Lecture Notes in Electrical Engineering 844

V. Bindhu
João Manuel R. S. Tavares
Ke-Lin Du *Editors*

Proceedings of Third International Conference on Communication, Computing and Electronics Systems

ICCCES 2021

Springer

Lecture Notes in Electrical Engineering

Volume 844

The book series *Lecture Notes in Electrical Engineering* (LNEE) publishes the latest developments in Electrical Engineering - quickly, informally and in high quality. While original research reported in proceedings and monographs has traditionally formed the core of LNEE, we also encourage authors to submit books devoted to supporting student education and professional training in the various fields and applications areas of electrical engineering. The series cover classical and emerging topics concerning:

- Communication Engineering, Information Theory and Networks
- Electronics Engineering and Microelectronics
- Signal, Image and Speech Processing
- Wireless and Mobile Communication
- Circuits and Systems
- Energy Systems, Power Electronics and Electrical Machines
- Electro-optical Engineering
- Instrumentation Engineering
- Avionics Engineering
- Control Systems
- Internet-of-Things and Cybersecurity
- Biomedical Devices, MEMS and NEMS

For general information about this book series, comments or suggestions, please contact leontina.dicecco@springer.com.

To submit a proposal or request further information, please contact the Publishing Editor in your country:

China

Jasmine Dou, Editor (jasmine.dou@springer.com)

India, Japan, Rest of Asia

Swati Meherishi, Editorial Director (Swati.Meherishi@springer.com)

Southeast Asia, Australia, New Zealand

Ramesh Nath Premnath, Editor (ramesh.premnath@springernature.com)

USA, Canada:

Michael Luby, Senior Editor (michael.luby@springer.com)

All other Countries:

Leontina Di Cecco, Senior Editor (leontina.dicecco@springer.com)

**** This series is indexed by EI Compendex and Scopus databases. ****

More information about this series at https://link.springer.com/bookseries/7818

V. Bindhu · João Manuel R. S. Tavares · Ke-Lin Du
Editors

Proceedings of Third International Conference on Communication, Computing and Electronics Systems

ICCCES 2021

Springer

Editors
V. Bindhu
Department of ECE
PPG Institute of Technology
Coimbatore, India

Ke-Lin Du
Department of Electrical and Computer Engineering
Concordia University
Montreal, QC, Canada

João Manuel R. S. Tavares
Departamento de Engenharia Mecânica
Faculdade de Engenharia
Universidade do Porto
Porto, Portugal

ISSN 1876-1100 ISSN 1876-1119 (electronic)
Lecture Notes in Electrical Engineering
ISBN 978-981-16-8861-4 ISBN 978-981-16-8862-1 (eBook)
https://doi.org/10.1007/978-981-16-8862-1

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

We are honored to dedicate the proceedings of ICCCES 2021 to all the participants, organizers and editors of ICCCES 2021.

Preface

This conference proceedings volume contains the written versions of most of the contributions presented during ICCCES 2021. The conference has provided a platform to share and exchange the recent developments in a wide range of topics including computational intelligence, machine learning, signal and image processing, electronic devices and systems, antenna and wave propagation, wireless communication networks and so on. The conference has been a good opportunity for participants coming from various destinations to present and discuss the state-of-the-art topics in their respective research areas.

ICCCES 2021 tends to collect the latest research results and applications on computing, communication and electronics. It includes a selection of 75 papers from 312 papers submitted to the conference from various universities and industries present across the globe. All the accepted papers were subjected to double-blinded peer-reviewing process by 2–4 expert referees. The papers are selected for its high quality and the relevance to the conference.

We ICCCES 2021 would like to express our gratitude and appreciation to all the authors for their valuable research contributions to this book. We would like to extend our thanks to all the referees for expressing their constructive comments on all the research papers. In particular, we would like to thank the organizing committee for their tireless hard work. Finally, we would like to thank the Springer publications for producing this volume.

Coimbatore, India	Dr. V. Bindhu
Porto, Portugal	Dr. João Manuel R. S. Tavares
Montreal, Canada	Dr. Ke-Lin Du

Contents

About the Editors

Dr. V. Bindhu received the B.E. degree in Electronics and Communication Engineering from Bharathair University, Coimbatore in 2002 and M.E. Degree in Applied Electronics from Anna University, Chennai in 2007 and Ph.D. Degree from Anna University, Chennai in 2014. She has 11 years of teaching experience and 6 years of research experience. Currently, she is a professor at PPG Institute of Technology, Coimbatore. Her area of interest includes Signal processing, and VLSI Design.

João Manuel R. S. Tavares graduated in Mechanical Engineering at the Universidade do Porto, Portugal in 1992. He also earned his M.Sc. degree and Ph.D. degree in Electrical and Computer Engineering from the Universidade do Porto in 1995 and 2001, and attained his Habilitation in Mechanical Engineering in 2015. He is a senior researcher at the Instituto de Ciência e Inovação em Engenharia Mecânica e Engenharia Industrial (INEGI) and Full Professor at the Department of Mechanical Engineering (DEMec) of the Faculdade de Engenharia da Universidade do Porto (FEUP).

João Tavares is co-editor of more than 60 books, co-author of more than 50 book chapters, 650 articles in international and national journals and conferences, and three international and three national patents. He has been a committee member of several international and national journals and conferences, is co-founder and co-editor of the book series *Lecture Notes in Computational Vision and Biomechanics* published by Springer, founder and Editor-in-Chief of the journal *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization* published by Taylor & Francis, Editor-in-Chief of the journal *Computer Methods in Biomechanics and Biomedical Engineering* published by Taylor & Francis, and co-founder and co-chair of the international conference series: CompIMAGE, ECCOMAS VipIMAGE, ICCEBS and BioDental. Additionally, he has been (co-)supervisor of several M.Sc. and Ph.D. thesis and supervisor of several post-doc projects, and has participated in many scientific projects both as researcher and as scientific coordinator.

Dr. Ke-Lin Du received the Ph.D. in electrical engineering from Huazhong University of Science and Technology, China, in 1998. He founded Xonlink Inc., Ningbo, China in 2014. He was Chief Scientist at Enjoyor Inc., Hangzhou, China from 2012 to 2014. He was on research staff at Department of Electrical and Computer Engineering, Concordia University, Montreal, Canada from 2001 to 2010. Prior to 2001, he was on technical staff with Huawei Technologies, China Academy of Telecommunication Technology, and Chinese University of Hong Kong. He visited Hong Kong University of Science and Technology in 2008. Currently, he is an Affiliated Associate Professor in the Department of Electrical and Computer Engineering at Concordia University from the year of 2011. His research area includes signal processing, wireless communications, and soft computing.

OTT (Netflix) Film Recommender System Using Data Mining

Deshak Bhatnagar

Abstract With the current COVID-19 situation on hand, the load of entertainment has completely fallen onto the OTT platforms for over a year now, and as a significant rise in viewership has been seen in the recent years and especially the past year, it is particularly important that these systems are used in best way. Along this increase a big issue becoming among the viewers is the quality and content of the films, documentary films, etc., which depends highly on some factors like choice of actors, genre, availability in different languages and most importantly reviews. So, to solve this issue, a decision was made to try new approach and put some new life into an old and trusted system, incidentally which has been one of the earliest implementations in the field of machine learning and is fondly known or called by the name of the recommender system or the recommendation system. This paper emphasizes on providing with an over-the-top media-service provider (OTT) film recommendation system which aims to deliver a personalized choice of the films as per the new trends among the viewers, the ratings, the reviews and sometimes as per the previous experiences of the user. The so-called recommender systems have been in use for more than a decade now, but the interesting thing with these machine learning-based systems is that there is always scope for some new features or improvements to be added or channelled in and further on when you think that it is perfect and requires no changes or additions to it, but in fact everything in machine learning can be designed and implemented or given a new direction as per your own perspective. The thought to simply redefine this system mainly grew because of a big issue that is the wastage of time in selecting appropriate films as per the choice, non-engaging films and in a bid to make people watch more quality and knowledgeable content. This paper reviews a host of previous work done in this field along with chalking out issues which can be improved along the way. This paper further talks about the methodology for conceiving the system and as well as sheds light on the implementation of the system using Python language and open-source tools. The paper further moves onto dictate about the applications of the system for OTT services and also take a bit of sneak peek into the applications beyond the OTT services.

D. Bhatnagar (✉)
Department of Computer Science, Birla Institute of Technology and Science-Pilani, Dubai International Academic City, Dubai, United Arab Emirates

V. Bindhu et al. (eds.), *Proceedings of Third International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 844,
https://doi.org/10.1007/978-981-16-8862-1_1

Keywords OTT · Machine learning · Recommender · Film · Platform · Dataset · Rating · Algorithm · Movies · Python · Over-the-top media-service provider

Introduction

The world is changing rapidly these days, and with such a deadly disease at our doorstep, we have suffered a lot recently which eventually has damaged our entertainment and fun seeking activities too. But with this COVID-19 situation being a long-term issue, many things have changed from how we used to know them. For an example, previously movie or film watching in a theatre or cinema hall was an event and more precisely a way of going out of the house and enjoying with the family or friends, but now with this new situation at front, these enjoyments have become a thing of past, so in order to bring some of this joy back, the over-the-top media-service provider (OTT) platforms have emerged as an antidote in recent years, frankly saying they have grown rapidly in past year and incidentally have saved us from being getting bored and helped us spend some quality time with the family. With it is being certain now that many new films or precisely saying every second film will be directly releasing on these over-the-top media-service provider (OTT) platforms it becomes extra important that there is a proper independent system in place apart from the currently present recommender systems provided by the over-the-top media-service provider (OTT) platforms in order to help the public with selection of good flicks to watch rather than the non-amusing ones. The system intends to work on the criteria of most loved films in particular genres and reviews.

A recommender system simply emphasizes on providing personalized choices as per viewers trend or sometimes as per the previous experience of user. The recommender systems have been in use for more than a decade now, but the interesting thing with machine learning is that there is always scope for some new features or improvements even when you think it is perfect. To resolve the issue of wastage of time and in a way to make people watch quality and knowledgeable content, the decision to take this problem was made.

The main aim or objective of this study is simply to provide a somewhat permanent solution to the problems encountered during the selection of the movies which is quite common these days due to influx of the films in a large quantity. The following are the main aim or objectives of this study and our listed below as follows:

- To prepare a recommendation system or recommender system focused on over-the-top media-service provider (OTT) platform films recommendation
- To prepare a fresh new batch of dataset comprises films currently available on the OTT platform including classics and favourites
- To reduce the amount of time wasted on going through the films which are not as per the taste of a particular individual, reviews or worthy of a watch during the selection process.

To filter non-engaging and boring content plus to address the issue i.e., wastage of time, and most important to improve the currently available recommender systems and to modify them in such a way that they are apt and up to date with the current form of movie offering that is OTT (Over-the-Top Media-Service Provider) Services. Plus, in order to help knowledgeable content, get noticed the decision to take this problem was made. The secondary goal apart from the above-mentioned objectives/aim is to motivate researchers and professionals to understand and extend their knowledge in this particular domain with the help of far more simple recommender systems like this.

There are few problems which I have scouted aside and would love to solve as an integral part of this project are as follows:

- Old and redundant dataset issues
- Implementation approach
- Scalability issues.

To solve this problem, we intend to use a set of open service tools from the Internet which will have a simple Python code embedded in them which will be further based on the famous content-centred filtering method. Though the above-mentioned Python program seems to be a simple one but is indeed quite effective and with a seemingly easy implementation to follow afterwards. As far as implementation goes, there are a few new and old methods, but the preferred one is still the content-based recommender system, but using collaborative filtering is also effective to use because of the fact that it has an automatic prediction paradigm on offer. We would like to address some of the pressing issues like, the current availability of Datasets, which have been a big concern in these systems so far, the main issues being that some of the Datasets used are far too old or have movies which are not available on the OTT streaming services currently or have been removed or simply do not amuse the public or are from genre which is least favourite among communities. Further the focus is also not just on the use of the already available datasets but to produce some of our own dataset which will be specifically based onto the movies available currently in the country at the disposal of the over-the-top media-service provider (OTT) services like Netflix, Amazon and a few more to mention. But it will more importantly cater to the younger or current generation including films which are termed as classics and favourites in their genre and as mentioned by people in the reviews.

This system in whole is particularly important to be developed at this stage because of the fact that there are currently not many systems, or we can say no system or zero systems which specifically focus on the over-the-top media-service provider (OTT) platform service film recommendation systems. Plus, it will lay a new foundation for the upcoming challenges which are expected to specifically cater for over-the-top media-service provider (OTT) platforms as more and more movies are being released and their own recommender systems do not seem to always put you in for a good flick and that is certainly from the personal experience of using various services for quite a period of time. Whereas this recommender system can in particular act as one of the keystone models for the rest of the recommender systems present or to be made in future simply because of it being predominantly focused completely on a particular service, which in my opinion has not been a case in most of the recommender systems as they cater to larger quantities of problems at once which indeed makes them a bit less efficient and slower as they grow older-and-older with time.

Literature Review

In [1], the primary objective is to accumulate multiple papers and evaluate them separately in order to identify then integrate the answers to the question what are recommender system actually. And what they simply do. Plus, it also explains it in length along with focus on providing a better recommender system then the currently present at the moment. In [2], the researchers carry out survey in bid to find about the different filtering techniques for data in different setups as well as it tries to provide with some different characteristics of the filtering techniques in order to understand them well enough. It starts with a common answer to the application of machine learning in [3] that it is applied in a broad manner in the recommender system. It also explains how the Amazon's e-buying mechanism works with respect to the user details along with it describes how the user experience is increased. The paper [4] analyses the recommendation system established on filtering the collaborative way. The two practices applied in recommender approach founded on the collaborative filtering which are totally item-centred and the user-driven methodologies. It is predominantly a survey-based paper in which surveys are conducted right from start to the very end and which tell about the artistic techniques for the recommender system. Plus, it also sheds light on the accurate results needed by the user which is especially important at large.

In, [5] the research is completely based on the concept of first follow, simply talking about the popularity and effectiveness of the Collaborative-filtering. It also talks about poor accuracy and scalability of the content-based filtering and uses the collaborative for creating the recommender. It also explores the angle that these collaborative filtering techniques can even work efficiently for wide ranging problems too. In [6], the research is simply based on the capability to recommend flicks to a newly signed up individual as well as the already existing individuals. It pits flick databases in order to collect all the valuable material knowledge present, such as status and desirability. These are required for the particular recommendation. The use of content-based filtering and collaborative-based filtering along with the hybrid filtering together is quite prominent. While the system reaps out the more precise recommendation with this system in comparison with these techniques being used separately. Further, it also adds a unique feature of proposing movies using the demographic which is quite intense and difficult to achieve. In [7], the research focus here is completely on a hybrid system for recommendation which is achieved by the help of tags and the ratings that are present at large. The experimentation results of the above-mentioned system fared well with outshining the three varieties of the collaborative-based filtering classifications which are named as follows here user-based collaborative filtering, model type collaborative filtering and last but not the least topic-based model collaborative filtering. In [8], the work is carried out on the capturing of the temporal preferences of the users in modelling and forecasting the desired movies. The intended approach gives an overview of more user-friendly outline which integrates the content attributes of all the flicks individually and furnish a suitable recommendation listing. In [9], a simple system recommender system based upon in the content-based filtering is explained rather than its counterpart which is collaborative. Plus, it does not involve the prediction-based system.

The recommender system in [10] is based upon the famously successful method on content-based filtering with use of multi-attribute networks which seamlessly mirror the attributes while calculating correlations. It also addresses the issues of sparsity too. In [11], the online experiments are conducted using the wide linear models and deep neural networks using an app store in Android systems to show that wide and deep learning has significantly increased app acquisitions compared to the separate models of the above-mentioned learning terms. In [12], the system is based on filtering specifically collaborative. The film list is sorted and uses the k-means algorithm. The recommender system data is generated using different types of the knowledges available. While they rely completely on PHP and Hadoop instead of traditional Python language. In [13], the focus shifts onto the use of deep learning methods for recommender system because of the fact the feature extraction plus representation learning is taught from start which is indeed quite a skill to have under-belly. Primarily, it also shares some models and recent activities in the field comprehensively. In [14], the tradition is restored again with use of content filtering with a twist of k-means algorithm being accompanied by k-nearest neighbour algorithm to aid the effectiveness of the system. It also describes the process along with implementation and use of each and every algorithm plus technique in detail. Also, it goes deep into the illustration of the framework used with describing and leaving hints for scope of improvement. In [15], both content-based filtering and collaborative-based filtering are explained in depth with full focus being towards the content-based filtering as it is the technique used further. Plus, it also provides with a testing dataset and training dataset for the users along with a functioning code of the recommender system.

Proposed Work

For the said implementation of the experiment that is OTT recommender system, we have used the Jupyter Notebook alongside Python 3.7. With the help of the following libraries NumPy, pandas which read the dataset and additionally matplotlib and seaborn were used for visualizations and plotting purposes. The dataset used has an array of over 45,000 movies with details like genre, date added, date released, etc. We only generate results for the top ten recommended movies with the help of vote count and vote average which by the help of rating algorithm give a new rating called as weighted ratings to determine the top recommended films. The ratings are further demonstrated in various forms like charts, graphs, scatter plots, etc. (Fig. 1).

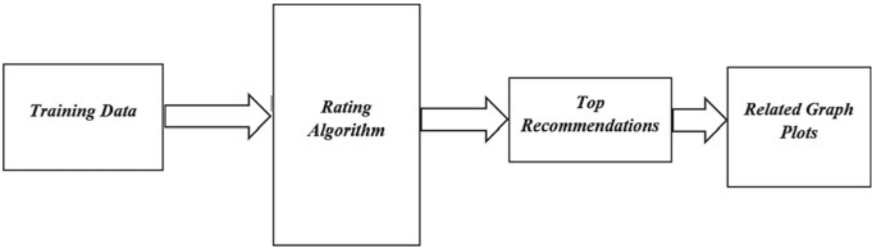


Fig. 1 OTT recommender system architecture

The development begins by the dataset sourced out directly from the foundation. The training data or simply data is read, plus saved in the system memory. Moving forward, in the second step, the rating algorithm is used which by the help of vote count and vote average generates a new type of total rating called rating. This rating further determines the top ten recommendations of the systems from 45,000 movies. The top ten recommendations are further shown in the form of bar plots and scatter plots which are used to show the vote count and vote average.

Rating Algorithm

The steps for the algorithm are as follows:

Inputs: vote_count, vote_average.

1 Read file by pandas library.
2 Read vote_count and vote_average from the file.
3 Calculate mean (C) of vote_averages.
4 Calculate quantile (m) of vote_counts.
5 Calculate:
    def rating(y):
    $g = y$['vote_average']
    $h = y$['vote_count']
    return $(h/(h+m) * g) + (m/(m+h) * C)$
6 Visualize the top ten recommended movies.

This algorithm simply gives the rating of each movie. The rating here measures the number of points or rate which tells how good a movie is rated or how good it is, plus determines how much likely someone is going to watch that particular movie or film. The algorithm starts with reading the dataset, then retrieves vote count and vote average of the movies, then calculates the mean and quantile for the both, respectively, and gives a new measure called as rating. Then, the following are visualized in the form of graphs and tables.

Result Analysis

The results for the above-mentioned implementation are carefully accumulated using an rather large dataset of over 45,000 movies with information like genre, date added, date released and vote count and vote average. The data is carefully checked and is read in float form only (for numeric). The results obtained are briefly explained below in a well-curated systematic order which is as follows:

1 The Top 10 Recommended Movies accumulated using the Algorithm stated above (Fig. 2).

It simply shows the ratings of the movies along with the total votes and average votes they got. While the rating changes with every change made in vote count and vote average. The rating can be defined as

$$\text{def rating}(y):$$
$$g = y\left['\text{vote\_average}'\right]$$
$$h = y\left['\text{vote\_count}'\right]$$
$$\text{return}(h/(h+m)*g) + (m/(m+h)*C)$$

2. A bar graph showing the ratings also known as total ratings of the top ten recommended movies. It simply shows the ratings of the movies along with the total votes and average votes they got. While the rating changes with every change made in vote count and vote average (Fig. 3).
3. A series of scatter plots in Figs. 4 and 5 show the relationship between the vote count and total rating, vote average and total rating, respectively, for better understanding of the respective components in smooth running of algorithm.

	title	vote_count	vote_average	r
0	The Shawshank Redemption	8358.0	8.5	7.903014
1	The Dark Knight	12269.0	8.3	7.894781
2	Fight Club	9678.0	8.3	7.806269
3	Inception	14075.0	8.1	7.766655
4	Pulp Fiction	8670.0	8.3	7.760416
5	The Godfather	6024.0	8.5	7.733254
6	Interstellar	11187.0	8.1	7.694655
7	Forrest Gump	8147.0	8.2	7.654238
8	The Lord of the Rings: The Return of the King	8226.0	8.1	7.579358
9	The Lord of the Rings: The Fellowship of the Ring	8892.0	8.0	7.530382

Fig. 2 Top recommendations

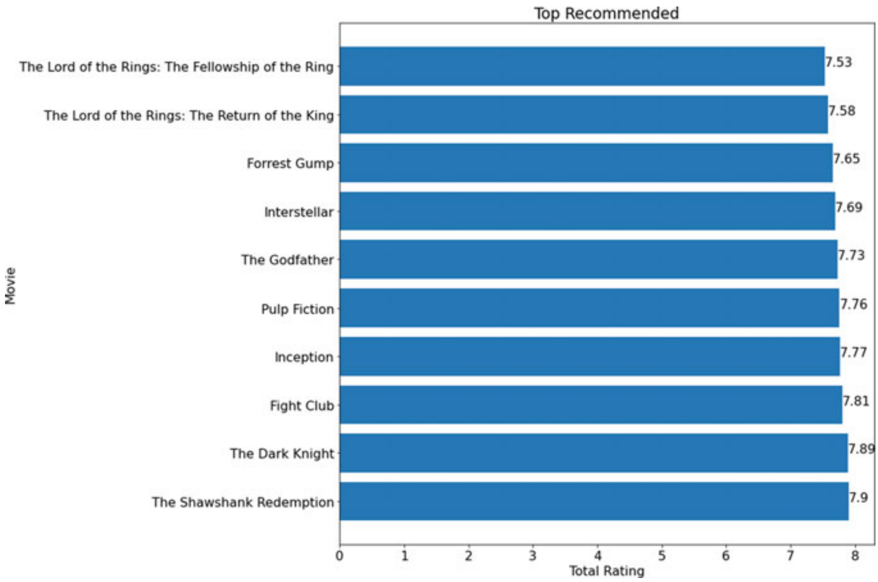


Fig. 3 Bar graph for top ten recommendation

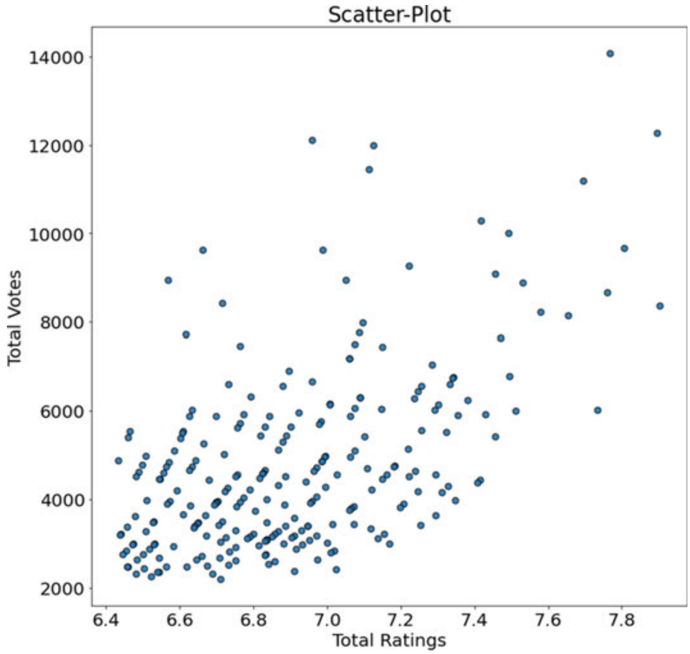


Fig. 4 Ratings versus vote count scatter plot

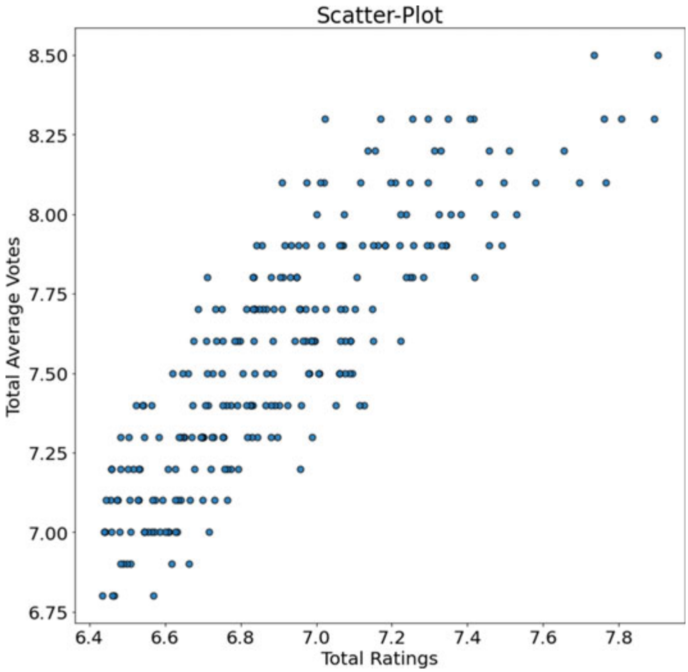


Fig. 5 Rating versus vote average scatter plot

Conclusion

The logic behind and use of recommender systems is quite viable, and certainly it is one of the useful and quintessentially critical technologies in today's environment. Frankly quoting we are using recommender system for each and everything right from movies to even selecting books or even clothes and is going to be every minute system from everyday system in just a blink. As we already know, it is a topic which is already quite used and experimented, but still there is a lot of juice of knowledge to be extracted from it. I certainly was able to do so with this new algorithm and experiment, and I was not only able to display the recommendations but also figured out a new way of visualizing the data to the user rather in the standard form than in graphical format which is astonishing and is not used much practically. Plus, it was quite interesting to formulate the algorithm and implementation which took a lot of time and work.

References

1. Alyari F, Navimipour NJ (2018) Recommender systems. Kybernetes
2. Raghuwanshi SK, Pateriya RK (2019) Recommendation systems: techniques, challenges, application, and evaluation. In: Soft computing for problem solving, Springer, Singapore, pp 151–164
3. Varga E (2019) Recommender systems. In: Practical data science with python 3. Apress, Berkeley, CA, pp 317–339
4. Khan BM, Mansha A, Khan FH, Bashir S (2017) Collaborative filtering based online recommendation systems: a survey. In: 2017 international conference on information and communication technologies (ICICT). IEEE, pp 125–130
5. Ponnam LT, Punyasamudram SD, Nallagulla SN, Yellamati S (2016) Movie recommender system using item based collaborative filtering technique. In: 2016 international conference on emerging trends in engineering, technology and science (ICETETS). IEEE, pp 1–5
6. Geetha G, Safa M, Fancy C, Saranya D (2018) A hybrid approach using collaborative filtering and content based filtering for recommender system. J Phys Conf Ser 1000(1):012101. IOP Publishing
7. Wei S, Zheng X, Chen D, Chen C (2016) A hybrid approach for movie recommendation via tags and ratings. Electron Commer Res Appl 18:83–94
8. Cami BR, Hassanpour H, Mashayekhi H (2017) A content-based movie recommender system based on temporal user preferences. In: 2017 3rd Iranian conference on intelligent systems and signal processing (ICSPIS). IEEE, pp 121–125
9. Aggarwal CC (2016) Content-based recommender systems. In: Recommender systems. Springer, Cham, pp 139–166
10. Son J, Kim SB (2017) Content-based filtering for recommendation systems using multiattribute networks. Expert Syst Appl 89:404–412
11. Kumar M, Yadav DK, Singh A, Gupta VK (2015) A movie recommender system: Movrec. Int J Comput Appl 124(3)
12. Cheng HT, Koc L, Harmsen J, Shaked T, Chandra T, Aradhye H, Anderson G, Corrado G, Chai W, Ispir M, Anil R, Shah H (2016) Wide and deep learning for recommender systems. In: Proceedings of the 1st workshop on deep learning for recommender systems, pp 7–10
13. Zhang S, Yao L, Sun A, Tay Y (2019) Deep learning based recommender system: a survey and new perspectives. ACM Comput Surv (CSUR) 52(1):1–38
14. Ahuja R, Solanki A, Nayyar A (2019) Movie recommender system using K-Means clustering and K-Nearest Neighbor. In: 2019 9th international conference on cloud computing, data science & engineering (confluence). IEEE, pp 263–268
15. Content-based filtering in machine learning. Retrieved from https://thecleverprogrammer.com/2021/02/10/content-based-filtering-in-machine-learning

Dietary Assessment and Nutritional Analysis Using Deep Learning

S. Madhumitha, M. Magimaa, M. Maniratnam, and N. Neelima

Abstract Nutrition management is an important feature in day-to-day life from prevention of obesity and chronic diseases related to food intake. To facilitate proper and adequate nutrition intake, suitable dietary assessment is essential. This dietary assessment method based on food recognition has the ability to detect the volume, based on the area of the image taken through computer vision. There has been many modern techniques developed recently on image-based dietary assessment. These techniques have shown trustworthy results on overcoming issues on epidemiology on dietary studies and multiple challenges. This model provides a detailed outline of how features are extracted from the image using computing algorithms, methodology used, and mathematical methods for image recognition. This model provides volume and mass estimation of the given food based on the area occupied by the food. From this, the detailed nutritional analysis was calculated for dietary assessment. The overall accuracy of 94.68% was achieved for classification of food which outperforms the existing methods.

Keywords Volume estimation · Food detection · Dietary assessment · Nutritional analysis

Introduction

The world is hit by one of the deadly viruses, and the entire world has gone into lockdown. One of the most common issues faced being at home is not having adequate physical exercise resulting in overweight. Excessive fat accumulation which might harm the health of an individual can be defined as overweight or condition of obesity [1]. A lot of health-related issues are born from merely having extra fat. Few issues

S. Madhumitha (✉) · M. Magimaa · M. Maniratnam
Department of Electronics and Communication Engineering, Amrita School of Engineering, Bengaluru, India

N. Neelima
Amrita Vishwa Vidyapeetham, Bengaluru, India
e-mail: n_neelima@blr.amrita.edu

V. Bindhu et al. (eds.), *Proceedings of Third International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 844,
https://doi.org/10.1007/978-981-16-8862-1_2

include cardiovascular disease and musculoskeletal disorders, and some types of cancer are serious effect of non-healthy diet.

Obesity is non-communicable, and yet it is one of the prime concerns of almost all health organizations [2]. The proposed model aims to give nutritional reports on the food which are consumed by the people, thus enlightening them on their dietary knowledge [3]. This analysis will help in understanding the essential requirements and harmful substances in the food that are being consumed. Now, with the onset of image processing and artificial intelligence, a lot can be done with minimum human intervention.

Food recognition and classification are the key factors in dietary assessment and can be done using different methods like geometry features, statistical features and methods based on machine learning. Multidimensional histogram can also be employed as a feature vector for classification [4]. This model uses Inceptionv3 as CNN model, to draw out a feature graph of the food image and classify the food image to the nearest food class. Each output from the CNN layer is a feature map, and it contains parameters linked to the layers. It further processes the image and trains it. The dataset used is Food-101, and it has 101. In this, mathematical models were used for image recognition. However, the complexity of the architecture and the time required are the major setbacks.

Aiming to reduce the complexity of the architecture, the proposed model identifies the food by extracting the features from the image using a computing algorithm. Main idea behind this model is to help individuals monitor their food intake and to have a calorie intake count. The aim of dietary assessment is to identify appropriate and useful areas of change in the individual's diet and lifestyle, thus aiming to improve a person's overall health.

Related Work/Literature Survey

Features can be compared with respect to their distance metrics which was discussed in [13].

Food Detection and Recognition

Food recognition is the crucial task when the images consist of multiple food categories. The model used in [5] detects food by demarcating between images with and without food. This further aids in food recognition and classification. Most image segmentation techniques are based on basic image properties like color and texture. However, the semantic segmentation is done under supervised learning model structure. Here, the model uses a feature graph for food detection.

Volume Estimation and Dietary Assessment

Volume may change based on density, so it is very hard to get the exact value. The inexactitude of the volume determined should be of prime concern to develop a model which is unbiased and closely accurate [6]. Some of the food items have a particular shape such as burger, cupcakes, and dumplings. So here, shape feature is extracted for each food item. The co-ordinates or the spatial points of the skin of the food can also be determined through HOG as done in [7]. The algorithm also sets the images through multistage processes including change in its colorscape, change to hue, saturation and value (HSV), thresholding the image through adaptive process and finally determining the mask and the skin of the food in the image. These images give the area covered by food in image, and based on the area covered, volume estimation is done.

One of the major hindrances in incorporating automated dietary assessment system into daily life is its efficacy and beneficial algorithm to obtain relevant data from the food images [8]. The study launched by The International Life Sciences Institute (ILSI) [9] shows that numerous technology-centered dietary assessment tools including mobile applications and Web applications were assessed in the period between 2011 and 2017. Based on the volume estimation, the calorie count is calculated. The database consists of the calories of each food per 100 gms, and using this, the calories present in the food are calculated. After the image is processed, it checks for the boundaries [10] and examines whether the image contains any food item in it. The efficiency of CNN with Food-101 is told to be 78.11% [11].

System Design

In this section, the overview and design based on deep learning system for food detection, classification and analysis on nutritional values present in each meal image is explained. Later, the details of every component of the model are discussed.

System Overview

In this part, model based on deep learning for food class detection and analysis of the nutritional information of each meal image is explained. As illustrated in Fig. 1, the main steps are clearly shown.

The model is used to identify and detect the food images from the given input, and it classifies them into the respective classes. Then, the food volume estimation methods are done to measure the portion size of food items. The final step is to use dietary assessment tools for nutritional analysis and to generate a detailed report for users derived from their food images.

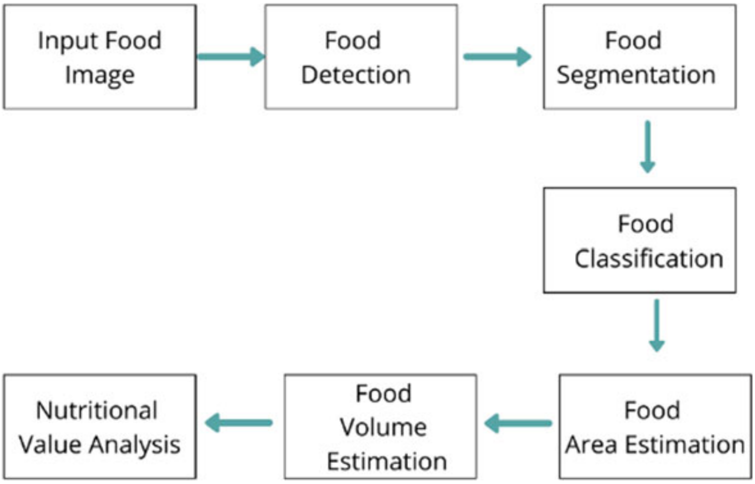


Fig. 1 Overview of the proposed model

Food Image Classification

As convolutional neural networks extract the different features from food images, it tends to have an effective way in food classification problem. In this model, the different types of meal images are casted as input to train the classifier, and a perfect classifier will be able to pinout and classify any food image class that has been included in the training process. The training process itself is a monotonous process as it requires a considerable amount of time to reach its intended final goal. As the classifier accuracy is affected mainly by the number of food images present and on standards of images used in pre-training, the categorization module uses feature maps for classification. This calculates the value for each class and thus classifying those food images implemented.

Deep Learning

Deep learning is a popular method of machine learning that helps in training the neural network in the more effective way. This mechanism extracts the features from the images and automatically passes through many connected layers. The fully connected layer is the final layer which helps in final classification of the input images.

Conventional neural network (CNN) is the most significant technique under deep learning owing to its exceptional performance from the replication of its biological counterpart of the human central nervous system [12]. This model uses the pre-trained model Inception which extracts the features from the input food image and forms the feature map of the defined regions. It later classifies the food image through the 64 weight layers including the full connected layers inside the network. The accuracy is based on the number of the iterations used. The images will be split into training sets and the validation sets in the random manner. The model is fine-tuned with the respective parameters according to the training images in our dataset. The images are taken from the Food-101 dataset. This contains 101 categories of the food which are consumed on daily basis by the people.

Implementation Details

Under the implementation of the model, the image is first resized to width = 299, height = 299. In order to generalize the system model, the Food-101 dataset is added to the CNN model while training. A size of $299 \times 299 \times 3$ is considered for the input sensor. 64 dense layers are given with ReLu activation function. The dropout rate is set to 0.2. Global average pooling 2D is a pooling operation that is used here which is designed to replace fully connected layers in classical CNNs. The first layer compresses or reduces the total input image into a single vector. The output of this step then passes through the next activation function which will produce the nonlinearities of the image in the element-wise manner. Then, down sampling is done using the pooling layer along with the dimensions including width and height from the previous layer volume. The ReLu layer works on the activation function, i.e., max $(0, x)$, and the fully connected layer FC computes the final values of the output that will be equal to number of the classes. So, the output from the final pooling layer is made into the vector which becomes the input for the fully connected layer, thus computing the final values.

Area and Volume Estımatıon

After the food item is detected and extracted from the image, feature extraction is the next step. Feature vector is extracted for precise training and testing of the model. The HSV color wheel contributes to high-quality graphics and thus is a very wise choice in image processing. Similarly the area, shape and size of the detected image are also analyzed from the features extracted. The image after each stage of the processing is saved for more methodical analysis of the final output. The volume of the particular food is computed by using trapezoidal volume estimation rule from the output of area of the food. The spatial points are taken using Delaunay triangulation methods, and the volume under the specific region is found. Delaunay tessellation in our model is in three dimensions, and an example of determining the Delaunay triangle is shown in Fig. 2.

The volume for finding the complex shape is difficult, and Delaunay triangulation method is used here by splitting the shape into several sections first. Wıth different

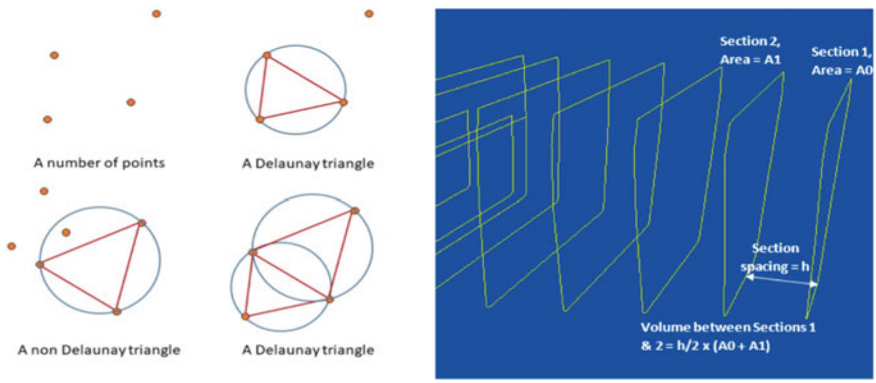


Fig. 2 Delaunay triangulation and trapezium rule

areas of section on either end, it will more or less make a trapezoid. The volume of the each section is found by its height and area. The average of the volume is calculated, and final value is evaluated. The formulas for finding the volume using the Delaunay triangulation and the trapezoidal rule are given in Eq. (1, 2 and 3).

$$\textbf{Volume of Section 1} : V1 = 0.5 \times h \times (A1 + A2) \tag{1}$$

$$\textbf{Volume of Section 2} : V2 = 0.5 \times h \times (A2 + A3) \tag{2}$$

$$\textbf{Total Volume} : V = 0.5 \times h \times (A1 + 2A2 + 2A3 + 2A4 \ldots + 2A9 + A10) \tag{3}$$

Dietary Assessment

After the food item is recognized, the model will analyze the nutritional values and gives a dietary assessment report. Calories, fats, carbohydrates, proteins and dietary fibers are displayed as a part of the analysis report. A database is created for the 30 classes of food and incorporated all the nutritional values for a normalized content. After the food is recognized, the area and volume are computed. With the previously calculated volume and the standard nutritional facts given in the database for normalized volume or portion of the specific food, the actual values are determined. These values will be present as dietary assessment as the final result from this project. This dietary assessment is a very integral tool used in multiple applications like hospitals, supermarkets, etc.

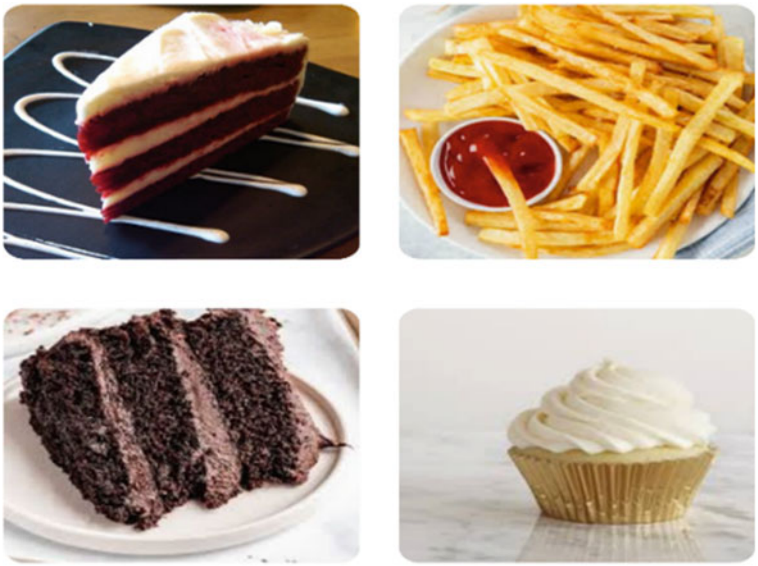


Fig. 3 Some sample images of food items in Food-101 dataset

Results and Discussion

Dataset

The Food-101 dataset consists of 101 classes each with 1000 images. 15 classes were taken for the experimental analysis with a total of 11,250 images for training and 3750 images for testing. Some of the sample images from the dataset are shown in Fig. 3.

Simulation Results

The training accuracy and the loss plot for the first five epochs in training are shown in Figs. 4 and 5, respectively.

Figure 6 represents the confusion matrix of the ten classes of food which was used to classify the given input image. After the image is processed through the many stages, this makes it suitable to find the area of the image. The image is segmented, the area is estimated, and it is shown in Fig. 7. The image is first transformed and processed through these multiple stages to estimate the final food image's area.

Figures 8 and 9 show the final volume estimated by the model and in-depth nutritional analysis and dietary assessment of the detected image taking into consideration the physiological data of the consumer, respectively.

