NLTK: the natural language toolkit. In: Proceedings of the COLING/ACL on Interactive Presentation Sessions (COLING-ACL '06). Association for Computational Linguistics, USA, pp. 69–72. http://doi.org/10.3115/1225403.1225421
12. Liu, B.: Sentiment analysis and subjectivity. In: Indurkhya, N., Damerau, F. (eds.) Handbook of Natural Language Processing, 2nd ed. Chapman & Hall, Boca Raton, FL (2010)
13. Hutto, C.J., Gilbert, E.E.: VADER: a parsimonious rule-based model for sentiment analysis of social media text. In: Eighth International Conference on Weblogs and Social Media. Presented at the Proceedings of the Eighth International AAAI Conference on Weblogs and Social Media, Ann Arbor, MI (2015)
14. Davidov, D., Tsur, O., Rappoport, A.: Enhanced sentiment learning using Twitter hashtags and smileys. In: ICCL-10 (2010)

Chapter 79
A Processing of Top-k Aggregate Queries on Distributed Data

Trieu Minh Nhut Le, Muoi Le Thi Be, and Dung Hoang Thi Ngoc

Abstract Top-k queries are crucial tools on data analysis, data mining, and decision making. Investigating top-k processing on distributed data has been interested by many researchers which eliminates all the redundancy answers. In this paper, we review and evaluate the process of top-k queries on distributed data. The algorithms of top-k queries on distributed data are based on the data access methods namely random and sequential access which will be discussed and analyzed for effectiveness. Then, the execution of two algorithms, best position algorithm and no random access, is about the data access modes being presented with specific examples. In experiments, the results show that the top-k aggregate queries using the sequential access better than using the random access in terms of the run-time execution. The effect algorithm is no random access (NRA) using sequential access, while best position algorithm (BPA) using the random access to calculate the results.

79.1 Introduction

In recent years, the huge increase in the number of data is a reason for overwhelming a large number of uninteresting answers being resource consuming. The top-k query is a selection of the k-most interest answers which attracts many researchers [1–3]. The k best answers of data are returned by the top-k query, which is a useful tool in many applications such as electronic markets, logistics (commercial shipping), prediction program [4, 5], and decision support system [6]. Data is stored on multiple data sources in various locations and different database management system. In managing distributed database system, users can query the data across multiple interconnection databases [7, 8]. One of the most important queries is the top-k aggregate query on distributed data. It is explained that supposing, dataset contain m lists of n items. Each item has local scores in their lists. The overall score of each item (ID) is calculated in all lists by the aggregate function. The top-k aggregate query returns the k highest overall scores of items, which has been studied and developed in many researches

T. M. N. Le (✉) · M. L. T. Be · D. H. T. Ngoc
Information Science Faculty, Sai Gon University, Ho Chi Minh City, Vietnam

T. Senjyu et al. (eds.), *IOT with Smart Systems*, Smart Innovation, Systems and Technologies 251, https://doi.org/10.1007/978-981-16-3945-6_80

such as threshold algorithm [9], best position algorithm [10], no-random algorithm [6, 11], and database transaction process [12].

On distributed data, access methods affect to data process of top-k queries of aggregation operators (such as COUNT and SUM). The random and sequential methods have described the data access. The random access denotes ability to access the data at random, while the sequential is accessed data in order. Akbarinia et al.[10] proposed the best position algorithm (BPA) for top-k queries to optimize the threshold algorithm (TA) by stopping earlier process, and the BPA processes data with being supported the random access method. While the others are being limited the access method which just using the sequential access, therefore, the paper [6, 11] proposed the no-random access algorithm (NRA) for this problem.

In this paper, we present and review some general concepts about top-k aggregate queries on distributed data and some specific examples to illustrate these concepts. Then, we will analyze and explain the best position algorithm and the no-random access algorithm with specific examples to demonstrate the random access method and sequential access method for top-k aggregate query on distributed data. Section 79.3 is discussion and evaluation about the effectiveness of these algorithms. In the experiment section, the generated data is applied to finalize the effectiveness of algorithms in terms of run time.

79.2 Preliminary

In this section, we present concepts about top-k aggregate queries on distributed data and an answers set of top-k aggregate queries on specific example.

Given the dataset (D), D includes a number of lists ($L_i = \{L_1, L_2, ..L_n\} \in D$, $i = \overline{1..n}$), each list has m tuples ($t_j = \{t_1, t_2, \ldots, t_m\}$, $j = \overline{1..m}$), for example, tuple t_{23} is at row 2 of list 3 in Table 79.1. The aggregated operator is one of monotone functions such as COUNT, SUM, AVERAGE, and so on. Overall score is calculated by applying aggregated functions ($f(t_{ji})$), which are used in selecting the top-k tuples in multiple resources. Users are commonly interested in the top-k answers from distributed dataset. In top-k query, the parameter k is defined by users.

The top-k query of aggregated operation on distributed data returns a set of k tuples having the k highest aggregated scores with the same identifier (ID) among dataset (D) [1, 6, 10, 13]. The aggregated function is a monotone function with variables from multiple sources having the same identified tuple, which is defined by users.

$$\text{top-}k(D) = \{\text{top}\{t_1, .., t_k\}, D\colon \forall t_a \in \text{Top}, \ \forall t_b \in (D\backslash\text{Top}) \Longrightarrow f(t_a) > f(t_b)\}$$

Table 79.1 Three sorted lists of the distributed data

Row	L1		L2		L3	
	ItemID	Score	ItemID	Score	ItemID	Score
1	A6	26	A3	24	A2	26
2	A8	24	A4	23	A5	25
3	A9	23	A14	21	A8	24
4	A3	22	A5	20	A4	21
5	A7	21	A7	19	A3	20
6	A4	19	A1	17	A6	15
7	A5	13	A8	16	A13	11
8	A1	10	A2	10	A1	10
9	A2	7	A6	9	A9	8
10	A11	6	A9	8	A7	7

where:

- D is a dataset containing n lists $L_i \in \{L_1, L_2, \ldots L_n\} = D$.
- $f(t)$ is an aggregated function (COUNT, SUM, AVERAGE, or mono-function).
- t is a multi-tuple which has the same identified number in n lists (D).

For the example in Table 79.1, dataset contains three sorted lists (L_1, L_2, L_3). The top-3 tuples with SUM function on Table 79.1 is queried by users. The process of this top-3 query returns three tuples(($A3$, 66), ($A8$, 64), ($A4$, 63)) or {$A3$, $A8$, $A4$} which have three highest scores {66, 64, 63}, respectively, where $A3$'s overall score equals SUM(22, 24, 20), $A8$'s overall score equals SUM(24, 16, 24), and $A4$'s overall score equals SUM(19, 23, 21) in dataset D.

For the process of top-k query on distributed data, it is assumed that the data being sorted in each list. The Naïve algorithm is looking at each tuple in sorted lists to calculate overall score using aggregated function and comparing with the others to obtain the top-k answers. In the middleware, the Naïve method is not efficient for applying on large databases. In the next section, we first study about the access and sequence modes, before implement the algorithms top-k query more suitable than Naïve method in middleware.

79.3 Processing of Top-k Query According to Access Methods

79.3.1 Random Access and Sequential Access

In this section, the top-k processing has been discussed due to the different access methods. In the middleware system, the process of top-k query is performed on number of lists which can be at various location or separated sources. In each list, data is sorted according to local scores. A score of the aggregated function is calculated by local scores in all lists having the same identifier. The top-*k* answer set contains the k highest scores of data tuples. There are two main access methods to select data scores namely random access and sequential access on sorted data list.

The sequence access is data being read or written sequentially by computer system, which means the data being accessed following in a logical order [6, 11]. The random access is the ability to read or write information anywhere in the database. It means that the computer system uses indexing to search and find the information [10]. It is clear that random access is more time consuming than sequential access method, when accessing a data unit [9, 14].