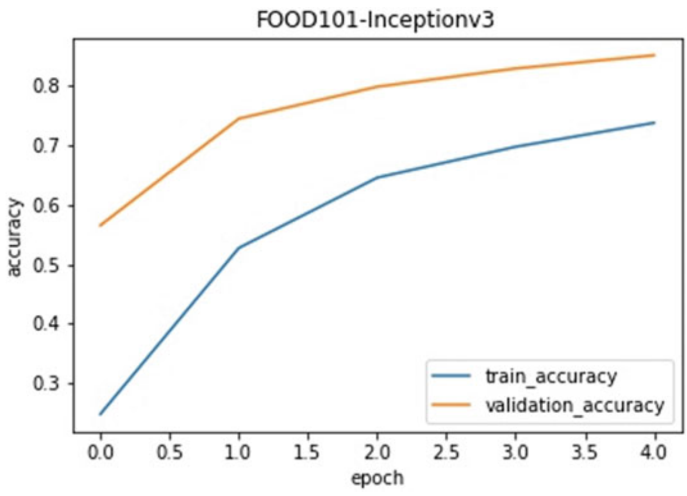


Fig. 4 Accuracy plot for ten classes

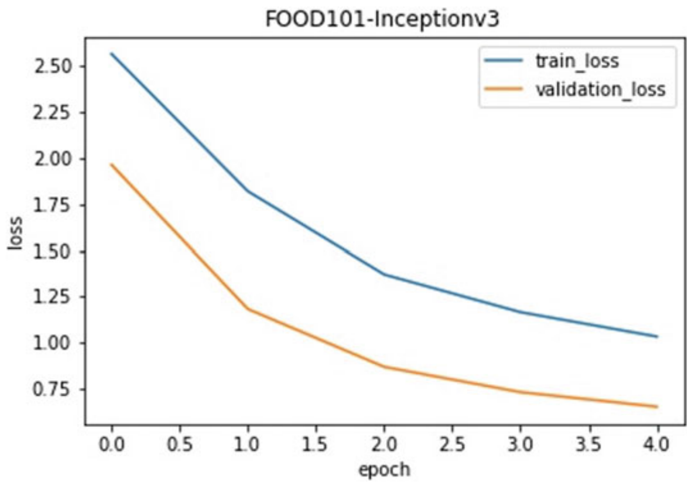


Fig. 5 Loss plot for ten classes

Conclusion

Obesity is non-communicable, and yet it is one of the prime concerns of almost all health organizations. One must monitor their individual dietary to maintain healthy lifestyle. Dietary assessment can play an important role in every one's individual life

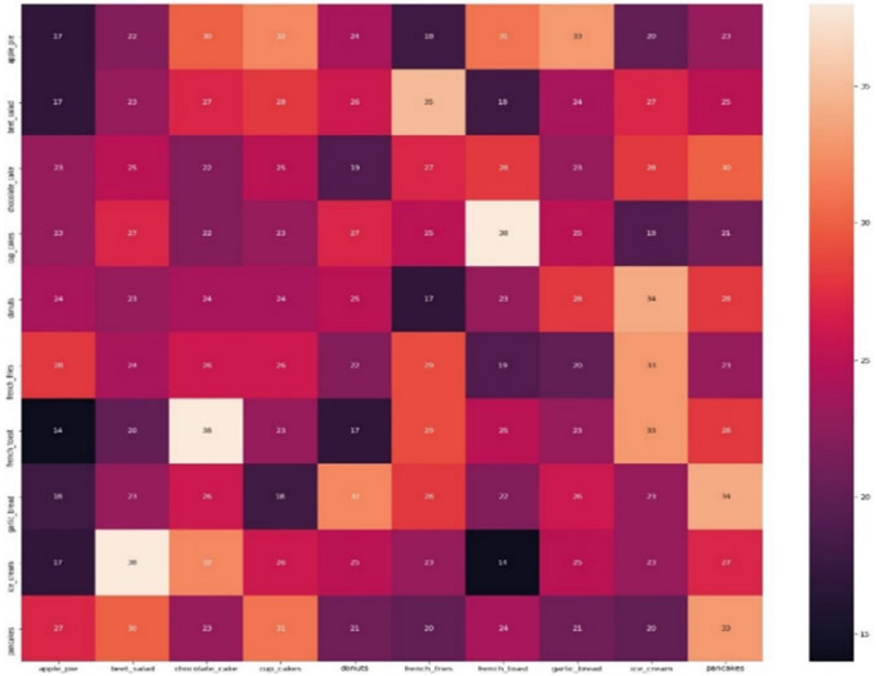


Fig. 6 Confusion matrix of ten classes

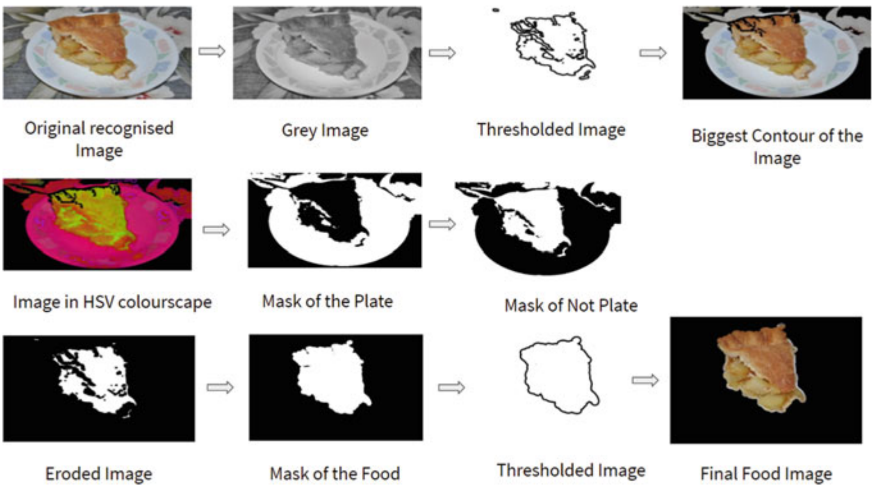


Fig. 7 Contour detection method for area estimation of food

Fig. 8 Volume determined from food image

```
Volume of the Food Image is: 217 cm^3
```

Fig. 9 Final output for dietary assessment for male category

which can assess every day's intake and calorie management. The proposed model classifies the good using deep learning approach and provides nutritional analysis and dietary assessment using volume estimation methods. Once the food is recognized, area and the corresponding volume are estimated. An accuracy of around 85% is obtained using this technique. A database is made using the dataset that is used and with normalized values of fat, proteins, calories and fiber of each food for 100 g.

With respect to food class, from the corresponding database of nutritional particulars, the calorie value is calculated, and the dietary output is obtained. This model can be used for health tracking and can be used for leading a healthy life and promoting self-owned diet.

References

1. Chang X, Ye H, Albert P, Edward D, Nitin K, Carol B (2013) Image-based food volume estimation. CEA13
2. He H, Kong F, Tan J (2016) DietCam: multiview food recognition using a multikernel SVM. IEEE J Biomed Health Inform 20(3):848–855
3. Aguilar E, Remeseiro B, Bolanos M, Radeva P (2020) Grab, Pay and eat: semantic food detection for smart restaurants. IEEE Tans Multimedia 1(1)
4. Yang S, Chen M, Pomerleau D, Sukthankar R (2010) Food recognition using statistics of pairwise local features. In: IEEE computer society conference on computer vision and pattern recognition, 2010
5. Bhavya sree B, Yashwanth Bharadwaj V, Neelima N (2021) An inter-comparative survey on state-of-the-art detectors- R-CNN, YOLO and SSD. Smart Innovation, Syst Technol 213:475–483
6. Madan K, Bhanu Anusha K, Pavan Kalyan P, Neelima N (2019) Research on different classifiers for early detection of lung nodules. Int J Recent Technol Eng 1037–1040
7. Raju JVVSN, Rakesh P, Neelima N (2019) Driver drowsiness monitoring system. Smart Innovation Syst Technol (SIST) 169:675–683
8. Eldridge A, Piernas C, Illner A-K, Gibney M, Gurinović M, Vries JD, Cade J (2019) 'Evaluation of new technology-based tools for dietary intake assessment—An ilsi europe dietary intake and exposure task force evaluation.' Nutrients 11(1):55
9. Lo FP-W, Sun Y, Qiu J, Lo B (2018) Food volume estimation based on deep learning view synthesis from a single depth map. Nutrients
10. Jiang L, Qiu B, Liu X, Huang C, Lin K (2020) DeepFood: food image analysis and dietary assessment via deep model. IEEE Access 8:47477–47489
11. Subhi MA, Ali SH, Mohammed MA (2019) Vision-based approaches for automatic food recognition and dietary assessment: a survey. IEEE Access 7:35370–35381
12. Lo FPW, Sun Y, Qiu J, Lo B (2020) Image-based food classification and volume estimation for dietary assessment: a review. IEEE J Biomed Health Inf 24(7):1926–1939
13. Neelima N, Seenivasa eddy E (2017) An efficient approach to CBIR using DWT and quantized histogram. In: International journal of ınnovative computing, ınformation and control, IJICIC (Scopus-Q1), vol 13(1), pp 157–166

Analyzing Game Theory Applications in a Layered Perspective for a Non-cooperative Environment with the Existence of Nash Equilibria in Various Fields of Research

S. Kanmani and M. Murali

Abstract Game theory aims at modeling actions among players or users who are dealing with different situations and facing different consequences in a common space. Primarily, game theory has been used in various fields like economics—for modeling competition between companies, wireless sensor networks, software-defined networks, auctions, renewable energy, etc. Game theory can be widely used in the field of networking where resource allocation and cooperation among network or terminal play a major role. Hence, formulations of games are used, and concrete solutions for players are inferred through the equilibrium concept. To arrive at a solution, players are categorized based on their behavior and computation of price of anarchy (PoA) and price of stability (PoS) which are performed accordingly to find the existence of Nash equilibrium and evaluating performance among them. This paper gathers various applications on game theory where equilibrium exists between players in a non-cooperative environment and in communication networks presenting them in the perspective of open systems interconnection (OSI) layers. Also, the paper focuses on the comparison of various fields of research in game theory. Topology plan of an organization centers around finding the network setup with the most ideal exhibition given few optimization models. With this analysis, games could be designed and modeled effectively and efficiently where the quality of service (QoS) enhances through protocols and emerging technologies.

Keywords Game theory · Nash equilibrium · Price of stability · Price of anarchy · OSI layers · Optimization models

S. Kanmani (✉) · M. Murali
Department of Computer Science and Engineering, SRM Institute of Science and Technology, Kattankulathur, Tamil Nadu, India
e-mail: kanmanis@srmist.edu.in

M. Murali
e-mail: muralim@srmist.edu.in

V. Bindhu et al. (eds.), *Proceedings of Third International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 844,
https://doi.org/10.1007/978-981-16-8862-1_3

Introduction

Game theory is a strategic decision-making concept [1], a more descriptive name for the discipline—decision theory. There are already two main reasons why games seemed to be a good area for exploring machine intelligence, offering organized tasks and not needing a huge amount of information. [1]. Game theory models situation where decision makers take actions that have mutual consequences. In view of networks, game theory concept acts as a tool among terminals, service providers and nodes to form cooperation dependency scheme [2]. Also, notion of game theory in recent years was used in networking concepts like routing, allocation of resources, etc. When nodes are connected and allowed to communicate in a dynamic communication network, they tend to share the data when they do not have conflict of interest [3]. Cost optimization, minimizing waiting time and traffic, increase in number of nodes and increase in nodes' bit rate are some of the important factors to consider while designing a dynamic communication network. In certain scenario, nodes might not cooperate while establishing communication between them [4]. Computer network topology architecture focuses on finding the network configuration with the best possible results given certain optimization parameters. To design a topology in this environment where nodes are non-cooperative, performance of the network can be considered with respect to price of link establishment, delay in path and congestion in path [5].

The first design consideration is price of link establishment. Here, considering the directional links along with allowing link costs where various values are assigned throughout the network leads to estimating the price of link establishment [6, 7]. Also, priority goes to analyzing the link establishment cost as a single point of deliberation before even analyzing the trade-off between them and the delay reflection. Link establishment differs with organization of graph topology—discussed in detail in Sect. 2.1. Since nodes in a network are modeled in various organizations, price of link establishment varies according to the topology on which network stands.

The second design consideration is path delay. After the routing paths are computed, hop count can be detected easily [8]. Counting the number of hops a particular node takes to traverse from its source to destination route results in path delay. Enhancing the feature of computing hop count with the existence of Nash equilibria in non-cooperative environment can be done with certain modification in Folk theorem [9].

The third design consideration is path congestion. It is coined as 'relaying extent.' If a particular node is being used as a relay by higher number of nodes, it is defined as higher relaying extent. This results in higher congestion rate in that node [6]. As a result of relaying extent, there exists poor bandwidth allocation and increase in packet loss. More emphasis is on choosing the type of game for every case and to find the utility functions elements. Among the games to be chosen, non-cooperative games are well suited for enhancing the concept of relaying extent. Non-cooperative games are games in which players make decisions independently.

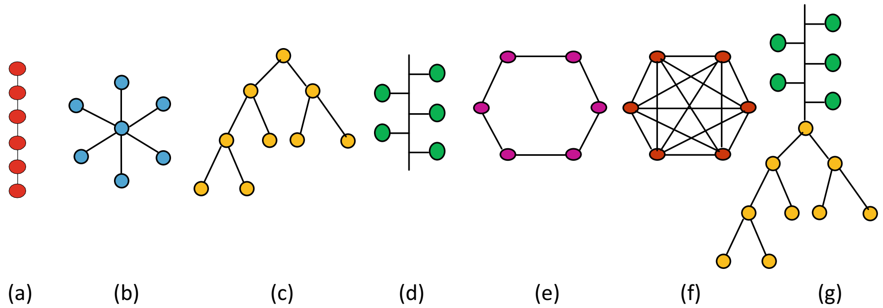


Fig. 1 Graph topologies—**a** path, **b** star, **c** tree, **d** bus, **e** ring, **f** mesh (fully connected) and **g** hybrid

Background

Graph Topologies

For n-player games with graphical architecture, the model or the game design changes its behavior with respect to the change in graph topology. When nodes are interacting in a network, a suitable or appropriate graph topology is selected in order to make the communication model efficient and effective. Figure 1 depicts the various graph topologies that are to be considered for nodes communication. Chain between the nodes (players) indicates the path or communication link of a network [10]. With these combinations of topologies, one needs to decide how appropriate the selected topology fits in a network.

Path topology (Fig. 1a) and bus topology (Fig. 1d) are a form of linear topology wherein path topology connections between nodes are done in a linear–unidirectional manner, whereas in bus topology there exists a connection between nodes in a network that share a common bus. In star topology (Fig. 1b), all the peripheral nodes are attached to a central node, where the central node receives the transmission data from a node and rebroadcasts the received transmissions to all the remaining nodes. Variation of star topology is tree topology (Fig. 1c) which handles the hierarchical flow of information among the nodes. In ring topology (Fig. 1e), each node is connected to exactly two branches or neighboring nodes. In mesh topology (Fig. 1f), each node is connected through a dedicated point-to-point link to every other node in a network. Hybrid topology (Fig. 1g) is a combination of two or more topological structure like bus topology and tree topology.

Game Theory

Modeling game theory applications in a non-cooperative environment in recent years is becoming a more challenging task. Reason behind selecting game theory strategic

Table 1 Mapping game theory model with networks [14]

Game element	Network element
Players	Customers, nodes or service providers
Strategies	An inference based on the actions of the player depending on network fields like handover, new call, packet loss, etc.
Actions	All kinds of actions performed by nodes to communicate with the network nodes successfully
Payoffs	Computed by utility functions u_i with quality of service (QoS) metrics like throughput, propagation delay, etc.

models for decision-making process is to find an alternate strategy for competing among players, if the existing strategy is not giving a desirable optimum solution or path [11]. In the recent game-theoretic models, networking concepts like flow control, allocation of bandwidth and routing have been imprinted to emphasize the importance of collaboration between game theory and networks. Researchers have been working on the Nash equilibria [12] of certain games where networking concepts can be resolved efficiently and effectively. With all the outcomes from these strategies, a big gap was there in the place of optimization. Topology design, with respect to networking concepts, lacks in optimization [1]. Thus, to fulfill the enforced optimum in Nash equilibria, price of anarchy (how efficiency decreases due to behavior of agent) and price of stability (best objective optimal value of level-k equilibrium) should be included in network game-theoretic model.

In a game, there exist finite set of players $P = \{1, 2, \ldots\ldots, p\}$ who select a strategy among $S = \{1, 2, \ldots.., s\}$ where utility u_i needs to be maximized. The utility function can be defined as u_i (s): $S \rightarrow A$ where A represents player's response to everyone's action. Payoff π_{ij} represents the payoff function assigned to player i for a particular action j [13]. Accordingly, a game model (G) can be defined as follows: $G = (P, S_i, A, \pi_{ij})$. While a big move was taken on game theory representations, time had come to redefine the mapping with respect to game theory and networks. Representation of mapping between game theory models and networks is listed in Table 1.

Nash Equilibrium

In game theory, Nash equilibrium plays a vital role in the place of solution concept involving more than one player, in which one player's change in strategy does not add to their gain or loss unilaterally. Players choose their own equilibrium strategy to maximize their payoffs. When a set of players choose their strategy and one player alone tends to choose a different strategy while other strategies are unchanged, then the present choices of strategies and the payoffs corresponding to it lead to a Nash equilibrium. Some of the applications of Nash equilibrium are prisoner's dilemma, currency crises, traffic flow, organizing auctions, natural resource management, penalty kicks in football, evacuation problems, strategies in marketing, etc. [15]. A strategy s_i is strictly dominated for a particular player i if there exists $s_i^{'} \in S_i$ such that,

$$u_i\left(s_i^{\prime}, s_{-i}\right) > u_i(s_i, s_{-i}) \forall s_{-i} \in S_{-i} \tag{1}$$

For a particular period of action or time, if a player always obtains a higher payoff regardless of the opponent turn or move, the corresponding strategy of that player is called as a strictly dominated strategy. In Eq. (1), consider there exists a set of strategies S_i for respective utility function u_i, where dominated strategy is to be identified based on the value obtained by utility function of a players' previously followed strategy and current strategy.

Related Works

A refined Nash equilibrium was proposed in repeated games [16], a level-k equilibrium, to enhance and trace computation of Nash equilibrium. Based on this level-k equilibrium concept, it is showed that the specific Pareto optimality of the feasible payoff profiles is Nash equilibrium payoff profiles. To compute level-k equilibrium, the minimax payoff profile is not necessary. In symmetric games, the game's result is not dependent on the player's identity [17]. All the evolutionary games are almost symmetric so that players in the population are intact with the same situation they evolved. According to level-k equilibrium and Folk theorem, Pareto analysis of payoff profiles must be evolutionary stable states in symmetric games. In evolutionary stable state, a slight disturbance cannot change the composition in genetics for a particular population [2, 14]. Apart from intrinsic disturbances like stochastic decisions, noise in game dynamics, mutation and crossover in the process of selection is also considered as disturbance. Replicator equation for dynamic evolutionary game is expressed as an equation.

$$\dot{x}_i = x_i(u_i(x) - \overline{u}(x)) \tag{2}$$

where $x = (x_1, \ldots\ldots, x_n)$ is a vector for the distribution of $i = 1, \ldots\ldots., n.$ in the population, $u_i(x)$ is the fitness function or payoff of type i and $\overline{u}(x)$ is the average population fitness

$$\overline{u}(x) = \sum_{j=1}^{n} x_j u_j(x) \tag{3}$$

Table 2 Representation of layers with game theory applications [18]

OSI layer	Application area	Game theory-specific application
Transport	Cell selection Admission or acceptance of call Load control	Inter- and intra-cell games Distribution of request among service providers Call acceptance depending on service provider and customer context Session termination based on service provider and customer context
Network	Routing	Forwarding and routing
Data link	Transmission medium and access control	Slotted ALOHA access Access control to interference channel
Physical	Allocation of spectrum Power control Cooperation in communication Multiple-input and multiple-output (MIMO) networks or systems	Sharing of spectrum Transactions on spectrum Code-division multiple access (CDMA) networks Encode–decode and forward cooperation Power management

There exist two contributions of researchers in evolutionary stable states. First, researchers proved the existence of level-k equilibrium, a subset of Nash equilibrium of repeated games, ubiquitous and traceable in symmetric games. Second, stability analysis that depends on evolutionary stable states does not suite evolutionary and repeated games.

Stability and optimization analyses are some of the major research perspectives in game theory. There exists cross-layer optimization for some combinations of frameworks in game theory which are used to solve problems in telecommunication domain [18]. As listed in Table 2, some application fields are organized under each layer of OSI [19, 20]. Processes such as load control and network delay cannot be explicitly allocated to a single layer in OSI. Few researchers focus on the control mechanism of cross-layer optimization techniques, thus involving multiple OSI layers. Earlier studies of game theory models in the context of networking investigated the network structure operating points.

Application Area in Hardware Layers of OSI

From the perspective of physical layer and the topology design in the application area, the performance is computed through a function to estimate signal-to-interference-noise ratio (SINR) that nodes or players receive. Each player or user's utility is diminishing in their power levels but increasing in their signal-to-interference-noise ratio [21]. If the power levels of all the users are fixed, then increasing a players or nodes power level will mirror the same in the players or nodes signal-to-interference-noise ratio level collectively. This concerns the power control application area with respect to CDMA networks [22] in game theory application. In game theory, there exist similar applications like CDMA where the objective is to minimize the overall transmission power, by adjusting the rate of transmission and the power control mechanisms.

With respect to data link layer, selfish users in game theory applications obtain an unfair share to access channels which in turn maximize their utility. This action of selfish users [6, 23] degrades the users' ability to access channels. Thus, need for medium access control should be aligned in a specific manner. Since many users try to transmit data as soon as possible, they act simultaneously resulting in access failure. Retransmission due to access failure may cost the users as well. To solve this issue to some extent, slotted ALOHA was introduced. Here, time is divided among users as slots following a specific synchronization method. All the users are aware of slot boundaries; when a particular user is in need of accessing a shared channel, he waits for slot boundary to start his transmission [24]. But the problem occurs when two or more users try to transmit data in a particular slot simultaneously [9].

Network layer functionalities include establishing routes and packet forwarding along those routes. Concept of game theory in network layer helps in deciding the better route through which packets can be forwarded and to decide whether or not to transmit the received packets. Game theory is a worthwhile concept in this context due to the reason where the nodes in a network need to decide individually to perform any action while considering the behavior of other nodes' actions [25].

In transport layer, the main part to address is congestion control mechanisms. Also, congestion avoidance needs to be considered. To control the network load, admission of new users should be restricted. Cell selection determines which base station (BS) is the optimal one to choose for a mobile terminal. Cell selection [26] is a two-tier game holding the first tier which is an inter-cell game where a mobile terminal chooses a cell from the computed payoff according to a selection strategy; whereas, in second tier which is an intra-cell game, mobile terminal chooses an appropriate source in the cell to realize the best payoff.

The gap that was found in the earlier research in the field of optimizing topology design was that load control and call admission control were not framed and addressed in an organized way. Since, topology design is completely based on dynamic network, the existing architecture or framework developed on Nash equilibrium is not optimum [12].

To broaden the research domain in game theory, fields like economics, software-defined networks (SDN), renewable energy, auctions, wireless sensors, manufacturing field, etc., are considered in analyzing the problems and finding out the researcher outcome or solution for those. Table 3 showcases the various fields in which game theory plays a vital role in problem solving.

In the following sections, discussions related to game theory challenges, incentive mechanisms, game theory applications and the model to bridge the gap are listed.

Table 3 Various fields of research in game theory—problem, solution and methodology

Domain	Reference number, author and year	Problem	Solution	Approach/method used
Manufacturing field	[27] (S. Bysko and J. Krystek 2019)	Car sequencing problem	Modify structure of car production line through optimization techniques like genetic algorithm (GA) or ant colony optimization (ACO)	Buffer slot assignment
		Color batching problem	ACO/heuristic computation	Buffer-out shuttle
Economics—auctions	[28] (M. Ilie, S. Ilie and I. Muraretu 2018)	Optimizing bid strategy	Three-tier crowd financing system	Semi-truthful strategy
	[29] (A. Byde 2003)	Multi-agent allocation problem	Strategy optimization	Revenue equivalence theorem
Renewable energy	[28] (M. Ilie, S. Ilie and I. Muraretu 2018)	Selecting appropriate energy source	Molten salt thorium reactor	Hydropower plant
		Nuclear waste	Modular fission reactors	Lossless electrical or thermal power transport or smart distributed grid
Software-defined networks (SDN)	[30] (M. Abderrahim, A. Ben Letaifa, A. Haji and S. Tabbane 2018)	Smart node placement problem	Mobile edge computing (MEC)	Embed MEC server in smart node
	[31] (B. Wang, Y. Sun and X. Xu 2019)	To produce efficient anomaly detection scheme (ADS)	Statistics or machine learning-based ADS	Strategy selection module
	[32] (B. P. R. Killi, E. A. Reddy and S. V. Rao 2018)	Effective network partitioning for controller placement	Clustering/partitioning/assignment heuristics	Spectral clustering K-means algorithm

(continued)

Table 3 (continued)

Domain	Reference number, author and year	Problem	Solution	Approach/method used
Wireless sensor networks (WSN)	[33] (I. Ben Abid and N. Boudriga 2013)	Behavior detection of WSN	Detect and solve selfishness in network	Coverage eligibility rule, coverage maintenance protocol
	[34] (A. Agah, S. K. Das and K. Basu 2004)	Security monitor in WSN	To embed cooperation, reputation and quality of service (QoS)	Compute equilibrium at strategy pair
	[35] (M. Abd, S. F. Majed Ai Rubeaai, K. Tepe and R. Benlamri 2015)	Nodes' lifetime depletion	To balance traffic load in network	To develop a three-dimensional game-theoretic energy balance (3DGTEB) protocol

Discussions

Game Theory Challenges

The usage of game theory in networks comes with certain challenges which include [36, 37]:

- *Nodes coherence:* Game theory is based on the supposition that players involved in a game are rational throughout the game or till the game ends, thus assuming players ask for their interest in a rational manner. But when dealing with mobile terminals or nodes, this rational approach toward the nodes is not guaranteed.
- *Nodes cooperation:* To maximize player's payoffs, in cooperative games, players tend to collaborate with other players. In some situations, co-players might choose to behave selfishly to improvise or optimize their profit. To handle this situation, certain incentive mechanisms are enforced for cooperation, and disincentive mechanisms are implemented for selfish or immoral behavior of the nodes.
- *Computation of payoff:* Player's perception toward their performance and satisfaction plays a key role in computing payoff. Utility function hereby interprets the players' game playing structure or game-theoretic framework, hence computing a payoff that maximizes the profit of the self.
- *Non-guaranteed existence of Nash equilibrium:* In game theory, a frequent check needs to be done to confirm whether all players reach Nash equilibrium or not. Even after reaching equilibrium, check has to be made to confirm the existence of multiple equilibria. In such case of multiple equilibria, most efficient one should be considered.

Incentives in cooperation

To guarantee optimization among nodes for a better possible solution, collaboration is required; but, cooperation among nodes cannot be guaranteed. In most of the situations, players or nodes try not to cheat or try not to act selfishly within a sharing or communicating environment, inversely they obtain optimal solution by sharing the resources uniformly [38, 39]. Some situation occurs where players tend to question why they should not act selfishly or try to cheat. To make it clear to the respective nodes, value of node punishment is computed either through means of reward or penalty. Resource offering nodes should be rewarded, and nodes that act selfishly should be kept isolated from the resource and network. There exist incentive mechanisms [40] to make players to cooperate and subside their selfish behavior. Reputation based and credit based are two important incentive mechanisms which build cooperation among players or nodes. Credit-based mechanism has been used in network routing formulations [41]. Nodes are credited with monetary benefits to pay off users for packet forwarding that they receive from other nodes. For retransmission, battery costs or packet loss, credit-based mechanism paves way to use this monetary benefit. Sprite is a well-known credit-based scheme to solve the routing problems of self-interested nodes [42]. This mechanism is more advantageous when nodes are connected in a large-scale network with selfish nodes. On the other hand, in reputation-based mechanism, the reputation is computed either in a central hub (centralized) or at individual node (decentralized). Reputation conveys willingness of players to contribute to entire network by resource sharing.

Game Theory and Next-Generation Networks

Heterogeneous networks are likely to have ubiquitous connection and have the common services for users. This accounts good quality of service (QoS) for users even when they migrate from one network to another. Since different networks are available simultaneously, heterogeneous services [43] support agents like transparent gateway and generic connectivity. As discussed, nodes cooperation is one of the major issues in self-organized networks. Including the concept of cooperation between nodes [44] brings a new form of assortment among nodes resulting in enhanced reliability in communication, coverage extension and minimized power consumption. But the fact is that even in the channel variation and effects of shadowing, nodes or terminals are less vulnerable and can transmit packets in lesser power levels to bring better throughput resulting in increased battery life. Game-theoretic models focus on these quality services of nodes and focus on rewarding or punishing users based on their behavior [8, 45].

Defining Objectives on Game Theory Applications

For a network to be arranged in a specific topology, objectives need to be framed for game theory-specific application based on the purpose and computation parameter. To accomplish this for game theory-specific application, price of anarchy and price of stability need to be computed and mapped across players or users as listed in Table 4. One important attribute to be considered for comparison is game type. There exist two types, non-cooperative and cooperative games. If players or nodes cooperate while playing an n-player game, then the need for analysis on selfish node behavior is not necessary since the set of protocols presented to play a game is accepted by all the players. So, this paper mainly focuses on non-cooperative games where the equilibrium varies based on the selfish behavior of players or nodes.

Table 5 lists the sample data from [dataset] including IP address, no. of players, hop count, price of anarchy (PoA) and price of stability (PoS) for transport layer. In Fig. 2, the comparison between price of stability and price of anarchy is discussed with respect to all the four layers listed: transport, network, datalink and physical layer. X-axis represents the hop count—number of hops a player takes to reach the

Table 4 Mapping application of game theory with computation parameters, price of anarchy (PoA) and price of stability (PoS), for specific players or users

Reference number, author and year	Game theory-specific application	Purpose	Computation parameter	Game type	Players/users
[1] (A. Nahir et al. 2014) [13] (A. R. Romano et al. 2019)	Inter- and intra-cell games	Cell that accomplishes service requirements	PoA	Non-cooperative	Service provider, user or terminal
[15] (C. W. Zaw et al. 2020) [2] (Correia A. D et al. 2019)	Distribution of request among service providers	Allocate requests to providers	PoS	Non-cooperative	Service provider
[5] (Fischer et al. 2014) [26] (H. Wang et al. 2018)	Acceptance of call depends on service provider and customer context	To choose if service request is beneficial for players	PoA	Non-cooperative	Service provider, User or terminal
[6] (Erik et al. 2012) [17] (Chernov et al. 2019)	Session termination depends on customer and service provider context	To choose if session termination affects players	PoA	Non-cooperative	Service provider, user or terminal
[21] (E. Nekouei et al. 2017)	Forwarding and routing	To decide forwarding of packets to players	PoA	Non-cooperative	Users or terminals
[7] (Farza et al. 2017) [25] (Paul, et al. 2012)	Slotted ALOHA access	Minimize collision through random access	PoS	Non-cooperative	Users or terminals
[9] (F. Salehisadaghiani 2014)	Access control to interference channel	Sharing access to interference channel	PoA	Non-cooperative	Users or terminals
[3] (M. Felegyhazi et al. 2006)	CDMA networks	To set transmission power with minimum interference	PoS	Non-cooperative	Users or terminals

(continued)

Table 4 (continued)

Reference number, author and year	Game theory-specific application	Purpose	Computation parameter	Game type	Players/users
[36] (K. Jin et al. 2020) [40] (Tim Roughgarden et al. 2015)	Transactions and sharing on spectrum	Spectrum distribution to maximize fairness	PoS	Cooperative/non-cooperative	Service provider, user or terminal
[44] (J. Li et al. 2016)	Encode–decode and forward cooperation	Allocating terminal strength	PoA	Cooperative/non-cooperative	Users or terminal
[45] (Y. Aziz-Alaoui et al. 2014)	Power management	Link power allocation	PoS	Non-cooperative	Links

Table 5 Sample data from [dataset] listing values of price of anarchy and price of stability for transport layer

Network serial/internet protocol (IP) address	No. of players	Hop count among players	Price of anarchy (PoA)	Price of stability (PoS)
103.40.196.17	122	61	0.988476	16.53
103.40.201.57	169	115	1.09346	40.57
103.40.198.121	217	178	1.678715	51.28
103.40.197.65	254	127	1.754702	61.09
103.40.202.74	69	51	2.069124	69.19
103.40.199.195	127	110	2.431693	76.95
103.40.203.212	177	102	1.126686	19.4
103.40.202.43	250	185	1.83927	51.66
103.40.198.12	70	57	2.00327	63.64
103.40.206.81	124	94	1.76251	62.26

destination from source—among the players connected in a network. Y-axis represents the number of players or users actively participating in a network. In transport layer, it is observed that as the number of players and hop count increases, price of stability increases and reaches a stable state, whereas price of anarchy diminishes from the initial phase where there exists selfish behavior of players or users [46]. These stages are observed in all the four OSI layers with same set of players or users. Also, this is where we justify why the analysis is progressing only in these four layers of OSI.

First, to have a recap, among the seven OSI layers, physical, datalink, network and transport are hardware layers, whereas session, presentation and application are software layers. Players' or users' behaviors affect only the hardware layers (layers through which data actually navigate), if there exist any selfish node, then it affects the data transportation in these hardware layers.

Second, in software layers, more focus is on how to present data to the other end, whereas in hardware layers, more focus is on data navigation. Once the data successfully pass through the hardware layers even with the existence of selfish behavior of certain nodes, topology selection and optimization over it are applied to enhance the design of communication networks.

Reasoning Analysis on Transport Layer

With due interest on games with cell types, the service requests are processed and allocated in transport layer by computing price of anarchy and price of stability (Table 4) in a non-cooperative environment. Also, to choose if a particular service request is beneficial for players is decided in transport layer. Since players are in a

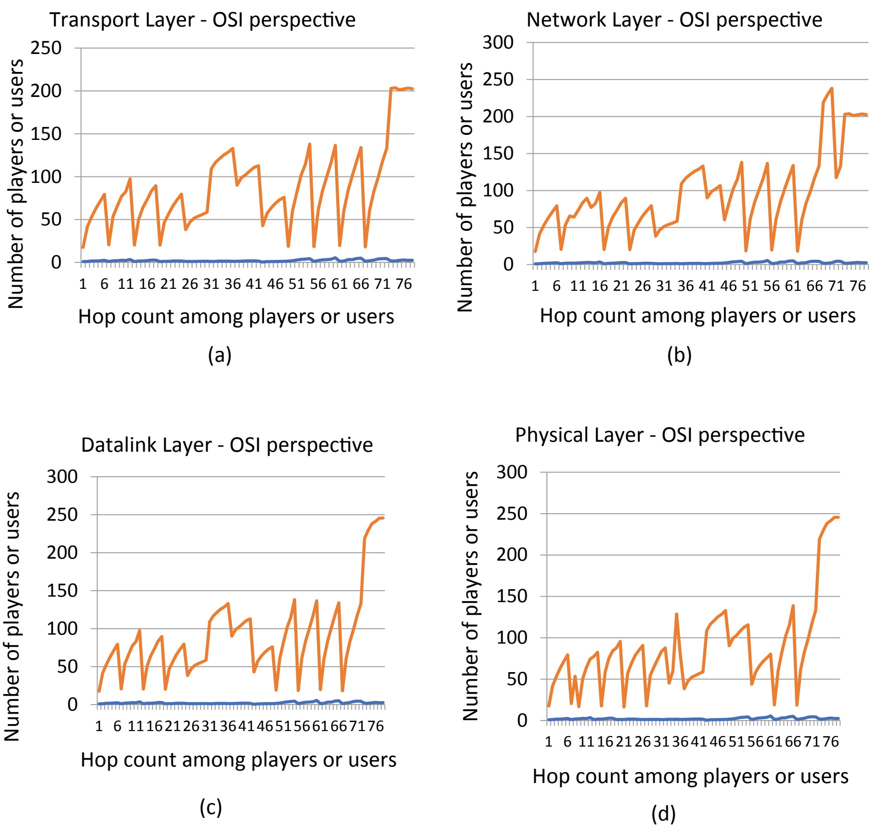


Fig. 2 [Dataset] represents the analysis on computing price of stability and price of anarchy for certain number of players with given hop count in physical, datalink, network and transport layer.

indicates Price of Stability

indicates Price of Anarchy

non-cooperative environment, session termination might affect a player's decision making. The problem in transport layer is formulated as shown in Eqs. 4 and 5:

$$\min_{s,f} P(s, f) = \sum_{l} f_l c_l(f, s) \tag{4}$$

$$z(f, s) = \min_{f} \sum_{l} \int_{0}^{f_l} c_l(x, s)\mathrm{d}x \tag{5}$$

where s is a vector which holds the values of control variables; f is a vector of volumes in traffic; l is the network link index; c is the cost of link travel time. Equation 4 represents the objective function of the authority of transport layer. Equation 5 represents the travelers or players objective function [47]. The authority forecasts the reaction of travelers to each viable set of control variables and selects the set that will make the travelers choose a pattern of paths that is optimal from the point of view of the authority.

Reasoning Analysis on Network Layer

In the perspective of network layer, a service request is sent to transport layer only if the forwarding and routing of packets are having promising path to deliver to accomplish this price of anarchy to be computed since players are in a non-cooperative environment as listed in Table 4. To perform the reasoning analysis, hidden nodes in the network should be computed [48]. In Fig. 3, the simulation topology was designed based on the percentage of total hidden nodes in the network to evaluate the protocol's performance for all possible scenarios. It is necessary to calculate the theoretical average number of hidden stations based on the point of view of each

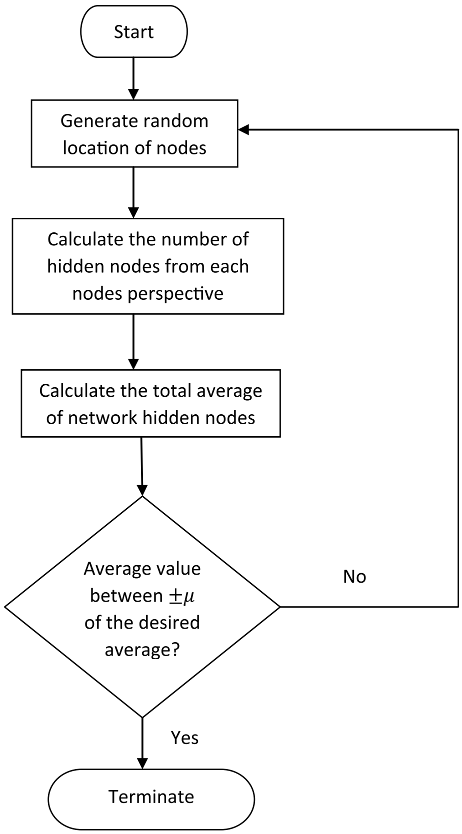


Fig. 3 Computation of network with hidden node percentage (μ)

station about the distribution of node locations around the access point (AP). A uniform spatial distribution is assumed in this analysis.

Reasoning Analysis on Datalink Layer

Datalink layer plays a major role in handling and protecting data where collisions among data that are moving in a path through a prescribed architecture are minimized through random access. Along with this, sharing the access to interference channels is also addressed. Thus, price of stability and anarchy should be computed for players handling a competitive environment as shown in Table 4.

Reasoning analysis on Physical layer

Sharing and allocation of spectrum in physical layer lay a strong base for fairness among players who share data in a network where several unknown players are connected with each other. When data packets are transmitted from one node to another, data strength gets degraded. To enhance this, encoding and decoding data before forwarding to another node were performed [49]. Fairness valuation is performed with the computed values of price of stability and anarchy (Table 4). Game theory is used not only to solve the physical layer protection problem, including the situation of communication subject reaction, but also situations in which various communication subjects communicate with each other (promote or limit).

Let Alice, Relay and Eve be game players; *a*, *r* and *e* be their strategies, respectively. In current physical layer protection approaches, the behavior of the player is compatible with the approach. In the model, *u* is defined as payoff and is frequently represented with a safe rate, probability of outage, mutual information, etc. [50]. The payoff of Eve is represented by the reduction in the safe rate, which is based on *zero-sum analysis.*

$$u_{sys} = \sum_{i=1}^{n} u_{A_i} + \sum_{k=1}^{l} u_{E_k} \tag{6}$$

where the payoffs of Alice and Eve are defined as u_{A_i} and u_{E_k} in Eq. 6, subject to $u_{E_k} \leq 0$ (represents damage is bigger) and u_{sys} is the secure payoff of the system.

Conclusion and Future Work

Game theory and its related work based on layered perspective give a broad picture about the challenges, existing opportunities, computing payoff matrix using Nash equilibrium and an action to be taken on incentive mechanisms. Continuous monitoring of behavior of nodes plays a vital role in moving forward toward incentive mechanisms. Issues related to computing price of anarchy and price of stability are addressed to enhance the strategy profile of players participating in a network. This, in turn, paved way to identify the research gap in load and admission control in optimizing topology design. To achieve this, modified Folk theorem needs to be implemented for further optimization or enhancement. Various optimization techniques are in the path of this research gap that are to be identified and used to accomplish the goal.