In real life, the various users' requirements of data access have been presented to express data sources and locations. It is assumed that some middleware applications are available for random access, while others are restricted to sequential access. First, assuming data is being ordered on score, and randomly accessed by its identifier. The best position algorithm (BPA) [10] for top-*k* query was proposed to optimize the calculation of threshold from threshold algorithm [9]. Secondly, assuming data is being accessed sequentially ordered on scores. The NRA algorithm was proposed to find the answer of this assumption. In the next section, we will present and analyze the two above algorithms to discuss and find the suitable method for some special cases.

79.3.2 The Best Position Algorithm Using both Sequence and Random Access on Sorted Data

First section is the implementation of the best position algorithm [10], and the pseudocode is presented as algorithm 18.1 and explained as follows.

It is assumed that D contains n sorted lists, denoted as $L_i \in D$. Each list (L_i) consists of m data items (also called tuples or objects). Each data item in a list (L_i) has a unique ID and a local score $(score_{pi})$ which is a non-negative real number. The $score_{pi}$ is a local score at row (p) in list (i). All lists are sorted in descending orders according to data item's local score. What expected of the output is a set of k data items having the highest overall score called $topkSet$. The overall score of each data item is computed by using aggregated functions on the local scores using random access in every list (line 7).

The best position algorithm also includes a variable called lambda (λ) that works like a threshold in threshold algorithm [9]. Initially, λ is set as infinite (line 1). A set $(accessSet)$ is used to store distinct tuples that have been seen, and it is initiated as an empty set (line 2). An answer set $(topkset)$ is used to store the k highest score tuples. First, for each list L_i (line 4), the $currentAccess$ set contains all current tuples in all lists (L), there are n members in $currentAccess$ set which cooperate with n List (L_i),respectively. $t_{position}$ is assigned by direct access to n sorted lists which is also

ALGORITHM 79.1: Best Position pseudo code

```
   Input: D, k, L = {L_1, ..., L_n}
   Output: topkSet
1  λ ← ∞;
2  topkSet ← {};
3  accessSet ← {};
4  foreach (list i{L_i ∈ L}) do
5      if (t_position NOT IN accessSet) then
6          currentAccess(p, i) ← t_position(ID, score_pi); // direct access in every list L_i
7          overallScore(ID) ← aggFunction(ID, score_p,i(1..n)); // using random access
8          if (topkSet.Count() < k) then
9              topkSet ← topkSet ∪ t(ID, overallScore);
10         else
11             if (Min(topkSet(ID, overallScore)) < overallScore(ID)) then
12                 topkSet ← topkSet\Min(topkSet);
13                 topkSet ← topkSet ∪ t(ID, overallScore);
14         accessSet ← accessSet ∪ t(ID);
15     while (t_position IN accessSet) do
16         λ ← aggFunction(currentAccess, score_p,i(1..n)); if (Min(topkSet) ≥ λ) then
17             Exit;
18         currentAccess(p, i) ← t_position(ID, score_pi);
19         t_position → next;
20 Return topkSet;
```

added into the *currentAccess* set (line 5). In line 6, the *ID* of $t_{position}$ is checked in the *accessSet* set. If it has not been accessed yet, the overall score of $t_{position}$ is calculated by aggregated function. Using random access to retrieve local scores from the others lists $(L \backslash L_i)$ with the same *ID* computes the aggregated function (line 7). From line 8 to line 12, the *topkSet* is assigned with tuple having the highest overall score among the ones in *accessSet*. In order to stop algorithm earlier, λ is calculated based on the aggregated function with *currentAccess*'s members (line 16). If λ is less than or equals to the minimum score of the tuples in *topkSet* (Min(*topkSet*)) (line 17), it means that the *topkSet* contains the k tuples with the highest overall score, the others cannot be in the *topkSet*, else, $t_{position}$ moves to the next row as a new position for repeat check (line 19, 20) or jump to the next list (line 5). Finally, *topkSet* returns to the users' query (line 21).

For specific example, we assume that the requirement is the top-3 query on data in Table 79.1. To illustrate the BPA, we consider the three sorted lists of distributed data shown on Table 79.1. Table 79.2 shows that the overall score and lambda (λ) are calculated in BPA. The BPA finds the data item ID $= A6$ with overall score 50 being SUM of all local scores by using random access to L_2 and L_3 (SUM(26, 9, 15) = 50), and (λ) equals ∞. At step 2, $A3$ is *ID* of current tuple with overall score (66) being SUM (22, 24, 20), and λ equals ∞. Step 3 has the current tuple ($A2$, 43). In this step, k is 3, so the assumed *topkSet* contains three tuples with *ID* $A6$, $A3$, $A2$, λ is sum of all scores of current tuples in all lists (SUM(26, 24, 26) = 76). Checking the stop condition, $\lambda(76)$ is greater than Min(*topkSet*) ($A2$, 43), the process jumps to step 4. The current access tuple is ($A8$, 64) and λ is 74 $(24 + 24 + 26)$. The *topkSet*

Table 79.2 Steps to calculate the overall scores in BPA.

Step	tuple $(ID, score)$	Overallscore	Lambda (λ)	$topkSet$
1	$(A6, 26)$	50	∞	$A6$
2	$(A3, 24)$	66	∞	$A3, A6$
3	$(A2, 26)$	43	76	$A3, A6, A2$
4	$(A8, 2)$	64	74	$A3, A8, A6$
5	$(A4, 23)$	63	73	$A3, A8, A4$
6	$(A5, 25)$	58	72	$A3, A8, A4$
7	$(A8, 24)$	Exit	71	$A3, A8, A4$
8	$(A4, 21)$	Exit	68	$A3, A8, A4$
9	$(A3, 20)$	Exit	67	$A3, A8, A4$
10	$(A6, 15)$	Exit	62	$A3, A8, A4$

is updated by replacing $A2$ with $A8$, so now it contains $A6$, $A3$, $A8$. The algorithm proceeds to calculating the overall score of $A4$ 63 in L_2 since the next tuple in L_1 ($A9$) has not been seen, λ is SUM ($A8$, $A4$, $A2$) = 73. The $topkSet$ sees another change as the overall score of $A4$ is greater than the minimum overall score value available in $topkSet$, which is $A6$, so it now contains $A3$, $A8$, $A4$. Since $A14$ in L_2 has not been seen, move on to $A5$ in L_3. The overall score of $A5$ is 58 and λ is 72. Continue to check the next tuples in L_3 from step 7 to 10, the data items $A8$, $A4$, $A3$, $A6$ have already been seen so their overall scores do not need to be computed again which results in a reduction in execution cost, λ is also computed in each step as seen in Table 79.2. The algorithm stops at step 10 as $\lambda = 62$ is now less than MIN ($topkSet$) = ($A4$, 63). As a result, the algorithm returns $topkSet$ with 3 tuples $A3$, $A8$, $A4$.

79.3.3 The No-Random Access Algorithm on Sorted Data

NRA algorithm requires the same input as BPA to process, which are a dataset (D) containing n sorted lists and k value representing the number of data items to be returned. The result is expected to be a set of k data items called $topkSet$. The no-random access algorithm is described with pseudocode algorithm 18.2 as follows.

When the algorithm initiates, $topkSet$ is declared as an empty set along with two other sets, $currentRow$ and $accessRow$ (line 1, 2, 3). $currentRow$ is used to store tuples of the same row that is currently seen using sequential access to all lists, whereas $accessRow$ contains tuples with unique IDs that have already been seen in dataset D. It stores ID, local score in each of list, the lower-bound and upper-bound score of a tuple.