References

1. Nahir A, Orda A, Freund A (2014) Topology design of communication networks: a game-theoretic perspective. IEEE/ACM Trans Netw 22(2):405–414. https://doi.org/10.1109/TNET.2013.2254125
2. Correia AD, Stoof HTC (2019) Nash equilibria in the response strategy of correlated games. Sci Rep 9(1):1–8. https://doi.org/10.1038/s41598-018-36562-2
3. Felegyhazi M, Hubaux J, Buttyan L (2006) Nash equilibria of packet forwarding strategies in wireless ad hoc networks. IEEE Trans Mob Comput 5(5):463–476, May 2006. https://doi.org/10.1109/TMC.2006.68
4. Fang Z, Huang L, Wierman A (2019) Prices and subsidies in the sharing economy. Perform Eval, p 136. https://doi.org/10.1016/j.peva.2019.102037
5. Fischer A, Herrich M, Schönefeld K (2014) Generalized Nash equilibrium problems—Recent advances and challenges. Pesquisa Operacional—Brazil Oper Res Soc 34(3):521–558, Sept 2014. https://doi.org/10.1590/0101-7438.2014.034.03.0521
6. Demaine ED, Hajiaghayi M, Mahini H, Zadimoghaddam M (2012) The price of anarchy in network creation games. ACM Trans Algorithms 8(2):Article 13, 1–13, April 2012, https://doi.org/10.1145/2151171.2151176.
7. Salehisadaghiani F, Pavel L (2017) Distributed Nash equilibrium seeking via the alternating direction method of multipliers. IFAC-PapersOnLine 50(1):6166–6171. https://doi.org/10.1016/j.ifacol.2017.08.983
8. Michalopoulou M, Mähönen P (2012) Game theory for wireless networking: is a Nash equilibrium always a desirable solution? In: IEEE 23rd international symposium on personal, indoor and mobile radio communications—(PIMRC), vol 1(1), pp 1249–1255, June 2012. https://doi.org/10.1109/PIMRC.2012.6362537
9. Salehisadaghiani F, Pavel L (2014) Nash equilibrium seeking by a gossip-based algorithm. In: IEEE conference on decision and control, Los Angeles, CA, pp 1155–1160, Dec 2014. https://doi.org/10.1109/CDC.2014.7039537
10. Dilkina B, Gomes C, Sabharwal A (2007) The impact of network topology on pure Nash Equilibria in graphical games. In: Twenty-second AAAI conference on artificial intelligence vol. 1, Vancouver, British Columbia, Canada, pp 42–49
11. Durand S, Garin F, Gaujal B (2019) Distributed best response dynamics with high playing rates in potential games. Perform Eval 129:40–59. https://doi.org/10.1016/j.peva.2018.09.007
12. Nash JF (1950) Equilibrium points in n-person games. Proc Natl Acad Sci 36(1):48–49. https://doi.org/10.1073/pnas.36.1.48
13. Romano AR, Pavel L (2019) Dynamic Nash equilibrium seeking for higher-order integrators in networks. In: European control conference (ECC). Naples, Italy, pp 1029–1035. https://doi.org/10.23919/ECC.2019.8796078

14. Etesami SR, Başar T (2018) Pure Nash equilibrium in a capacitated selfish resource allocation game. IEEE Trans Control Netw Syst 5(1):536–547. https://doi.org/10.1109/TCNS.2016.2631440
15. Zaw CW, Tran NH, Saad W, Han Z, Hong CS (2020) Generalized Nash equilibrium game for radio and computing resource allocation in co-located MEC. In: IEEE international conference on communications (ICC). Dublin, Ireland, pp 1–6. https://doi.org/10.1109/ICC40277.2020.9149267
16. Das R, Goswami S, Konar A (2019) Relationship between Nash equilibria and Pareto optimal solutions for games of pure coordination. In: International conference on computing, communication and networking technologies (ICCCNT). Kanpur, India, pp 1–7, https://doi.org/10.1109/ICCCNT45670.2019.8944817
17. Chernov AV (2019) On some approaches to find Nash equilibrium in concave games. Autom Remote Control 80(5):964–988, May 2019, https://doi.org/10.1134/S0005117919050138
18. Gadjov D, Pavel L (2019) A passivity-based approach to Nash equilibrium seeking over networks. IEEE Trans Autom Control 64(3):1077–1092. https://doi.org/10.1109/TAC.2018.2833140
19. Banner R, Orda A (2007) Bottleneck routing games in communication networks. IEEE J Sel Areas Commun 25(6):1173–1179. https://doi.org/10.1109/JSAC.2007.070811
20. Martínez-Cánovas G, Del Val E, Botti V, Hernández P, Rebollo M (2016) A formal model based on game theory for the analysis of cooperation in distributed service discovery. Inf Sci 326:59–70. https://doi.org/10.1016/j.ins.2015.06.043
21. Nekouei E, Alpcan T, Nair G (2017) Nash equilibrium approximation under communication and computation constraints in large-scale non-cooperative games. In: Asian Control Conference (ASCC), Gold Coast, QLD, pp 2083–2088, https://doi.org/10.1109/ASCC.2017.8287496
22. Hao W, Wang W, Wen C, Li Z (2018) Game theoretical security detection strategy for networked systems. Inf Sci 453:346–363. https://doi.org/10.1016/j.ins.2018.04.051
23. Komali RS, MacKenzie AB, Gilles RP (2008) Effect of selfish node behavior on efficient topology design. IEEE Trans Mob Comput 7(9):1057–1070. https://doi.org/10.1109/TMC.2008.17
24. Shi C-X, Yang G-H (2019) Distributed Nash equilibrium computation in aggregative games: an event-triggered algorithm. Inf Sci 489:289–302. https://doi.org/10.1016/j.ins.2019.03.047
25. Krstic P, Başar M, Frihauf T (2012) Nash equilibrium seeking in noncooperative games. IEEE Trans Autom Control 57(5):1192–1207, May 2012, https://doi.org/10.1109/TAC.2011.2173412
26. Wang H, Deng J, Wang C, Sun W, Xie N (2018) Comparing competition equilibrium with Nash equilibrium in electric power market. CSEE J Power Energy Syst 4(3):299–304. https://doi.org/10.17775/CSEEJPES.2015.01140
27. Bysko S, Krystek J (2019) A game theory approach for solving the new concept of car sequencing problem. In: 5th international conference on control, automation and robotics (ICCAR). Beijing, China, pp 531–535, https://doi.org/10.1109/ICCAR.2019.8813372
28. Ilie M, Ilie S, Muraretu I (2018) Applying game theory economics to clean renewable energy source implementation in urban areas. In: International conference on applied and theoretical electricity (ICATE). Craiova, pp 1–6, https://doi.org/10.1109/ICATE.2018.8551394
29. Byde A (2003) Applying evolutionary game theory to auction mechanism design. In: EEE international conference on E-commerce, Newport Beach, CA, USA, pp 347–354, https://doi.org/10.1109/COEC.2003.1210270
30. Abderrahim M, Ben Letaifa A, Haji A, Tabbane S (2018) A game theory-based effective network management in SDN networks. In: 32nd international conference on advanced information networking and applications workshops (WAINA). Krakow, pp 390–395. https://doi.org/10.1109/WAINA.2018.00117
31. Wang B, Sun Y, Xu X (2019) Loose game theory based anomaly detection scheme for SDN-based mMTC services. IEEE Access 7(1):139350–139357. https://doi.org/10.1109/ACCESS.2019.2943056

32. Killi BPR, Reddy EA, Rao SV (2018) Cooperative game theory based network partitioning for controller placement in SDN. In: 10th international conference on communication systems & networks (COMSNETS). Bengaluru, pp 105–112, https://doi.org/10.1109/COMSNETS.2018.8328186
33. Ben Abid I, Boudriga N (2013) Game theory for misbehaving detection in wireless sensor networks. In: The international conference on information networking 2013 (ICOIN). Bangkok, Thailand, pp 60–65, https://doi.org/10.1109/ICOIN.2013.6496352
34. Agah A, Das SK, Basu K (2004) A game theory based approach for security in wireless sensor networks. In: IEEE international conference on performance, computing, and communications. Phoenix, AZ, USA, pp 259–263, https://doi.org/10.1109/PCCC.2004.1394995
35. Abd M, Rubeaai SFMA, Tepe K, Benlamri R (2015) Game theoretic energy balancing routing in three dimensional wireless sensor networks. In: IEEE wireless communications and networking conference (WCNC). New Orleans, LA, USA, pp 1596–1601, https://doi.org/10.1109/WCNC.2015.7127706
36. Jin K, Khalili MM, Liu M (2020) Games on networks with community structure: existence, uniqueness and stability of equilibria. In: American control conference (ACC). Denver, CO, USA, pp 4664–4670, https://doi.org/10.23919/ACC45564.2020.9147822
37. Ahmadyan SN, Etesami SR, Poor HV (2016) A random tree search algorithm for Nash equilibrium in capacitated selfish replication games. In: IEEE 55th conference on decision and control (CDC). Las Vegas, NV, pp 4439–4444, https://doi.org/10.1109/CDC.2016.7798943
38. Law LM, Huang J, Liu M (2012) Price of anarchy for congestion games in cognitive radio networks. IEEE Trans Wireless Commun 11(10):3778–3787. https://doi.org/10.1109/TWC.2012.083112.120371
39. "Nash equilibrium—A strategy of games theory," Madhavi Garikaparthi 4(7):1316–1318, July 2014, 10.37118
40. Roughgarden T (2015) The price of anarchy in games of incomplete information. ACM Trans Econ Comput 3(1):1–20. https://doi.org/10.1145/2737816
41. Mihai-Alexandru S, Noémi G, Ioana LR (2017) Approximation of Nash equilibria and the network community structure detection problem. PLoS ONE 12(5):1–24. https://doi.org/10.1371/journal.pone.0174963
42. Wang X, Xiao N, Xie L, Frazzoli E, Rus D (2015) Analysis of price of anarchy in traffic networks with heterogeneous price-sensitivity populations. IEEE Trans Control Syst Technol 23(6):2227–2237. https://doi.org/10.1109/TCST.2015.2410762
43. Vives X (1990) Nash equilibrium with strategic complementarities. J Math Econ 19(3):305–321. https://doi.org/10.1016/0304-4068(90)90005-T
44. Li J, Kendall G, John R (2016) Computing Nash equilibria and evolutionarily stable states of evolutionary games. IEEE Trans Evol Comput 20(3):460–469. https://doi.org/10.1109/TEVC.2015.2490076
45. Zhang Y, Aziz-Alaoui M, Bertelle C et al (2014) Local Nash equilibrium in social networks. Nat Sci Rep 4, 6224, 1(1):8, Aug 2014. https://doi.org/10.1038/srep06224
46. Malik F, Hanawal MK, Hayel Y, Nair J (2021) Revenue sharing on the Internet: A case for going soft on neutrality regulations. Perform Eval, p 145, https://doi.org/10.1016/j.peva.2020.102152
47. Hollander Y, Prashker JN (2006) The applicability of non-cooperative game theory in transport analysis. Transportation 33(1):481–496. https://doi.org/10.1007/s11116-006-0009-1
48. Najafabadi SHEM, Constantinou CC (2013) Game theoretic approach to medium access control in wireless networks. IEEE Wirel Commun Netw Conf (WCNC) 1(1):872–877. https://doi.org/10.1109/WCNC.2013.6554678
49. Zhao Y, Kou G, Peng Y, Chen Y (2018) Understanding influence power of opinion leaders in e-commerce networks: an opinion dynamics theory perspective. Inf Sci 426:131–147. https://doi.org/10.1016/j.ins.2017.10.031

50. Xiao W, Huang K, Luo X, Hong Y (2013) Study on physical layer security with game theory. In: IEEE 4th international conference on software engineering and service science, vol 1(1):697–700. Dec 2013, https://doi.org/10.1109/ICSESS.2013.6615402
51. [Dataset] Turcanu I, Engel T, Sommer C (2020) Dataset for adaptive content seeding for information-centric networking under high topology dynamics. IEEE Dataport, 16 July 2020 https://dx.doi.org/10.21227/946z-y844

Essay Scoring Systems Using AI and Feature Extraction: A Review

S. Vijaya Shetty, K. R. Guruvyas, Pranav P. Patil, and Jeevan J. Acharya

Abstract An automatic essay grader is a system that scores and evaluates essays. Although automatically generated scores by machines are not considered apt and as accurate as grades generated manually, the systems have been evolving from time to time and have been showing significant accuracy and reliability. The perceived subjectivity of the grading process is considered as one of the main difficulties of grading essays. It has been claimed by many researchers that variations in grades awarded by human evaluators are due to the subjective nature of the essay evaluation which is considered as a main source of unfairness to students. Whereas automated essay evaluation systems are consistent in the way they score the essays and moreover saves time and cost as well. Since now-a-days essays are being considered as an important testing tool for assessing academic accomplishment, integration of new ideas and ability to recall qualities in a student, there is a need and a huge potential for such automated evaluation and grading systems. A deep dive into the characteristics of the automated essay grading system would help researchers to analyze the features based on which essays are graded such as style, semantics and content of the essays. This paper describes and compares different methods based on machine learning, artificial intelligence and natural language processing that can be adopted to evaluate and score essays written by students.

Keywords Automatic essay grader · Machine learning · Artificial intelligence · Accuracy · Reliability

Introduction

In the past few years, there is a bloom in the development in technologies related to the Internet, the online conduction of examinations has obtained great importance in the educational institutions because of the advanced processes where students learn, prepare and attend the examination.

S. Vijaya Shetty (✉) · K. R. Guruvyas · P. P. Patil · J. J. Acharya
Nitte Meenakshi Institute of Technology, Bangalore, India
e-mail: vijayashetty.s@nmit.ac.in

V. Bindhu et al. (eds.), *Proceedings of Third International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 844,
https://doi.org/10.1007/978-981-16-8862-1_4

During the grading process, teachers have to go through a large number of essays. It is always been a matter. This job is relatively lengthy and costly, and sometimes takes a few hours. In addition, grading is often dependent on the teacher's interpretation of the answer given and the required details highlighted in the teacher's response scheme. Teachers' mental health and physical stability will be affected by the labeling task. The invention of an automated system for grading these essays was therefore a significant process. There were several discussions on the usefulness of a computer being used to rate an article, the most controversial was that it is impossible for a computer to have the same comprehending capabilities similar to that of a human reader and hence cannot be able to provide a grade that respects much complex facets of a written document. Determining the tacit sense of amorphous content, such as written text is also considered a challenge encountered by areas of data processing, and a number of remedies were suggested in the area of text mining. Some of these approaches were able to recognize items in text, but the challenge of connecting them and the establishment of relationship to support interpretation and ambiguity was a more complicated issue.

Automatic Essay Grading Systems

As on today essays are regarded as an important tool to evaluate the abilities of a student to recall, creative interpretation of ideas, integrating and organizing them into an article [1]. Now-a-days not only examiners evaluate essays but also machines known as automated essay graders are used to grade essays and moreover manual grading is tedious and likely to be allowing inconsistencies and incorrectness. The latent semantic analysis is used as a data reclaim method in systems performing automated essay grading [2–4].

Over the last few years, plenty of study has been performed on methodologies concerning automatic evaluation of essays, but little has been done to take semantic coherence, syntax and sentiments of the essays into consideration.

The area of automated grading of essays is fairly new, and in its beginning years, still it is considered to have a history of 50 years. A previous program developed for automated notation was in 1966 by Ellis Page, and the system was named as "The Project Essay Grader" (PEG). This system is the root for all other AEG systems. In this model they have used multiple linear regression method techniques for evaluating and predicting the result of the essay. They trained the model with 100–400 essay datasets and predicted the score of unevaluated essays. The main weaknesses of this system were the contextual characteristics of the essay [5, 6].

To solve the issue in PEG, in 1998 "Intelligent Essay Assessor" (IEA) was developed by Landauer, Foltz and Laham. IEA used one of the "Latent Semantic Analysis" (LSA) concepts which means that the meaning of a sequence of text relies on each and every word used along with its meaning and modifying any word would have an effect on the meaning of the passage in a way or another [7].

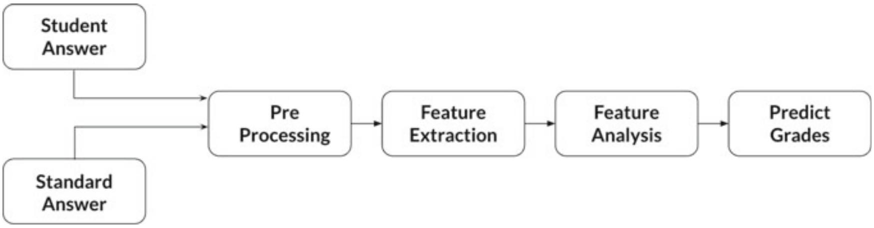


Fig. 1 Flow chart of an automatic essay grader

In 1998, Jill Burstein developed a new system called "E-Rater". He used natural language processing technique and computations using statistics to train the system. He used 270–465 essay datasets to train the system. The system was created by keeping assumptions that there will not be much difference between good essays and poor essays. This system provided automated constructive feedback on things like grammar, mechanics, style and organization which help to improve the writing skills. E-Rater graded essays with 87% accuracy [8].

In 1998, Scott Elliot developed the artificial intelligence (AI)-based automatic scoring system called "IntelliMetric". The architecture contained a combination of artificial intelligence, statistical tools and natural language processing. This system acted and thought like humans. The evaluation was primarily based on five dimensions: Focus and unit, development, organization and structure, punishment structure and mechanics and conventions [8, 9].

In 2000, Heinz Dreher and Robert Williams developed a model called "MarkIt". They used "Multiple Linear Regression and Vector space computation" technology to build the model and using this model they trained the system with 50–100 essays. The system used verb phrases, clauses and noun phrases from the essay. By matching the original word meaning with the meaning of the word in the text, giving it an index number, the text could then be numerically represented. These were consequently used in a classification method to predict the score of a test using multilinear regression [9].

The fundamental flow structure of an automated essay scoring system is shown in Fig. 1 where a model is provided with an input student essay and a set of pre graded essays for learning based on which it is preprocessed to focus only on the essential parameters and eliminate the obvious ones and followed by which it performs feature extraction of those essential parameters and thereby analyses them to award grades and a relevant feedback to the student.

Literature Survey

Previously Annapurna and Dinesh's work in [1] demonstrated the building of a system that was capable of taking handwritten essays as input in image format and a grade was provided as output on the scale of 0–5 showing an integration of automated essay grading and optical handwriting recognition. They used a network consisting of multidimensional long short-term memory and convolution layers to transcribe the hand- written essays. Connectionist temporal classification was used as the loss function and the resulting grades obtained were compared to that graded by human graders. The correlative accordance between the two ways of grading showed a 0.88 Quadratic Weighted Kappa score. A drawback was that in the results indicated through this optical handwriting recognition system, transcription errors were depicted and showed that it could perform well for an automated essay grader system, and prior to that a system for essay grading in Indian context was proposed by Siddhartha and Samia [2] where the system graded English essays and also provided response in such a way that the students could interpret the errors made and if the essay was either influenced by a local language or not. It also depicted some of the approaches used lately for automated assessment of essays such as Project Essay Grader, Intelligent Essay Assessor, Electronic Essay Rater, Intelligent Essay Marking System and used these as a bedrock for the new framework which had the ability to identify local languages hidden in the text and how much it affected the essay and allowed resubmissions with corrections. But this frame- work failed to capture the psychological status of the student writing and submitting the essay which could have been done using psychometric models.

Wang et al. [10] presented a novel automatic essay scoring system using NLP and deep learning technologies which encoded any given essay into serial implantations and harnessed a bi-directional long short-term memory to fetch the semantic details and built a system to learn to focus on valid information correctly like important keywords and sentences in an article and after prediction of the result it provided an evidence for the grades. This model had a matching degree of 0.83 between the predictions and manual grades. The improvements included adding sentiment analysis since the attitude to- ward a topic could be either for or against it and to identify the positive and negative emotions of students and evaluating if their argument provided valuable support. A survey performed by Rudragouda and Syed [3] discussed the different tools that could be used to evaluate short answers and explained them along with their drawbacks. Here different users had tried different tools, and each discussed their tool using machine learning algorithms.

Fragulis's work [4] described an online dynamic examination application plugin named O.D.E.S which was integrated with WordPress that had an ability to assist two types of users, student and the teacher, where teacher had the ability to create, update, modify and remove questions and view students' responses, whereas students took random tests and received their grades. It categorized questions into MCQ and long answer type and auto grading was performed for MCQ but long answer grading had to be done by teachers. This approach created a UI for teachers as well as students to create an examination environment.

Harnit's work [5] proposed a system that incorporated verification of grammar and consistency of surface level as a function of rules, as well as semantic similarity of sentences. They used spatial relationships using graphs and formed unique features by generating relationships within the essay's content, as well as found the semantic similarity between every argument in the essay. The algorithm used 23 outstanding characteristics with high predictive power. They realized that fewer features helped to get rid of redundant data, so forecasts were based on characteristics that were highly robust with noisy data. Prediction of grades was performed by neural networks and the mutual agreement measured between grades awarded by a human grader and that by the system measured with the aid Quadratic Weighted Kappa (QWK) was 0.793. The singular value decomposition (SVD) technique was used to decompose a word by text matrix generated by latent semantic analysis. But it was seen that existing automatic essay grading systems based on latent semantic analysis could not obtain great level of performance to be regarded as an exact substitute for a human grader, where latent semantic analysis is considered a technique in NLP that is used to analyze relationships between a set of text documents by generating a set of concepts related to the documents, hence Monjurul and Latiful [11] designed an automated essay grading system using generalized latent semantic analysis which focused on term vectors rather than dual document term representation that made n-gram through document matrix rather than word by document matrix and was assessed by detailed representation and their experimental results depicted that the system outperformed the existing systems by gaining 89–90% of accuracy, whereas in Akeem's study [12], two methods of detecting similarity in text were used to compare answers submitted by students with the model answers from human graders. The statistics referred to the distance between student's answer with that of the predefined model answer and the generated grades referred to the scores allotted based on the distance of similarity by both the techniques and the model evaluator compared the way in which marks was assigned based on textual similarity to each student's answers by considering the English grammar syntax including the consideration of spelling mistakes by students.

Another project by Abdulaziz [6] presented a hybrid method to design an automatic essay grader that had two components: Writing feature analysis software based on NLP and a neural network-based grader that focused on a series of previously graded essays that assessed the responses of students and awarded a score. It was tested using the datasets from information science and computer science essays by students at University of Mansoura and the findings obtained indicated a consensus with professors in the region of 70–90% showing that the initiative would be a valuable method for the automated evaluation of student essays and would lead to a significant reduction in the cost of essay grading. Meanwhile Ashwin Sathyanarayana [13] presented an approach to prevent grading bias with respect to gender, race, ethnic background and previous performances which was accomplished by prescriptive analysis. Birpal [14] presented a study summarizing the current mechanisms and analyzed the performance of systems that performed automated grades generation for essays, characteristics designed using the corpus-based methods and NLP was found as a remarkable section of the artificial intelligence structure.

According to Rudragouda and Syed Zakir [15] the construction of machine- driven scoring of short type answers usually followed a series of five steps, particularly the creation of a dataset, linguistic communication process, building the model, grading and assessment. The primary necessity for placing and conductivity of communication and referral information sets were unbroken and prepared for the assessment. Then, mistreatment linguistic communication process statistics were designed on the text annotations and grades were evaluated from the marks generated. The design of automated short answer grading system was unfolded over varied implementation methodologies or methods like mapping of ideas from student answers to corresponding model answers, fetching of data from answers to check the closeness between student's answer and the model answers. Govinnage, Deenuka and Weerasinghe's [7] research presented a novel approach using vector space models and NLP (to handle variations in student's essays) methods to automated essay scoring by building a semantic space using same student's answers and providing scores based on similarity in semantics or a significant deviation from the response of the model answer in the student's response.

The research of Mustapha, Idris and Abdullah [16] described the design of the system that assigned grades to associate essays written by students in the history subject, the grading system was constructed upon blocks performing arts of preprocessing documents and questions, making ready model answers and then distribution of scores.

Ratna, Artajaya and Adhi [17] discussed the proposal of a GLSA based AEG system and its performance was compared to the standard latent semantic analysis (LSA)-based system.

The investigations in [18] were focused on the relative usefulness of the use of n-gram extracted features, the mixture of these features and the integration of attribute fetch methods was used in the construction of an automatic essay scoring system. The dissertation focused on modifying the principal component analysis by adding the features of n-grams as an input into it. The marking scheme documents were transcribed into electronic documents. The documents' square measure was preprocessed for the removal of stop words. The documents' square measure was pictured during a vector house model in the form of a document term matrix. The changed principal analysis part was accustomed to cut back the meagerness of the matrix and circular functioned similarly compared the document vector of the solution script of every student with the document vector of the marking scheme. The model was observed to have high positive correlation and when manual scores were compared to those of machine scores, less mean absolute error was seen.

The study in [19] attempted to create a novel linguistic tool feature for NLP-based automated essay scoring systems and a connection analysis that consisted of 2000 graded essays. The RBPRP quantitative relation was extracted that supported the statistical analysis because the formalized style of the thematic feature was seen to own correlational statistics with the manual grades and therefore the outcomes urged that the linguistic tool feature improved the performance of the initial automated essay scoring system. Razon and Vargas in [20] developed a method called concept indexing (CI) that focused on dimensional reduction and latent semantic analysis (LSA) was used to evaluate the output of CI K-means and CI-Fuzzy C-means and found that CI had the ability to replace LSA by analyzing content-based essays.

In [21] Zupanc and Bosnic suggested an enhancement of the original scheme, which had new semantic attributes. Sequential sections of the essay were translated into semantic space and adjustments were calculated between them to approximate text coherence and the method obtained considerably had better grading accuracy than 8 other state-of-the-art automated essay assessment systems. This method failed to detect the correctness of the truth and whether the assertions in the article were valid and consistent.

Researchers in [22] developed a system in such a way that it could mechanically grade the short answers. Typically, totally different individuals wrote short answers directly thus grading was created in line with the factors that short answers written ought to match the means of the first answer outlined because of the answer key. Equal marks for the essays were conjointly provided and that they were provided support to the linguistic relation extracted from the sentences. In this paper, the typical score distinction between AMS-SAE and the human grades was around 0.049, 0.028. Key points compared were function word resolution, negation and morphology, filling within linguistics gaps and their matching. All answers were mechanically mapped to the mapping module. Then the system tried to match the key words and also it worked by tagging a part of speech and knowledge of verbs and noun phrases within the text data. This algorithmic rule tried to match the pattern and if it matched then the answers were correct. In this paper machine-controlled text marker (ATM) was used. ATM used syntax and linguistics analyzer. Machine-controlled marks depicted four forms of variable degrees. They were generation of single word, single worth creation, generation of a brief instructive series of words and representation of patterns in the information. Students then wrote answers to the queries within the kind of Malay text and so supported the similarity with marking scheme marks. AMS-SAE algorithmic rule was taken and used here, for the essays, grammatical relations were extracted and associated degree supported this and an algorithmic rule was generated and accustomed to allot marks.

Arikat in [23] suggested that automated essay grading (AEG) programs currently available used various methods to extract basic written dimensional attributes to evaluate written text. Various neuro-fuzzy methods were used and attempted to address the issue of the short AEG article. The main concept behind this method was to classify the keywords and each had four correspondent words based on unique constraints. The inputs were analyzed by creating two generalized models, namely, the ANN back propagation optimization technique and the subtractive clustering technique. In addition, another model was created to integrate a greater number of keywords by handling twice the number of inputs of the earlier model, an exclusive result with unique limitations was obtained by incorporating neuro fuzzy and subtractive clustering techniques. Greater than 900 databases were used, representing issues relating to the field of IT. To calculate the agreements among the real and the estimated grades of the models the correlation coefficient was used, and two error measurements, that included the root mean square error and mean absolute percentage error, were used to denote the accuracy of the system models. The mean results of all the three models demonstrated the reliability and suitability of the system models designed to handle the problem of the AEG and showed that the results were positive.

Mustapha and Abdullah in [24] suggested that there should be a number of methods to review essays. Approaches varied from mathematical methods such as vector model analysis or LSA, natural language analysis such as phrase types and readability by psychological calculation.

In [8] GLSA vectors were used to increase performance of multiple tasks such as synonym testing, text sorting and clustering relative to conventional LSA methods. The hypothesis was that the grading system based on GLSA should have higher consistency than the essay grading system based on LSA, i.e., reduced disparity between human graded grades and computer-based machine grading, higher Pearson product moment correlation threshold, and marginally poorer processing time efficiency.

In [25] Groza and Szabo analyzed the mechanism of textual encroachment, which was influenced by theories drawn from the domain ontology. Textual implication checked whether the validity of the hypothesis had been accompanied by a text. Textual involvement was enacted in order to equate the student's response to a sample response from ontology. The solutions were validated against many essays published by students in the field of chemistry. The NLP method was used for automated essay scoring in various contexts which enacted the textual link to compare written subject with the standard of the human grader. All these specifications had been created from a given domain ontology in natural language. One of the major advantages was that users were able to choose multiple ontologies for processing text in different domains. Trust in the evaluation, however, depended on the accuracy of the textual system of the enclosure, which relied on the training of domain datasets. The first step of authors was to succeed in configuring unstructured text to allow computer reading and analysis. It was achieved by the UIMA; a basic annotator, UIMA was able to interpret each word in a paragraph, which in turn allowed to be parsed into various sections of speech or called entities. The contents could also be viewed in an XML format, which encouraged data analysis. Using the annotator, organizations such as the "Home News Network" could be tokenized. Thus, the key concept of the next stage was to represent these annotations within a context that would enable more useful details to be extracted. The central points of the short article could be quickly identified. There were six entities in the passage, and a lot of things happening around them. It was, therefore, important to classify the tokens as parts of speech for future study. Thus, coupled with the labeling rubric, it was possible to assess the degree to which the essay pertained to the subject. The scope of this project was carried out in the WA marking heading as set out in the National Assessment Program. Lam et al. [9] discussed a number of schemes to award grades to essays, along with a description of few key problems that remained. A brief review of meta data annotation and text mining techniques had also been presented, outlining the necessary principles along with the feasibility of employing these methodologies to accelerate the action of grading answers as an automatic process. The use of meta data annotation methods was also illustrated for a small part of the text. Normally, the question contained fewer words than the text itself. However, in this scheme, it was thought that the question was the student response part, whereas the documents were the model response by the teacher. With respect to this, if both had a similar match then it was assumed that the paper would have the same span of text as the question. Text in the query and the document had to go through the same preprocessing document, where the protocol could be separated primarily into three actions: stemming, functional removals and identification of features, prior to being sent to the grading system to decide the score.

Rating was calculated based on the proximity of the text in the document to the text query with the help of nearest neighbor algorithm. The ultimate result was a score applied to a non-graded student response.

The limitations of the above-described models include the inability to perform sentiment analysis to judge the approach of the students toward a topic as either positive, negative or neutral. Incorporation of such an approach can help to guide students to adopt to the correct approach with respect to a particular topic by proving a feedback for the same. An efficient feedback mechanism with information specifying the syntactical errors performed by the students while writing the essay may be helpful for the students. A plagiarism check can be incorporated in the systems to monitor any plagiarized contents n the submitted essay from either the open-source articles or from other fellow students.

As a comparison between artificial intelligence and feature extraction techniques, both the methods are expected to perform with equal efficiency but since the AI approach follows a supervised learning technique, it provides grades based on the similarity measures between a model and the submitted answer by comparing the two, whereas the feature extraction approach performs feature engineering and hence awards grades based on the features extracted from the submitted answer and since it is prone to noise, it appears to be complex at certain situations unlike the AI method which is highly robust to noise.

Table 1 shows a comparative analysis between various essay evaluators discussed above. It depicts the techniques and methodologies adopted to, for building the system and their corresponding advantages and limitations. The limitations can be considered as a driving factor for the future systems to enhance the characteristics of the automatic essay grading systems with respect to accuracy, valuable feedback, capture psychological status and capacity to evaluate large sized essays.

Conclusion

In this paper, we have presented some of the approaches to build an automated essay grading system. They can be classified into two broad categories that include feature extraction using natural language processing techniques, latent semantic analysis and generalized latent semantic analysis and a neural network approach using artificial intelligence that is awards grades to essays through the estimating similarity measures by comparing the model and the submitted answer.

The paper also summarizes some of the previous implementations of an automated essay grading systems such as the E-Rater that used statistics to train the system and gained an accuracy of 87%, AEG, PEG that used multiple linear regression techniques for prediction but failed to extract the contextual characteristics of the essay, IntelliMetric that used a combination of AI, statistical methods and NLP to perform grading and Markit that used vector space computation along with multiple linear regression technique.

Table 1 Comparison of various essay grading systems based on their methodologies used, advantages and limitations

Essay grading system	Methodologies used	Advantages	Limitations
Project essay grader (PEG)	Multiple linear regression	Predicts scores of unevaluated essays	Contextual characteristics of the essay
Intelligent essay assessor (IEA)	Latent semantic analysis	Captures transitivity relations, goes beneath the essay surface to capture and quantify its deeper semantic content	Does not provide feedback to improve the quality of writing
E-rater	Natural language processing and statistics	Provides feedback on grammar, mechanics, style and organization and has accuracy of 87%	Assumes there is no much difference between good and poor essays
Intellimetric	Artificial intelligence, statistical tools and natural language processing	Acts and thinks like humans to award accurate grades, web-based tool	Incapable of comparison and hence unable to grade essays that it hasn't seen examples of
MarkIt	Multiple linear regression and vector space computation	Used web phrases, clauses and noun phrases from the essay to match the meaning of the words	Does not provide feedback regarding the quality of the essay
Grading in Indian context	Artificial intelligence, natural language processing, linear regression	Provides feedback to interpret errors due to influence of local languages	Fails to capture the psychological status of the student writing and submitting the essay
Online dynamic examination application plugin	WordPress	Assists two types of users, students and teachers and distinguishes between MCQ and long answer type questions	Long answer type had to be graded manually by teachers
Toward the use of semi-structured annotators for automated essay grading	Nearest neighbor algorithm	Provides various schemes for awarding grades	Essay grading works only when question contains fewer words than the meta text
Enacting textual entailment and ontologies for automated essay grading in chemical domain	NLP	Users can choose between multiple ontologies for processing text in different domains	Annotations used XML format

(continued)

Table 1 (continued)

Essay grading system	Methodologies used	Advantages	Limitations
Subtractive neuro-fuzzy modeling techniques applied to short essay auto-grading problem	Back propagation optimization technique and subtractive clustering	Model was able to achieve the accuracy and handle the problem of AEG	Limitations due to neuro-fuzzy and subtractive clustering techniques

A flow chart indicating the stages of a traditional automated essay grading system is demonstrated that consists of student's answer, a pre graded model answer, a preprocessing phase, a feature extraction stage in case of NLP or matching of answers in case of an AI approach and a final grading stage along with a valuable feedback.

Through the survey conducted it can be concluded that automated essay grading systems operate as an important tool in the educational institutions, where students are graded based on the quality of their essay which is determined by focusing only on important keywords, phrases and sentences through feature extraction and text mining. The various systems also enable the detection of spelling and grammatical errors and provides feedback to students regarding the same.

Based on the insights provided over the currently existing systems the main limitations observed are the detection of attitude of the students toward a particular topic, providing easily understandable feedbacks and a guess on the student's elegancy in English language. The future works include exploring enhanced characteristics to improve the results and make the system more advanced by incorporating mechanisms to predict the mental status of a student using psychometric models or sentiment analysis, also judge a student's fluency in English through the vocabulary and style of language adopted while writing the essay and exploit the feedback features of grading systems to provide a much valuable and meaningful comments to students for their better performance in the future.

References

1. Sharma A, Jayagopi DB (2018) Automated grading of handwritten essays. In: 16th international conference on frontiers in handwriting recognition
2. Ghosh S, Fatima S (2008) Design of an automated essay grading (AEG) system in Indian context. In: Conference paper in international journal of computer applications. Dec (2008)
3. Patil RG, Ali SZ (2018) Approaches for automation in assisting evaluator for grading of answer scripts: a survey. In: 4th international conference on computing communication and automation (ICCCA)
4. Fragulis GF, Lazaridis L, Papatsimouli M, Skordas IA, O.D.E.S.: An online dynamic examination system based on a CMS wordpress plugin. Laboratory of Web Technologies and Applied Control Systems Dept. Of Electrical Engineering Western Macedonia Univ. of Applied Sciences, Kozani, Hellas

5. Janda HK, Pawar A, Du S, Mago V (2019) DaTALab, Department of computer science, Lakehead University, Thunder Bay, ON P7B 5E1, Canada Corresponding author: Vijay Mago, Syntactic, Semantic and Sentiment Analysis: The Joint Effect on Automated Essay Evaluation, Received 10 July 2019, Accepted 25 July 2019, date of publication 5 Aug 2019, date of current version Aug 20
6. Shehab A, Elhoseny M, Hassanien AE, Faculty of computers and information, mansoura university, egypt faculty of computers and information, Cairo University, Egypt. The scientific research group in egypt (SRGE), A hybrid scheme for automated essay grading based on LVQ and NLP techniques
7. Perera GR (2015) University of Colombo, school of computing, Colombo, Sri Lanka, Deenuka N. Perera University of Colombo, school of computing, Colombo, Sri Lanka, A. R. Weerasinghe University of Colombo, school of computing, Colombo, Sri Lanka, "A dynamic semantic space modelling approach for short essay grading". In: (2015) fifteenth international conference on advances in ICT for emerging regions (ICTer)
8. Ratna AAP (2013) Electrical engineering department university of Indonesia Depok, Indonesia, Henry Artajaya electrical engineering department university of Indonesia Depok, Indonesia and Boma Anantasatya Adhi Electrical Engineering Department University of Indonesia Depok, Indonesia, "GLSA based online essay grading system". In: (2013) IEEE international conference on teaching, assessment and learning for engineering (TALE)
9. Lam HW (Sean), Dillon T, Chang E (2010) Digital ecosystems and business intelligence institute curtin business school curtin university of technology Perth, Australia, "Towards the use of semi-structured annotators for automated essay grading". In: 4th IEEE international conference on digital ecosystems and technologies (IEEE DEST 2010) © (2010). IEEE
10. Wang Z, Liu J, Dong R (2018) Intelligent auto-grading system. In: Proceedings of CCIS (2018)
11. Islam MM, Latiful Hoque ASM (2010) Department of CSE, Bangladesh university of engineering & technology, Dhaka, Bangladesh (BUET), automated essay scoring using generalized latent semantic analysis. In: Proceedings of 13th international conference on computer and information technology (ICCIT 2010). Dhaka, Bangladesh, 23–25 Dec 2010
12. Olowolayemo A, Department of cognitive science, faculty of cognitive science and human development, University Malaysia Sarawak, Malaysia. Santhy David Nawi Dept of cognitive science, faculty of cognitive science and human development, University Malaysia Sarawak, Malaysia. Teddy Mantoro department of computer science, faculty science and technology, Sampoerna University Jakarta, Indonesia. Short answer scoring in english grammar using text similarity measurement
13. Satyanarayana A (2019) Computer systems technology New York City college of technology (of the city university of New York) Brooklyn, US, Reneta Lansiquot English New York city college of technology (of the city university of New York) Brooklyn, US and Christine Rosalia curriculum and instruction hunter college (of the city university of New York) New York, US, Using prescriptive data analytics to reduce grading bias and foster student success, (2019) IEEE frontiers in Education Conference (FIE)
14. Kapoor BSJ, Nagpure SM, Kolhatkar SS, Chanore PG, Vishwakarma MM, Kokate RB (2020) An analysis of auto- mated answer evaluation systems based on machine learning. In: Proceedings of the fifth international conference on inventive computation technologies (ICICT-2020)
15. Patil RG (2018) Research scholar, SECAB institute of engineering and Technology, VTU Belagavi, Vijaypur, INDIA faculty, school of computer engineering and technology MIT academy of Engineering, Alandi (D.) Pune, INDIA and Syed Zakir Ali Professor, computer science department SECAB institute of engineering and technology, VTU Belagavi Vijaypur, INDIA, "Approaches for automation in assisting evaluator for grading of answer scripts: a survey". In: (2018) 4th international conference on computing communication and automation (ICCCA)
16. S.M.F.DS. Mustapha Department of Artificial Intelligence, University of Malaya, Kuala Lumpur, Kuala Lumpur, Malaysia, N. Idris Department of Artificial Intelligence, University of Malaya, Kuala Lumpur, Kuala Lumpur, Malaysia, R. Abdullah Department of Artificial Intelligence, University of Malaya, Kuala Lumpur, Malaysia, "embedding information retrieval

and nearest-neighbour algorithm into automated essay grading system". In: Third international conference on information technology and applications (ICITA'05)
17. Ratna AAP (2013) Electrical engineering department, university of Indonesia, Depok, Indonesia, Henry Artajaya electrical engineering department, university of Indonesia, Depok, Indonesia, Boma anantasatya adhi electrical engineering department, university of Indonesia, Depok, Indonesia, "GLSA based online essay grading system". In: Proceedings of (2013) IEEE international conference on teaching, assessment and learning for engineering (TALE)
18. Oduntan OE (2016) Department of computer science, the federal polytechnic, Ilaro, Ogun State, Nigeria, Stephen Olatunde OLABIYISI department of computer science and engineering, LAUTECH, Ogbomoso, Nigeria, Ibrahim Adepoju ADEYANJU department of computer engineering, federal university Oye-Ekiti, Ekiti state, Nigeria and Ibrahim Adepoju ADEYANJU department of computer engineering, federal university Oye-Ekiti, Ekiti state, Nigeria, "A modified principal component analysis approach to automated essay-type grading" In: FTC 2016—future technologies conference 2016 6–7 Dec (2016). San Francisco, United States
19. Wang H, PLA university of foreign languages Luoyang, China, Keliang Zhang PLA university of foreign languages Luoyang, China, Yali Li The ThinkIT speech laboratory, institute of acoustics, CAS Beijing, China and Yong Hong Yan The ThinkIT speech laboratory, institute of acoustics, CAS Beijing, China, "A new linguistic feature for automated essay scoring
20. Razon1 AR, Vargas MLJ, Guevara RCL, Naval Jr PC, 11 Department of computer science, university of the philippines electrical and electronics engineering institute, University of the philippines communication arts department, university of the philippines integrated school. In: Automated essay content analysis based on concept indexing with fuzzy c-means clustering
21. Zupanc K, Bosni´c Z (2014) Faculty of computer and information science university of ljubljana Ljubljana, Slovenia, "Automated essay evaluation augmented with semantic coherence measures". In: (2014) IEEE international conference on data mining
22. Aziz MJA (2009) Computer science department, FTSM Universiti Kebangsaan Malaysia., Fatimah Dato' Ahmad computer science department, Universiti Pertahanan Malaysia and Abdul Azim Abdul Ghani and Ramlan Mahmod Faculty of Computer Science and Information Technology, Universiti Putra Malaysia, "automated marking system for short answer examination (AMS-SAE)". In: 2009 IEEE symposium on industrial electronics and applications (ISIEA 2009), 4–6 Oct 2009, Kuala Lumpur, Malaysia
23. Arikat YM (2012) Technology and applied science department Al-Quds Open University Jericho , Palestine, "Subtractive neuro-fuzzy modeling techniques applied to short essay auto- grading problem. In: (2012) 6th international conference on sciences of electronics, technologies of information and telecommunications (SETIT)
24. Mustapha SMFDS, Idris N, Abdullah R (2008) Department of artificial intelligence faculty of computer science and information technology university of Malaya, 50603 Kuala Lumpur, Malaysia, "Embedding information retrieval and nearest-neighbour algorithm into automated essay grading system". In: Proceedings of the third international conference on information technology and applications (ICITA'05) 0–7695–2316–1/05 $20.00 © (2005) IEEE
25. Groza A, Szabo R (2015) Intelligent system group, department of computer science, technical university of Cluj-Napoca, Romania, "Enacting textual entailment and ontologies for automated essay grading in chemical domain". In: CINTI 2015, 6th IEEE international symposium on computational intelligence and informatics. Budapest, Hungary, 19–21 Nov 2015

Development of Real-Time Violence Detection with Raspberry Pi

J. Krishna Chaithanya, Mohammand Alisha, S. Manish Sagar, and K. Raghuram

Abstract The detection of real-time object has drawn an increased interest in surveillance strategies, and it is one of the applications of CNNs. This project has focussed on the detection of fire and pistols in places that are tracked by cameras and fires in the home, business explosions, and wildfires are all major headaches that have negative implications for the environment. Mass shooting and violence caused by guns are also on the upward push in certain components of the sector. Such type of incidents is time touchy and may purpose a huge loss to lifestyles and assets. Hence, the proposed system is designed with YOLO v3 that detects perfectly by analysing the video or live feed frame to frame to discover such type of situations in real time and send an alert to the authorities through an email and mobile message. The model has been working well on data sets like IMFDB and Fire Net with accuracy more than 83%. Experimental output satisfies the aim of the proposed design is implemented on Raspberry Pi and that is tested with unique situations, and its detection is also very fast, and it can be installed indoor and outdoor.