As the name implies, NRA only do sequential access throughout the whole dataset, which means it accesses each row of all sorted lists in D one by one (line 4). The row of tuples that it currently accesses to is temporarily stored in $currentRow$,

ALGORITHM 79.2: No Random Access pseudo code

Input: $D, k, L = \{L_1, ..., L_n\}$
Output: $topkSet$
1 $topkSet \leftarrow \{\}$;
2 $currentRow \leftarrow \{\}$; `// using sequential access to all lists`
3 $accessRow \leftarrow \{\}$; `//` $(ID, score_L 1, score_L 2..., score_L m, score_{lowBound}, score_{upBound})$
4 **foreach** $(Row_i\ in\ \{L_1...L_n\})$ **do**
5     $currentAccess(p, i) \leftarrow \{t_1(ID_1), t_2(ID_2), ..., tn(ID_n)\}$;
6   **foreach** $(t_c(ID)\ in\ currentRow)$ **do**
7   **if** $(t_c(ID) \notin accessRow)$ **then**
8             $accessRow \leftarrow accessRow \cup t(ID)$;
9   **else**
10             $Updates(accessRow(t_c(ID, score))$;
11   **foreach** $(t_c(ID)\ in\ accessRow)$ **do**
12         $Updates(accessRow(t_c(ID, score_{LowBound}, score_{upBound}))$;
13   **repeat**
14         $lowBound \leftarrow MAX(all\ t(score_{lowBound}) \in accessRow)$;
15         $upBound \leftarrow MAX(all\ t(score_{upBound}) \in accessRow)$;
16   **if** $(lowBound \geq upBound)$ **then**
17             $topkSet \leftarrow t(ID, lowbound) \in accessRow$;
18             $accessRow \leftarrow accessRow \backslash t(ID, lowBound)$;
19   **until** $(lowBound \geq upBound)$;
20   **if** $(topkSet.Count() \geq k)$ **then**
21   **Exit**;
22 **Return** $topkSet$;

including their IDs and all local scores (line 5). Next, check whether each tuple (t_c) in $currentRow$ is presented in $accessRow$ or not (line 7). If the tuple (t_c) is not in $accessRow$ set, it is added to $accessRow$ with its ID and current local score (line 8). Otherwise, the tuple's local score in a new list is updated in the $accessRow$ (line 10). Then, the algorithm proceeds to updating the lower-bound score and upper-bound score of each tuple in $accessRow$ (line 12). The lower-bound score of a tuple is computed by applying aggregate function on seen local scores, while the upper-bound score is computed by applying aggregate function on the scores in $currentRow$ and the deviation (difference between the seen local score and the last seen scores in the corresponding ranked lists). After updating all tuples in $accessRow$, the algorithm finds the tuple with the highest lower-bound score ($lowBound$) and the maximum upper-bound value ($upBound$) (line 14). If $lowBound$ is greater than or equals to $upBound$, the tuple with the $lowBound$ value is added to $topkSet$ and also removed from $accessRow$ (line 16, 17, 18). Simply put, if the lower-bound score of a data item is not below the upper-bound score of all other data items, it is bound to be the next top-k tuple. The process repeats if $lowBound$ is greater than or equal to $upBound$ (line 19). The algorithm continues until the number of $topkSet$ reaches k value, it then terminates and returns $topkSet$ as a result (line 20, 21, 22).

For specific example, top-3 results of data on Table 79.1 processing using NRA algorithm. Table 79.3 illustrates the no-random access algorithm (NRA) as follows.

To illustrate NRA, we use the same dataset of three sorted lists shown on Table 79.1. Table 79.3 describes the $score_{lowBound}$ and $score_{upBound}$ of each data item

Table 79.3 Steps to calculate the upper-bound and lower-bound in NRA

ID	Step1		Step2		Step3		Step4		Step5		Step6		Step7		Stop
	LB	UB	LB	UB	LB	UB	LB	UB	LB	UB	LB	UB	LB	UB	Stop
A6	26	76	26	74	26	71	26	67	26	65	41	58	41	53	
A3	24	76	24	73	24	71	46	67	66	65	Top-1				
A2	26	76	26	73	26	70	26	68	26	66	26	62	26	55	
A8			24	72	48	69	48	65	48	63	48	56	64	51	Top-2
A4			23	72	23	70	44	66	44	64	63	57	63	57	Top-3
A5			25	72	25	69	45	67	45	65	45	61	58	54	
A9					23	68	23	64	23	62	23	55	23	50	
A14					21	68	21	64	21	62	21	55	21	45	
A7									40	60	40	53	40	48	
A1											17	51	17	41	
A13													11	40	

through each step of the algorithm. In step 1, NRA sees the first data item in each list ($A6$, $A3$, $andA2$), and they are added to the $currentRow$. After that, the tuples in $currentRow$ are assigned to the $accessRow$ set due to they have not contained in it. Now, each tuple stored in $accessRow$ is updated the $score_{lowBound}$ and $score_{upBound}$. The $score_{lowBound}$ of each data item is calculated by summing the score of data item with same ID stored in the $accessRow$. It means that we only see the local score of A6 in L_1, the lower bound of $A6$ is 26. Next, the $score_{upBound}$ is computed by summing all scores of tuples in $currentRow$ and deviation between considering tuple in $accessRow$ and tuple in $currentRow$ at the same List of considering tuple. It is explained that $score_{upBound}$ of $A6$ is computed as Sum((26, 24, 26), 0) = 76 without any depth due to the first accessed row. Applying the same to the other tuples, we have $A6$ (26, 76); $A3$ (24, 76); $A2$ (26, 76). As can be seen in step 2, the $score_{lowBound}$ of lasted accessed tuples in $accessRow$ do not change because there is no tuple with same ID appears. In addition, data items with $ID = \{A8, A4, A5\}$ have $score_{lowBound}$ equal to $\{24, 23, 25\}$. On the other hand, the $score_{upBound}$ of tuples in $accessRow$ are updated due to the existing of deviation. The $score_{upBound}$ of the access tuple ($A6$) is sum of the all scores of current tuples $\{A8, A4, A5\}$ (72) and the deviation between $A6$(26) and $A8$(24) in $A6$'s list (L_1) is 2 (26-24) so that $score_{upBound}$ of $A6$ is calculated by Sum((24, 23, 25), 2) = 74, likely to the $score_{upBound}$ of the others. Continue to step 3, the $currentRow$ is now consisting $\{A9, A14, A8\}$. It is noticeable that tuple $A8$ has already existed in $accessRow$ so that the $score_{lowBound}$ is recalculated by adding new score to previous score which is 48 ($24 + 24$). Moreover, because $A8$ having deviation between $A8$ (L_1 of $accessRow$) and $A9$ (L_1 of $currentRow$) being 1 ($24 - 23$). Therefore, $score_{upBound}$ of $A8$ is Sum((23, 21, 24), 1) = 69. Similar to step 3, $currentRow$ is now consisting $A3$, $A5$, and $A4$ in step 4 and the $score_{lowBound}$ of existed tuples in $accessRow$ will be updated as $A3$ (46), $A5$ (45), $A4$ (44). In this step, due to the existing of deviation of $A8$ in L_1 (between $A8$ and $A3$) = 2, the $score_{upBound}$ of $A8$ is Sum((22, 20, 21), 2) = 65. Step 5 the $currentRow$ $\{A7, A7, A3\}$ is retrieved, the $score_{upBound}$ and $score_{lowBound}$ are updated as previous steps in which the MAX(all $t(score_{lowBound}) \in accessRow \geq$ MAX(all $t(score_{upBound}) \in accessRow)$ (66 = 66) so that the data item A3 with $\max_{lowBound} = 66$ is assigned to $topkSet$ $A3$ and also removed from $accessRow$. After that, MAX($lowBound$) and MAX($upBound$) in $accessRow$ are assigned again. Due to the condition is not satisfied, it moves to step 6. When the calculation is finished, we can see the $\max_{lowBound} = 63$ and $\max_{upBound} = 62$ so that the data item with that $\max_{lowBound}$ which is $A8$ is assigned to $topkSet$ consisting $\{A3, A8\}$. Doing the similar step until step 7, The algorithm returns $topkSet$ consisting $\{A3, A8, A4\}$ due to having enough k tuples in $topkSet$, NRA stops at step 7.