Keywords Surveillance system · Violence · YOLO · Raspberry Pi · Deep learning · Video processing · Neural networks · Object tracking

Introduction

Recently, in the field of action recognition, the use of function descriptors around points of interest has become common. This method analyses interventions with gradients, intensity or other local features, and uses the video sequence. The tolerance for posture, occlusion, illumination or deformation has been increased. Current approaches, on the other hand, typically include 3D descriptors in high-resolution

J. K. Chaithanya · M. Alisha · S. M. Sagar (✉) · K. Raghuram
Department of ECE, Vardhaman College of Engineering, Hyderabad, India

J. K. Chaithanya
e-mail: J.krishnachaithanya@vardhaman.org

V. Bindhu et al. (eds.), *Proceedings of Third International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 844,
https://doi.org/10.1007/978-981-16-8862-1_5

video applications on many scales for space-temporal analysis. Function descriptor-based approaches usually use the famous word bag. The performance is only a histogram, which shows the distribution of words as frequencies. The representation of the bag of words generates a vocabulary with k-media clustering, for instance, to obtain the histogram. Section 4 describes the whole process. The purpose of this paper is to analyse the performance of contemporary acts for the identification of hardships in photographs, films, or video surveillance footage. Many past practises in appreciation of acts concentrate on basic human activities such as walking, jumping, or waving hands. Violent behaviour detection has been less studied considering its possible utility. Whereas, many well-studied data sets are available for action identification, large data sets have not been made accessible for violent acts. In this work, we present a weapon-based data set using two of the currently best available action detection methods and MO SIFT to evaluate the effectiveness of combat detection. A detector of violence is immediate relevant both to the domain of monitoring and to online video ranking and tagging. In institutions like the schools, jails, and nursing facilities, the primary role of large-scale supervision programmes is to alert authorities of possible dangers. The demand for automated rating systems processing large numbers of videos uploaded to websites is growing. And not only detection sending alert at right time is more important to prevent the violence, so we have included an alert with captured image of violence.

Literature Review

Celik et al. [1] proposed a work focussing on fireplace recognition in video sequences. The device compares foreground item information with statistical shade statistics for the fireplace. Adaptive history subtraction methods are employed to extract foreground data, and the statistical fire colour model is utilised to identify whether or not the foreground item is a flame candidate. The statistical hearth shade variant is made up of three policies.in keeping with the expense of an RGB pixel's crimson issue must be bigger than the mean of all reddish additives in the image. According to the first rule, the price of the red colour value of a pixel must be greater than the price of the green component, which must be greater than the price of the blue component, according to the following principle. The final rule takes into account the ratio of crimson, blue, and inexperienced components. Trinath et al. [2] for this type of problem, he suggested an IOT-based solution, and he integrated temperature and smoke sensors in their device. The most significant disadvantage of this technology is that the sensors are expensive and delicate, and they can easily be damaged by a variety of natural conditions. Satellite TV for pc-based constructions can disclose a large area; however, the resolution of satellite TV for pc imaging is inadequate [3]. A fire is only recognised after it has grown rather large; therefore, it cannot be spotted in real time. The cost of such structures is likewise high [4]. Weather conditions like as cloudy or rain severely reduce the accuracy of satellite-based woody area hearth detection systems because to limitations resulting from the long scanning length and limited resolution of satellites. Method for portable gun detection is based on R-CNN [5]. The model should be able to identify and classify the any type of guns like shotguns and rifles. However, in order for this approach to locate guns, it must be handled by humans only and not by any other means. The use of scale integrated firearms technology (SIFT) is a visible gun detecting system. In, some other previous works include sounding indicators for the detection or colour detection of signals such as blood. Here, important applications are noted, especially in the monitoring sector, where there are no audio and video greyscale. Though explosions blood and exhaust can provide useful indicators of violence in action films; in real-world surveillance videos, they are rare. In this article, we focus on identification for violence in some environments. Kanehisa [6] specified about the applications of YOLO algorithm to create a fire detection system; he made a data set based on a fire movie data set Grega et al. [7] proposed an image processing and machine learning-based technique for automatically detecting risky situations in CCTV systems. Sliding window approaches, fuzzy classifiers, and canny detectors were used to detect firearms and knives in video. The authors' data set and detection technique were made publicly available [8]. Sungheetha [9] told about the 3D image processing where human–robot interaction allows the objects identified from the environment to be exchanged between humans and robots. The environment is more typically used to create 3D images. Robots, on the other hand, have a hard time detecting these images. For this, we use a 3D to 2D conversion procedure to ensure accuracy. Humans will be able to identify the thing and receive the results. Kayastha [10] has mentioned about interviews she have taken and told about mental illness and behaviour of the person leading to mass shooting. Verma and Dillion [11] made a gun detection model using deep convolutional net (DCN} a state of art faster region-based CNN model for automatic gun detection from clustered scenes. Valanarasu [12] done a work on personality (which is combination of behaviour and emotions) which have a distinct set of characteristics that are influenced by the environment and biological factors, so she have developed machine learning model to predict the personality of a person with comparatively high accuracy. These type of models are HR management and various departments. We also gone through implementing deep learning models on Raspberry Pi and installing libraries in it and to make a portable model to detect and make alert system.

Existing System

An approach for the fusion of an audio-visual sensor is proposed and used for automated detection of violence. Although there are a number of meanings of violence, we consider this paper to be any behaviour that affects others. Multi and unimodal evaluations were obtained by people who provided 3-point scales of aggression. To forecast the multimodal labels from the unimodal labels, there are no specific algorithms. In order to find the structure in the specific technique, we propose an intermediary stage. These characteristics are known to us, and we find a collection of five that affect the fusion process. We use the state-of-the-art features of low-level audio and video to prevent audio and video violence, and we also anticipate the three functionalities we meet. In comparison with traditional fusion techniques, such as function and decision level fusion, we showed the important positive impact that meta-functions can add in the prediction of the multimodal marks. There are models for detecting guns and fire separately but not in the same model; in our model, we have used both fir and gun data set, and there are existing systems which are made only for detection but not implemented any prevention method or alert system to inform to any authorities such as giving an alert mail or message; so, in this system, we have tried to implement alert system so that it is useful for the society.

Proposed System

The suggested test makes use of the YOLO (you look only once) version3 model, which is a neural network-based approach for real-time object recognition, as well as Darknet-based deep learning framework, an open-source neural community in C. Because it provides real-time detection with great accuracy, YOLO v3 is the preferred choice. Darknet 53 employs a topology that consists of 53 convolutional layers, each of which is accompanied by batch normalisation layers and Leaky ReLU activation, resulting in a completely convolutional network (FCN). The total layers employed for the detection venture are 106, making the model bigger than earlier versions. To avoid losing low-stage functions, the model does not use pooling. Layers which are unsampled are additionally stacked on top of the preceding levels to aid in the discovery of small things by preserving minor details. YOLO recognises things in photographs better than sliding window and area suggestion-based algorithms because it receives every element about the entire photograph and the item by taking a look at the complete photo. In the application of photograph classification, the image is divided into n boxes along with confidence ratings are expected to be employed.

YOLO v3 is the top choice because it allows for real-time violence detection without sacrificing high precision. Full layers used for detection are 106, which makes this version bigger than prior versions. To compensate for the lack of low-stage capabilities, the version does not use pooling. The unsampled layers are additionally concatenated with the preceding levels to aid in the discovery of tiny gadgets. By maintaining the tiniest functions not as appealing as the sliding window and inspiration for the location because it receives every detail about the entire shot and the item by seeing the entire shot.

YOLO detects devices in an image fully. Grids are used to split the image. The usage of picture categorisation and localisation on each grid cell should result in N-boundary bins with self-assurance ratings. YOLO detects objects on three scales which are different, from small to large, in layer ranges 82–106. Larger things are recognised using a 13×13 layers; medium objects are recognised using a 26×26 layer as it is good, and smaller objects are recognised using a 52×52 layer [13]. Our data set consists of 3000 images of different types of guns from UGR data set [14]. The data set contains a variety of guns with different angles, positions, and orientations and also 500 images of fire also taken from Google and from [15] FireNet data set to detect fire too, and apart from this, we have also taken some CCTV images of person holding guns and also movie scenes from guns movies database [8] to make the data set more better, so that we could get better accuracy, as well as and that system we are going to implement on Raspberry Pi so that we can make a portable model of our system which is easy for connections and can be kept anywhere and easy to use. And after detecting, we are going capture the image and send an alert message to the mail through this we can know the image of the person too, and to send SMS, we are going to use Twilio API interface to get mobile alerts.

Hardware Description

Block Diagram

The block diagram of proposed system is shown in Fig. 1; it is a basic hardware model which contains a power supply (5 V, 2.4 Amp) which is given through micro-USB cable, a 5 MP Raspberry camera to capture the video and it connected to camera port on the Raspberry Pi and a monitor connected to HDMI port which is used to display the result.

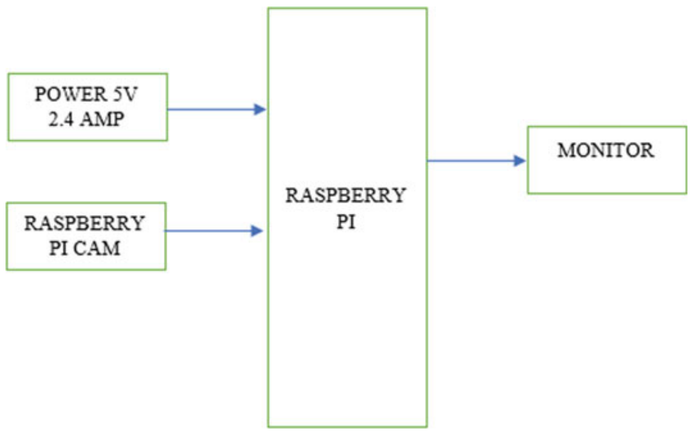


Fig. 1 Block diagram

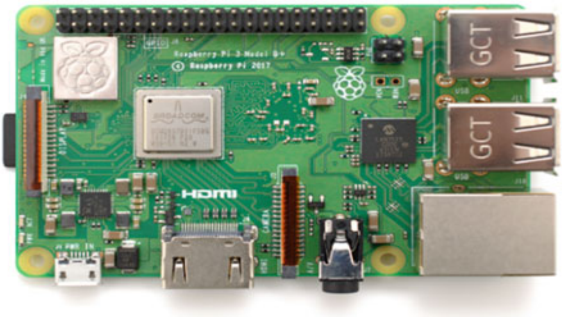


Fig. 2 Raspberry Pi

Raspberry Pi

Raspberry Pi is made by the Raspberry Pi Foundation. The Raspberry Pi project originally focussed on the promotion in schools and developing countries of basic computer science education. The model was popular rather than anticipated, with uses such as robotics sold outside its target market. It is widely used in many fields, for example weather surveillance, due to the low cost, modularity and open design. The HDMI and USB devices are typically used by computers and electronical hobbyists. The Raspberry Pi Foundation established a new entity, called Raspberry Pi trading and installed Eben Upton as chief executive officer in charge of technology development after its release. The foundation moto was to promote basic computer science education in schools and making countries as an education charity. One of the British best-selling computers is the Raspberry Pi. More than 30 million boards were sold as of December 2019. In Pencoed, Wales, most Pis are produced in Sony, whilst other are produced in China and Japan. To start with Raspberry Pi, we should firstly install Raspbian OS in the SD card; after installing the OS, we should the insert into its slot present in the Raspberry Pi USB slots and with the help of HDMI cable connect Raspberry Pi to a monitor; now, perform all the necessary start-up steps and your OS is ready to use. To send the files to the Raspberry Pi, use Win SCP to transfer the files by connecting the IP address (Fig. 2).

Raspberry Pi Camera

You can use the Pi camera module to take good HD videos and photos. It is easy to use; advanced users can provide a great deal for you to gain your knowledge. This Pi cam libraries can also be used to create your camera effects. The units are equipped with a 5 mp, 1080 p 30, 720 p 60 with a fixed focus camera as well as a capture. And it contains a ribbon cable which we need to connect to the CSI port of the Raspberry Pi, and it has many libraries and have built for this use, which includes Pi camera library. After connecting the hardware, we have to configure the Pi to enable the camera,

Fig. 3 Pi camera

and then, we have to use the command "sudo raspi-config" to open the configuration window. Then, under interfacing options enable camera to access, after that finally reboot the Pi and your camera module is ready to use. Then, you can make the Pi to take photos or record videos using simple Python scripts. These are mainly used in home security, and in detecting objects, image processing, and machine learning, the camera module is popular (Fig. 3).

Software Description

Raspberry Pi OS is a Debian-based Raspberry Pi operating system. It is designed especially for Raspberry Pi. Mike Thompson and Peter Green were developed as an independent project for the first Raspbian version. In June 2012, the initial construction was completed. Raspberry Pi OS is highly optimised for the Raspberry Pi range of compact ARM CPU single-board computers. It runs on all Raspberry Pi with the exception of the Pico microcontroller. The Raspberry Pi OS, along with a special topic, uses a modified LXDE with an open-box window manager. A copy is provided with the distribution and the version of the web browser chromium and Wolfram Mathematical algebra and the Minecraft version, called the Minecraft version. After the OS installation completed, the SD card is inserted in Raspberry Pi loaded the Raspbian OS, and then, installation of the necessary libraries is done in Pi terminal with sudo commands; after this, once to run the Python model, the Pi cam turns on and sends the live video feed for the image processing, and if any violence is detected, it inform that violence is detected.

Once the model detects the violence, we have made a Python function to capture the image and send that email to the authorities with alert message with captured image, and apart from this, we also used Twilio to get mobile alert (Twilio messaging is an API to send and receive SMS, MMS, OTT messages globally. It uses intelligent sending features to ensure messages reliably reach end users wherever they are), once you get registered with Twilio it gives a number, and an API token or key that can be placed in our function to send the SMS.

YOLO Detection

YOLO, otherwise you best look once, is a popular actual-time item detection set of rules. YOLO combines what was once a multi-step process; the use of a single neural community to carry out both category and prediction of bounding containers for detected items. As such, it is far heavily optimised for detection overall performance and can run much quicker than walking separate neural networks to locate and classify objects one at a time. It does this by using repurposing traditional picture classifiers for use the regression venture of figuring out bounding bins for objects. This newsletter will most effective examine YOLO v1, the primary of the numerous iterations this structure has long past via. even though the following iterations feature severa upgrades, the simple idea at the back of the structure stays the identical. YOLO v1, known as simply YOLO, can perform quicker than actual-time item detection at 45 frames in step with 2d, making it a tremendous preference for applications that require real-time detection. It seems at the entire image at once, and handiest as soon as—subsequently the name you only appearance once—which permits it to capture the context of detected objects. This halves the wide variety of false-fine detections; it makes over R-CNNs which study special elements of the image separately. Moreover, YOLO can generalise the representations of diverse gadgets, making it extra relevant to a selection of new environments.

YOLO v3 is the top choice because it allows for real-time violence detection without sacrificing high precision. Full layers used for detection are 106, which makes this version bigger than prior versions. To compensate for the lack of low-stage capabilities, the version does not use pooling. The unsampled layers are additionally concatenated with the preceding levels to aid in the discovery of tiny gadgets. By maintaining the tiniest functions not as appealing as the sliding window and inspiration for the location, because it receives every detail about the entire shot and the item by seeing the entire shot. YOLO detects devices in an image fully. Grids are used to split the image. The usage of picture categorisation and localisation on each grid cell should result in N-boundary bins with self-assurance ratings. YOLO detects objects on three scales which are different, from small to large, in layer ranges 82–106. Larger things are recognised using a 13×13 layers; medium objects are recognised using a 26×26 layer as it is good, and smaller objects are recognised using a 52×52 layer (Fig. 4).

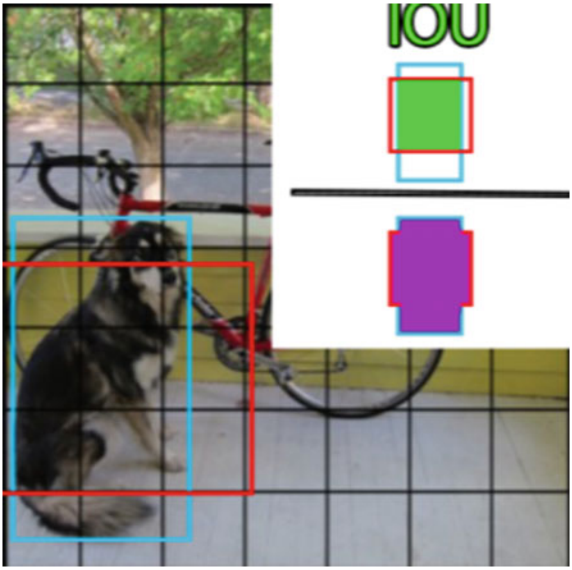


Fig. 4 S × S grid on input

Table 1 Training parameters

Parameters	Description
Max batches	4000
Steps	3200, 3600
Filters	21
Classes	2

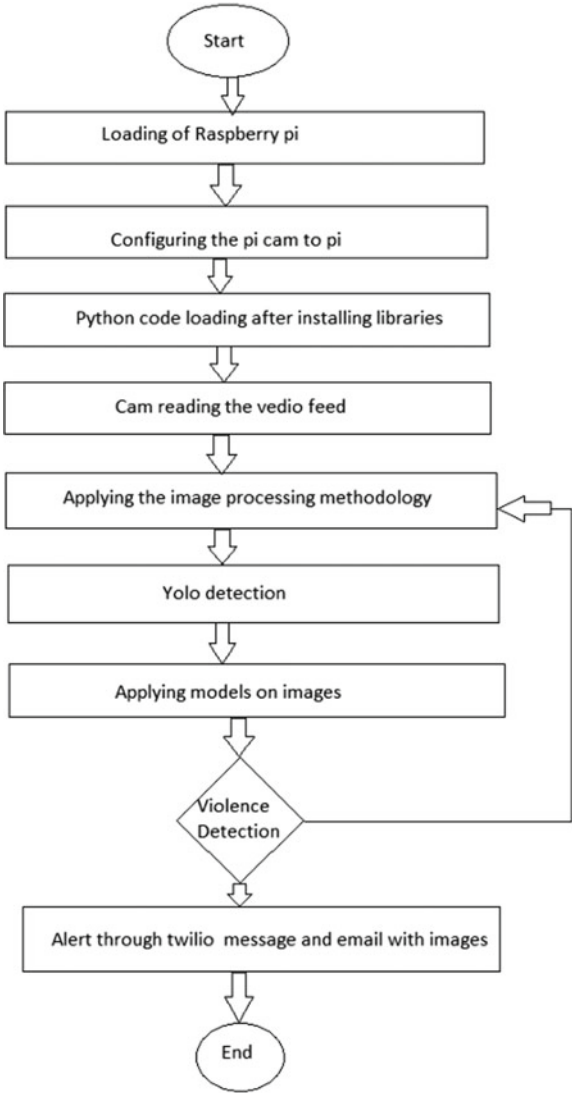
We made a custom data set of 3000 images using UGR data set, FireNet data set, and various images from Google; it contains different types of guns, fire, and CCTV footages, and we have labelled the images in data set; guns are represented by label 0 and by label 1, and our model is trained for the parameters of YOLO as shown in the Table 1, and after the model is trained using the 4000 iterations, the loss we got is 0.2864, and the model also performed good when we give video input well as it does frame-by-frame processing.

Flow Chart

- Start
- Include the necessary libraries and define pins
- Define the check for the Wi-Fi connectivity.
- Read the video from local system and analysis using image processing technics from the users account

- Pi camera capturing the live video and sending to the model
- Simultaneously, checking if the violence detected.
- Sending the alert through captured image through email and message through Twilio API
- Stop (Fig. 5).

Fig. 5 Flow chart



Results and Discussions

The working model of the proposed system is designed and implemented, and the protype of the proposed model is shown Fig. 6; it performs on both video feed as well as live camera feed. We gave different types of videos, images of different guns, fire, etc. the model could detect accurately, and we got accuracy about 0.89 and precision about 0.82 but sometimes the model struggled, but in most of the cases the model showed us the good performance and accuracy. Figure 7 shows us model detecting fire; this video was taken from a movie; Figures 8 and 9 show us the model detecting various guns, and Figure 10 shows us the model detecting knife, and this was a live feed taken from the Pi camera.

Our one more aim of this proposed system to send the alert message to the concerned authorities once the violence is detected; Figure 11 shows us the email alert with a pic captured after the violence detection, and the function could be able

Fig. 6 Prototype of proposed model

Fig. 7 Detecting burning car

Fig. 8 Detecting gun

Fig. 9 Detecting knife in live camera

to capture the image accurately; this will able to know face of the person and will make job of cops easy, and Figure 12 shows us the mobile SMS which is sent with the help of Twilio API.

Conclusion

A low cost, efficient, and deep learning-based solution for the real-time violence detection and recognition of weapons in surveillance videos are based on YOLO algorithm. The system is tested with 3000 images data set which showed all kinds of weapons and also fire in one frame. Raspberry Pi is one chip single-board processor

Fig. 10 Detecting riffle in video feed

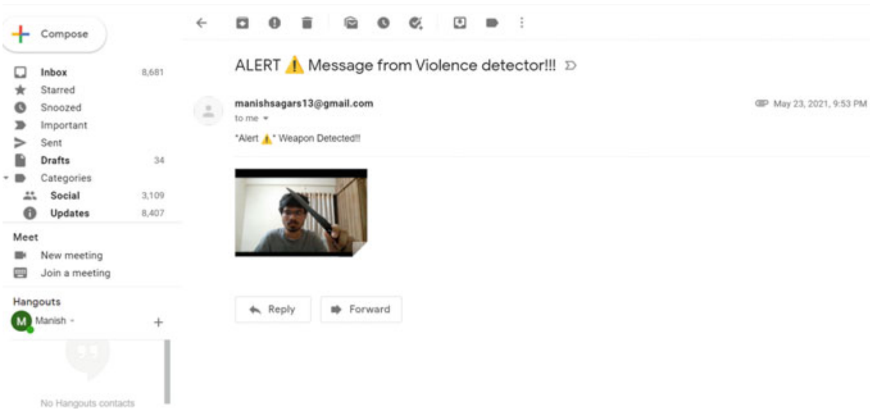


Fig. 11 Alert mail with captured image

with 8 GB RAM, and it also have Bluetooth and Wi-Fi technology. It is useful in running the programme based on object detection and also violence detection and gives a very good accuracy, and one more advantage is in this we have used Raspberry Pi, so we can cut cost as its price is low, and because of its portable nature, we can do much more advancement in this module. We can communicate to the higher authorities or police man to react when the violence has happened and help them to prevent the violence and as model also captures the image of the person it is easy to catch that person if he escapes also. We can also make the robots using Raspberry Pi and detect the weapons using these robots in the army boarders and highly dangerous areas where a person can't go. These type of model helps in prevention of violence and taking the right steps at right time, Though this system is very useful for this

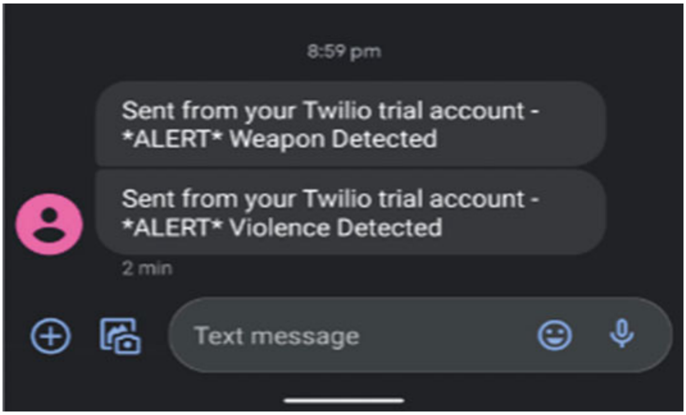


Fig. 12 Alert SMS on mobile

police department, we can't completely rely on this system for the security; at the end, this is a software it can also make mistakes and may lead to false information.

Acknowledgements We sincerely thank the faculty of Department of Electronics and Communication Engineering, Vardhaman College of Engineering, Hyderabad for their help in reviewing of the manuscript and constructive suggestions made at various levels of the research work and thank the management for providing the facilities to carry out our research work at #3021 Lab physically/remotely (Centre of Excellence for IoT).

References

1. Celik T, Demirel H, Ozkaramanli H, Uyguroglu M (2006) Fire detection in video sequences using statistical color model. In: IEEE international conference on acoustics speech and signal processing proceedings. Toulouse, pp II–II
2. Kotiyal S, Thapliyal H, Ranganathan N (2014) Circuit for reversible quantum multiplier based on binary tree optimizing ancilla and garbage bits. In: 2014 27th international conference on VLSI design and 2014 13th international conference on embedded systems. Mumbai, pp 545–550
3. Toreyin BU, Dedeoglu Y, Cetin AE (2005) Flame detection in video using hidden Markov models. In: IEEE international conference on image processing. Genova, pp II–1230
4. Li Z, Nadon S, Cihlar J (2000) Satellite-based detection of Canadian boreal forest fires: development and application of the algorithm. Int J Remote Sens 21(16):3057–3069
5. Celik T, Demirel H, Ozkaramanli H (2007) Fire and smoke detection without sensors: image processing based approach. In: Proceedings of 15th European signal processing conference. Poland, 3–7 Sept 2007
6. Kanehisa R, Neto A (2019) Firearm detection using convolutional neural networks. In: Proceedings of the 11th international conference on agents and artificial intelligence, vol 2, pp 707–714
7. Grega M, Lach S, Sieradzki R (2013) Automated recognition of firearms in surveillance video. In: 2013 IEEE international MultiDisciplinary conference on cognitive methods in situation awareness and decision support (CogSIMA). San Diego, CA, pp 45–50

8. Guns movies database, Katedra Telekomunikacji AGH. [Online]. Available: http://kt.agh.edu.pl/grega/guns/. [Accessed: 30 Mar 2020]
9. Sungheetha A, Sharma R (2021) 3D image processing using machine learning based input processing for man-machine interaction. J Innovative Image Proc (JIIP) 3(01):1–6
10. Kayastha R (2016) Preventing mass shooting through cooperation of mental health services, campus security, and institutional technology
11. Verma GK, Dhillon A (2017) A handheld gun detection using faster R-CNN deep learning. In: Proceedings of the 7th international conference on computer and communication technology—ICCCT2017
12. Valanarasu MR (2021) Comparative analysis for personality prediction by digital footprints in social media. J Inf Technol 3(02):77–91
13. UGR Handgun Dataset, Weapons detection—soft computing and intelligent information systems. [Online]. Available: https://sci2s.ugr.es/weapons-detection. [Accessed: 19 Mar 2020]
14. Fire-Gun Dataset, Kaggle, 18 Mar 2020. [Online]. Available: https://www.kaggle.com/parthmehta15/fire-gun
15. FireNet dataset, GitHub, 11 Dec 2019. [Online]. Available: https://github.com/arpit-jadon/FireNet-LightWeight-Network-forFireDetection. [Accessed: 19 Mar 2020]
16. LabelImg Annotation Tool, GitHub, 30-Jan-2020. [Online]. Available: https://github.com/tzutalin/labelImg. [Accessed: 30 Mar 2020]

Secure Optical Image Encryption and Data Hiding Technique in Compression Domain Using Dual Key-Based Bit Swapping

L. Anusree and M. Abdul Rahiman

Abstract Over the last twenty years or more, researchers have suggested a significant variety of optical information security systems that take advantage of the inherent benefits of optics to outperform existing digital security methods. Along with encryption, hiding secret data plays added advantages in various applications such as defense, biomedical diagnostics, and other security-related imaging applications. In this work, a new technique is proposed along with secret data hiding in the compressed bitstream. American Standard Code For Information Interchange (ASCII) representation is used directly to hide the secret text into the image using least significant bit (LSB) replacing technique. To evaluate the performance, various performance measures such as correlation coefficient (CC) of 0.99, peak signal to noise ratio (PSNR) of 28.53, mean square errors (MSE) of 0.0056, structural similarity index measure (SSIM) of 0.99, mean absolute errors (MAE) of 0.015, and root mean square error (RMSE) of 0.056 values are calculated. From the acquired result, it is quite obvious that this work delivers 0.2% high performance when compared to existing optical encryption methods.

Keywords Optical information · Data hiding · Image compression · Lossless compression · ASCII · Bit swapping

Introduction

Convert bit sequences with high coding capacity and efficient authentication are required in modern high-security encryption and optical communication. As photonic information technology advances, secret tags with high density and security are urgently needed and have piqued curiosity [1]. Information security is becoming

L. Anusree (✉)
LBSITW, Thiruvananthapuram, Kerala, India

M. A. Rahiman
LBSCST, Thiruvananthapuram, Kerala, India

V. Bindhu et al. (eds.), *Proceedings of Third International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 844,
https://doi.org/10.1007/978-981-16-8862-1_6

increasingly important in a variety of applications, such as secure communication routes, secure identification, and safe data storage. Optical approaches have been developed in recent years as prospective alternatives in a variety of specific applications. Because of the exceptional capabilities of optical methods, there is a growing interest in using optical methods to build more complex information security solutions. The groundwork for today's optical encryption technologies was laid by Refregier and Javidi. The recorded speckle pattern is white noise, and no information is shown graphically. To encode an image, double random phase encoding (DRPE) was originally suggested by applying two uncorrelated random phase-only masks within the input image plane and the Fourier transform domain. The plaintext may be transformed into stationary text using the DRPE. Many researchers have been inspired by the DRPE system and have further refined it by adopting it [2].

Integral imaging is a true three-dimensional (3D) imaging technology based on an integrated photographic method that allows us just to record numerous two-dimensional (2D) images from a 3D scene using a lenslet array. These 2D images are referred to as elemental images since they contain the direction and brightness information of a 3D scene [3]. In addition to random phase encryption, characteristics of a light field, including spatial ranges and polarized phases, can be employed as encoding and decoding keys. A technique of space-based optical encryption in which the plaintext picture pixels are randomly relocated to various depth ranges prior to diffractive light field transmission [4].

Deep learning method is also used for optical encryption [5]. Depending on the input, noise following decoding might result in the loss of minor details, prompting the addition of approaches in which the input is encrypted in parts, making use of the fact that each piece is simpler and less impacted by noise [6]. Individual parts can be multiplexed into a single ciphertext, which can result in a significant improvement in decryption quality when recovered. Yet, because each component must be processed separately, the time and effort necessary to access the encrypted data are increased [7].

In this work, a new technique is proposed to perform optical image encryption in the compressed bitstream. The encryption is performed using dual key-based bit patch swapping. Also, the secret text is inserted in the image using LSB replacing technique.

The remainder of this work is structured as follows: Section 2 describes the optical encryption techniques that have been published. Section 3 discusses the suggested optical encryption technology. Section 4 explains the outcome and discussion of the suggested approach, comparative research, and analysis. Finally, Sect. 5, explored the conclusion.

Literature Survey

Previously, a large number of works were presented to implement the optical encryption approach. These various strategies aim to minimize design complexity by improving the algorithm's architecture. This section contains some of the previously suggested efforts for performing optical encryption implementation.

Shuming Jiao et al. a realistic technique for optically encrypting and decrypting a grayscale image using QR codes is provided, which is compatible with standard QR code producers and readers. A grayscale image is converted to a decimal number sequence, which may then be translated to numerous QR codes. A numerical simulation depicts the full process of encrypting and decrypting a grayscale image using this approach. The findings show that a grayscale image may be effectively encrypted and lastly retrieved in a noise-free manner using the suggested approach [8].

Yonggang Su et al. proposed an optical encoding based on the full trinary tree structure for multiple color images. The approach uses encryption modules as branched nodes and raw image components of color as leaf nodes. Each encryption modulesencodes three input images into a complex function, which is subsequently encrypted to a real-value image using phase truncated Fresnel transformations and random amplitude-phase masks. This encryption approach may encrypt several color images into a real-value grayscale cipher image while separating the encryption and decryption routes for each color image. This encryption method enables high-security authority management among several users [9]. Furthermore, the suggested system has benefits such as strong resilience against diverse attacks and good encryption efficiency. Furthermore, when the quantity of plain color photos grows, the excellent quality of the decrypted color images can be retained with embedding capacity enhancement using a hybrid technique [10].

Kang Yi et al. proposed optical encryption using ghost imaging and public key cryptography. The RSA public key algorithm is utilized to overcome the key distribution problem in this system. In addition, the cost of setting up security channels is decreased. In the event of fewer ciphertexts, the use of the CS method provides good quality plaintext reconstruction. This system combines the benefits of the RSA public key algorithm with ghost imaging (GI) encryption techniques to provide security and ease for efficient data transmission [11]. The simulation results validate the method's viability. It has a high level of resistance to statistical analysis and repeated attacks, as well as a high level of resilience [12].

Lina Zhou et al. proposed diffractive imaging optical encryption is subject to learning-based attacks. An opponent can recover anonymous plaintexts from provided ciphertexts using a machine learning assault. In this technique, end-to-end learning is utilized to derive a superior mapping link between plaintexts and ciphertext. Unauthorized users can use trained supervised learning to retrieve unknown plaintexts from provided ciphertexts without the need for optical encryption key retrieval or estimation [13]. As optical experimental and simulation results show, the proposed learning approach is practical and useful for assessing the susceptibility of optical encryption systems with embedding capacity enhancement using a hybrid technique. The trained learning model's ubiquity is also proven, as is the reality that the machine learning model taught to use a database is powerful enough to target multiple databases [14].

Kang Yi et al. proposed a ghost imaging-based camouflaged encryption technique is suggested. A light source lights a camouflaged image with certain modulated patterns during the encryption process, and a produced sequence of intensity is communicated to the recipient as ciphertext. Authorized receivers and prospective eavesdroppers obtain the ciphertext [15]. Only authorized receivers with keys may acquire the hidden image, but eavesdroppers can only obtain camouflaged images if they steal both the ciphertext and the keys. The camouflaged image hides the secret image, which preserves its security by confounding the eavesdropper [16]. GauravVerma et al. proposed based on the phase retrieval method and the phase truncated Fourier transforms (PTFT) scheme, an asymmetric cryptosystem for optical image encryption utilising biometric keys, which tackles the issue of key distribution with increased security in the optical encryption process [17].

It is obvious from the preceding discussion that numerous works have previously been offered to undertake to increase the robustness. The primary disadvantages of prior efforts are their poor quality and lack of security. The major purpose of this effort is outlined in the parts below:

1. To improve the security of optical encryption.
2. To preserve the quality of the image.
3. To create a versatile methodology for developing real-time applications.

The details of the implementation of the proposed work are given below sections.

Secure Optical Image Encryption and Data Hiding Technique in Compression Domain Using Dual Key-Based Bit Swapping

In this work, first, read the image and resize it into 128 × 128. Using bitplane slicing extract LSB. In lossless compression extracting bitstream. Then convert bitstream to matrix blocks. Using two secret keys swap the position of bit randomly. Then it can be converted into the bitstream. The bitstream converts to matrix patch and reverse swapping using two secret keys. After swapping bits converted into the bitstream and decompress the image. Using bitplane slicing extract LSB and ASCII number of image. Finally, receive the secret text behind in the image. Figure 1 shows the secure optical image encryption and data hiding technique in compression domain using dual key-based bit swapping. Fig. 2 showswork flow diagram of proposed optical encryption.

Bıtplane Slicing

Bit plane slicing is a way of expressing an image by using one or more byte bits for each pixel. Only most significant bit (MSB) may be used to represent a pixel,

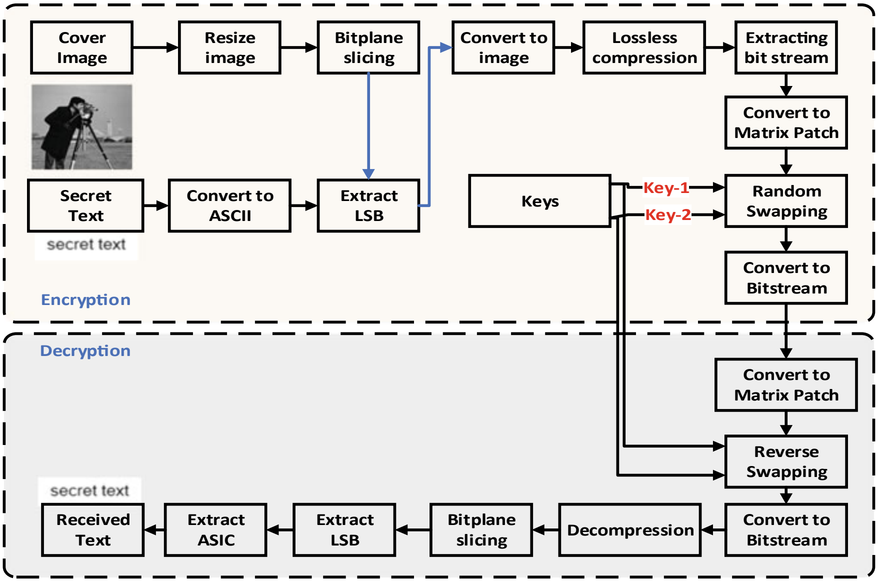
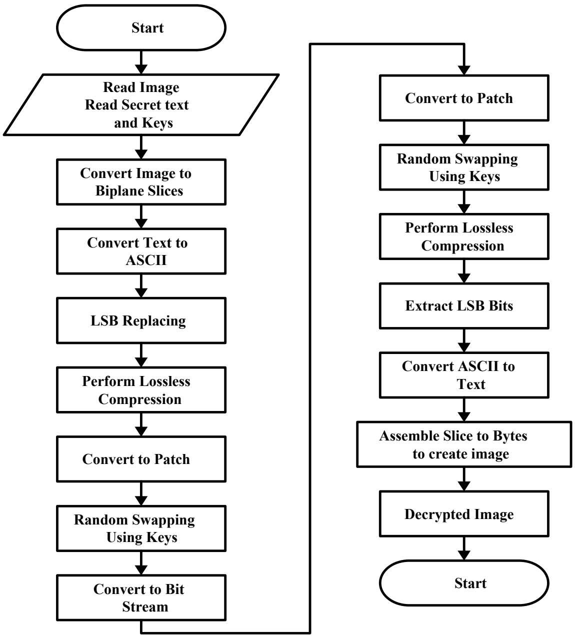


Fig. 1 Secure optical image encryption and data hiding technique in compression domain using dual key-based bit swapping

Fig. 2 Work flow diagram of proposed optical encryption



reducing the original gray level to a binary picture. Converting a grayscale image to a binary image is one of the three major aims of bit plane slicing. The supplied image is a 3-bit image because it has a maximum gray level of 7. Then convert the image to binary and separate the bit planes.

Extract LSB

The LSB is the lowest bit in a series of binary integers, positioned at the far right of a string. The least significant bit is also known as the rightmost bit in positional notation. It is the inverse of the MSB, which has the greatest value in a binary number with multiple bits, as well as the number closest to the right. The importance of a bit in a multi-bit binary number diminishes as it approaches the least significant bit. Because it is binary, the most important bit can only be 1 or 0 [18].

Algorithm 0.1 shows the algorithm of an image encryption. Take the input image (I) with secret text (x) and two keys such as $k1$, $k2$. Firstly resize the input image (I_1) and slice bitplane wise (I_{slices}) and then extract LSB named as (I_{lsb}). Let take ASCII value of the text (X_{asci}) and replace the LSB of the ASCII as I_2.Then do lossless compression of the image (I_{stram}). Then convert the stream into patches (I_{sets}). Random swapping method is used for image encyption with $k1$, $k2$ keys and I_{sets}. Finally get the secret image $I_{\text{secret.}}$

Algorithm: 1

Input:	Input Image i, Secret Text x, Keys k1, k2
Output:	Encrypted Image
1	I_1 = Resize image(I)
2	I_{Slices} = Bitplane Slicing(I_1)
3	I_{lsb} = Extract LSB(I_{Slices})
4	X_{asci} = ASCII(x)
5	I_2 = LSB Replacing(X_{asci},I_{lsb})
6	I_{stram} = Lossless Compression(I_2)
7	I_{sets} = Convert to Patch(I_{stram})
8	I_{secret} = Random swaping(I_{sets},k_1,k_2)

Losslessn Bit Compressıon

Lossless compression is a type of data compression algorithm that allows the compressed data to be completely rebuilt from the original data. Lossy compression, on the other hand, allows just an approximation of the original data to be reconstructed, but with significantly enhanced compression speeds. Lossless compression is employed when the original and decompressed data must be the same, or when changes from the original data would be detrimental. In this work, Huffman coding method can be used.

Each character, symbols and numerical value has distict ASCII, binary codes seperately. Table 1 shows the distinct binaryvalues of some characters. This method generates binary tree nodes. These may be kept in a standard array, the length of which is defined by the amount of symbols. A node can either be a leaf or an inner node. To begin, every nodes are leaf nodes that give the symbol, the weight of the symbol, optionally, a link to a parent node, which simplifies comprehending the code from a leaf node. Internal nodes have a weight, two-child node connections, and a relationship to a parent node that is optional. Following the left child is denoted by bit '0,' whereas following the right child is denoted by bit '1.' A fully grown tree can contain up to n leaf nodes and n-1 inner nodes.

A Huffman tree is used to generate the most optimal code lengths by eliminating superfluous symbols. The procedure starts with leaf nodes, which store the probability of the symbols that signify. The algorithm then selects the two nodes only with smallest probability and constructs a new internal node with these two as children. The weight of the new node is set to the sum of its children's weights. Then, repeat the process on the new internal node and the other nodes (excluding the two leaf nodes) so there is only one node left, which is the root of Huffman tree. Figure 2 shows compressed bitstream.

Decompression is actually turning a stream of prefixed code to single byte values, which is typically accomplished by visiting the Huffman tree node to nodes for every bit from the input stream is received. The Huffman tree should be rebuilt before this can happen. When character frequency is pretty consistent, the tree can be rebuild and therefore utilized each time, just at cost of certain compression performance.

Table 1 Distinct binary values of some characters

Character	Binary code	ASCII code
E	01000101	069
e	01100101	101
M	01001101	077
m	01101101	109
p	01110000	112

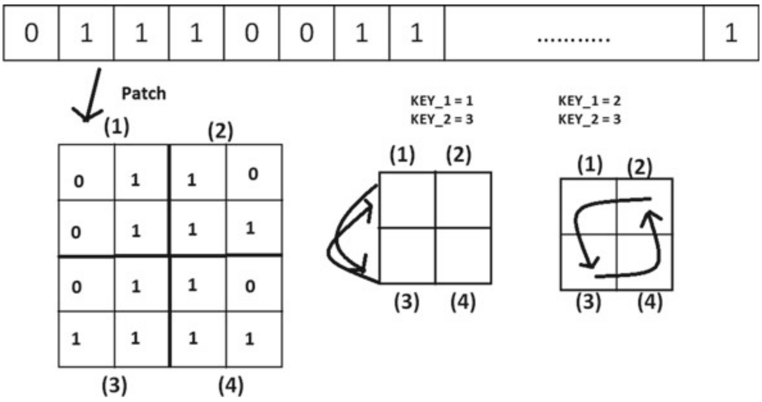


Fig. 3 Compressed bit stream

Bitstream

A bitstream is a bit sequence, sometimes referred to as a binary sequence. A bytestream is a collection of bytes. Because each byte is typically an 8-bit value, the terms octet stream and octet stream are frequently used interchangeably. Because one octet can be encoded as a series of 8 bits in a variety of ways, there is no distinct and direct conversion between bytestreams and bitstreams (Fig. 3).

Results and Dıscussıon

This section discusses the outcomes of numerical simulations performed to determine the efficacy and robustness of the suggested solution. This work is done by MATLAB R2020b using a computer with CPU Intel(R) Core(TM) i5-3320 M CPU @ 1.60 GHz,1 0.60 GHz, and 4 GB of RAM. The dataset images with size 256 × 256 pixel as input sample used as the cameraman and mandrill. Figure 4 sample images from natural image database from Kaggle.

PSNR

The PSNR is the proportion of the signal's maximum potential strength to the power of completely corrupted input, which influences the accuracy of which it is represented [19].

$$\text{PSNR} = 20.\log_{10}\text{MAX}_{PY} - 10.\log_{10}\text{MSE} \tag{1}$$

Fig. 4 Sample images used for image encryption

where MAX_{PY} represents a maximum image pixel value.

CC

The CC is a graphical representation of a type of correlation, which is a statistical relationship between these two variables. The variables may be two columns from a given set of data or two components of a quantitative probability distribution with the good distribution [20].

$$\mathrm{CC}(K, k) = \frac{M\{[K - M(K)][k - M(k)]\}}{M\left\{[K - M(K)]^2\right\}M\left\{[k - M(k)]^2\right\}} \tag{2}$$

Here K and k represent the plain image and decrypted image.

SSIM

SSIM is a perspective paradigm that treats image loss as a perceived shift in structural details while often integrating core visual effects including intensity of light masking and intensity masking concepts [21].

$$\mathrm{SSIM}(i, j) = \frac{\left(2k_i k_j + r1\right)\left(2l_{ij} + r2\right)}{\left(k_i^2 + k_j^2 + r1\right)\left(l_i^2 + l_j^2 + r2\right)} \tag{3}$$

where k_i is the average of i, k_j is the average of j, k_i^2 variance of i, k_j^2 variance of j, l_{ij} the covariance of i and j, $r1$ and $r2$ are two variables to stabilize the division with weak denominator.

MSE

The MSE is a measure of an estimator's consistency; it is often non-negative, with values closest to zero being greater. The distinction between the original and decrypted images is represented by MSE [22].

$$\text{MSE} = \frac{1}{Px*Px}\sum_{i=1}^{Px}\sum_{j=1}^{Px}|\hat{I}(i,j) - I(i,j)|^2 \tag{4}$$

where *Px***Px* denotes the number of image pixels, $\hat{I}(i,j)$, $I(i,j)$. signify original image values, decrypted image values, and at that pixel value (i, j).

RMSE

The RMSE is used to calculate the residuals' standard deviation. Residuals are a metric about how far apart the data points are all from the regression line; RMSE is a measure of how evenly distributed these residuals are. In other words, it denotes how dense the data is along the best fit line.

$$\text{RMSE} = \sqrt{E - K} \tag{5}$$

where E is the expected value and K is known results.

MAE

MAE is a statistical aessment of error among matched data representing the same phenomenon. Comparisons of predicted vs observed future time vs starting time, as well as one measuring technique versus another.

$$\text{MAE} = \frac{\sum_{I=1}^{N}|Y_I - X_I|}{N} \tag{6}$$

where N is the number of predictions, Y_I is the vector of observed values of the variable, X_I is the predicted values.

From Table 2 shows the better comparative performances are CC, PSNR, and MSE in previous works. This work has higher CC, PSNR, and lower MSE with compared to other previous methods. Double random polarization encoding (DRPO) returns only the PSNR values like 17.54, MAE of 0.05, and RMSE of 0.12. Trinary

Table 2 Comparative performance of previous works

Method	CC	PSNR	MSE	SSIM	MAE	RMSE
DRPO [6]	0.76	17.54	0.12	0.74	0.05	0.12
TTS [9]	0.65	–	0.10	–	–	0.26
GI [11]	0.97	–	–	0.93	0.43	0.11
DI [13]	0.63	33.17	–	–	–	0.101
CGI [15]	0.78	–	0.09	–	0.16	–
This work	0.99	35.53	0.0056	0.99	0.015	0.056

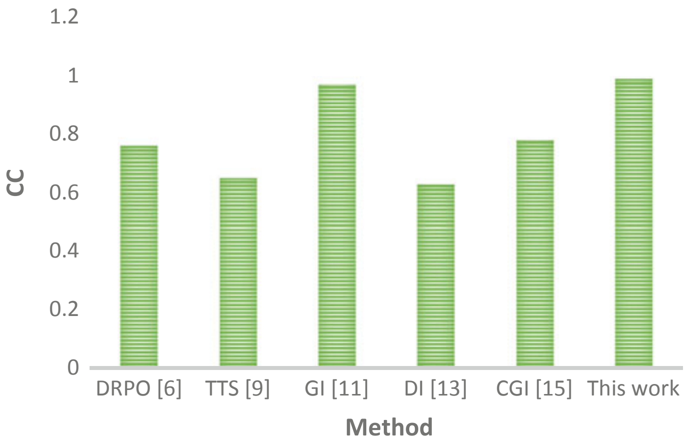


Fig. 5 Comparative performance of CC

tree structure (TTS) has CC of 0.65. The GI method returns the CC of 0.97. The diffractive imaging (DI) method has 33.17 PSNR only. Compressive ghost imaging (CGI) has 0.78 CC only. Figure 5 shows comparative performance of CC and Fig. 6. Comparative performance of PSNR. Fig. 7 shows comparative performance of PSNR with noise density.

Conclusion

This work experimentally demonstrates optical encryption in the compressed bitstream. A dual key-based binary patch swapping technique is used in this work to perform image encryption. Secret data hiding is performed LSB replacing technique followed by ASCII conversion. To validate the feasibility and effectiveness of the suggested technique, simulations, and optical tests were carried out. The suggested optical encryption technique can improve their security while also holding the potential of additional advancements in optical encryption cryptanalysis. From the obtained

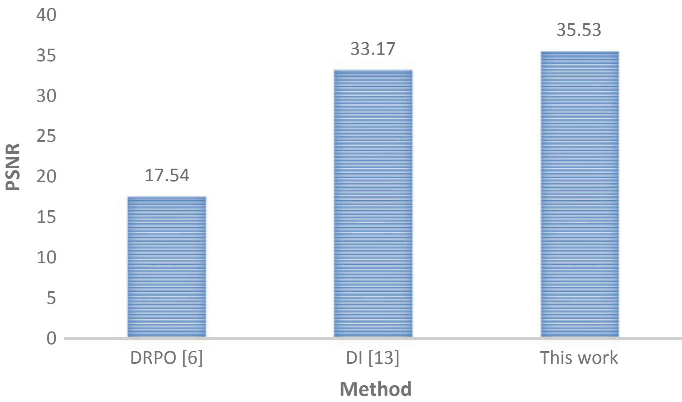


Fig. 6 Comparative performance of PSNR

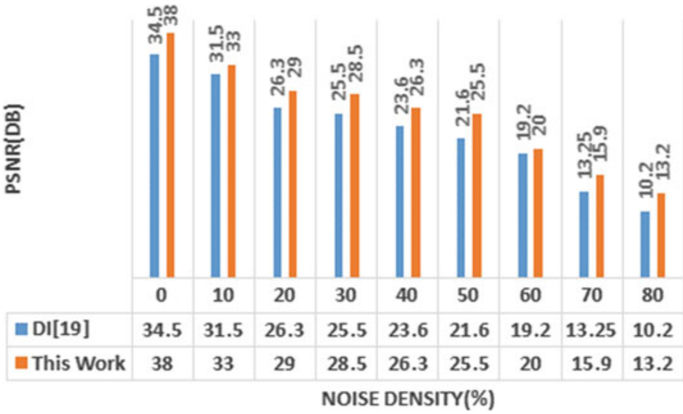


Fig. 7 Comparative performance of PSNR with noise density

result, it is very clear that this work provides good performance when compared to conventional optical encryption standards.

References

1. Gao, Z, Wang K, Yan Y, Yao J, Zhao YS (2021) Smart responsive organic microlasers with multiple emission states for high-security optical encryption. Nat Sci Rev 8(2):nwaa162
2. Zhou L, Xiao Y, Chen W (2019) Machine-learning attacks on interference-based optical encryption: experimental demonstration. Opt Express 27(18):26143–26154
3. Li X, Zhao M, Xing Y, Li L, Kim S-T, Zhou X, Wang Q-H (2017) Optical encryption via monospectral integral imaging. Opt Express 25(25):31516–31527

4. Haoxiang W, Smys S (2021) Big data analysis and perturbation using data mining algorithm. J Soft Comput Paradigm (JSCP) 3(01):19–28
5. Ranganathan G (2021) A study to find facts behind preprocessing on deep learning algorithms. J Innovative Image Proc (JIIP) 3(01):66–74
6. Jiao S, Gao Y, Lei T, Yuan X (2020) Known-plaintext attack to optical encryption systems with space and polarization encoding. Opt Express 28(6):8085–8097
7. Zea AV, Barrera JF, Torroba R (2017) Innovative speckle noise reduction procedure in optical encryption. J Opt 19(5):055704
8. Jiao S, Zou W, Li X (2017) QR code based noise-free optical encryption and decryption of a gray scale image. Opt Commun 387:235–240
9. Su Y, Tang C, Gao G, Fan G, Lei Z, Tang S (2017) Optical encryption scheme for multiple color images using complete trinary tree structure. Opt Lasers Eng 98:46–55
10. Manoharan JS (2012) Embedding capacity enhancement using a hybrid technique for medical images. Eur J Sci Res 75(1):25–38
11. Yi K, Leihong Z, Dawei Z (2018) Optical encryption based on ghost imaging and public key cryptography. Opt Lasers Eng 111:58–64
12. Xi S, Nana Y, Wang X, Ying M, Dong Z, Zhu Q, Wang W, Wang H (2019) Optical encryption method of multiple-image based on θ modulation and computer-generated hologram. Opt Commun 445:19–23
13. Zhou L, Xiao Y, Chen W (2020) Vulnerability to machine learning attacks of optical encryption based on diffractive imaging. Opt Lasers Eng 125:105858
14. Kumar P, Joseph J, Singh K (2016) Double random phase encoding based op-tical encryption systems using some linear canonical transforms: weaknesses and counter-measures. In: Linear canonical transforms. Springer, New York, pp 367–396
15. Yi K, Leihong Z, Hualong Y, Mantong Z, Kanwal S, Dawei Z (2020) Camouflaged optical encryption based on compressive ghost imaging. Opt Lasers Eng 134:106154
16. Liu J, Bai T, Shen X, Dou S, Lin C, Cai J (2017) Parallel en-cryption for multi-channel images based on an optical joint transform correlator. Opt Commun 396:174–184
17. Verma G, Liao M, Dajiang L, He W, Peng X, Sinha A (2019) An optical asymmetric encryption scheme with biometric keys. Opt Lasers Eng 116:32–40
18. El-Shafai W, Almomani IM, Alkhayer A (2021) Optical bit-plane-based 3D-JST cryptography algorithm with cascaded 2D-FrFT encryption for efficient and secure HEVC communication. IEEE Access 9:35004–35026
19. Chen W, Chen X (2014) Iterative phase retrieval for simultaneously generating two phase-only masks with silhouette removal in interference-based optical encryption. Opt Commun 331:133–138
20. Liu J, Bai T, Shen X, Dou S, Lin C, Cai J (2017) Parallel encryption for multi-channel images based on an optical joint transform correlator. Opt Commun 396:174–184
21. Mohammed EA, Saadon HL (2019) Sparse phase information for secure optical double-image encryption and authentication. Opt Laser Technol 118:13–19
22. Sui L, Lu H, Ning X, Wang Y (2014) Asymmetric double-image encryption method by using iterative phase retrieval algorithm in fractional Fourier transform domain. Opt Eng 53(2):026108

Analysis of Machine Learning Algorithms for Predicting the Suitable Career After High School

Albin Thomas, Alen K. Varghese, P. L. Alex, Bobby Joseph Mathews, and L. K. Dhanya

Abstract Choosing the right career is not an easy task in anyone's life. Peer pressure, parental pressure, and personal interests may all cause mental conflict, making the selection process frustrating. As a result, people make poor career choices and end up working in positions that they despise. To address this problem, this paper investigates the various machine learning algorithms and their applications and analyses if they are suitable for the task. The model's dataset includes a wide range of career paths chosen by a diverse group of people, including working professionals, job seekers and PSC, and IAS aspirants. This paper employs a number of well-known machine learning algorithms in the area of artificial intelligence. The aim of this paper is to compare the performance of the algorithms and decide which algorithm is better suited for the task.

Keywords Career · Machine learning algorithm · Artificial intelligence · Prediction · Naïve Bayes · Decision tree · Logistic regression · Artificial neural network · K-nearest neighbor · AdaBoost · Multilayer perceptron · Support vector machines with RBF kernel

A. Thomas (✉) · A. K. Varghese · P. L. Alex (✉) · B. J. Mathews · L. K. Dhanya
Department of Computer Science, Mar Baselios College of Engineering and Technology, Trivandrum, India
e-mail: albinthomas.17cs014@mbcet.ac.in

P. L. Alex
e-mail: alexpl.17cs016@mbcet.ac.in

A. K. Varghese
e-mail: alenkvarghese.17cs015@mbcet.ac.in

B. J. Mathews
e-mail: bobbymathews.17cs036@mbcet.ac.in

L. K. Dhanya
e-mail: dhanya.lk@mbcet.ac.in

V. Bindhu et al. (eds.), *Proceedings of Third International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 844,
https://doi.org/10.1007/978-981-16-8862-1_7

Introduction

A career is an important decision an individual has to make in his lifetime. It is actually a life-defining moment in an individual's life. So, it is of paramount importance that an individual chooses his/her career wisely. It has the potential to take an individual to the golden stairway of success or through the ghettos of dire desperation.

In the current scenario, the testimonials often heard are mostly about the frustrating present, remorse in choosing the current career, but the most frequent one is self-pity about the lack of choices the affected had during his/her academic years. In India, it was found, in multiple surveys, that 93% of the students is aware of just seven career paths. "Our research on over 10,000 students across India suggests an alarming ignorance of the career options available to students today," said Prateek Bhargava, founder and CEO of Mindler, who has recently launched a platform for career counseling in tier II and tier III cities across India [1]. Another worrying factor, in this case, is that the youth do not possess the kind of skills the new industries require.

Why are Indian's fixated on the "Prime 7" career choices, even though there are about 250 career choices that can be pursued in India, which are both new and much-in-demand skills [1]? There can be a number of factors, but the predominant factors include peer pressure, family prestige, and lack of knowledge on emerging career opportunities.

This problem cannot be isolated to a specific country. A survey was conducted for 3000 students, and it was found that the career choice of half of them was influenced or forced by their parents. To be more precise, 69% of students confirmed that their choice of university was influenced by their parents. More interestingly, it is also found that if given the chance, 20% of the students currently enrolled at universities may choose a different course, and 18% of students regret their current choice of degree. The study further shows that for 18% of the 1,805 respondents, the main reason for their disappointment was their lack of initial research in that area of study [2].

Following the establishment of the field of artificial intelligence at a conference at Dartmouth College in 1956, it has risen to prominence as an autonomous empirical discipline in the computer science research spectrum [3]. Artificial intelligence has numerous benefits, and because of its adaptability, AI is used for a variety of applications, including performance comparison for judicial cases [4], employee attrition prediction [5], employee churn prediction [6], and more sophisticated applications such as human behavior prediction in smart home environments [7]. The arrival of machine learning has been a boon to several industries such as medicine, finance, and data analytics. A computer programme is assigned to perform some tasks in machine learning, and it is said that the machine has learned from its experience if its measurable performance in these tasks improves as it gains more and more experience performing these tasks [8]. So, the machine takes decisions and does predictions/forecasting based on data. The advantages of ML include easy interpretation of data and the ability to handle large multi-dimensional data. Machine learning can classify available data into groups, which are then defined by rules set by analysts. When the classification is complete, the analysts can calculate the probability of a fault. ML can also be used for predictive analysis in problems with cases dealing with multiple outputs. This paper explores the working of different machine learning algorithms in predicting the career of students who have graduated high school. A comparison between the working of each algorithm is performed, and based on that comparison, this paper analyses which algorithm is best for the given application.

Literature Survey

The right career will provide much greater satisfaction than one which is forced to choose. It helps an individual to have a quality work life. Career recommendations and current job predictors are utilized on a number of platforms such as Cocubes and AMCAT. These tools only analyze the technical knowledge and do not delve deeply into an applicant's emotional intelligence factor [9]. A study was conducted by Sripath Roy et al. [9] on computer science students which focused on finding the right career path those students should opt based on their skills and interests. Study points out that the right career will help them in improving their performance and also direct them toward a targeted career goal. The proposed paper also mentions that the impact this would have on the time of their recruitment where their roles can be easily decided based on their performance. The study also introduces a recommendation system which helps in assessing prerequisite knowledge and skills of candidates. For a better career path prediction, the proposed paper considered various factors including abilities of students in sports, academics, and their hobbies, interests, competitions, skills, and knowledge. The study considered 36 parameters of computer science candidates for predicting the optimum path they should choose. The final job roles which include Database administrator, Business process analyst, Developer, Testing manager, Networks manager, and Data scientist came to be a fixed 15 in number. They collected data such as student's academic scores in various subjects, specializations, programming, and analytical capabilities, memory, personal details like relationship, interests, sports, competitions, hackathons, workshops, certifications, books interested, and many more since these factors play a vital role in the perfect career area. The typical sources were data of employees of different organizations, LinkedIn API, college alumni database, and some amount of data randomly generated for enhancing the dataset. The proposed paper formulates a dataset comprising nearly 20,000 records with 36 features. Clearly, the career chosen by working professionals and their path through which their academics progressed will have the maximum relationship among all the data. The skills and performance of individuals show the degree to which they can grow or their potential which may directly reflect their present profession and designation. The aim of the above research was to predict which position a recruited candidate would be ideally suited for in a computer science setting. For implementation, this work employs three algorithms: SVM, decision Tree, and XGBoost.

The study proposed by Rangnekar [10] introduces a career prediction model to find the optimal career for a student based on their personality, aptitude, and other general information. This includes mining of data of students through their Facebook accounts, in order to assess their personality based on the content posted and have introduced a Facebook Graph API for this purpose. Study used Myers-Briggs type to analyze the psychological aspects and hence determine the student's personality and used a dataset of 8775 entries. Personality data had four label I and E, N, and S, T and F, J and P with 90–10 split of data for training and testing. Secondly, an one-hour aptitude test in the topics of Mathematics, English, General Science and Logic each topic having a maximum of 20 questions with predefined mark ranges determines the aptitude scores. The preprocessing step involved label encoding to map each career to numerical test values with 80–20 split of data for training and testing. And finally, the background information is used to analyze if the student has enough resources to pursue a predicted career. One-hot encoding was used to split labels, and it was used to cut down entries from 480 to 261 for a single student making a dataset having a 80–20 split for training and testing. K-nearest neighbors and stochastic gradient descent were used for all three datasets, and additionally, logistic regression was used for both personality and background data. The proposed paper also used random forest for background data to assign weights for features. From the training and testing of all aptitude data, this technique got accuracy of 73.74% and 81.035% for KNN and stochastic gradient 120 descent, respectively. The personality data was obtained 81%, 73.75%, and 121 71.53% for KNN, stochastic gradient descent and logistic regression, respectively. Finally, the accuracy of 57.10 and 60.09 before and after adaptive boosting using logistic regression was obtained for background data. The study concluded that aptitude, personality, and background are excellent factors to be considered for career prediction.

However, the projects have a range of flaws. To begin with, the dataset considered in the reference paper [9] is only around 18,000 people, which is very small and incapable of drawing a realistic picture. Another finding in this study was that about 36 vectors were chosen; but upon closer inspection, some of them were inserted unnecessarily to boost the number of features. Take, for example, relationships, interests, types of books read, and so on, which can be summed up into a single parameter by evaluating the candidate's emotional quotient (EQ) score. While it is true that an EQ score cannot accurately capture the depth of a person's interests or mind, it is also worth noting that an EQ score provides ample rich information about parameters such as relationship, interests, and so on and can provide a general overview of a candidate's mindset. Reference paper [10] uses a Facebook API in order to assess the student's personality and a candidate always need not be active on Facebook. The problem with using Mayer's–Briggs test to assess personality is that this test is often labeled as inaccurate and is often said to give inconsistent results. So, the major drawback is that the personality of the candidate labeled need not be true. Also, the proposed paper uses three major algorithms and does not delve deep into other algorithms.

Proposed System

This research is intended to assist a high school graduate in choosing a career path that will enable him or her to be successful. This is a more in-depth examination in which the applicant is not limited to a specific environment or group. This review compares up to eight algorithms to find the one that is ideally suited to the problem statement. This study aims to forecast the career paths of students with a variety of specializations, such as technology, commerce, or the humanities. The proposed analysis's goal is to forecast best algorithm capable of predicting career trajectories with great accuracy and at a lower cost of execution. This analysis differs from prior studies [9–11] in a number of ways:

1. Conventional approaches analyze the relationship using online APIs such as Facebook [10] to determine the EQ scores. The O.C.E.A.N test, which takes a more clinical and psychological approach, is used in this investigation to assess emotional wellness.
2. Various studies in the field of psychology [12] found that IQ was one of the many important elements associated with career success. As a result, the proposed analysis employs a questionnaire to test IQ score of the individual and collect the result.
3. The memory test employs a puzzle and is based on a real-time memory scoring system. The memory score is based on the fact that the faster you solve the given puzzle, the better your memory.
4. Previous studies on the same issue statement [9, 10] used mostly three algorithms. The suggested study differs from its predecessors in that it evaluates the performance of six different algorithms.

The dataset for this study contains approximately 8 lakh rows, with 23 feature vectors selected. The intellect (IQ) and emotional intelligence (EQ) factors are included in this study. Socioeconomic considerations are also taken into account in the data collection. According to a study conducted by the Carnegie Institute, a candidate's personality can be gleaned in large part by following these steps, which is a critical factor in a candidate's performance (Fig. 1).

A. **Data Collection**

Data are an essential component of every machine learning algorithm that is utilized for training and testing [7]. Dataset was obtained via data collection and generation. The information was collected through a Google form. The form was a questionnaire, about subjects like high school grades, and a set of questions to assess intelligence and emotional quotients. Some of the data obtained were collected from working-class individuals, who work in a broad diaspora of organizations; some data were obtained submitting the form in a free online response collection platform called SurveyTandem, and a certain amount of data has been generated randomly and using synthetic data generation methods. Dataset was chosen with the diversity of data and data type in mind to be evaluated during the implication of the approaches used throughout the test [13].

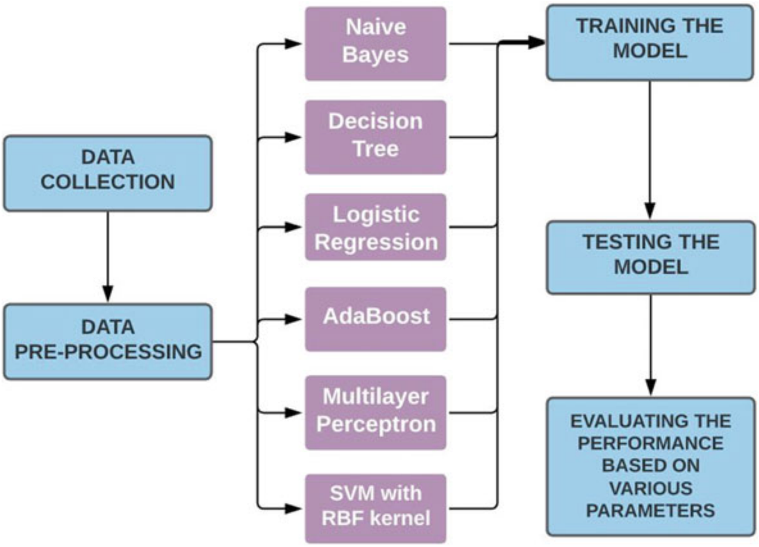


Fig. 1 Process flow diagram of proposed system [9]

The rows contain the data of each working-class professionals. The columns consist of different attributes which are considered critical in predicting the career such as intelligence quotient [IQ], the academic scores of 6 subjects in their high school grade, emotional quotient [EQ] test using O.C.E.A.N test, memory test, satisfaction in the present career, interests, certification in any area, leadership quality, ability of self-learning, team working skills, socioeconomic factors, work-life balance, and eligibility in any certification/competitive exam. The data were collected from working professionals who previously studied in science, commerce, humanities batch after their matriculation. The initial dataset consists of 100,000 rows and 23 features. The entire responses in this dataset were restricted to the science dataset only. The new, enlarged dataset contains a total of 8 lakh data, with 8 lakh rows and 23 columns. The majority of the data obtained was from the science stream, from which four fifty thousand data were obtained. The responses from commerce and humanities stream were about 2 lakh and 1 lakh 50 thousand, respectively. The dataset also consists of the prominent student attributes which include information about the students' capabilities and tastes.

B. **Data Preprocessing**

Data cleaning and data transformation are the two major sub-phases of the data preprocessing system. This phase's output is fed into the machine learning algorithms.

	IQ	PHYSICS	CHEMISTRY	MATHS	BIOLOGY	COMPUTER	LANGUAGE	OPENNESS	CONSCIENTIOUSNESS	EXTRAVERSION
0	138	98	92	88	95	0	86	80	80	20
1	103	48	62	68	0	73	60	20	40	60
2	84	39	36	31	0	31	46	60	60	60
3	136	79	77	81	99	0	86	100	100	80
4	95	62	61	46	64	0	69	60	60	60
5	138	77	96	85	65	0	99	80	100	80

(a)

SATISFACTION	SOCIO-ECONOMIC-FACTORS	M-SCORE	TEAM-WORK	COMPETIVE-EXAM-WINNER	LEADERSHIP-SKILL	WORK-LIFE-BALANCE	CERTIFICATIONS	SELF-LEARNING-CAPABILITY	CAREER
90	0	2	1	0	1	1	0	1	1
51	2	0	0	0	0	0	0	0	2
35	2	0	0	1	1	1	0	0	0
84	1	2	1	0	1	1	0	1	1
64	2	0	0	0	0	0	0	1	2
99	1	1	1	0	1	0	0	1	0

(b)

Fig. 2 Figures **a** and **b** show a portion of the dataset (after data cleaning and transformation) that was utilized for training and testing. Figure **a** depicts a portion of the data rows and columns, while Figure **b** depicts the remaining columns

1. **Data Cleaning**

   Data cleaning is done by removing many redundant data that occurs in the dataset. Unwanted columns such as Name and Age are also removed from the dataset since they don't directly help in prediction.

2. **Data Transformation**

   Data are transformed suitably to optimize performance and execution time. Data are scaled, and categorical data are converted to numerical data (Fig. 2).

C. **Machine learning Algorithms**

1. **Naive Bayes**

   Naive Bayes is a popular classification technique which classifies examples based on the probability of chances that are likely to be occurred [6]. Researchers working in the domain of event prediction in a variety of fields have intensively examined probabilistic approaches such as Naive Bayes. Statistical models outperform more exact mathematical models in terms of prediction accuracy, especially when the variable being predicted varies with a certain degree of randomness with regard to numerous variables [5]. As generalization is adopted in SVM, so there is less probability of over fitting. It can scale up with high dimensional data [8]. Naive Bayes exhibited admirable performance, as it was able to provide an accuracy of 46% when the initial science dataset of 1 lakh rows was used. This improved drastically to 83.6, 85.87, and 83.89% for science, commerce, and humanities datasets, respectively, and is the fastest of all the algorithms taken into consideration. Naive Bayes had a precision of

0.86%. Since the dataset had a high dimensionality, the data could be easily separated as clusters, for easier evaluation.

2. **Decision Tree**

   This is one of the popular learning techniques in machine learning which is often compared with the new algorithms that are being developed [6]. Construction of decision tree and classification in a straightforward manner were a comparatively slow process. The algorithm performed at a varying pace in the three different datasets. However, it gave an accuracy of 87.9, 95.02, and 86.5%, respectively for the three batches of datasets. This algorithm achieved an average precision of 0.88 which indicates a better test condition.

3. **Logistic Regression**

   After proper preprocessing of data, the analysis showed that the multinomial logistic regression algorithm has given an accuracy score of 84.77, 93.89, and 83.61%, respectively, for the three datasets. Logistic regression was slightly slower than Naive Bayes, but faster than all other algorithms and gave a precision of almost 0.84.

4. **Artificial Neural Network**

   Artificial neural network (ANN) is a deep learning model that was taken into consideration. When the algorithm was performed on a dataset of size 1 lakh, it performed with an accuracy of around 70%, but the downside of using this algorithm was it was comparatively time-consuming than the machine learning algorithms. Due to this factor, ANN was not considered for further analysis.

5. **KNN**

   KNN classifier is known as lazy learner in machine learning community. It never learns from the data and do not build any models [6]. K-nearest neighbor algorithm performed in an excellent manner when the dataset was of a relatively small number. When the process was conducted on a dataset of size 1 lakh, it was observed that KNN had an accuracy of about 52%. But, as the size of the dataset increased, the algorithm took excessive amounts of time, and hence, the performance of this algorithm for a larger dataset was not taken into consideration for further analysis.

6. **AdaBoost**

   Adaptive boosting, known short as Adaboost, is a machine learning meta-algorithm. This algorithm was able to generate 83.08, 90.53, 82.91% on science, commerce, and humanities datasets, respectively, and it was observed to have a code execution time of fewer than 2 min. Considering the satisfactory performance of the algorithm, the algorithm can be considered for real-world modeling.

7. **Multilayer Perceptron**

   A class of feed-forward ANN, multilayer perceptron was also considered for evaluation. This algorithm performed better than its ANN counterpart, in terms of both accuracy and performance. The accuracy shown by the model varied according to the different batches of the dataset. MLP showed an accuracy of 92.18, 97.52, and 91.24% in science, commerce, and humanities datasets, respectively. Despite its exceptionally high accuracy, the discouraging factor of the model (to be considered for real-world application) was that the time, taken by the model to execute the code was greater than 30 s, and took almost 209 s in case of, that is, greater 2 min.

8. **SVM with RBF Kernel**

A support vector machine is a kind of classification technique, where the data points are separated by a line in case of linear SVM and a hyperplane in case of non-linear SVM [6]. Support vector machines (SVMs) are a simple algorithm that belongs to the field of supervised machine learning techniques and are used for classification and regression. Because this study involves numerous classes, it employs a kernelized SVM model. The default kernel, radial basis function (RBF), is employed in this case. This kernel function calculates the similarity between two data points or the similarity between points in a two dimensional feature space. The similarity between two locations in the modified vector space in the RBF kernel is an exponentially declining function of the vectors' distance from the original input space. The kernel function is expressed as follows:

$$K(x, x') = \exp\left(-\gamma \|x - x'\|^2\right) \tag{1}$$

where x, x' are feature vectors in some input space. $||x - x'||^2$ is the squared Euclidean distance between the two feature vectors, and γ is a constant parameter.

For the different professional pathways, this algorithm was able to produce an accuracy of roughly 87, 96, and 87%, respectively. This method performed just as well as the MLP and decision tree algorithms in terms of precision, recall, and F1-scores. The primary advantage of this kernelized model of SVM over decision tree is that it is less prone to over fitting, making it a viable technique for this research.

9. **Model Validation**

The training and testing phase were done for each of the algorithms by splitting the dataset into 80:20 ratio. 80% of data is used for training, and 20% of data is used for testing. It is used to estimate the accuracy of a model on unseen data.

10. **Setting**

We used Jupyter Notebook for working with the datasets and analysis of all algorithms [14]. In KNN, the number of neighbors is set to three.

11. **Output**

The results of the analysis are mainly three labels for each category: three science labels, three commerce labels, and three humanities labels. The output depends on the sector of specialization of secondary school pupils. The class labels for the sciences, commerce, and humanities are 0, 1, and 2. Labels are categorized differently under different categories: in science, 0 signifies engineering, 1 for medicine, 2 for other career areas than engineering or medicine. Likewise, in the commerce stream, 0 stands for Bachelor of Business Administration (BBA), 1 for CA, 2 for career areas other than BBA or CA. Again, within the humanities category, 0 stands for sociology, 1 for business/law, and 2 for careers other than those mentioned.

Result and Analysis

For training and testing of all models, we used Jupyter Notebook. The proposed study tested on an eight-lakh dataset of text type with twenty-three features. Four lakh fifty thousand data were from the scientific stream, two lakhs from commerce, and one lakh from humanities. From the science dataset alone, three lakh sixty thousand data points were used for training purpose, and remaining ninety thousand was used for testing purpose. Similarly, one lakh sixty thousand data points were used for training in the commerce, while the remaining forty thousand was used for testing. Similarly, in humanities, eighty thousand data points were utilized to train the algorithms, with the remaining twenty thousand used for testing. The best algorithm suited for the problem statement is found by using these datasets on each of the six algorithms.

This study chooses the most advantageous algorithm based on primarily two parameters:

(1) Accuracy
(2) The amount of time spent

The confusion matrix can be used to explain the algorithms' performance.

A confusion matrix is a table that shows how well the prediction model has performed. The number of predictions made by the model that are classified correctly or erroneously is represented by each entry in the confusion matrix. In a confusion matrix, there are several metrics to consider:

True Positive (**TP**): The predicted value has a positive polarity, and the actual value also has a positive polarity.

False Positive (**FP**): The predicted value has a positive polarity; however, this is not the case.

False Negative (**FN**): The predicted value has a negative polarity, whereas the actual value is the polar opposite.

True Negative (**TN**): The predicted value has a negative polarity, and the actual value has a negative polarity as well.

The following qualities of a prediction model can be defined using these metrics: Accuracy: The fraction of total samples that the classifier properly classified as positive. It is the ratio of the sum of true positives and true negatives to sum of true positives, true negatives, false positives, and false negatives.

Precision: The fraction of all positive classes were accurately predicted from the overall positive class. It is the ratio of true positives to sum of true positives and false positives

Recall: It's the proportion of accurately predicted positive samples to all positive samples. It is the ratio of true positives to sum of true positives and false negatives.

F1-Score: It combines precision and recall into a single picture. It is the harmonic mean of precision and recall. It is calculated as twice the ratio of product of precision and recall to sum of precision and recall.

If:

(1) The accuracy is high, and
(2) It takes less time to forecast the result(s),

An algorithm is considered favorable/suitable for the task.

Performance comparison of the various models was carried out. The accuracies of the different models are shown in the tables below which show the results from the initial dataset of 100,000 rows (Table 1).

All algorithms considered were able to produce results, and SVM with RBF kernel was found out to be the best predictor with an accuracy of 90%. The dataset consisted of values related to the science stream only.

These are the results from the new, enlarged dataset. This dataset was further optimized for performance. Hence, the time taken for generating results was also taken into consideration (Tables 2, 3, and 4).

Tables below illustrates summary of the confusion matrices of each algorithm for each stream which is used for analysis (Table 5).

There were 4 Lakh 50 thousand datasets in the Science data stream. The table above displays the confusion matrix for all of the algorithms in a tabularized format. The rows represent the algorithm, while the sub-rows numbered 0, 1, 2 represent the predicted values, and the sub-columns represent the actual values. 0 denotes engineering; 1 denotes medicine, and 2 denotes specialties other than engineering and medical (Table 6).

There were 2 lakh datasets in the commerce data stream. The table above displays the confusion matrix for all of the algorithms in a tabularized format. The rows

Table 1 Comparison of accuracy and other parameters of science dataset

No	Algorithm	Accuracy (%)
1	SVM with RBF kernel	90
2	Decision tree	69
3	ANN	63
4	Multilayer perceptron	59.5
5	KNN	52
6	Adaboost	48.6
7	Logistic regression	48
8	Naïve Bayes	46

Table 2 Comparison of accuracy of science dataset

No	Algorithm	Accuracy (%)	Time taken (seconds)	Precision	Recall	F1-Score	Support
1	Multilayer perceptron	92	209	0.92	0.92	0.92	90,000
2	Decision tree	87.9	33	0.88	0.88	0.88	90,000
3	SVM with RBF kernel	87	1.21	0.87	0.87	0.87	10,000
4	Logistic regression	84.77	2.1	0.85	0.85	0.85	90,000
5	Naïve Bayes	83.6	0.4	0.84	0.84	0.84	90,000
6	AdaBoost	83	143	0.83	0.83	0.83	90,000

Table 3 Comparison of accuracy of commerce dataset

No	Algorithm	Accuracy (%)	Time taken (seconds)	Precision	Recall	F1-score	Support
1	Multilayer perceptron	97	53	0.92	0.92	0.92	90,000
2	Decision tree	96	316	0.95	0.95	0.95	40,000
3	SVM with RBF kernel	95.02	14	0.96	0.96	0.96	10,000
4	Logistic Regression	93.89	1.3	0.94	0.94	0.94	40,000
5	Naïve Bayes	90	50	0.86	0.86	0.86	40,000
6	Adaboost	85.87	0.19	0.83	0.83	0.83	90,000

Table 4 Comparison of accuracy of humanities dataset

No	Algorithm	Accuracy (%)	Time taken (seconds)	Precision	Recall	F1-score	Support
1	Multilayer perceptron	91	61	0.91	0.91	0.91	30,000
2	Decision tree	87	361.8	0.86	0.86	0.86	30,000
3	SVM with RBF kernel	86.5	9.5	0.88	0.87	0.87	10,000
4	Logistic regression	83.89	0.15	0.84	0.84	0.84	30,000
5	Naïve Bayes	83.61	0.7	0.85	0.84	0.84	30,000
6	Adaboost	82	40	0.83	0.83	0.83	30,000

Table 5 Confusion matrix for science data stream

Confusion matrix		Science		
		0	1	2
Multilayer perceptron	0	27,724	846	1561
	1	936	27,548	1447
	2	1073	1120	27,745
Decision tree	0	26,739	1389	2003
	1	1341	26,622	1968
	2	1974	2104	25,860
SVM with RBF kernel	0	2713	260	326
	1	172	3036	208
	2	90	217	2978
Logistic regression	0	24,605	2468	3058
	1	2746	25,677	1508
	2	954	2989	25,995
Naïve Bayes	0	22,937	3237	3957
	1	2262	25,968	1701
	2	955	2526	26,457
Adaboost	0	24,277	2815	3039
	1	1768	25,883	2280
	2	2498	2823	24,617

represent the algorithm, while the sub-rows numbered 0, 1, and 2 represent the predicted values, and the sub-columns represent the actual values. 0 denotes BBA; 1 denotes CA, and 2 denotes specialties other than engineering and CA (Table 7).

There were 1 lakh datasets in the humanities data stream. The table above displays the confusion matrix for all of the algorithms in a tabularized format. The rows represent the algorithm, while the sub-rows numbered 0, 1, and 2 represent the predicted values, and the sub-columns represent the actual values. 0 denotes sociology; 1 denotes LLB, and 2 denotes specialties other than sociology and LLB.

Discussion

This section describes the noteworthy observations made, while the machine learning algorithms were running.

A. **Effect of number of features**

The analysis started with 11 feature vectors and a data set of 25,000 rows. The dataset included entries in the science class alone. When the analysis was performed on this dataset, classification algorithms like KNN and SVM behaved

Table 6 Confusion matrix for commerce data stream

Confusion matrix		Commerce		
		0	1	2
Multilayer perceptron	0	27,646	649	1836
	1	984	27,372	1575
	2	855	1002	28,081
Decision tree	0	12,607	167	524
	1	162	12,951	325
	2	511	327	12,426
SVM with RBF kernel	0	3316	50	14
	1	21	3242	52
	2	187	65	3053
Logistic regression	0	12,620	380	298
	1	141	12,770	527
	2	777	327	12,160
Naïve Bayes	0	11,226	427	1645
	1	93	13,000	345
	2	2320	621	10,323
Adaboost	0	24,277	2815	3039
	1	1768	25,883	2280
	2	2498	2823	24,617

well, with a high accuracy score. But, as the size of the feature vectors increased, the performance of both algorithms started to decline. The reason behind the performance drop of SVM is because adding additional feature vectors tend to increase the algorithms generalization error. In KNN, as the size of the features increases, it lead to the problem of over fitting. But, in the case of SVM, this problem was rectified when RBF kernel was used.

B. **Effect of number of dataset points**

On small datasets, classification algorithms like KNN and SVM perform admirably. However, as the number of data points grew larger, efficiency of most of the algorithms' deteriorated. In the case of KNN, as the number of points in the dataset grows, calculating the distance between all of the vectors in the dataset becomes a computationally expensive operation. That is why KNN is no longer favored in this particular problem statement.

Naive Bayes, decision tree, and SVM were the algorithms that performed admirably and consistently as the dataset grew larger. Since it makes assumptions about the freedom of the features in the dataset, Naive Bayes is capable of handling large datasets. This is the reason why Naive Bayes is capable of handling massive datasets. The decision tree is capable of managing large dataset, since a series of decision trees are constructed in parallel on tractable size training datasets which

Table 7 Confusion matrix for humanities data stream

Confusion matrix		Humanities		
		0	1	2
Multilayer perceptron	0	8906	246	808
	1	168	9310	511
	2	320	645	9086
Decision tree	0	8724	390	846
	1	376	8813	800
	2	939	803	8309
SVM with RBF kernel	0	2920	67	319
	1	97	2846	368
	2	173	239	2971
Logistic regression	0	8660	396	904
	1	562	8091	1336
	2	607	1140	8304
Naïve Bayes	0	7993	456	1511
	1	495	8114	1380
	2	140	847	9064
Adaboost	0	8771	224	965
	1	240	8678	1071
	2	1272	1355	7424

are a subset of the original data. Each learned decision tree will be reduced to a set of rules; conflicting rules are resolved, and the resulting rules are combined into a single set. Cross-validation studies on a dataset show that this technique is viable. [15]. The dataset had to be partitioned into different batches of 50,000 for SVM to work properly.

On deeper examination, it was observed that decision trees attained 100% training accuracy, indicating that this model was over-fitted.

Conclusion

This present study explores the working of different machine learning algorithms in career prediction. Considering the factors which are most critical for a successful career, the model tries to predict the most suitable career, which an individual can confidently choose, with surety that he/she can succeed in the career predicted.

The evaluation of the performance of each algorithm on the dataset can be done in the terms of both accuracy and time taken for execution. This paper evaluates that, on the basis of accuracy, SVM with RBF kernel, decision tree, and multilayer perceptron (MLP) performed the best. When execution time is taken as the deciding factor, it was observed that Naive Bayes performed the best, while algorithms such as multilayer perceptron (MLP) and Adaboost took almost one minute to generate results. It is important to note that SVM with RBF kernel took a great amount of time to produce results.

This research can be further improved by adding more real-time data into the data pool, hence making the system more accurate. Also, this model can be developed in to a layman friendly application by integrating the system on to the Web, using tools such as Python Flask.