79.3.4 Evaluation the Two Above Algorithms

There are many studies that have been investigated to find the answer set of top-k query in relation database. The Naïve algorithm runs step by step from the first tuple to the last tuple (m tuples). In each step, the overall scores are calculated from all local scores (n list) in lists to select the top k answers. Therefore, it is time consuming to select the results using the Naïve algorithm. The complexity of Naïve method is $O(m * n)$.

There are two effective algorithms of top-k aggregate queries using the random access. It is true that the best position algorithm is more effective than threshold algorithm and Naïve algorithm [10] which uses the threshold value by calculating the local score of currently accessed tuples in each list, and this prevents the process from getting the result early. The main contribution of best position algorithm is to calculate the position to stop the algorithm earlier. To process this algorithm, the scores of all items in each list are being sorted and being supported with random access to all lists. For the reason, we present the best position algorithm for the system supporting the random access.

For another assumption of only supporting the sequences access in some systems, it means that the top-k aggregate process is being provided with only sequential access on all lists. In this case, the no-random access algorithm [6, 9, 11] is presented. The NRA also calculates the lower-bound and upper-bound values to stop the algorithm earlier when the lower-bound equals the upper-bound. In that point, at least one answer in top-k set will be obtained.

To compare two algorithms, best position algorithm and no-random access algorithm, firstly, we consider the complexity of random access and sequences access. It is true that, the sequence access is better than random access when a data unit is read in data system [10]. As we can see the pseudocode of best position algorithm and no-random access algorithm, both algorithms are a little difference on complexity. It seems that the best position algorithm is bit better than no-random access algorithm, however the random access is more complexity than sequence access. Therefore, when the number k of top-k queries increases, the time complexity of best position technique is increasing due to the rise of using random access. Thus, the no-random access algorithm is a little better than the best position algorithm in terms of the complexity Which will be proved in the next section.

To conclude this evaluation, due to servers, sources, or techniques of management system, random access and sequential access are supported in middleware system. For the random access on sorted data, the best position algorithm is applied, while for the sequential access on sorted data, the no-random access algorithm is processed.

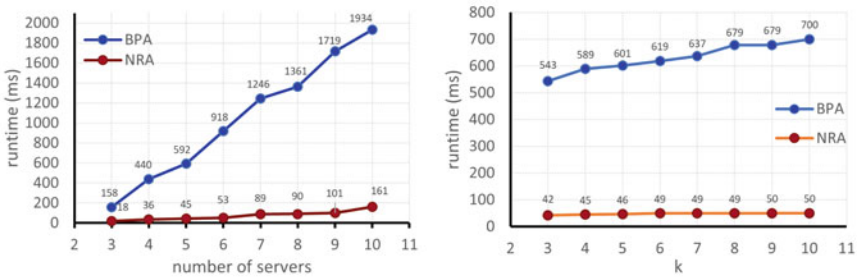
79.4 Implementing Experiments

In this section, we implement the best position algorithm and no-random access algorithm over synthetic data to evaluate the effectiveness of these algorithms. The experiments were conducted on PC with an Intel Core i5 6300U CPU 2400 GHz (4CPU), 16 GB RAM, and 240-GB SSD, running Windows 10 Professional Operating System. The implementation codes used C sharp language on database being stored in SQL server. To implement for these algorithms, the random access is coded as a SQL command to read data with item ID, while the sequence access is read directly on data. Therefore, the random access is more time consuming than the sequence access.

79.4.1 Synthetic Data

We generated a dataset with 10 relations, each relation contained over 600 records having two main attributes (ID, score). The dataset was generated following the normal distribution on score from 100 to 100,000. The aggregated function in this query is SUM function. The algorithms were executed on this generated dataset. We obtained the result as follows:

In Fig. 79.1, we analyze the performance of BPA and NRA. Figure 79.1a shows the number of servers that was increased to compare the run time of both algorithms. Overall, the run time of NRA is faster than the run time of BPA. The reason is that BPA using both random access and sequential access is more time consuming than NRA using only sequential access. Similarly, in Fig. 79.1b, we run both algorithms by increasing the value of k from 3 to 10, and the result shows that NRA is outperforming BPA on run time. As we can see, run time in both methods very lightly increases about 1 millisecond to 50 ms when increasing 1 value of k, so the run time is nearly no effect to the both algorithms according to the value k. Finally, NRA method is more efficient than BPA method on run time due to sequential access which is faster than random access on database management system.