References

1. Chakrabarthy R (2020) 93% Indian students aware of just seven career options: what are parents doing wrong? India Today, 5 Feb 2019. Accessed on: 24 Oct 2020. [Online]. Available: https://www.indiatoday.in/education-today/news/story/93-indianstudents-aware-of-just-seven-career-options-what-are-parentsdoing-wrong-1446205-2019-02-04
2. Career Hunter (2017) The shocking cost of making a wrong career choice, PR Newspaper, 14 Feb 2017. Accessed on: 24 Oct 2020 [Online]. Available: https://www.prnewswire.com/newsreleases/the-shocking-cost-of-making-a-wrong-caeer-choice613686493.html
3. Tian J (2020) The human resources development applications of machine learning in the view of artificial intelligence. In: 2020 IEEE 3rd international conference on computer and communication engineering technology (CCET). Beijing, China
4. Liu Z, Chen H (2017) A predictive performance comparison of machine learning models for judicial cases. In: 2017 IEEE symposium series on computational intelligence (SSCI). Honolulu, USA
5. Ray AN, Sanyal J (2019) Machine learning based attrition prediction. In: 2019 global conference for advancement in technology (GCAT). Bangalore, India
6. Sisodia DS, Vishwakarma S, Pujahari A (2017) Evaluation of machine learning models for employee churn prediction. In: 2017 international conference on inventive computing and informatics (ICICI). Coimbatore, India
7. Zhang W, Li W (2019) A deep reinforcement learning based human behavior prediction approach in smart home environments. In: 2019 international conference on robots & intelligent system (ICRIS). Haikou, China
8. Ray S (2019) A quick review of machine learning algorithms. In: 2019 international conference on machine learning, big data, cloud and parallel computing (COMITCon). Faridabad, India
9. Sripath Roy K, Roopkanth K, Uday Teja V, Bhavana V, Priyanka J (2018) Student career prediction using advanced machine learning techniques. Int J Eng Technol 7(2.20):26–29. http://dx.doi.org/10.14419/ijet.v7i2.20.11738
10. Rangnekar RH, Suratwala KP, Krishna S, Dhage S (2018) Career prediction model using data mining and linear classification. In: 2018 fourth international conference on computing communication control and automation (ICCUBEA), pp 1–6. https://doi.org/10.1109/ICCUBEA.2018.8697689
11. Mathew V, Chacko AM, Udhayakumar A (2018) Prediction of suitable human resource for replacement in skilled job positions using Supervised Machine Learning. In: 2018 8th international symposium on embedded computing and system design (ISED). Cochin, India
12. Sternberg R, Grigorenko E, Bundy D, Merrill-Palmer (2001) The predictive value of IQ. Quarterly. 47:1-41. https://doi.org/10.1353/mpq.2001.0005
13. Choudhary R, Gianey HK (2017) Comprehensive review on supervised machine learning algorithms. In: 2017 international conference on machine learning and data science (MLDS). Noida, India

14. Fahad PK, Pallavi MS (2018) Prediction of human health using machine learning and big data. In: 2018 international conference on communication and signal processing (ICCSP). Chennai, India
15. Hall L, Chawla N, Bowyer K (1998) Decision tree learning on very large data sets. In: Proceedings of the IEEE international conference on systems, man and cybernetics, vol 3. pp 2579–2584. https://doi.org/10.1109/ICSMC.1998.725047

An Enhanced Approach to Predict Re-occurrences of Breast Cancer Using Machine Learning

M. Sheshikala, Dadi Ramesh, Sallauddin Mohmmad, and Syed Nawaz Pasha

Abstract The most affected cancer among women is breast cancer (BC). After skin cancer, the most affected cancer in the world is BC. The term re-occurrence is used for BC, which affects after successful treatment. The prediction of re-occurrence BC can reduce the chances of death. In this paper, we proposed a machine learning approach and neural network model to calculate recurrence chances of BC. We used the Wisconsin dataset of the UCI machine learning repository in our approach, implemented an artificial neural network (ANN), and compared the results with the support vector machine. Our model outperformed when compared to other models with an accuracy of 95% with support vector machine.

Keywords Breast cancer · Re-occurrence · Machine learning · Prediction · Artificial neural network

Introduction

The term "cancer" refers to one of many diseases characterized by the development of abnormal cells that divide uncontrollably due to genetic mutation in DNA and can then destroy normal body tissue. The cancer cells can invade the nearby tissues and spread to other parts of the body and are the second leading cause of death in the world. Breast cancer is one of the most common types of cancers among women worldwide, and it is increasing 2 percent annually. Although the average age of breast cancer is 40–50 years [1], it has also occurred in 25-year-olds, and the age is declining yearly. Breast cancer is the most common cancer in women, affecting about 10% of all women at some stages of their life. This disease occurs when people

M. Sheshikala (✉)
Head of School of Computer Science and Artificial Intelligence, SR University, Warangal, Telangana, India

D. Ramesh · S. Mohmmad · S. N. Pasha
School of Computer Science and Artificial Intelligence, SR University, Warangal, Telangana, India

V. Bindhu et al. (eds.), *Proceedings of Third International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 844,
https://doi.org/10.1007/978-981-16-8862-1_8

are exposed to carcinogenic materials by inhaling, eating, drinking, and exposing to the workplace environment.

In the current generation, cancer has become a common disease. Reaching over to different types of cancers, breast cancer is most commonly seen in most women. The main part of our approach to the problem is the re-occurrence [2] of cancer. As cancer is a disease that can re-occur, the chances to affect the person who has already been affected in the past are higher. As it is priorly known that, prediction of cancers [3–5] has already been existing, we have come up with the idea to predict the rate at which cancers' re-occurrence is usually seen. There exist different types of cancers where the rate of re-occurrence of breast cancer in women is high because of the various biological factors.

According to tumor [6] features, whether it is malignant or benign, BC can be diagnosed. We can classify re-occur or none, and re-occurrence can be predicted [7] by checking some features like the numbers of lymph, area of the tumor and original tumor, etc. Many researchers worked on the early prediction of BC with symptoms and worked on its classification using machine learning. Artificial intelligence has grown and solved problems by learning critical patterns that were not possible by human beings.

To find out the chances of re-occurrence of breast cancer, our project deals with its prediction. Early prediction of breast cancer [8] is one of the most crucial works. As such, there remains a need to remove cancer to reduce re-occurrence. Since re-occurrence within five years of diagnosis is correlated with the chance of death, being able to understand and predict re-occurrence susceptibility is critical.

Literature Review

Understanding the impact of recurrent breast cancer is vital to alert the patients before themselves, so that they take necessary precautions and reduce the risk of livelihood. However, many of the studies only concentrated on predicting cancer and its treatment; researchers have paid very little attention to the concept of recurrence of cancer.

Some researchers used machine learning techniques such as regression, support vector machines, and artificial neural networks to predict 5-, 10-, and 15-year breast cancer survival. They studied many breast cancer patients and used tumor size, age, mitotic count, cell type, etc., as their input variables. They proved that machine learning could be a tool for predicting the recurrence of diseases.

Chakra deo et al. [9] have derived model for breast cancer analyzed with Wisconsin Prognostic Breast Cancer Dataset (WPBC). Multiple linear regression and SVM are implemented for the classifier with association of K-fold cross-validation and decision tree for better result gaining. Authors proved that the more accuracy of the result accounted by using SVM–radial basis function. As per the author's research, the result is more accurate compared with decision tree results when implemented the SVM–RBF. Mandeep Rana et al. [10] implemented the model Wisconsin Prognostic and Diagnostic Datasets. In this paper, they have used the four classification techniques such as Naive Bayes algorithm, support vector machine, nearest neighbor, and logistic regression. The main intention of the authors for this research was to find the difference between benign and malignant breast cancer. By implementing of this model, authors also predicted the recurrence of cancer. They finalize that KNN will perform well in training and tested dataset when data is small amount. However, when the parameters and dataset are large amount, the SVM with Gaussian Kernel has given best result. When increased the number of iterations then required to train the model more they present. Sayed et al. [11] have implemented a model with Wisconsin Breast Cancer Database (WBCD) dataset. The pretrained samples of data and other efficient collection of dataset in a supervised learning are extracted using deep Boltzmann machine (DBM). After this process, the results are classified with DNN and finally found the two groups, namely benign group or malignant group. Here, the tumors were predicted and implemented on R-Language. From their research, author achieved their proposed module, but they have not presented the proper implementation of DBM feature extractions. Mosayebi et al. [12] analyzed on the frequency of occurrence of cancer, and predictive model has been developed with multilayer neural network, Bayesian, LVQ, random forest, and KPCA-SVM. As implementation of several algorithms, in that authors produced the more accuracy by using KPCA-SVM. The KPCA process is implemented after the preprocessing of data. In this paper, authors implemented the model with several different kinds of ML-based algorithms but have said that accuracy of the system is gained by the KPCA-SVM. Nivaashini and Soundariya RS [13] have done their research to find the recurrence of the cancer. To develop the model, authors have utilized Wisconsin Diagnostic Dataset for the model. To perform the classification on trained datasets, they have implemented the C4.5 decision Tree, SVM, and Naïve Bayes for their research contribution in a significant way. They showed the model effectively to find lymph node status.

The algorithms implementation has done good validation for recurrence. Obaid et al. [14] explained the concept of evaluation process of several machine learning algorithms to predict the breast cancer. In this paper, mostly SVM, KNN, and decision tree-based algorithms were compared with result. Here, authors collected dataset from breast cancer (Diagnostic). Finally, they concluded that in finding of tumors and breast cancer, SVM works better under the ROC curve. Here, the authors not explained the concept of why other algorithm not performed well.

Ganggayah et al. [15] have done their research to predict the survival capacity of breast cancer patient. Here, they used the University Malaya Medical Centre dataset for their research. This model is constructed with decision tree, random forest, SVM, and logistic regression. Here, the random forest method used for clustering the data and also have given ranking system in the implementation. For the process of validation, authors implanted the decision tree. Here, authors have utilized the several machine learning-based algorithm feature to find the survival capacity among the many people. Ahmad et al. [16] have mainly focused on recurrence of breast cancer. In their research, they have implemented the Iranian Center for Breast Cancer (ICBC) dataset for their model. They have progressed their model with decision tree, SVM, and ANN [17, 18]. From their research, they got different results for different algorithms, but all the results are mostly near to each other. Apart from that, SVM has showed the better interpretation with low error rate and high accuracy on the result. Osman et al. [19] proposed a new concept to identify and predict the breast cancer which is hybrid SVM with two-step clustering. Here, proposed one provides the more accuracy on perdition and reduces the misclassification. To predict the malignant and benign tumors in this research, authors have implemented the two-step algorithm. Here, this hybrid method enhanced the key technique and brought the accuracy up to 99.1% on UCI-WBC dataset. Maglogiannis et al. [20] implemented the system to find automated breast cancer detection and utilized the WPBC dataset. Here, author implemented SVM for comparison and Bayesian classifier to classify the data. Here, they implemented radial basis function as key technique to achieve the better results. As per the results showed in the paper, they have achieved their proposed model but not satisfactory to current situation. Table 1 illustrates the comparative analysis of machine learning model.

Proposed Method

Data Preparation

We used Breast Cancer Wisconsin (Prognostic) Dataset from the UC machine learning library which has 578 instances for 34 features, out of which 34 are essential input features, and one is the target attribute that tells whether cancers re-occur or not (R/N). Out of 34 features like time, tumor size, and lymph node status are external features. The remaining 30 are internal features of the tumor cell nucleus. They are radius of tumor, area of tumor, perimeter of tumor, fractal dimension ("coastline approximation"—I), smoothness, texture (standard deviation of grayscale values), compactness (perimeter2/area—1.0) and symmetry, etc.

We have done data preprocessing to select useful features for training through a heatmap and removed the attributes with null values. We applied label encoding to convert all dataset values to numeric because we can train with only numerical values for a machine learning model.

Table 1 Comparison of all machine learning models

Reference	Dataset	Algorithms	Key technique	Best approach	Accuracy	Precision	Scope
1	Wisconsin Prognostic Breast Cancer Dataset (WPBC)	SVM, decision tree	Radial basis function kernel, or RBF kernel	RBF-SVM	97.93	93.36	Breast cancer recurrence prediction
2	Wisconsin Prognostic and Diagnostic Datasets	KNN SVM	Gaussian kernels	SVM-Gaussian kernel	93.54	–	Breast cancer recurrence prediction
3	Wisconsin (prognostic) dataset	Decision tree, SVM, DNN	Decision tree	Decision tree	94.70	92.23	Breast cancer diagnosis
4	Iran cancer research center	Random forest, C5.0, Bayesian, SVM, KPCA-SVM	KPCA	KPCA-SVM	98.24	–	Frequency of accuracy of cancer
5	Wisconsin Diagnostic Dataset	Naive Bayes, C 4.5 decision tree, and SVM, DNN	Deep Boltzmann machine	DNN	99.73	98.61	Breast cancer risk detection
6	Wisconsin Breast Cancer (diagnostic)	Support vector machine (SVM), decision tree (DT), K-nearest neighbor (KNN)	Linear kernel function	SVM	97.9	–	Classification of Wisconsin breast cancer
7	University Malaya Medical Centre	MLP-ANN, Decision tree, Weighted random Forest	Multilayer-perceptron-based artificial neural networks (MLP-ANN);	Decision tree weighted random forest	81.80	95.0	Predicting factors for survival of breast cancer patients

(continued)

Table 1 (continued)

Reference	Dataset	Algorithms	Key technique	Best approach	Accuracy	Precision	Scope
8	The Iranian Center for Breast Cancer (ICBC)	Decision tree (C4.5), SVM, and ANN	Tenfold cross-validation	SVM	95.0	–	Predicting breast cancer recurrence
9	UCI-WBC dataset	Hybrid support vector machine (SVM), and the two-step clustering technique	Two-step clustering technique	Hybrid SVM	99.10	98.20	Automatic diagnostic method for cancer
10	WDBC WPBC	SVM, Bayesian classifiers	Radial basis function	SVM	97.54	–	Automated breast cancer diagnosis and prognosis

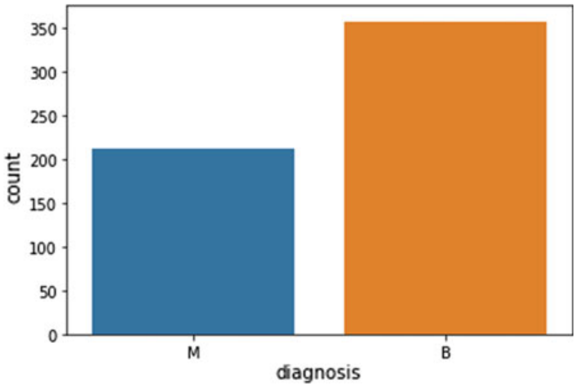


Fig. 1 Number of malignant (M) and benign (B) samples

We used one-hot encoding to normalize the data row-wise and column-wise. This can solve problems raised by label encoding. The dataset is divided into train and test in the ratio of 80:20 sing sklearn library with random sampling (Fig. 1).

Proposed Approach

We implemented two models one is support vector machine (SVM), and the other is ANN. The SVM is a machine learning model used for classification and regression. We used for classification to classify re-occurrence chances like yes or no. In SVM, we used the radial basis function (RBF) kernel (Gaussian Kernel) for polynomial classification. It calculates the squared Euclidean distance between origin and some point of data. The formula for the RBF kernel is shown in Eq. (1).

$$k\left(x, x^{1}\right) = \frac{\mathrm{e}^{\left(\left\|x-x^{1}\right\|\right)^{2}}}{2\sigma^{2}} \tag{1}$$

x, x^1 are feature vectors and σ is free parameter.

We also implemented ANN to train the BC dataset, which simulates a biological neuron for computations. We implemented ANN as shown in Fig. 2. With three layers, the first layer contains 30 input dimensions, and the activation function is rectified linear unit (Relu). The second layer with 20 nodes and activation function is Relu. In the third layer, its deciding layer is with a single node, the activation function is sigmoid. All three layers are densely connected. We used the ADAM optimizer for changing weights.

The Relu activation function is linear will give the output when it is a positive number, otherwise giving zero as output. The sigmoid is a nonlinear function in the output layer used as a binary classifier in the range of (0, 1).

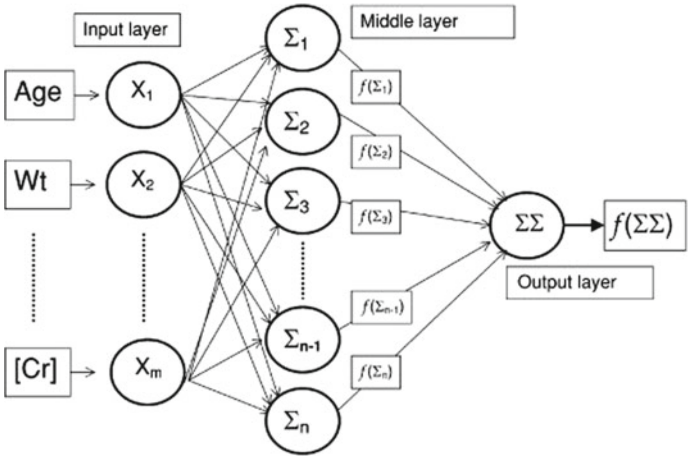


Fig. 2 ANN with three layers, input layers with Relu activation function and output layer with sigmoid activation function

The ADAM optimizer will change the weights unlike RMSprop, Adagrade. To calculate the loss, we used binary cross entropy loss function. ANN model is trained for 100 epochs, the accuracy we got average of 96.49.

Results and Analysis

Table 2 shows the results of RBF-SVM on the BC dataset, we got an average accuracy of 95%, and with the ANN model, we achieved an average of 96.49. Our proposed model outperformed when compared to other machine learning models. The proposed ANN model's training and testing accuracy is high, and the precision and recall showed better. The F1-score we got is 0.96, better than all the machine learning models. As shown in Fig. 3 confusion matrix, we got a high true positive rate and less false-negative rate. From the results given in Table 3, ANN model results are better than SVM RBF kernel algorithm.

Table 2 Results of SVM on RBC

	Precision	Recall	F1-score	Support
0	0.93	1.00	0.96	110
1		0.87	0.93	61
Accuracy			0.95	171
Macro avg	0.97	0.93	0.95	171
Weighted average	0.96	0.93	0.95	171

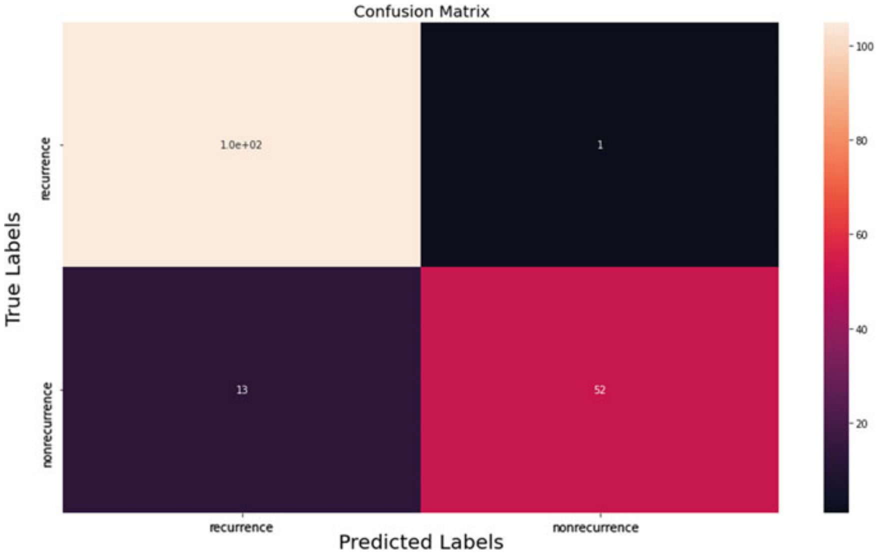


Fig. 3 Confusion matrix of ANN model

Table 3 Comparison of SVM and ANN accuracy

	SVM	ANN
Accuracy	0.95	0.9649
F1-score	0.93	0.95

S. No	Dataset	Algorithms	Key technique	Best approach	Accuracy	Precision	Scope
1	Wisconsin Prognostic and Diagnostic Datasets	KNN SVM	Gaussian kernels	SVM-Gaussian kernel	93.54	–	Breast cancer recurrence prediction
2	Wisconsin (prognostic) dataset	Decision tree, SVM, DNN	Decision tree	Decision tree	94.70	92.23	Breast cancer diagnosis
3	University Malaya Medical Centre	MLP-ANN Decision tree weighted random forest	Multilayer-perceptron based artificial neural networks (MLP-ANN);	Decision tree weighted random forest	81.80	95.0	Predicting factors for survival of breast cancer patients
4	The Iranian Center for Breast Cancer (ICBC)	Decision tree (C4.5), SVM, and ANN	Tenfold cross-validation	SVM	95.0	–	Predicting breast cancer recurrence

(continued)

(continued)

S. No	Dataset	Algorithms	Key technique	Best approach	Accuracy	Precision	Scope
5	**Proposed model**	**SVM, ANN**	**Neural network with multiple layers**	**ANN**	**96.49**	**93.0**	**Predicting breast cancer recurrence**

Conclusion

Re-occurrence of BC can happen any time in cancer patients who are affected first time. So, if we accurately diagnose the BC re-occurrence chances, we can prevent the losses. This paper has proposed two types of models, SVM and ANN, and compared the results. Based on results, the ANN model outperformed in predicting the re-occurrence of BC. We are concluding that ANN is better in prediction and classification because its multiple layers can also find critical computations on big datasets. With ANN, we got 96.49 of testing accuracy, so we suggest that using ANN we can predict BC at early stages to prevent re-occurrence chances after surgery.

References

1. McGuire A, Brown JAL, Malone C, McLaughlin R, Kerin MJ, Cicenas J (2015) Effects of age on the detection and management of breast cancer. Cancers
2. Jasmine, Leder K (2012) Rare events in cancer recurrence timing. In: Proceedings title: proceedings of the 2012 winter simulation conference (WSC), pp 1–10
3. Vijayakumar T (2019) Classification of brain cancer type using machine learning. J Artif Intell 1(02):105–113
4. Pandian AP (2019) Identification and classification of cancer cells using capsule network with pathological images. J Artif Intell 1(01):37–44
5. Manoharan S (2020) Early diagnosis of lung cancer with probability of malignancy calculation and automatic segmentation of lung CT scan images. J Innovative Image Proc (JIIP) 2(04):175–186
6. Karuppusamy P (2020) Hybrid manta ray foraging optimization for novel brain tumor detection. J Soft Comput Paradigm (JSCP) 2(03):175–185
7. Richter AN, Khoshgoftaar TM (2012) Predicting cancer relapse with clinical data: a survey of current techniques. In: 2016 IEEE 17th international conference on information reuse and integration (IRI) (2016): 369–376. Foo
8. Vijayakumar T (2019) Neural network analysis for tumor investigation and cancer prediction. J Electron 1(02):89–98
9. Chakradeo K, Vyawahare S, Pawar P (2019) Breast cancer recurrence prediction using machine learning. In: 2019 IEEE conference on information and communication technology, IEEE, pp 1–7
10. Rana M, Chandorkar P, Dsouza A, Kazi N (2015) Breast cancer diagnosis and recurrence prediction using machine learning techniques. In: IJRET: International Journal of Research in Engineering and Technology eISSN, pp 2319–1163

11. Sayed S, Ahmed S, Poonia RK (2017) Holo entropy enabled decision tree classifier for breast cancer diagnosis using wisconsin (prognostic) data set. In: 20 17 7th international conference on communication systems and network technologies (CSNT) (2017), pp 172–176
12. Mosayebi A, Mojaradi B, Naeini AB, Hosseini SHK (2020) Modeling and comparing data mining algorithms for prediction of recurrence of breast cancer. PloS one 15(10):e0237658
13. Nivaashini M, Soundariya RS (2018) Deep Boltzmann machine-based breast cancer risk detection for healthcare systems. Int J Pure Appl Math 119:581–590
14. Obaid OI, Mohammed MA, Ghani MKA, Mostafa A, Taha F (2018) Evaluating the performance of machine learning techniques in the classification of Wisconsin Breast Cancer. Int J Eng Technol 7(4.36):160–166
15. Ganggayah MD, Taib NA, Har YC, Lio P, Dhillon SK (2019) Predicting factors for survival of breast cancer patients using machine learning techniques. In: BMC medical informatics and decision making 19(1):1–17
16. Ahmad LG, Eshlaghy AT, Poorebrahimi A, Ebrahimi M, Razavi AR (2013) Using three machine learning techniques for predicting breast cancer recurrence. J Health Med Inform 4(124):3
17. Ramesh D, Shashikala, Harshavardhan A, Mahesh D (2021) COVID-19 and SARS Virus predictions from chest X-ray images using a deep learning model. In: Bhateja V, Satapathy SC, Travieso-González CM, Aradhya VNM (eds) Data engineering and intelligent computing. Advances in intelligent systems and computing, vol 1. Springer, Singapore. https://doi.org/10.1007/978-981-16-0171-2_39
18. Sheshikala M, Dadi R, Swamy P, Rajanala V (2020) A survey paper on convolution neural network in identifying the disease of a cotton plant
19. Osman AH (2017) An enhanced breast cancer diagnosis scheme based on two-step-SVM technique. Int J Adv Comput Sci Appl 8(4):158–165
20. Maglogiannis I, Zafiropoulos E, Anagnostopoulos I (2009) An intelligent system for automated breast cancer diagnosis and prognosis using SVM based classifiers. Appl Intell 30(1):24–36

Image Processing-Based Switching Matrix for PV Array Under Partial Shading Conditions

D. Manimegalai and M. Karthikeyan

Abstract The photovoltaic (PV) system will have the impact of decreasing its efficiency because of partial shadowing, and it can be overcome by using reconfiguration method. In this paper, an analysis of shadow moving and the Sudoku pattern-based reconfiguration method are proposed. The proposed approach gives a new configuration with radiation values incorporated in to PV system. This method had two sections: flexible and static, which are arranged with the help of switching matrices. The optimization of PV arrays is done by using image processing, to acquire the information of shadows. Also, in this paper, both the Total-Cross-Tied (TCT) and Sudoku pattern connections are used in the test. The camera is used to monitor the shadows on the PV system, and the reconfiguration is applied according to a predefined time threshold. The tests are performed in real-time, and the proposed system gives the increase in average energy extraction for TCT and Sudoku pattern configuration.

Keywords Image processing · Shadow moving · Switching matrix · Reconfiguration method · Sudoku pattern

Introduction

The solar, wind, rainfall, tides, geothermal heat, and other renewable natural resources are used to generate renewable energy. The photovoltaic (PV) system, which is based on the photovoltaic effect, is the most direct way to convert solar radiation into electricity. The photovoltaic effect is a physical phenomenon that involves the conversion of optical electromagnetic radiation's energy into electrical energy. Mostly, the performance of the PV system is affected by the external factors which include deteriorations on the panel surface (e.g., stains, spots, residuals, scratches),

D. Manimegalai (✉)
Department of Electrical and Electronics Engineering, Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Chennai, Tamil Nadu, India

M. Karthikeyan
Department of Electrical and Electronics Engineering, Vel Tech Multi Tech Dr. Rangarajan Dr. Sakunthala Engineering College, Chennai, Tamil Nadu, India

V. Bindhu et al. (eds.), *Proceedings of Third International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 844,
https://doi.org/10.1007/978-981-16-8862-1_9

clouds casting shadows on the panels and full or partial shadows cast by nearby objects. An ideal PV system is expected to overcome these problems by a flexible system and increase the energy extraction with less cost. Advanced PV systems follow several methods to diminish the aforementioned negative effects and seek to maximize the performance and also cost ratio.

The partial or full shadows on the panels are the main source of the loss in available power in PV systems [1, 2]. Even though there has been a significant boost in the amount of energy mustered from the PV panels, the effort to prevent the factors hindering the performance falls short. Many methods can lessen the negative effects in PV arrays [3]. These include the maximum power point tracker (MPPT), returned energy architecture, and reconfiguration techniques proposed in the literature [1–4]. The MPPT usually comes first among these methods. The MPPT is maintained by utilizing a DC-DC converter and controlling the algorithm and the current and voltage values obtained from the panels [2, 4]. The main goal of MPPT is to attain MPP by tracking the terminal voltage and current. To attain MPP, the DC-DC converter and a MPPT algorithm between the PV module and load are employed [4, 5]. Various MPP techniques are proposed in the literature. The most prominent methods are Hill Climbing (HC) [4], Perturb and Observe (P&O) [6, 7], and incremental conductance [8]. A more efficient hybrid structure is formed in one of the proposed methods by using these techniques [9]. The major problem with the MPPT methods is partial shading. In this case, the global peak point (GP) in the power-voltage (P–V) curve cannot be determined. Consequently, the PV system produces less energy.

Another approach to overcome the negative effect of full or partial shading on PV panels is reconfiguration [10, 11]. In short, reconfiguration is the rearrangement of the panels in PV arrays within a presented topology (e.g., Series–Parallel, Total-Cross-Tied). This process is usually carried out by using a switching matrix [12]. The effects of any type of shadow on the panels could be eliminated by changing the electrical connections without any need of a rearrangement of the panel positions. Thus, the shadows on the panels can be relocated to any position by only altering the connections. Therefore, the amount of energy produced increases. Furthermore, the problems coming from the shading of the panels are mitigated in this way. Hence, it facilitates the detection of the GP more accurately and precisely.

Nguyen et al. proposed an adaptive reconfiguration method in [13]. Two different configuration approaches are presented in this paper. A bubble sort algorithm is used in the first approach to find the optimal panel layout before applying the adaptive array switching. In the second approach, the current panels' power levels are estimated using a reference model of power levels from each fixed panel row. Karakose et al. utilize an image processing algorithm in [13, 14]. A camera constantly monitors the PV array, and shadows casting on panels are detected with this approach. Then, radiation values of the panels are evaluated. The optimal panel layout is determined by basing on the radiation values, and it is realized by altering the switching matrix accordingly. Liu et al. considers the shading degree for the reconfiguration. To accomplish this, Liu et al. determine the shadows formed on the rows in the first stage. The shading degrees in these rows are then determined by a sorting algorithm. In the second stage, the switching matrix connects an adaptive part to a fixed part of the PV arrays, based on the shading-degree control algorithm [15]. El-Dein et al. proposed a new mathematical formula for a reconfiguration [16]. Moreover, the branch and bound algorithm were used to determine the optimal layout. The fully or partially reconfigurable arrays are considered in the simulation studies for this paper. It has been observed that the proposed method extracts between 10 and 20% of the energy available. Cheng et al. has performed a fuzzy logic based reconfiguration process [17]. The shading degrees are used in this paper and the optimal panel layout is obtained by using fuzzy logic. It has been observed that the system provides fast and accurate responses in different environmental conditions.

Karakose et al. presents a fuzzy partitioning based on the reconfiguration technique in another Karakose paper [18]. Fuzzy values are used to express the short circuit current and radiation values of the panels. These are divided into five intervals. Storey et al. [19] proposed a dynamic cell configuration approach for an optimal energy output from the PV arrays. The system in this study has shown an average increase of 22.6% energy, as compared to the traditional approaches. Karakose et al. [20] utilizes the clonal selection algorithm for the reconfiguration. The calculation of the optimal panel layout is provided for energy extraction under the partial shading conditions in this paper.

In conventional MPPT techniques, PV system experiences a significant loss in their performance under full or partial shading conditions. This effect can be overcome by using this reconfiguration method, and it has the following benefits:

- The energy extraction from PV system are increases
- Performance of the system increases with reduced cost.
- The effects due to partial shading can be eliminated.

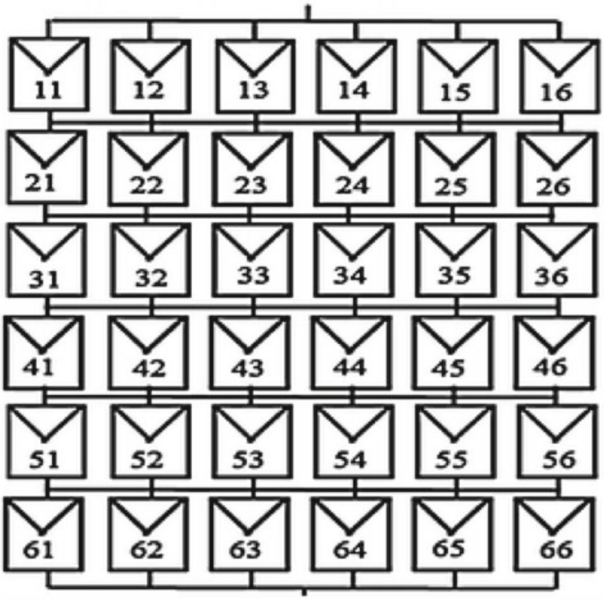
In this paper, the proposed reconfiguration methods are described in Sect. 2. The optimization using image processing method and the optimization algorithm are examined in Sect. 3. The experimental setup and results are presented in Sect. 4. The conclusion is discussed in Sect. 5.

Reconfiguration Method

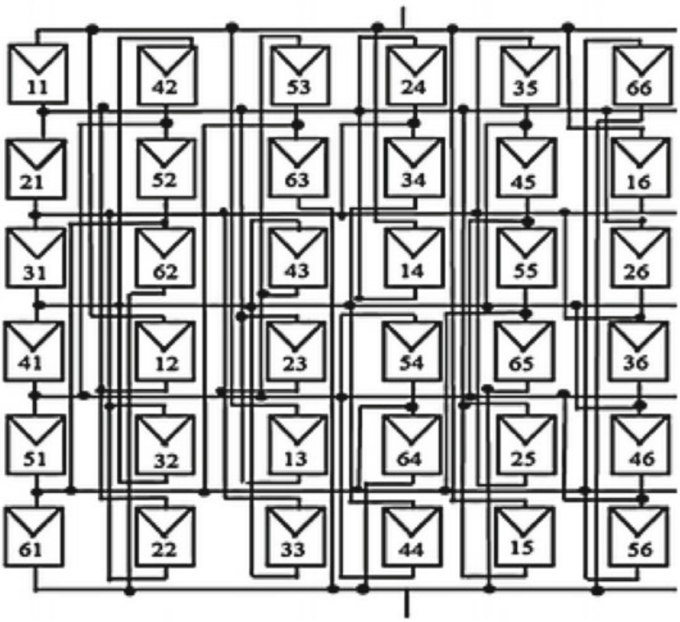
The proposed reconfiguration method reduces the effect due to partial shading occurs on PV arrays. Due to this, the amount of energy extracted from the system can be improved. In this paper, we consider two types of connections: 1. Total-Cross-Tied (TCT) and 2. Sudoku Pattern.

Total-Cross-Tied (TCT): When all solar cells are a connected parallel, they are comprised of modules. These modules comprise the TCT type by connecting the series. The TCT connection type is presented in Fig. 1a.

Sudoku Pattern: Sudoku as depicted in Fig. 1b inspired its name from the Sudoku game, which is an improvement of TCT configuration. Sudoku based on changing the row numbering of each module in a 6×6 array to improve the overall system's power output under partial shading conditions. Furthermore, the column number



(a)



(b)

Fig. 1 Connection layout of **a** TCT and **b** Sudoku pattern remains the same as in the TCT topology, while the row in which the module is physically placed is adjusted according to the Sudoku layout. As a result, as in a TCT arrangement, the electrical connections of the modules remain unchanged, while the physical positions of the modules are the only ones that change [21–23].

The physical placement of the modules inside the PV array is changed to disperse the shade effect fairly equitably throughout the PV array. To accomplish this, the subarrays are wired as static reconfiguration in the first phase of the suggested technique. In the suggested method, the physical position of the modules is determined as follows:

$$Y_{ij} = X_{kj}(i = 1, 2, \ldots, m \text{ and } j = 1, 2, 3, \ldots n) \tag{1}$$

$$k = (i + (j - 1)d) = \begin{cases} k, & k < m, \\ k - m, & \text{otherwise} \end{cases}. \tag{2}$$

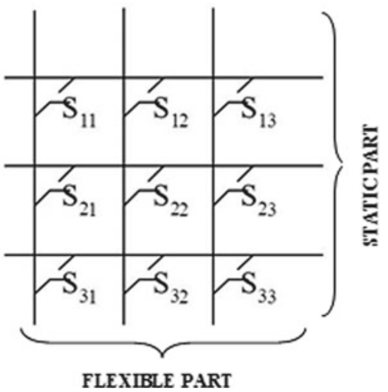
The distance of separation d for a mxn subarray is an index that indicates the degree of module variation in each row and may be calculated as

$$d = \text{cell}(\sqrt{m}) \tag{3}$$

where m and n are the number of the previous row and column of the module, respectively. Also, d is the size of the subarray rows and columns, and reflects the number of the new rows in the module (Fig. 2).

If the physical position of the modules results in an existing value of "k," a value of "1" is added to the offset. The left column's first-row indices are determined by adding "d" to the previous index. If the result of this subtraction is an index that already exists, the new index will be calculated by adding an offset of "1" to the final value. If the result of this subtraction is an index that already exists, the new index will be calculated by adding an offset of "1" to the final value.

Fig. 2 Switching matrix



The characteristics of the PV module are mostly affected due to the radiation values. In this, the output current equation is given by,

$$I_{\text{Ph}} = \frac{G}{G_n}(I_{scn} + K_i(T - T_n)) \tag{4}$$

where I_{ph} is a short circuit current of a PV cell; G_n is a nominal radiance value, T_n is the nominal temperature, and K_i is the temperature coefficient of the short circuit current. The proposed reconfiguration method contains image processing and the radiation values. It has the following advantages:

- Sensors are not used to measure current and voltage values
- Simple in implementation
- It can implement in all environments.

Proposed System

The proposed reconfiguration has two basic steps:

Step 1: To detect the shaded portion using an image processing algorithm and calculate the degree of shade.

Step 2: To decide the optimal panel layout using artificial immune system.

The camera will monitor the panel continuously and analyze the shadow whether it is temporary or permanent. If it is a permanent shadow, the radiation and the proportion of its area are calculated and taken into consideration; otherwise, it will be ignored. The information of the calculated radiation and region are fed into the control and decision-making module, to check the feasibility of the optimal PV layout. The general flowchart of the proposed system is shown in Fig. 3.

Image Processing for Shadow analysis

The moving shadow analysis stage consists of two sections: (1) A camera continuously monitoring the PV arrays and (2) An image processing module detecting a full or partial shading area. The image frames acquired from the camera are processed on Raspberry Pi kit. The shadows are categorized in this stage.

The shadows are categorized according to time spent on the monitored region. If the shadow stays longer than an experimentally predefined threshold time, then it is considered a permanent shadow. As such, it is a necessary decision, because cloudiness varies during the daytime and clouds could cover the sun for a very short period of time. After this categorization, the features of the shadow are extracted. The contour of the shadow is determined by the Canny edge detection algorithm. The panels on which the shadow falls are then determined. The intensity of the shadow

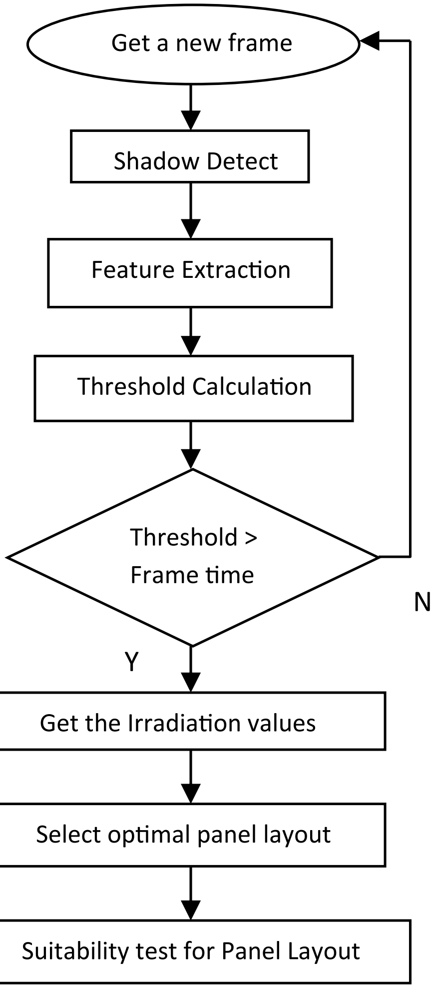


Fig. 3 Block diagram of proposed system

and the size of its area are calculated. Lastly, the radiation values are evaluated from the information derived. In the conducted studies, the radiation values are completely obtained by the image processing algorithm. The shadows occurring on the panels are detected and the gray level of these shadows in the application is revealed.

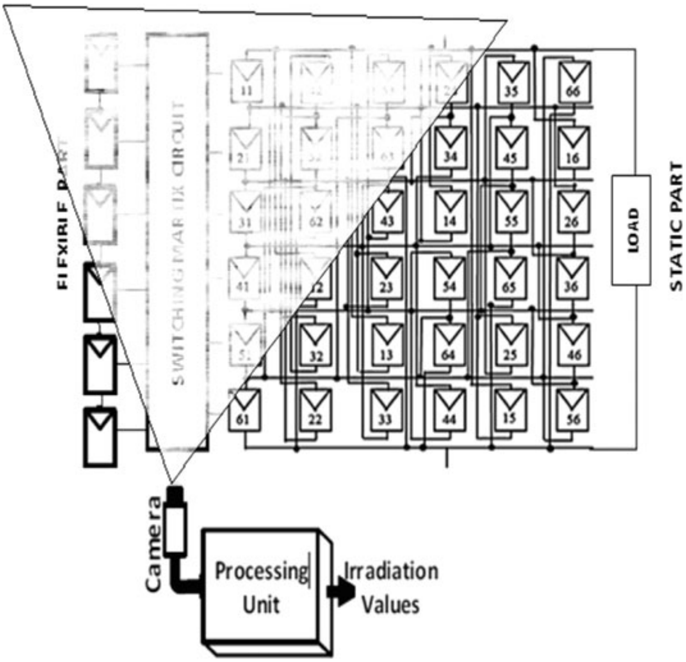
Equation (4) in Sect. 2 is used for the calibration and verification of the irradiation values obtained from the application. As it is mentioned in the beginning of the section, the detection of the shadows and analysis process are carried out by the image processing algorithm used in this paper. Firstly, an image frame is captured from the camera. Then, the image frame is converted into the HSV color space. Several morphological methods, such as erosion and dilation, are subsequently applied. A noise reduction is performed on the image frames afterward. Contour detection takes place next. It is carried out on the denoised frames. A Canny edge detection algorithm is used during the contour detection phase. It is reported that the Canny yields the best result among the other available edge detection algorithms (e.g., Sobel, Prewitt). The shadow is categorized by its level of darkness. The panels on which the shadow falls are determined in the next stage of the algorithm. Lastly, the total image frame time from the camera and the predetermined threshold time in the proposed approach are compared. This threshold value is a user-defined. Thus, it determines whether the shadows are temporary or permanent. The reason for the realization of this process is to avoid energy losses by continuously changing the state of the panels. This is because, during short time changes, an instantaneous change in the shadow values (sudden movement of the clouds, flying objects, etc.) will result in a certain loss of energy.

As initially mentioned, this threshold used in the system is completely user-defined and is defined as 9 min for this study. In the measurements, the image taken and the elapsed time for processing it is approximately 600 ms. In this process, one image per second is taken. This process is carried out 540 times for 9 min. If the time calculated is higher than a threshold value, the shadow is considered a permanent shadow. The radiation and shadow area information derived are saved and later sent to the sorting and matrix control. Otherwise, the capturing of a new image frame will be conducted. The proposed algorithm and the flowchart of image processing are demonstrated in Fig. 4.

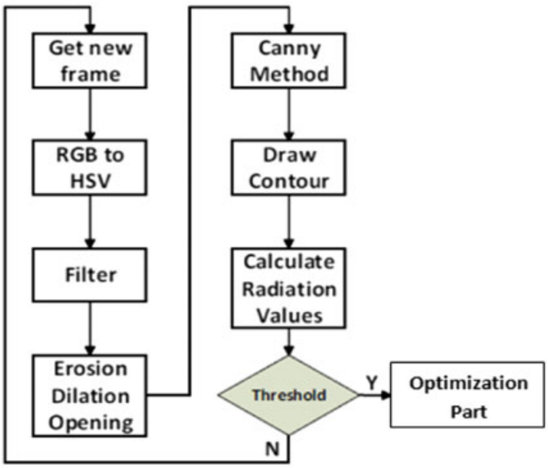
The image processing part is mainly comprised of six primary steps. These steps are as follows.

- *RGB to HSV conversion*: The first of these steps is converted into the HSV color space of the received image frame. The gray level color conversion is generally preferred in the literature. The HSV color space is an effective tool in revealing the shadows. HSV stands for hue, saturation, and value and is also often called HSB (B for brightness). In this paper, these three values are scaled between 0 and 255, and each value is determined individually.
- *Filter*: After the HSV conversion, a filter is applied to the image in the second step, and an image frame is converted to a binary image. This conversion process is necessary for the erosion and dilation operations.
- *Erosion and Dilation*: The erosion and dilation operations are two basic morphological operations used to reveal the details in the image frame. If these two processes are performed in sequence, the noises in the image will be removed, and the image will remain the same
- *Opening:* After the erosion and dilation operations, the opening process is applied to the image frame in the fourth step. The opening process is useful to remove small objects. It is a combination of the erosion and dilation operations. The erosion operation is followed by the dilation operation. The opening procedure is presented in Eq. (5).

$$\text{Opening} = \text{Dilation}(\text{Erosion}(\text{frame})) \tag{5}$$



(a)



(b)

Fig. 4 **a** Block diagram of proposed system and **b** Flowchart of image processing

- *Canny Edge Detection and Draw Contour*: The shadows can be identified by using all of these steps. The boundaries of the shadow are detected by using the Canny algorithm.

Optimization-Based Reconfiguration

The second part of the paper is the implementation optimization-based reconfiguration. The radiation values obtained from the image processing are fed into the optimization algorithm. Therefore, the optimum panel arrangement is determined. The clonal selection algorithm is used as the optimization method. The algorithm takes the radiation values and the shaded region information as input parameters from the image processing part. The system that calculates the optimum arrangement of the panels by using these values proceeds with the sorting procedure and the matrix control. The purpose of this step is to choose the best combination of possible new panel layouts that will occur when the switching matrix is used.

The new panel layout obtained using a clonal selection algorithm may not be possible to be implemented with the system. To avoid this, a decision and control module are added to the optimization module. The optimization module calculates the best yield among all the possible locations of the switching matrix module. Then, the control and decision module takes this new panel layout as input parameters and tests the applicability of these possible combinations. If the new order is confirmed, the "true" value is produced and a new configuration performs. Otherwise, this module producing a "false" value returns to the optimization module and the panel. The block diagram of the arranging and matrix control is shown in Fig. 5.

The implementation of the optimization module in the paper is compatible with both the TCT and Sudoku connection types. As is described in Sect. 2, for the TCT connection type, gathering the shadows in different regions into the same rows

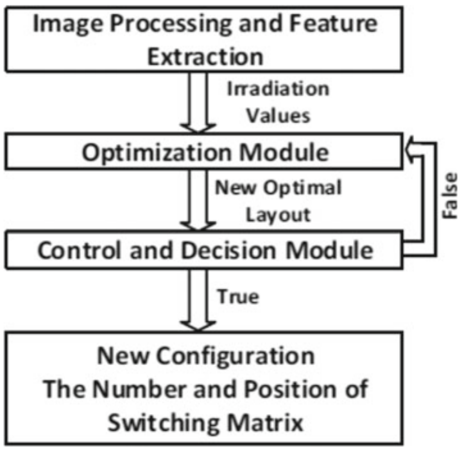


Fig. 5 Block diagram of arranging and matrix control

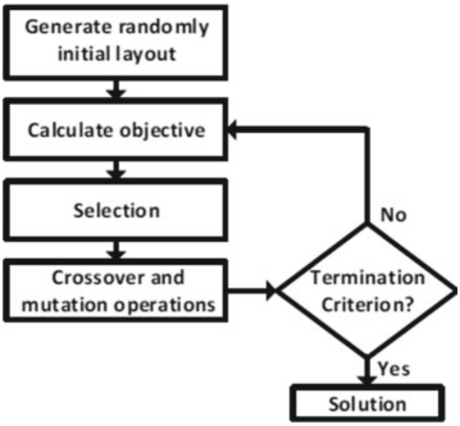


Fig. 6 Flowchart of clonal selection algorithm

increases the energy extraction. Conversely, for the Sudoku connection type, the even distribution of the shadows over the system is expected. In other words, every row is expected to have an equal or its close to equal radiation level.

The clonal selection algorithm is used as the optimization method in this paper. To carry out the optimization process, the radiation values are taken into account. Figure 6 is a block diagram illustrating the use of the clonal selection algorithm in its operation.

The desired connection scheme can be obtained without the need to change any place in the PV systems by using the switching matrix. Theoretically, all connection variants may be obtained by applying a switching matrix between each module. This is not very practical in terms of both the complexity and the costs. Thus, one switching matrix is used in this paper, and the position of this matrix in the system is constant. The schematic diagram of the switching matrix used in the system is presented in Fig. 7.

Experimental Results

Experiments were performed on a real PV system to verify the output of the proposed method. In this study, a PV system was established that was 6×6 in size. The connections were performed according to the TCT and Sudoku connection types. The system basically consists of four units: (1) the camera for the monitoring PV system, (2) the 6×6 PV panel system, (3) Raspberry Pi kit that performs image processing and control procedures, and (4) a switching matrix. The tests were conducted under the different shading conditions and temperatures. The applications were simulated in the MATLAB platform, and the results were observed. A real experiment environment was then created for the reconfiguration in the PV systems. The validations of the results obtained from the simulation were then performed. A block diagram

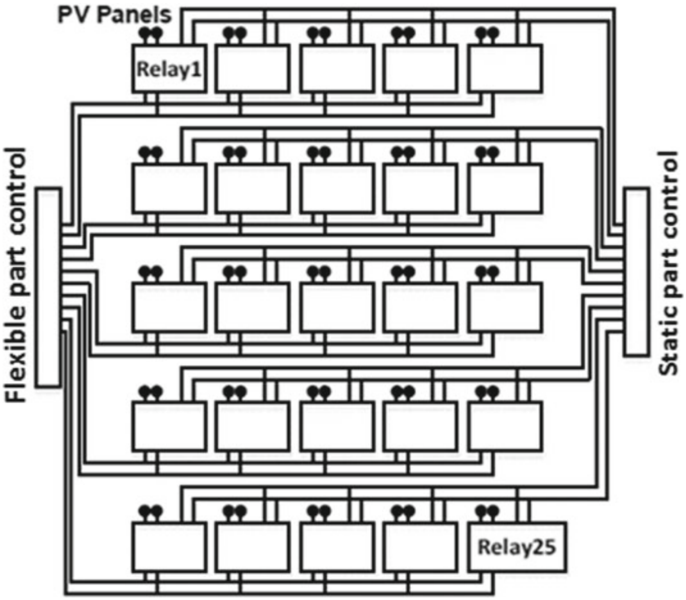


Fig. 7 Schematic diagram of switching matrix

illustrating the relationship between the test equipment is provided in Fig. 8. The system has the capacity to produce a total of approximately 1 kW. The PV panels used in the system are Lorentz LC80-12 M models.

Shadow Detection Using Image Processing

The PV system with image processing applications is constantly monitored by using a camera. The FL2G—13S2M—C 1394 model cameras are used to monitor the system. The features of the camera of the image processing are presented in Table 1.

Fig. 9 gives the outcomes obtained as a result of the image processing obtained for TCT and Sudoku connection types.

Reconfiguration of PV Array

In the reconfiguration part of this paper, the sorting and matrix control processes are performed by using radiation values obtained from the image processing. In this section, the optimization and control-decision module are used for the reconfiguration. The optimization module takes the radiation values of the panels as input parameters and yields an optimal panel layout as the output parameters. After this

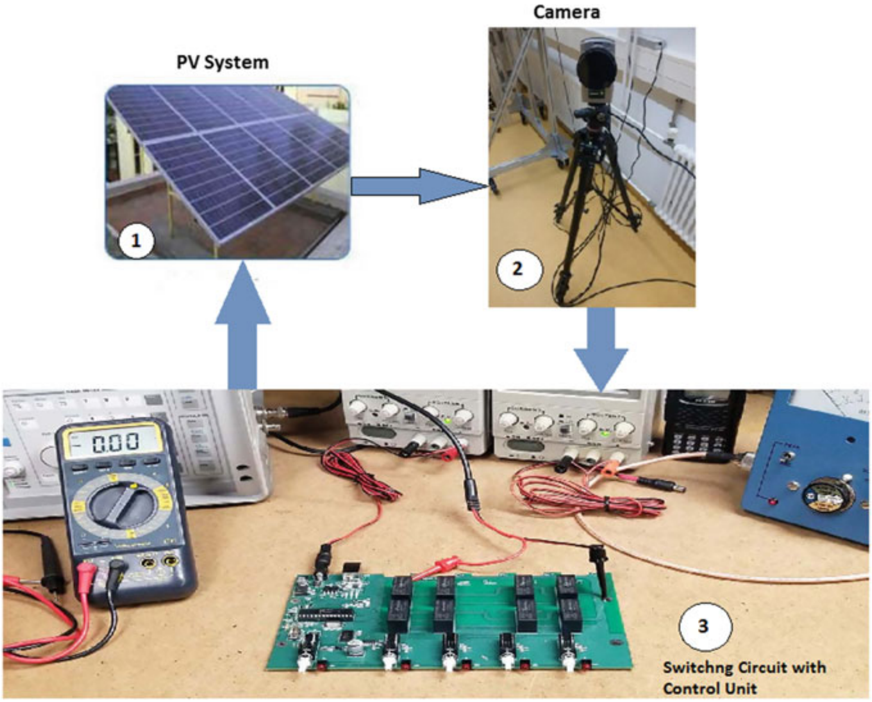


Fig. 8 Testing equipment

Table 1 Characteristics features of camera

Description	Specifications
Model name	FL2G—13S2M—C 1394
Company	Pint grey research, Inc
Resolution	1288 × 964
Frame rate	30 FPS
Megapixels	1.3 MP
Chroma	Mono
Sensor name	Sony ICX445
Sensor type	CCD
Pixel size	3.75 μm
Lens mount	C-Mount
Image processing Power requirement dimensions	Gamma, lookup table, hue, saturation, and sharpness 8-30 V 29 mm × 29 mm × 30 mm

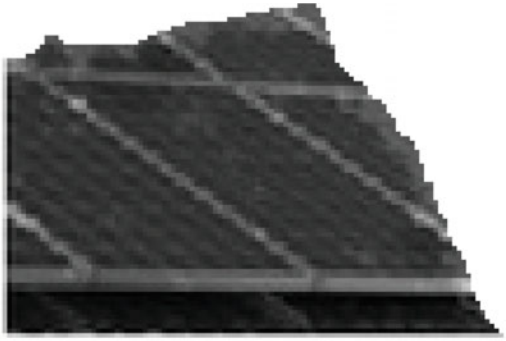
process, the control and decision module are activated. The control and decision module determine the optimal panel layout practically, whether it is provided or not, and if it is provided. The change of the power-voltage curves can be obtained from the PV system for both TCT and Sudoku configurations. The schematic block diagrams of the new panel layout are presented in Fig. 10 for the TCT connection type and Fig. 11 for Sudoku type. Also, the energy amount changes for this type of connection are presented in Table 2.

The system is operated by a practical selection of the threshold time. The threshold time value is user-defined and may be determined in seconds or minutes. These values, determined as 9 min in this paper, and the status of the shading occurring within this time were examined. Thus, there was an average increase of about 11% and 8% for the TCT and Sudoku connection layouts, respectively. The proposed method mainly consists of three modules. These are image processing, optimization

(a) Real PV System

(b) Conversion of RGB to HSB

Fig. 9 Output obtained from ımage processing



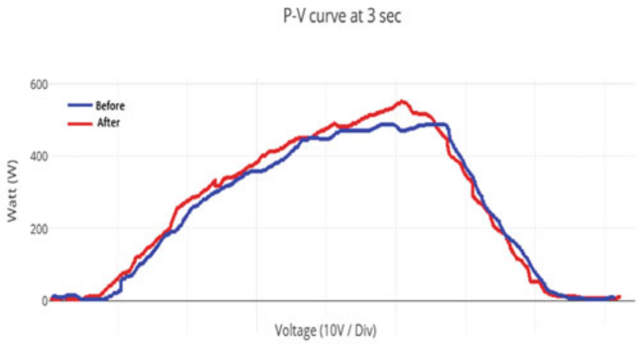
(c) Grey Level Detection

(d) Shadow Detection

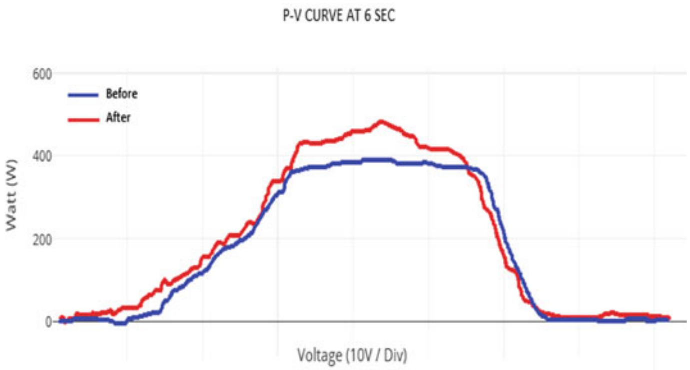
Fig. 9 (continued)

and combination, and the switching matrix module, respectively. The operations performed on these modules are carried out on the control unit. The average time consumption of all these modules is given in Table 3. These time values may be better depending on the processor type.

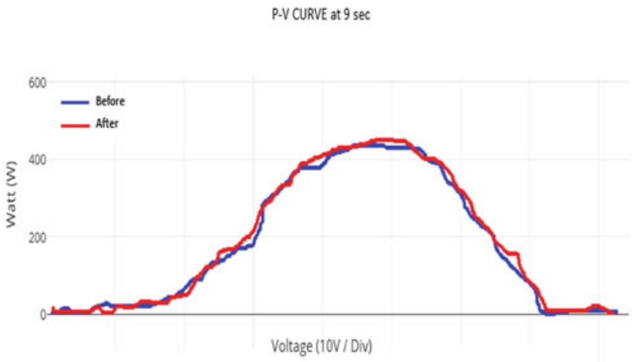
Limitation of this work is that the camera has to monitor the PV array continuously and it will not show the correct result based on the permanent shadow like building, trees, etc., and moving shadows like clouds, birds, etc. To overcome these limitations, in future, and to differentiate this shadow types, the artificial neural network (ANN) can be used to do it.



(a) Reconfiguration process at 3 sec

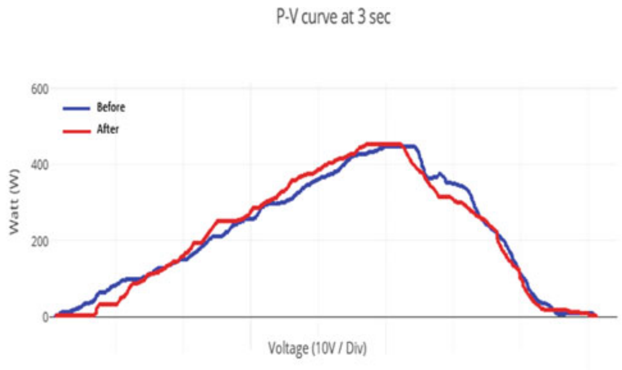


(b) Reconfiguration process at 6 sec

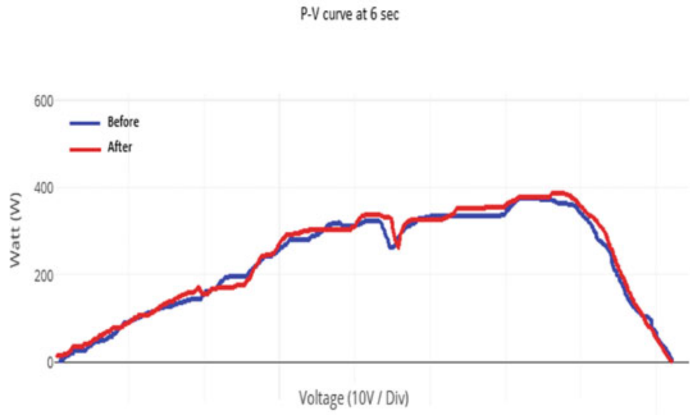


(c) Reconfiguration process at 9 sec

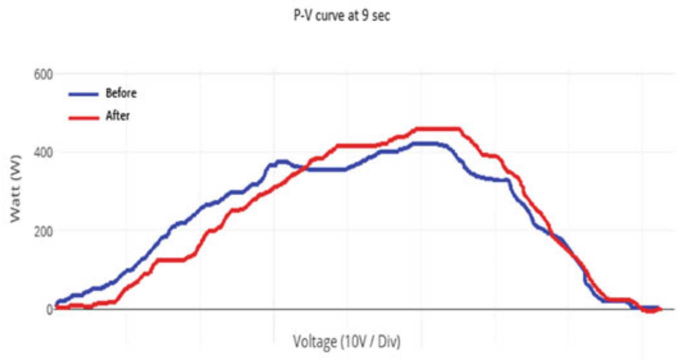
Fig. 10 Results of the reconfiguration for the TCT connection type



(a) Reconfiguration process at 3 sec



(b) Reconfiguration process at 6 sec



(c) Reconfiguration process at 9 sec

Fig. 11 Results of the reconfiguration for the Sudoku pattern connection type

Table 2 Power variation for different connections

TCT				Sudoku			
Scenario	Before	After	Percent (%)	Scenario	Before	After	Percent (%)
Figure 10a	472 W	543 W	13.1	Figure 11a	424 W	436 W	2.7
Figure 10b	396 W	464 W	14.6	Figure 11b	365 W	396 W	7.8
Figure 10c	423 W	444 W	4.7	Figure 11c	412 W	478 W	13.8

Table 3 Computational time for the proposed approach

Modules	Time consumption (~ms)
Image processing	102
Optimization and combination	64
Switching matrix	22
Total	188

Conclusion

In this WORK, a new method has been proposed to determine the full or partial shading situation and decrease the negative effects of these shadings. The shading occurring at any time during the day on PV arrays is detected by using image processing based on a shading analysis and obtained values for the shading area. The next step is to determine whether the shading is temporary or permanent. If it is temporary, the process of a new frame is carried out; otherwise the optimization module is put into operation. The optimization module designed in this study can be used with both TCT and Sudoku connection types. The optimization module takes the radiation values as an input parameter. It then determines the best configuration. This configuration has been tested in combination with the availability of the module. After the process ends in the combination module, the new connection structure is applied on the switching matrix circuit and PV system, respectively. The proposed approach has four basic advantages: (1) it can be applied to both the TCT and Sudoku connection types; (2) the effects of shadings can be eliminated, and there is no need for sensors for this process; (3) the system works according to the threshold value. Because of that, it does not commission unnecessary switching matrix circuits, and it does not cause energy consumption; and (4) the efficiency of the method is validated with the experimental data, which yields a significant energy increase.

References

1. Patel H, Agarwal V (2008) MATLAB-based modeling to study the effects of partial shading on PV array characteristics. IEEE Trans Energy Convers 23:302–310
2. Dolara G, Lazaroiu C, Leva S, Manzolini G (2013) Experimental investigation of partial shading scenarios on PV (photovoltaic) modules. Energy 55:466–475

3. Manimegalai D, Meenakshi S (2016) Energy harvesting from solar cells under electric lighting sources in indoors. ARPN J Eng Appl Sci. 11(19):11395–11402. ISSN 1819–6608
4. Rawat R, Chandel SS (2013) Hill climbing techniques for tracking maximum power point in solar photovoltaic systems-a review. Spec Issue Int J Sustain Dev Green Econ (IJSDGE) 2(1):90–95
5. Femia N, Petrone G, Spagnuolo G, Vitelli M (2005) Optimization of perturb and observe maximum power point tracking method. IEEE Trans Power Electron 20(4):963–973
6. Bashar A, Smys S (2021) Integrated renewable energy system for stand-alone operations with optimal load dispatch strategy. J Electron 3(02):89–98
7. Femia N, Petrone G, Spagnuolo G, Vitelli M (2007) Predictive and adaptive MPPT perturb and observe method. IEEE Trans Aerosp Electron 43(3):934–950
8. Safari, Mekhilef S (2011) Simulation and hardware implementation of incremental conductance MPPT with direct control method using Cuk converter. IEEE Trans Ind Electron 58(4):1154–1161
9. Jiang LL, Nayanasiri DR, Maskell DL, Vilathgamuwa DM (2013) A simple and efficient hybrid maximum power point tracking method for PV systems under partially shaded condition. In: The 39th international annual conference of the ieee industrial electronics society (IECON 2013). Vienna, Austria, pp 1513–1518, 10–13 Nov 2013
10. Dos Santos P, Vicente EM, Ribeiro ER (2011) Reconfiguration methodology of shaded photovoltaic panels to maximize the produced energy. Braz Power Electron Conf (COBEP) 700–706, Sept 2011
11. Spagnuolo G, Petrone G, Lehman B, Ramos Paja CA, Zhao Y, Orozco Gutierrez ML (2015) Control of photovoltaic arrays: dynamical reconfiguration for fighting mismatched conditions and meeting load requests. Ind Electron Mag 9(1):62–76
12. Balato M, Manganiello P, Vitelli M (2014) Fast dynamical reconfiguration algorithm of PV arrays. In: 9th international conference on ecological vehicles and renewable energies (EVER). Monte Carlo, pp 1–8, Mar 2014
13. Nguyen D, Lehman B (2008) An adaptive solar photovoltaic array using model-based reconfiguration algorithm. IEEE Trans Ind Electron 55(7):2644–2654
14. Karakose M, Baygin M (2014) Image processing based analysis of moving shadow effects for reconfiguration in PV arrays. In: IEEE international energy conference (ENERGYCON), pp 724–728, May 2014
15. Shakya S (2021) A self monitoring and analyzing system for solar power station using IoT and data mining algorithms. J Soft Comput Paradigm 3(2):96–109
16. Liu Y, Pang Z, Cheng Z (2010) Research on an adaptive solar photovoltaic array using shading degree model-based reconfiguration algorithm. In: Control and decision conference (CCDC). Chinese, pp 2356–2360, May 2010
17. El-Dein MZS, Kazerani M, Salama MMA (2013) Optimal photovoltaic array reconfiguration to reduce partial shading losses. IEEE Trans Sustain Energy 4:145–153
18. Cheng Z, Pang Z, Liu Y, Xue P (2010) An adaptive solar photovoltaic array reconfiguration method based on fuzzy control. Intell Control Autom (WCICA). Jinan, pp 176–181
19. Karakose M, Baygin M, Baygin N, Murat K, Akin E (2014) An intelligent reconfiguration approach based on fuzzy partitioning in PV arrays. In: IEEE international symposium on innovation in intelligent systems and applications (INISTA). Alberobello, Italy, pp 176–181, June 2014
20. Storey JP, Wilson PR, Bagnall D (2013) Improved optimization strategy for irradiance equalization in dynamic photovoltaic arrays. IEEE Trans Power Electron 28:2946–2956
21. Karakose M, Firildak K (2015) A shadow detection approach based on fuzzy logic using images obtained from PV array. In: 6th international conference on modeling, simulation, and applied optimization (ICMSAO). Istanbul, Turkey, pp 1–5

22. Manimegalai D, Karthikeyan M, Vijayakumar SC (2018) Maximizing power output of partially shaded photovoltaic arrays using sudoku configuration. ARPN J Eng Appl Sci. 13(1):124–133, Jan 2018, ISSN 1819–6608
23. Karthikeyan M, Manimegalai D (2018) Sudoku pattern based solar PV array for maximum power extraction using genetic algorithm. J Adv Res Dyn Control Syst 10(12-Special Issue):559–564, Aug 2018

A Deep Learning Implementation of End-to-End Image Denoising Steganography Model

Raksha Ramakotti and Surekha Paneerselvam

Abstract Steganography is a methodology that essentially aims in hiding a secret data in a fitting carrier, essentially in such a way that the hidden data does not attract any attention toward it. The postulation in steganography is if the secret feature is perceptible, then so is the point of attack. Also, clean images when subjected to prolonged transmission, improper image acquisition or conditioned to multiple feature changes lead to image tarnishing due to unwanted noisy pixels. This proposes to be a major threat in steganography as the presence of noise itself could be a reason for the secret image to be revealed. In this work, a bi-objective end-to-end stacked denoising stenography model is implemented. The results are estimated by evaluating the average peak signal-to-noise ratio (PSNR) and structural similarity index (SSIM) of the trained images using COCO dataset. The stability of the proposed model is estimated using the average loss obtained by training the model for multiple iterations.

Keywords Image steganography · Image denoising · Convolutional neural networks · COCO dataset · Batch size · Peak signal-to-noise ratio · Structural similarity index

Introduction

With the expeditious proliferation in the public availability of digital information and with the Internet becoming more and more common, the necessity of acquiring a secure means of communication for sensitive data has become a great challenge. The work on avoiding plagiarism, data eavesdropping, morphing and data alteration has become the need of the hour [1]. Primarily, three such techniques have come into

R. Ramakotti (✉) · S. Paneerselvam
Department of Electrical and Electronics Engineering, Amrita School of Engineering, Amrita Vishwa Vidyapeetham, Bengaluru, India
e-mail: bl.en.p2ebs19009@bl.students.amrita.edu

S. Paneerselvam
e-mail: p_surekha@blr.amrita.edu

V. Bindhu et al. (eds.), *Proceedings of Third International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 844,
https://doi.org/10.1007/978-981-16-8862-1_10

existence, i.e., cryptography, watermarking and steganography [2, 3]. Though the former techniques have been comprehensively put into practice, they portray a few shortcomings, i.e., regardless how robust the cryptographic-encryption algorithm is, it provides a scope of being decoded. In case of watermarking, the capacity of information chosen is limited by the application. It is steganography that potentially and effectively bridges these gaps. The advantage of steganography when compared to other methods is that the trace of secret information is totally camouflaged by the carrier. To formally define, steganography is the art of covered or hidden writing where the goal is to covertly communicate a digital message [4]. In computer vision and image processing, one of the major challenges faced is handling images that are tampered with noise [5]. Digital images, when subjected to prolonged exposure of environment, multiple feature changes or random variations of brightness and color information, lead to degradation of the image quality. In the context of steganography, the presence of noise itself, therefore, could be a reason for the relevance of the secret image. Hence, it is important to adapt a robust denoising technique. The challenge however is that not only does it have to identify and remove the noise, but conduct this process without compromising on the image reconstructed quality.

Literature survey reveals many image processing techniques that have been used to accomplish the steganography technique. Broadly, these methods can be categorized into three, namely spatial domain, transform domain and machine learning techniques. Spatial domain aims to represent a grayscale image as a 2D matrix or a color image as a 3D vector of 2D matrices. Popular algorithms include least significant bit (LSB), edge-based data embedding and random pixel embedding methods [6, 7]. The major transform domain techniques imparted for steganography include discrete cosine transform (DCT) [8–11], discrete Fourier transform (DFT) [12] and discrete wavelet transform (DWT) [10, 11].

For noise removal, the conventional methods adapted though seem to be effective, however, convey certain disadvantages. Firstly, based on the nature of the image, the adapted methodology's parameters have to be set manually, each and every time. Secondly, the process of image denoising is mathematically and computationally complex and expensive. Also, in order to ensure that the background image is not lost, some amount of preprocessing technique may have to be adapted [13]. However, with the development of artificial intelligence, the above challenges can be bridged [14].

Machine learning algorithms are computational models that learn to make decisions without a monitored instruction by learning through a training data. Literature survey suggests that these models are broadly used in classification and regression [15, 16]. Deep learning is a subset of machine learning that mimics the workings of the human brain in processing data for use in detecting objects, recognizing speech and making decisions, to name a few [17–20]. The edge presented by deep learning is that it evinces the ability to learn without human supervision, drawing from data that is both unstructured and unlabelled. Convolutional neural network (CNN) is a subset of neural network that is used to understand the objectives related to images. The handshake of neural networks to perform steganography has proven to improve the performance, robustness and efficiency of secret image communication [21–23]. In [24], the authors present a comparative study on LSB substitution technique and

CNN-based architecture. Wu et al.'s whole processing pipeline consists of two almost identical neural network structures responsible for encoding and decoding.

Image denoising has also been performed using neural networks in the recent times. A survey of the recent models providing state-of-the-art results is presented in [14]. Autoencoders with skip connections are used in the denoising model architecture showed in [25]. Multiple objectives of denoising and data classification are done in [26]. In [27], a comparative analysis on various convolutional models for rain streak modeling has been addressed showcasing the working of autoencoders and residual networks. Other related work can be referred from [13, 28–30].

In this work, which is an extension of [31], the prospective of noisy images in the context of steganography is explored. This has led to proposing two objectives: one to design a stacked denoising model and two to implement a robust steganography model. The dataset used is COCO and is subjected to a mixed additive Gaussian noise with a variances of 0.0001, 0.0005, 0.001, 0.005, 0.01 and 0.05. The criteria of mixed denoising are explored to evaluate the stability of the models. The performance is estimated based on peak signal-to-noise ratio (PSNR) and structural similarity index (SSIM). The following sections consist of methodology, where the description of the models is presented, and in the analysis and results section, the model performance based on PSNR, SSIM and loss is evaluated. The paper culminates with conclusion and future scope.

Methodology

In this section, a brief description of the proposed bi-objective model is elucidated.

Proposed Model Architecture

The aim of this paper is to present a bi-objective sequential denoising steganography model. As shown in Fig. 1, a denoised model is stitched with a steganography model for this purpose. The details are presented in the following sections.

Input Images The images are treated in their digital format and represented as $(m \times n \times p\times)$, where $(m \times n)$, represents the image size and p the number of color channels. The dataset chosen for training the model is COCO, i.e., common object in context [32]. This color dataset consists of objects captured from everyday scenes. This dataset being open source was particularly chosen as it introduces the concept of generalization among images through non-iconic images. A subset of 1000 images was selected for the analysis and was resized to a standard size of (128×128). In the context of image processing, normalization refers to changing the range of the pixel values. An image I can be represented as

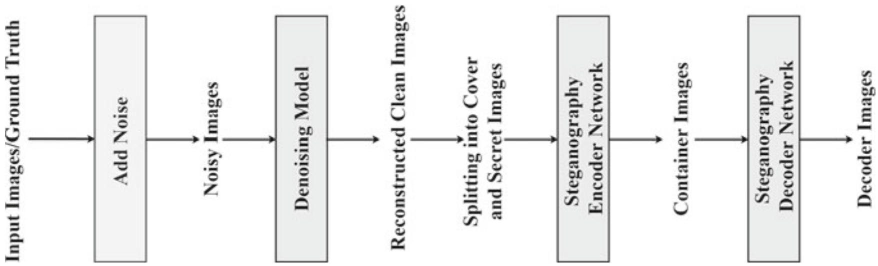


Fig. 1 End-to-end denoising steganography model

$$I = \sum_{i=0}^{M} \sum_{j=0}^{N} P(i, j) \tag{1}$$

where M and N are equal to 255 and P represents the pixel value at the ith and jth coordinate in the image surface.

Since the range is large, image normalization is adapted to reduce the computational complexity. Thus, Eq. 1 can be reframed as

$$I_{\text{norm}} = \sum_{i=0}^{1} \sum_{j=0}^{1} P(i, j) \tag{2}$$

Nature of Noise Noise is a synonym of the unwanted signal that does not contain any useful information. Noise on an image can simply be translated to random variation of brightness or color information in images. More often, its impact has always been posed as a major challenge in the image processing domain as it abases the constitution of the image, thereby posing serious threats to image quality and the information that it contains. As pointed out in the previous section, it can be adapted during processes such as image acquisition, subjecting the image to heavy coding, transmitting the image in an unpredictable channel and during subjective processing. In this work, the clean COCO dataset is subjected to a Gaussian noise, which usually arises as a threat during image acquisition. The nature of such a noise is treated to be additive in nature, i.e.,

$$x' = x + n \tag{3}$$

where x' is the noisy image, x is the ground truth, and n is the Gaussian noise.

The probability density function of a Gaussian variable z is given by

$$\text{PDF}(z) = \frac{1}{s\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2s^2}} \tag{4}$$

where m represents mean and s represents variance. Here, the noise's level is defined with zero mean and mixed variances 0.0001, 0.0005, 0.001, 0.005, 0.01 and 0.05.

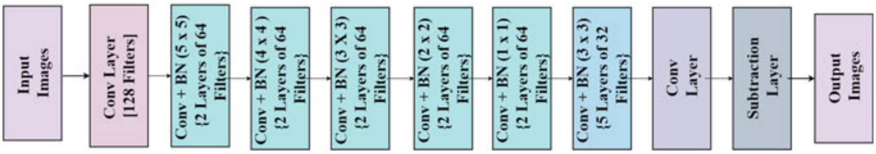


Fig. 2 Architecture of VFRN model

Mixed noise is induced to address a real-time scenario where images are not subjected to same noise level, thereby avoiding overfitting of the model. The resultant image x' is a noisy image, which is the actual input to the denoising models.

Proposed Denoising Model: Varied-Filter Size Residual Network (VFRN)

As observed in Fig. 2, the VFRN aids in treating noise as a residual problem. Here, instead of constructing the clean image directly, the residue, or in other words, the noisy distribution is determined by the model. Then, it subtracts this residue with the noisy image in order to obtain the clean image. The mathematical representation is showed in Eq. 5,

$$r = x' - f(n) \tag{5}$$

where r is the reconstructed clean image, x' is the noisy input image, n is the noise quantity, and $f(n)$ is the function of the noisy pattern obtained by the model. The objective of the model is to present the reconstructed image r very similar to that of the ground truth or the clean image x. The model fabricated to achieve this is known as denoising convolutional neural network (DnCNN). The DnCNN has a convolutional layer that aids in extracting the important features of the image. The model also has sandwiched domino layers of convolution and batch normalization. The input images are divided into several parts which are called mini-batches, for a single epoch. The batch normalization layer normalizes a layer input by subtracting the mini-batch mean and dividing it by the mini-batch standard deviation. As observed in Eq. 6, for the kth mini-batch, the batch normalized output $b\hat{n}k$ is given by

$$b\hat{n}k = \frac{bnk - E[bnk]}{\sqrt{Var[bnk]}} \tag{6}$$

where $E[bnk]$ represents the expectation of the kth mini-batch and the denominator represents its standard deviation. The usage of batch normalization alleviates the effect of internal covariance shift, which is caused due to effect of randomness in the parameter initialization and input data during training process. This process thus standardizes inputs to a layer for each mini-batch, thereby stabilizing the learning process reducing the number of training epochs. The problem of model overfitting is resolved, and the process of model learning is accelerated [20]. The residual pattern obtained is then subjected to a subtraction layer in order to obtain the reconstructed clean image.

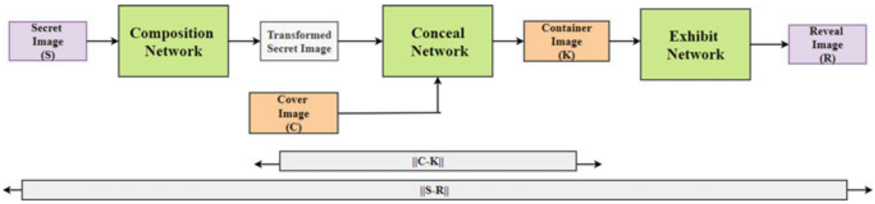


Fig. 3 End-to-end steganography model architecture

The noisy input image is fed to a single layer containing 128 filters of (3 × 3) filter size. Post that, the model is subjected to a 15 layers of convolution + batch normalization layers. The first ten layers consist of 64 filters where the filter sizes vary from (5 × 5) to (1 × 1). The larger filters such as (3 × 3), (4 × 4) and (5 × 5) aim in capturing the bigger picture, thereby providing a substantial understanding of the background image. Smaller filter sizes aid in capturing the intricate details of the image. The subsequent five layers consist of 32 filters of a uniform (3 × 3) filter size, which is followed by a convolution layer and a subtraction layer. The output of the latter layer presents the reconstructed clean image. These images of the VFRN model are then split into 500 cover (C) and secret images (S) in order to train the subsequent steganography model. It is to be noted that the VFRN is trained to minimize the mean0-squared error between (MSE) r and x, i.e.,

$$\mathrm{MSE} = \frac{1}{MN}\sum_{M,N}[r(m,n) - x(m,n)]^2 \tag{7}$$

Proposed Steganography Model The end-to-end process of steganography is obtained from our previous work [31]. As shown in Fig. 3, the encoder has two networks, composition and conceal networks. The composition network is trained to extract the features of the secret image and present it to the conceal network in such a way that there is slight to nil trace of the secret image on the carrier image. The latter aims in actually indulging in the process of steganography. The model's objective is to present a container image perceivably similar to that of the cover image. The decoder consists of a single exhibit network whose task is to remove the carrier image without distorting the structural integrity of the secret image. Strikingly, the container image is the input to the exhibit network, and the output image is to be as similar as the secret image. The loss equation that is used to train the steganography model is retained from [31] and is given by,

$$L(C, K, S, R) = ||C - K|| + ||S - R|| \tag{8}$$

The details of the composition network, which is also shared by the other two networks, are depicted in Fig. 4. Though the network borrows its architecture heavily from [31], a slight modification of introducing TanH activation function in the last

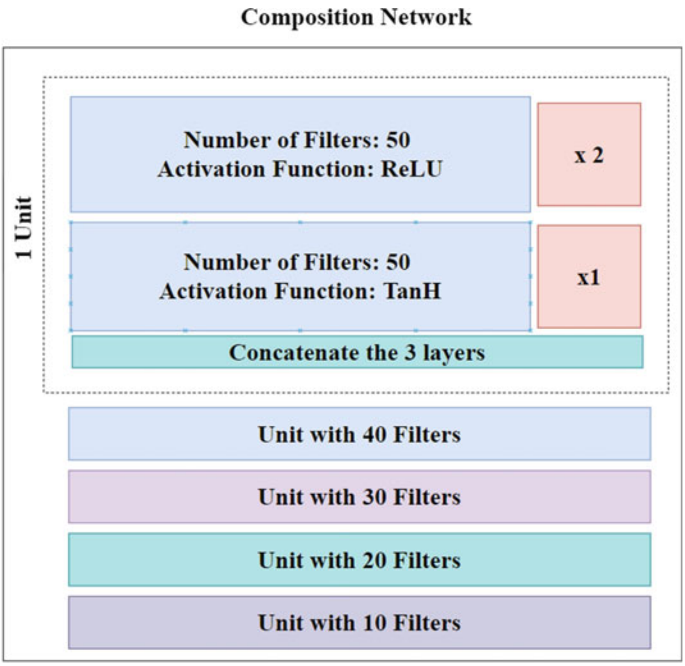


Fig. 4 Composition network architecture

layer of each unit has been done to boost the model's performance. The first two layers of the unit retain ReLU activation function. Their equations are illustrated below, and it is to be noted that x is a general variable in Eqs. 9 and 10 and does not stand for ground truth image.

$$f(x) = ReLU(x) = 0, x \tag{9}$$

where n ∈ positive integers I+

$$f(x) = \mathrm{TanH}(x) = \frac{e^{2x} - 1}{e^{2x} + 1} \tag{10}$$

Implementation and Results

The evaluation metrics used to estimate the performance of the model architectures are presented in this section. Also the description of the model implementation and results is showcased. All the models were implemented in Python 3 script in Google Colab.

Evaluation Metrics

In applications that involve images, the quantifying evaluation metrics used to determine the performance of the model are PSNR and SSIM. The details are provided in the following subsections.

Peak Signal-to-Noise Ratio The peak signal-to-noise ratio (PSNR) [33] represents the ratio between the maximum power value of a signal and the power of distorting noise that affects the quality of its representation. It will be measured in decibels. The higher the PSNR, the better the quality of the reconstructed image and is given as,

$$\mathrm{PSNR} = 10\log_{10}(\frac{\mathrm{MAX}_{y_{ij}}^2}{\mathrm{MSE}}) \tag{11}$$

where MSE stands for mean-squared error and the numerator term represents the maximum possible pixel value of the image. When the pixels are represented using 8 bits per sample, this is 255.

Structural Similarity Index The structural similarity index (SSIM) [34] is an enduring metric that quantifies the degradation of an image quality due to extensive image processing or modeling. SSIM is given by,

$$SSIM(I, O) = \left(\frac{(2m_I m_O + c_1)(2s_{IO} + c_2)}{(m_I^2 + m_O^2 + c_1)(s_I^2 + s_O^2 + c_2)}\right) \tag{12}$$

where,

- m_I m represents the average of the input image I
- $m_O$represents the average of the output image O
- s_I represents the variance of the input image O
- s_O represents the variance of the output image O
- s_{IO} represents the covariance of the images I and O

Steps Incorporated to Validate the Model's Stability

Performance Analysis of the Steganographic Model

In order to validate the stability of the proposed model depicted in Fig. 1, a couple of steps have been adapted to ensure that the model does not throw any overshoots and undergoes saturation. It is to be noted that

- PSNR_R is the PSNR of the reconstructed denoised images when compared to the ground truth.
- SSIM_R is the SSIM of the reconstructed denoised images when compared to the ground truth.

- PSNR_C1 is the PSNR of the container/cover images when compared to the reconstructed denoised images.
- PSNR_C2 is the PSNR of the container/cover images when compared to the ground truth.
- PSNR_S1 is the PSNR of the reveal/secret images when compared to the reconstructed denoised images.
- PSNR_S2 is the PSNR of the reveal/secret images when compared to the ground truth.
- SSIM_C1 is the SSIM of the container/cover images when compared to the reconstructed denoised images.
- SSIM_C2 is the SSIM of the container/cover images when compared to the ground truth.
- SSIM_S1 is the SSIM of the reveal/secret images when compared to the reconstructed denoised images.
- SSIM_S2 is the SSIM of the reveal/secret images when compared to the ground truth.

The sequence of steps adopted to estimate the stability of the proposed model are as follows: 1. Perform the experiment for a couple of iterations
2. Compute the respective cover and secret images' absolute losses and calculate their average losses, i.e., compute the averages of :

i. PSNR_C2-PSNR_C1
ii. SSIM_C2-SSIM_C1
iii. PSNR_S2-PSNR_S1
iv. SSIM_S2-SSIM_S1

3. Repeat the experiment for steganography model subjected to clean images and modified hyper-parameters and compute the average losses, i.e.,

i. PSNR_C2-PSNR_C0
ii. SSIM_C2-SSIM_C0
iii. PSNR_S2-PSNR_S0
iv. SSIM_S2-SSIM_S0

The parameters that were kept constant throughout the experiment are tabulated in Table 1.