(a) Runtime versus Number of servers (b) Runtime versus k

Fig. 79.1 BPA versus NRA

79.5 Conclusion

Top-*k* query on distributed data is a key issue on middleware system. In this paper, we first review concepts of top-*k* queries. Next, we present two current efficient algorithms, including best position algorithm and no-random access algorithm with specific examples. Then, we discussed and evaluated two algorithms with using the random access and sequential access modes to find the suitable method for some special cases. Our experimental result showed and confirmed that NRA using only sequential access is faster than BPA using both random and sequential access. It is explained that the random access costs more run time than sequential access. In future work, various techniques to extract information in data discovery and data mining will be broadly studied. In expanding of data such as uncertain data, probabilistic data, and fuzzy data, it is crucial to study the top-*k* queries on probabilistic distributed data with random access and sequential access. We also investigate and apply these techniques to get better performance in query processing and optimizing algorithms.

References

1. Le, T.M.N., Cao, J., He, Z.: Top-*k* best probability queries and semantics ranking properties on probabilistic databases. DKE J. **88**, 248–266 (2013)
2. Hua, M., Pei, J., Zhang, W., Lin, X.: Ranking queries on uncertain data: a probabilistic threshold approach. In: The ACM SIGMOD International Conference on Management of Data, pp. 673–686. Association for Computing Machinery (2008)
3. Le, T.M.N., Cao, J., He, Z.: Answering skyline queries on probabilistic data using the dominance of probabilistic skyline tuples. Inf. Sci. **340–341**, 58–85 (2016)
4. Ross, R., Subrahmanian, V.S., Grant, J.: Aggregate operators in probabilistic databases. J. ACM **52**, 54–101 (2005)
5. Lei, M., Zhang, X., Yang, J., Fang, B.: Efficiently approximating top-*k* sequential patterns in transactional graphs. IEEE Access **7**, 17–32 (2019)
6. Ilyas, I.F., Beskales, G., Soliman, M.A.: A survey of top-*k* query processing techniques in relational database systems. ACM Comput. Surv. **40**(4) (2008)

7. Zhou, L., Chen, Y., Li, T., Yu, Y.: The semi-join query optimization in distributed database system. In: ITCS 2012 (2012)
8. Shanbhag, A., Pirk, H., Madden, S.: Efficient top-*k* query processing on massively parallel hardware. SIGMOD, vol. 4, pp. 1557–1570. Association for Computing Machinery (2018)
9. Fagin, R., Lotem, A., Naor, M.: Optimal aggregation algorithms for middleware. J. Comput. Syst. Sci. **66**(4), 614–656 (2003)
10. Akbarinia, R., Pacitti, E., Valduriez, P.: Best position algorithms for top-*k* queries. In: The 33rd International Conference on VLDB, pp. 495–506. VLDB Endowm (2007)
11. Gong, N., Sun, G., Yuan, J., Zhong, Y.: Efficient top-*k* query algorithms using *k*-skyband partition. In: Infoscale (2009)
12. Cowling, J., Liskov, B.: Granola: low-overhead distributed transaction coordination. In: The USENIX Conference, p. 21. USENIX Association (2012)
13. Gao, J., Agarwal, P.K., Yang, J.: Durable top-*k* queries on temporal data. Proc. VLDB Endow. **11**(13), 2223–2235 (2018)
14. Li, F., Yi, K., Jestes, J.: Ranking distributed probabilistic data. In: The ACM SIGMOD International Conference on Management of Data, pp. 361–374. Association for Computing Machinery (2009)

Author Index

T. Senjyu et al. (eds.), *IOT with Smart Systems*, Smart Innovation,
Systems and Technologies 251, https://doi.org/10.1007/978-981-16-3945-6