PSNR_R and SSIM_R are computed for all the 1000 images, whereas the other metrics are computed for 500 images. From Table 2, it can be observed that the average PSNR_R is 28.63 dB and SSIM_R is 0.87. However, due to the computational limitation, since a lesser number of images were used to train two complex objectives using stacked networks, when the reconstructed images are split into 500 cover and secret images and are subjected to steganography, PSNRs less than 20 dB and SSIM less than 0.9 were achieved. However, in order to justify the stability of the model, the evaluation metric losses mentioned Sect. 10.3.2 must be less. The significance is that the image distortion between the output and input images is minimal, which is the ultimate agenda for any computer vision problem statement. These losses are presented in Table 3.

Table 1 Constant parameters for the implementation

Parameters	Values
Train image count	1000
Image Size	(128 × 128 × 3)
Mean of Gaussian noise	0
Batch size	10
VFRN epochs	50
Steganography epochs	100
Loss	Mean-squared error
Optimizer	Adam
VFRN model parameters	616,707
Steganography model parameters	1,027,626

Table 2 Metrics tabulation for COCO color dataset

Iteration	PSNR_R (dB)	SSIM_R	PSNR_C1 (dB)	PSNR_C2 (dB)	SSIM_C1
1	22.96	0.76	15.63	17.31	0.78
2	32.37	0.92	16.08	17.29	0.82
3	29.70	0.88	17.43	17.44	0.83
4	29.49	0.90	15.47	16.30	0.79
Average	28.63	0.87	16.15	17.09	0.81
Clean	N.A.	N.A.	N.A.	18.38	N.A.
Itreration	SSIM_C2	PSNR_S1 (dB)	PSNR_S2 (dB)	SSIM_S1	SSIM_S2
1	0.79	13.21	15.03	0.76	0.60
2	0.81	18.51	17.83	0.87	0.87
3	0.77	17.56	17.75	0.83	0.87
4	0.78	18.57	17.67	0.88	0.86
Average	0.79	16.96	17.07	0.84	0.80
Clean	0.85	N.A.	18.25	N.A.	0.82

From Table 3, it can be observed that all the PSNR losses are minimum and the differences are lesser than 1.3 dB. In most cases, the average of absolute difference between the respective reconstructed outputs of steganography and that of denoising models (i.e., $\|PSNR\_X2 - PSNR\_X1\|$) are lower than the average of absolute difference between the respective reconstructed outputs of steganography and that of the ground truth ($\|PSNR\_X2 - PSNR\_X0\|$), where X can be substituted as S or C. The SSIM losses are negligible too, proving the fact that the cognitive construction of images is pretty similar to the ground truth. The highest obtained loss is 0.06 which

Table 3 Tabulation of mean of the absolute metrics differences for COCO color dataset

Absolute diff.: mean	(M_C1, M_C2)	(M_C2, M_C0)	(M_S1, M_S2)	(M_S2, M_S0)
Metric (M): PSNR (dB)	0.94	1.29	0.11	1.18
Metric (M): SSIM	0.02	0.06	0.04	0.02

is observed between $\|SSIM\_C2 - SSIM\_C0\|$. From the above observations, it is undoubtedly provable that the constructed denoising steganography model is stable considering the fact that considering that the models are trained for two complex objectives.

Conclusion and Future Scope

In this work, two sequential objectives are addressed, i.e., to perform denoising and steganography. The models were trained using 1000 COCO images, and the results showcased minimal distortion between the input and the reconstructed images. The overall PSNR loss was less than 1.3 dB, and SSIM loss was less than 0.07 proving that the proposed model is stable. This work's restrain was computational limitation. Therefore, it can be extended by training the models on high-performance computers. Also, GANs and ensemble techniques can be adopted to improvise the model's performance. The models can also be trained by subjecting the images to other noise types and explore the working of the model for real-time noisy images.

References

1. Zou, Y., Zhang, G., & Liu, L. (2019). Research on image steganography analysis based on deep learning. Journal of Visual Communication and Image Representation, 60, 266–275
2. Cheddad, A., Condell, J., Curran, K., & Mc Kevitt, P. (2010). Digital image steganography: Survey and analysis of current methods. Signal processing, 90(3), 727–752
3. Al-Yousuf, F. Q. A., & Din, R. (2020). Review on secured data capabilities of cryptography, steganography, and watermarking domain. Indonesian Journal of Electrical Engineering and Computer Science (IJEECS), 17(2), 1053–1059
4. Pund-Dange S (2018) Steganography: a survey. In: Cyber security. Springer, pp 327–333
5. Tian, Chunwei, Yong Xu, and Wangmeng Zuo. "Image denoising using deep CNN with batch renormalization." Neural Networks 121 (2020): 461–473
6. Hussain, M., Wahab, A. W. A., Idris, Y. I. B., Ho, A. T., & Jung, K. H. (2018). Image steganography in spatial domain: A survey. Signal Processing: Image Communication, 65, 46–66
7. Wang, Z. H., Chang, C. C., & Li, M. C. (2012). Optimizing least-significant-bit substitution using cat swarm optimization strategy. Information Sciences, 192, 98–108

8. Kadhim, I. J., Premaratne, P., Vial, P. J., & Halloran, B. (2019). Comprehensive survey of image steganography: Techniques, Evaluations, and trends in future research. Neurocomputing, 335, 299–326
9. Rachmawanto, E. H., & Sari, C. A. (2017). Secure image steganography algorithm based on dct with otp encryption. Journal of Applied Intelligent System, 2(1), 1–11
10. Sari, W. S., Rachmawanto, E. H., & Sari, C. A. (2017). A Good Performance OTP encryption image based on DCT-DWT steganography. Telkomnika, 15(4), 1987–1995
11. Kumar V, Kumar D (2010) Digital image steganography based on combination of DCT and DWT. In: International conference on advances in information and communication technologies. Springer, Berlin, Heidelberg, pp 596–601
12. Mandal JK, Khamrui A (2011) A genetic algorithm based steganography in frequency domain (GASFD). In: 2011 international conference on communication and industrial application, Kolkata, pp 1–4
13. Tian, Chunwei, Yong Xu, Lunke Fei, Junqian Wang, Jie Wen, and Nan Luo. "Enhanced CNN for image denoising." CAAI Transactions on Intelligence Technology 4, no. 1 (2019): 17–23
14. Tian C, Fei L, Zheng W, Xu Y, Zuo W, Lin C-W (2020) Deep learning on image denoising: an overview. Neural Netw
15. Raksha R, Surekha P (2020) A cohesive farm monitoring and wild animal warning prototype system using IoT and machine learning. In: 2020 international conference on smart technologies in computing, electrical and electronics (ICSTCEE). IEEE, pp 472–476
16. Se S, Vinayakumar R, Kumar MA, Soman KP (2016) Predicting the sentimental reviews in Tamil movie using machine learning algorithms. Indian J Sci Technol 9(45)
17. Zhao, Z. Q., Zheng, P., Xu, S. T., & Wu, X. (2019). Object detection with deep learning: A review. IEEE transactions on neural networks and learning systems, 30(11), 3212–3232
18. Zhang, Z., Geiger, J., Pohjalainen, J., Mousa, A. E. D., Jin, W., & Schuller, B. (2018). Deep learning for environmentally robust speech recognition: An overview of recent developments. ACM Transactions on Intelligent Systems and Technology (TIST), 9(5), 1–28
19. Ganesan S, Krichen M, Alroobaea R (2020) Robust malware detection using residual attention network
20. Poorna SS, Reddy MRK, Akhil N, Kamath S, Mohan L, Anuraj K, Pradeep HS (2020) Computer vision aided study for melanoma detection: a deep learning versus conventional supervised learning approach. In: Advanced computing and intelligent engineering. Springer, pp 75–83
21. Baluja S (2017) Hiding images in plain sight: deep steganography. In: Advances in neural information processing systems, pp 2069–2079
22. Rahim R, Nadeem S (2018) End-to-end trained CNN encoder-decoder networks for image steganography. In: Proceedings of the European conference on computer vision (ECCV)
23. Zhang, R., Dong, S., & Liu, J. (2019). Invisible steganography via generative adversarial networks. Multimedia tools and applications, 78(7), 8559–8575
24. Wu, P., Yang, Y., & Li, X. (2018). Stegnet: Mega image steganography capacity with deep convolutional network. Future Internet, 10(6), 54
25. Bajaj K, Singh DK, Ansari MA (2020) Autoencoders based deep learner for image denoising. Proc Comput Sci 171:1535–1541
26. Zhang, Jinsong, Yi Zhang, Lianfa Bai, and Jing Han. "Lossless-constraint denoising based auto-encoders." Signal Processing: Image Communication 63 (2018): 92–99
27. Akaash B, Aarthi R (2021) An analysis of rainstreak modeling as a noise parameter using deep learning techniques. In: Advances in computing and network communications, pp 465–477. Springer
28. Nishio M, Nagashima C, Hirabayashi S, Ohnishi A, Sasaki K, Sagawa T, Hamada M, Yamashita T (2017) Convolutional auto-encoder for image denoising of ultra-low-dose CT. Heliyon 3(8):e00393
29. Haque KN, Yousuf MA, Rana R (2018) Image denoising and restoration with CNN-LSTM encoder decoder with direct attention. arXiv:1801.05141
30. Menon N (2017) A survey on image steganography. In: 2017 international conference on technological advancements in power and energy (TAP Energy). IEEE, pp 1–5

31. Ramakotti R, Paneerselvam S (2021) An analysis and implementation of a deep learning model for image steganography. In: International conference on ubiquitous communications and network computing. Springer, Cham, pp 210–224
32. Lin T-Y, Maire M, Belongie S, Hays J, Perona P, Ramanan D, DollÃÂir P, Lawrence Zitnick C (2014) Microsoft coco: common objects in context. In: European conference on computer vision. Springer, Cham, pp 740–755
33. Hore A, Ziou D (2010) Image quality metrics: PSNR vs. SSIM. In: 2010 20th international conference on pattern recognition. IEEE, pp 2366–2369
34. Wang, Z., Bovik, A. C., Sheikh, H. R., & Simoncelli, E. P. (2004). Image quality assessment: from error visibility to structural similarity. IEEE transactions on image processing, 13(4), 600–612

Analysis on Contribution of Cryptography and Steganography in Protecting Information in Diverse Environments

Muzammil H. Mohammed

Abstract Information is a key parameter in any type of application that must be treated with the utmost care at all times when it is used for any purpose. Security measures for this information are critical when dealing with anonymous or illegitimate users of web-based services and applications. Information security can be dealt with any techniques, which are specifically protected with a great extent of algorithms that are available under network security. Cryptography is the most succsessful technique used for protecting the information to any domain specific application, which contains huge information. This research work has analyzed various algorithms of cryptography and steganography for securing the data. Finally, this research work has analyzed the aforementioned strategies by combining steganography along with cryptography for obtaining good results in securing the sensitive information.

Keywords Information security · Cryptography · Steganography · Cryptography algorithms · Types of steganography

Introduction

Cryptography is the study of developing novel methods that allow data to be moved in a secure structure so that only the intended recipient can recover it. Contingent upon the key, cryptography is arranged into two groups. The encryption and decoding cycles of Symmetric Key Encryption [SKE] utilize a similar key.

This methodology is basic yet compelling, anyway key conveyance is the most major problem that should be given a research attention. Unbalanced Key Encryption utilizes two numerically related keys: For encryption, the user require a public key and private key. It is notable that these boundaries can't be utilized to improve the calculation due to the strategy's unstructured engineering. For encryption and

M. H. Mohammed (✉)
Department of Information Technology, College of Computers and Information Technology, Taif University, Taif, Saudi Arabia
e-mail: m.muzammil@tu.edu.sa

V. Bindhu et al. (eds.), *Proceedings of Third International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 844,
https://doi.org/10.1007/978-981-16-8862-1_11

unscrambling, practically all algorithms utilize keys, which may address execution time, proficiency, and other execution measurements [1–3].

Security design involves a variety of numerical concepts associated with control access, validation, approval, information honesty, and privacy components that are implemented by using modern cryptography techniques. It oversees security convention apparatuses and their applicability in current innovation. Encryption improves cryptographic items, for example, e-banking, e-exchanging, ATM cards, and email. Correspondence security, cloud security, and data security will intensely depend upon encryption algorithms. The investigation of private square algorithms dependent on selected frameworks is introduced in this paper. The primary objective is to survey the exhibition of the most generally utilized SKC algorithms as far as security, enigmatic adaptability, validation, reliability, strength, and versatility to distinguish calculation restrictions [4–6].

Cryptography and steganography are the notable and normally utilized strategies, which are used for controlling the information to encode or cover its reality. Steganography is the workmanship and study of imparting in a technique that disguises the correspondence's essence. Cryptography scrambles a correspondence with the end goal that it can't be decoded; Steganography covers the message so it's anything but apparent. Notwithstanding the way that the two advancements give security, research are gaining increasing attention in integrating cryptography and steganography into one framework for acheiving improved secrecy and security [7, 8].

The transmission of data by means of the web may include touchy individual information that can be blocked. Besides, numerous web applications and sites request clients to round out structures that contain delicate individual data. For example, telephone numbers, locations, and visa numbers. Clients may request private and secure correspondences for an assortment of reasons, for example, shielding privileged intel from programmers, while it is communicated over an open channel; henceforth, secrecy and information respectability are required to forestall unapproved access and utilization. Two algorithms are utilized in the picture steganography framework: one for implanting and one for extraction [9, 10]. The network security issues and the respective implementations have been discussed in [11–15]. These issues can be addressed by using the proposed methodology.

Figure 1 is depicting the various procedures involved in the process of data protection. Depending on the volume of data handled by different applications, we can choose any one of the method given above selected for securing the data. Cryptography will be chosen according to the needs of data protection and similarly the steganography is also used for average kind of application. But when an application belongs to be more important, it can be suitable to add both crypto and stegnography combined method.

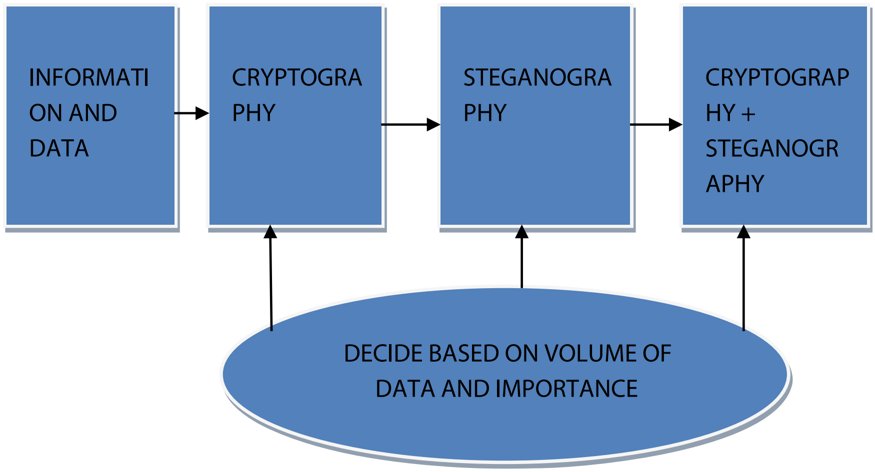


Fig. 1 Overview of concepts in data security

Methodology

A block chipher partitions the information into equivalent length squares and afterward encodes each square by utilizing a novel arrangement of numerical capacities known as key. This calculation encodes and translates information by utilizing a symmetric key instrument, which implies it and utilizes similar key on the two closures. Another benefit of the proposed approach is that it forestalls Brute-power attacks on the encryption text.

Further, the proficiency will be considered as far as by and large CPU clock and cycle time, just as force use, in every one of the three algorithms viable. The memory is worried about the measure of memory required for the whole encryption and unscrambling technique. Execution assessment on the above determined boundaries at the above chose symmetric will be displayed in the proposed research [1–3].

It's anything but a more customary way to deal with security than different finalists by selecting a more prominent security edge. As per the creator of snake, 16 emphasess are adequate against all known strategies for assault, yet it is reached out to 32 cycles as a defend against future cryptanalysis forward leaps. Snake much of the time adjusts imperceptibly to keep away from crash assaults. Twofish-this SKC is impervious to related key assaults, in any event, when utilized related to a slide assault and a connected key differential assault. It tends to be used to dispatch any connected key assault on the grounds that there are no frail keys. In the event that information move isn't secure, security perils to data are continually present, since programmers are ceaselessly hoping to take critical data for improved execution whenever progressed [4–6].

Cryptography and steganography are not exactly the same thing. For quite a while, information concealing procedures have been as often as possible used to send covered up secret messages. For PC clients, guaranteeing information security is a huge errand. Money managers, experts, and home clients all have delicate data they need to hold protected back from intrusive eyes. Despite the fact that the two innovations give security, it's anything but a decent practice to consolidate Cryptography and Steganography to make extra layers of security. The code text is then implanted in a picture or other media utilizing the stego key. At the point when these two strategies are consolidated, the information inserted will be safer [7, 8].

Unadulterated steganography is a method that simply utilizes the steganography approach, without the utilization of some other procedures. It is endeavoring to hide data inside the cover transporter. Secret key steganography is a sort of steganography that consolidates secret key cryptography and steganography. The expression "public key steganography" alludes to a blend of public key cryptography and steganography. The point behind this sort is to utilize a public key method to scramble the restricted information and afterward shroud the encoded information inside the cover transporter [9, 10].

Implementation

The proposed model has considered a small class room maintenance application with the help of NS2 simulator to analyze the cryptography and steganography. Figures 2 and 3 depict the comparision results of efficiency represented in terms of percentage.

Finally, this research work has checked the result of same class room application to the mixture of both techniques to prove combination will be the best one rather than implementing separae procedure. The result compared against mixture with cryptography and mixture with steganography. The result obtained from software is yielding best outcome when it is being supplied as single technology and the same comparison given in Figs. 2 and 3.

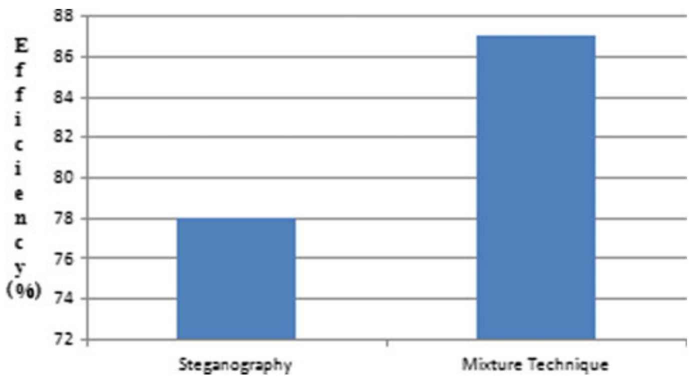


Fig. 2 Comparison of steganogrphy with mixture

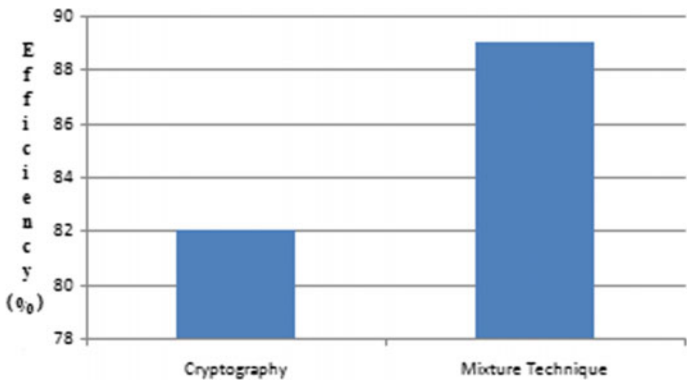


Fig. 3 Comparison of cryptography with mixture

Conclusion

The data pertaining to any application needs more privacy due to the sensitive issue in internet. Suppose this data is revealed or if it is known by illegitimate users, it will lead to drastic damages to the whole application. We need to give concrete mechanism, which will not allow leakages of data by any external or internal intruders are highly participated or used these applications. The strong protection against this kind of misbehavior in network can be prevented with the help of algorithms or any other procedures like cryptography, steganography, and mixture of it.

References

1. Gutmann P (2004) Cryptographic security architecture: design and verification. Springer
2. Verma S, Choubey R, Soni R (2012) An efficient developed new symmetric key cryptography algorithm for ınformation security. Int J Emerg Technol Adv Eng 2(7). ISSN 2250-2459
3. Chatterjee D, Nath J, Dasgupta S, Nath A (2011) A new symmetric key cryptography algorithm using extended MSA method: DJSA symmetric key algorithm. Published in 2011 ınternational conference on communication systems and network technologies, 978-0-7695-4437-3/11 $26.00 © IEEE
4. Pandey KK, Rangari V, Sinha SK (2013) An enhanced symmetric key cryptography algorithm to ımprove data security. Int J Comput Appl 74(20):29–33
5. Prajapat S et al (2014) Sparse approach for realizing AVK for symmetric key encryption. Int J Recent Devel Eng Technol 2(4):15–18
6. Mewada S, Sharma P, Gautam SS (2016) Classification of efficient symmetric key cryptography algorithms. (IJCSIS) Int J Comput Sci Inf Secur 14(2)
7. Bharti P, Soni R (2012) A new approach of data hiding in images using cryptography and steganography. Int J Comput Appl 58(18):1–5
8. Francis N (2015) Information security using cryptography and steganography. Int J Eng Res Technol (IJERT) 3(28). ISSN: 2278-0181. Published by, www.ijert.org NSRCL-2015 Conference proceedings, Special Issue

9. Petitcolas FAP et al (1999) Information hiding—a survey. Proceedings of the IEEE, Special Issue on Protection of Multimedia Content 87(7):1062–1078
10. Saleh ME, Aly AA, Omara FA (2016) Data security using cryptography and steganography techniques. (IJACSA) Int J Adv Comput Sci Appl 7(6)
11. Suthir S, Janakiraman S (2018) SNT algorithm and DCS protocols coalesced a contemporary hasty file sharing with network coding influence. J Eng Res 6(3):54–69
12. Suthir S, Janakiraman S (2018) A survey of fast file sharing system in network. Int J Eng Devel Res 5(2):1298–1304
13. Subbiah B, Sriram S, Manoharan R, Obaidat M, Suresh KC (2020) Selection of ıntermediate routes for secure data communication systems using graph theory application and grey wolf optimization algorithm in MANETs. In: IET networks, onlinefirst April 2020, https://ietresearch.onlinelibrary.wiley.com/doi/full/https://doi.org/10.1049/ntw2.12026
14. Balaji S, Sasilatha T (2018) An efficient routing approach for detection of syn flooding attacks in wireless sensor networks. EAI Endorsed Trans Energy Web Inf Technol 5(20), e7:1–6. https://doi.org/10.4108/eai.12-9-2018.155562. ISSN: 2032-944X.
15. Balaji S, Sasilatha T (2019) Detection of denial of service attacks by domination graph application in wireless sensor networks. Cluster Comput J Netw Softw Tools Appl 22(6):15121–15126

Segmentation and Sentiment Word Categorization Using Feature Extraction—A Novel ASFW Framework

S. Ashika Parvin and M. Sumathi

Abstract Sentiment analysis (SA) is essential for classifying people's thoughts about whatever they submit as reviews online. Because the content on these media is unstructured, the segmentation of the sentiment word, which is critical for detecting attitudes, must be done properly to overcome the problem of missing data, which can lead to erroneous criticism classifications and render the SA approach useless. This study provides a novel framework called ASFW for automatically segmenting the sentiment word in order to categorise the sentiment of "reviews". This framework contains a pre-processing technique, feature extraction with characteristics such as terms of presence and frequency (TPF), parts of speech (POS), opinion words and phrases (OWP), and negations, and word segmentation using the RBDT algorithm. Experiments show that our proposed framework ASFW is successful and efficient in segmenting the words necessary for sentiment classification without incurring data loss, with 92% accuracy and a time complexity of 0.0008 ms. Furthermore, with a time complexity of only 0.0006 ms, the classification of sentiment words according to the category has obtained excellent accuracy of 94%.

Keywords Sentiment analysis (SA) · Feature extraction · Terms of presence and frequency (TPF) · Parts of speech (POS) · Opinion words and phrases (OWP) · Negations

Introduction

Web 2.0s fast growth has resulted in users creating huge volumes of content online in the shape of views, blogs, tweets and other forms. This cornucopia of data contains consumer feedback on events, goods and people. It gives businesses and organisations new ways to better understand their customers, improve product quality and increase their competitiveness. The study of people's views, sentiments, opinions, perceptions and emotions based on social media data is known as sentiment analysis, or opinion

S. Ashika Parvin (✉) · M. Sumathi
Sri Meenakshi Government Arts College for Women(A), Madurai, India

V. Bindhu et al. (eds.), *Proceedings of Third International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 844,
https://doi.org/10.1007/978-981-16-8862-1_12

mining. These topics are more likely to be discussed in reviews. Examining product reviews on the Internet to determine how a thing makes you feel or what your overall impression is known as sentiment analysis in reviews. The goal of sentiment analysis is to find new points of view, figure out what emotions they represent, and then categorise them according to their polarity. Sentiment analysis is a method that may be used to classify data.

Opinion extraction, analysis, categorization, scalability and summarization are all aspects of an opinion mining system. The initial step in the opinion mining process is to extract all of the views once the user has loaded them all. Each new phase is dependent on the preceding one, and the outcome of the previous step is utilised as an entry in the following step. The first stage is opinion gathering, which lays the groundwork for all subsequent stages. If extraction proves to be ineffective, the method for mining opinions will get modified, and all the extracted characteristics will be not topic-related.

The first stage in sentiment classification is to pre-process the text; this approach turns unstructured Internet information with noise into a format that can be classified. Tokenization, stop word removal, lower case conversion and number removal are all part of the pre-processing procedure. The dataset was pre-processed in one of our earlier papers, which may be found here [1]. The sentiment term is then segmented using approaches like term presence and frequency (TPF), parts of speech (POS), opinion words and phrases (OWP), and negations. The RBDT technique is then used to prune the words necessary for sentiment categorization, with no missing of data. The segmented words are then categorised into any one of the five categories: interested, uninterested, sad, happy and angry. The results are then analysed for the accuracy of segmentation and categorization of sentiment word individually by the features mentioned above with the accuracy of the ASFW framework. Our proposed framework worked better in segmentation and categorization of sentiment words with high accuracy rates with no missing of data.

This study focuses on the problems that emerge while using SA techniques, and an innovative framework is proposed as a solution to the problems stated below. When assessing sentiments, data contains a lot of false information that has to be eliminated. The problem here is deciding which data should be preserved and which should be discarded. The answer might be that all of the data is useful, but just a few are crucial for classifying the clients' feelings. Rather than having all of the information and erroneously categorising the sentiments, just those facts that are irrelevant should be eliminated from classification. Due to a lack of information, inappropriate emotions may be expressed.

As a result, the major goal of this study is to ensure that data is not ignored throughout the categorization process in order to avoid misclassifications of sentiments. The ineffectiveness of the SA process can be attributed to a misunderstanding of emotions. To correct the error, ASFW framework has been proposed that pre-processes, segmenting the sentiment word from the features individually by gathering all of the sentiment words from the classification table and pruned using the RBDT algorithm with high weight assigned words to accurately classify the necessary sentiment word rather than having all of the words together. The comparative analysis of results are tabulated in the experimentation section that shows the accuracies of segmenting and categorising the sentiment words individually by the features with the accuracy of the ASFW framework.

The other sections of the study are organised as follows: secion 'Methodology for Literature Review' deals with existing literature approaches in the SA field, section 'Proposed Methodology' with suggested method, section 'Experimental Analysis' with observational data and section 'Conclusion' with the research's conclusion.

Methodology for Literature Review

This section provides an overview of relevant feature extraction work in sentiment analysis. We looked at over sixty articles and divided them into categories based on their major approaches and contributions. The methodology of the survey is divided into three categories: (1) methodology adopted for conducting the survey and (2) relevant works in the field of SA. This section presents the major feature extraction and modification procedures and approaches identified in the cited papers. Finally, (3) illustrates the motivation to choose this research area, and the answer has been justified.

Methodology of Conducting the Survey

The creation of a clear grasp of our aims should precede a deep dive into the literature. The goal of this study is to demonstrate why it is critical to properly divide the sentiment word. As a result, in section 'Conclusion', we have clearly framed the questions and described their repercussions in full. The following are the questions that have been formulated:

1. Why is word segmentation necessary?
2. What algorithms were utilised to do the word segmentation?
3. How effective are the results obtained through the use of algorithms?

We compiled a list of all connected publications that answered our research questions using appropriate journal articles. Conference proceedings, books, chapters, survey studies and other sorts of papers were examined as part of the study. The search terms were "segmentation in sentiment analysis", "word segmentation technique in sentiment analysis", "feature extraction in sentiment analysis", or a combination of these terms. Alternative additions include the following: (1) publication in peer-reviewed journals; (2) the text is written in English; (3) all articles published in journals in the previous 6 years (from 2016 to 2021).

Text Segmentation

Text segmentation is an important part of natural language processing. Depending on the level of granularity, the task might be described as segmenting a text into subject pieces or a statement into primitive discourse units. We concentrate on state-of-the-art feature extraction paradigms for sentiment analysis in this paper.

Feature Extraction: With the growth of social media such as social networks, blogs and presentations, there is a need to evaluate people's feelings. Sentiment analysis is a difficult undertaking since there is so much data available online. It is the study of people's feelings about something in terms of positive, negative, or neutral polarity. The difficulty is determining the meaning of the word, which might vary depending on the context. To determine sentiment polarity, feature extraction and identification are required. The purpose of this research [2] is to compare feature extraction approaches used in sentiment analysis. Using Cornell movie review datasets, UCI sentiment labelled datasets, and Stanford movie review datasets, [3] investigated the performance of feature extraction techniques TF-IDF (Term Frequency-Inverse Document Frequency) and Doc2vec (Document to Vector), effectively classifying the text into positive and negative polarities by using various pre-processing methods such as eliminating stop words and tokens. The classifiers are used to train and verify the features extracted from the text sentences using feature extraction techniques. Support Vector Machines, Logistic Regression, K-Nearest Neighbours, Decision Tree and Bernoulli Naive Bayes. People are more interested in frequent product features, often known as hot features. In text mining, association rule mining or frequent pattern mining is done using Apriori algorithm and is frequently utilised. In conventional information retrieval and text classification tasks, term frequency has long been deemed crucial. However, Pang-Lee discovered that term presence is more relevant than term frequency in sentiment analysis, that is, feature vectors with binary values indicating whether a word exists (value 1) or not (value 0).

In this work, [4] looked at the influence of two characteristics on the SS-Tweet dataset of sentiment analysis: TF-IDF word level and N-Gram. The performance of sentiment analysis using TF-IDF word level (Term Frequency-Inverse Document Frequency) features is 3–4 percent higher than using N-gram features. Analysis is performed using six classification algorithms (Decision Tree, Support Vector Machine, K-Nearest Neighbour, Random Forest, Logistic Regression and Naive Bayes) and F-score, accuracy, precision and recall performance. Kumar and Abirami [5] conducted an experimental research for several feature extraction or selection approaches accessible for o inion mining tasks in this work. There are four steps to this experimental research. First and foremost, data was gathered from publicly available sources. Second, utilising tools to extract words, pre-processing procedures are applied automatically (parts of speech). Third, the material is subjected to various feature selection or extraction procedures. Finally, an empirical investigation is conducted to examine sentiment polarity using various characteristics.

Chandrasekaran et al. [6]'s goal is to give an overview of current advancements in assessing multimodal attitudes (text, audio and video/image) that include human–machine interaction, as well as the problems that come with doing so. This paper presents a thorough examination of emotional datasets, feature extraction algorithms, data fusion approaches and the efficacy of various classification strategies. This study [7] discusses the efficacy of Web-based learning for pupils. Feedback is one of the most important aspects of any educational or learning system; it is useful to learning if it is handled correctly. İt is investigated how machine learning techniques such as Logistic Regression (LR), Support Vector Machine (SVM), Naive Bayes (NB), and Decision Tree (DT) may be used to Web-based learning, with a focus on sentiment in student feedback. Count Vector (CVr) or Bag of Words (BoW) and Term Frequency and Inverse Document Frequency (TF-IDF) Vector are two types of Feature Extraction Techniques (FETs) that is focused on.

Motivation and Justification

In the twenty-first century, social media sites can collect massive amounts of data. People communicate their thoughts and opinions with the rest of the world via social media platforms such as Twitter, Facebook, and others. In this competitive climate, companies utilise the same technique to understand their market position. Thoughts or assessments are used to communicate feelings, which helps businesses come up with a variety of more enticing offers. Unwanted records may also exist in the views, which must be pre-processed in a timely manner. Text segmentation or word segmentation is necessary after data pre-processing to identify the correct sentiment word, and only then can the mis-classification problem be rectified. As a consequence, I suggest that word segmentation has to be improved because it causes sentiment misclassifications.

An unique approach is presented in this study with the objective of resolving the research problem of missing crucial data while segmenting a word from a sentence. The recommended approach takes considerable care in data segmentation, guaranteeing that no data is lost. This framework considers all sorts of words, short and long, and employs the aforementioned features to generate all weighted words, which are then put through the RBDT algorithm [1] to get segmented sentiment words, which are then classified into one of five categories. The section III delves into the structure in further depth.

Proposed Methodology

The suggested sentiment word segmentation phases are depicted in Fig. 1, which are utilised to carry out the whole segmentation procedure. The pre-processing and word split phases have been completed in the author's earlier research works and

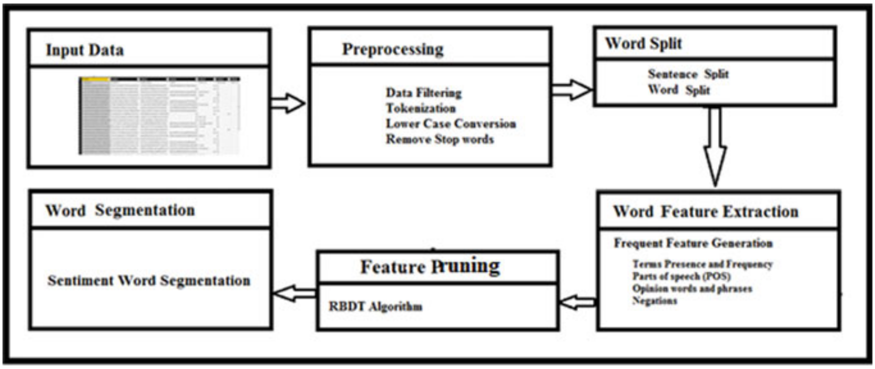
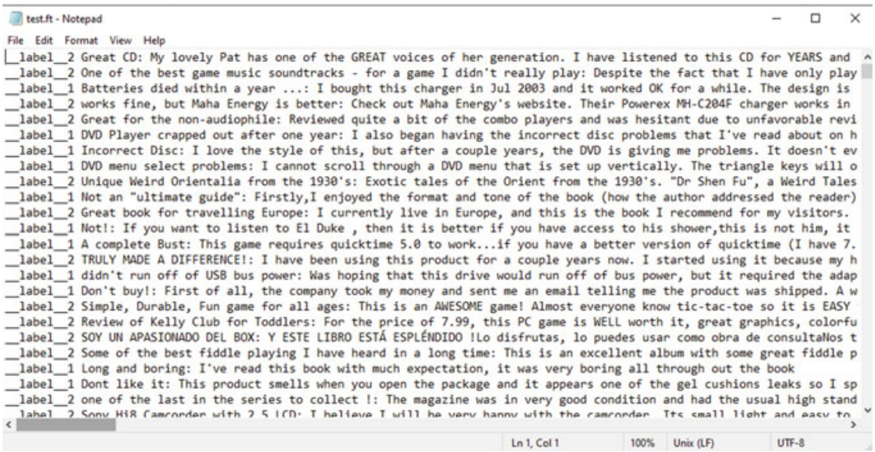


Fig. 1 Framework of the proposed sentiment word segmentation

the reference is cited in [1]. The proposed methodology starts with the word feature extraction phase.

Data Description

The sentiment word segmentation and categorization dataset was taken from Kaggle, an open source platform. It contains 10,000 records of Amazon reviews, as well as data for training and testing. The link to the webpage may be found here: https://www.kaggle.com/bittlingmayer/amazonreviews. A sample dataset before pre-processing is shown in Fig. 2.



```
__label__2 Great CD: My lovely Pat has one of the GREAT voices of her generation. I have listened to this CD for YEARS and
__label__2 One of the best game music soundtracks - for a game I didn't really play: Despite the fact that I have only play
__label__1 Batteries died within a year ...: I bought this charger in Jul 2003 and it worked OK for a while. The design is
__label__2 works fine, but Maha Energy is better: Check out Maha Energy's website. Their Powerex MH-C204F charger works in
__label__2 Great for the non-audiophile: Reviewed quite a bit of the combo players and was hesitant due to unfavorable revi
__label__1 DVD Player crapped out after one year: I also began having the incorrect disc problems that I've read about on h
__label__1 Incorrect Disc: I love the style of this, but after a couple years, the DVD is giving me problems. It doesn't ev
__label__1 DVD menu select problems: I cannot scroll through a DVD menu that is set up vertically. The triangle keys will o
__label__2 Unique Weird Orientalia from the 1930's: Exotic tales of the Orient from the 1930's. "Dr Shen Fu", a Weird Tales
__label__1 Not an "ultimate guide": Firstly,I enjoyed the format and tone of the book (how the author addressed the reader)
__label__2 Great book for travelling Europe: I currently live in Europe, and this is the book I recommend for my visitors.
__label__1 Not!: If you want to listen to El Duke , then it is better if you have access to his shower,this is not him, it
__label__1 A complete Bust: This game requires quicktime 5.0 to work...if you have a better version of quicktime (I have 7.
__label__2 TRULY MADE A DIFFERENCE!: I have been using this product for a couple years now. I started using it because my h
__label__1 didn't run off of USB bus power: Was hoping that this drive would run off of bus power, but it required the adap
__label__1 Don't buy!: First of all, the company took my money and sent me an email telling me the product was shipped. A w
__label__2 Simple, Durable, Fun game for all ages: This is an AWESOME game! Almost everyone know tic-tac-toe so it is EASY
__label__2 Review of Kelly Club for Toddlers: For the price of 7.99, this PC game is WELL worth it, great graphics, colorfu
__label__2 SOY UN APASIONADO DEL BOX: Y ESTE LIBRO ESTÁ ESPLÉNDIDO !Lo disfrutas, lo puedes usar como obra de consultaNos t
__label__2 Some of the best fiddle playing I have heard in a long time: This is an excellent album with some great fiddle p
__label__1 Long and boring: I've read this book with much expectation, it was very boring all through out the book
__label__1 Dont like it: This product smells when you open the package and it appears one of the gel cushions leaks so I sp
__label__2 one of the last in the series to collect !: The magazine was in very good condition and had the usual high stand
```

Fig. 2 Sample dataset of Amazon reviews

Word Feature Extraction

The following are the feature extraction techniques used in this paper explained briefly in the following sub-sections, and the outputs are displayed and discussed for each phase.

TPF: Feature-based sentiment analysis includes feature extraction, sentiment prediction, sentiment classification and optional summarization modules. Feature extraction deals with the prediction of product features that consumers have remarked on, sentiments related to the products are obtained by the sentiment prediction by assessing the polarity of the sentiments into either positive or negative, and finally, the work of compiling the results obtained from the previous two phases are done by summarization module. It is the most basic approach to express features, and it is widely used in both information retrieval and sentiment analysis. As features, it counts the frequency of individual sentences or a list of n sequential phrases in the shape of a n-grams [8]. In traditional text classification jobs, it is regarded as critical. However, term presence is more significant than term frequency in sentiment analysis since the presence of a single phrase can occasionally affect the polarity of the entire sentence. The words are either binary weighted or called frequency weights [9]. In addition, frequency weights are used to reflect a word's relative importance. This characteristic is used in our suggested technique to calculate both the existence and frequency of a word before assigning a weight to it. Similarly, the whole dataset was analysed, and the number of words counted by this characteristic was tabulated for future analysis. The input data is pre-processed and TPF feature is applied to the review. The terms that occur in high frequency which are important for the sentiment classification are being extracted and the output for this phase is shown in Fig. 3.

POS: To assess a phrase for sentiment properly, it must be split down into components utilising various sub-processes, including POS-tagging, as briefly shown here. Part of speech tagging is identifying the most important elements of a text, such as verbs, pronouns, adjectives and adverbs to retain sentence structure and make it obvious which part of speech the word belongs to. After the tokenization procedure, but before the removal of any words (stop words removal), the POS must be tagged. Finding descriptive words/adjectives in the material since they are key point of view identifiers. The meaning of subjective words must be captured in order to properly classify texts by emotion. Adjectives (words like "excellent", "horrible", "awful", "amazing" and so on) are often used to forecast mood. Subjective text analysis and SA research support this viewpoint. Nouns, verbs and adverbs are other open parts of speech (POS) categories that have been related to SA. Each word's syntactic role in a phrase has an impact on how it is interpreted. Components of speech are another name for syntactic roles. The verb, the noun, the pronoun, the adjective, the adverb, the preposition, the conjunction and the interjection are the eight parts of speech in English. Nouns and pronouns, for example, are typically devoid of sentiment in our proposed technique. When used as an adjective, distinguishing words that may be used in multiple parts of speech, for as "enhanced" as a verb, may have a distinct

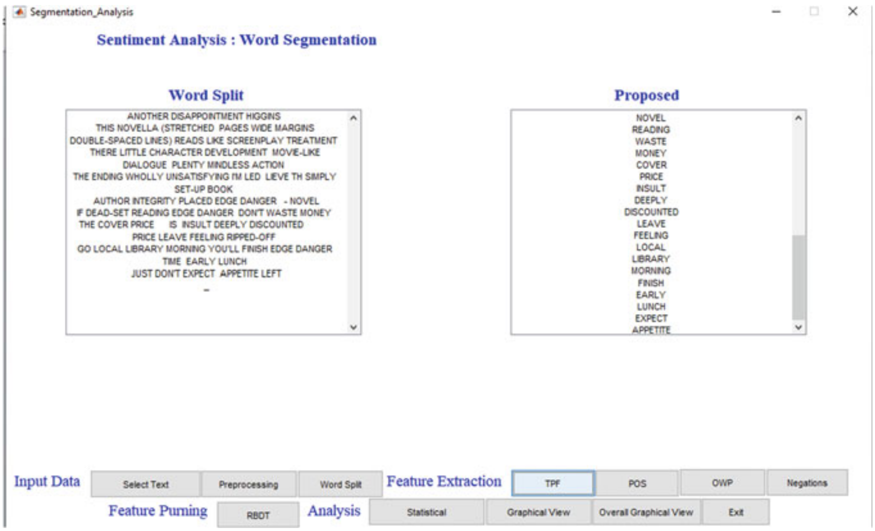


Fig. 3 Output of the TPF phase

amount of sentiment, thus these circumstances should be treated with caution. The final word count for this feature is recorded for future trimming. The words that are extracted from this phase after considering all the eight parts of speech is shown in Fig. 4.

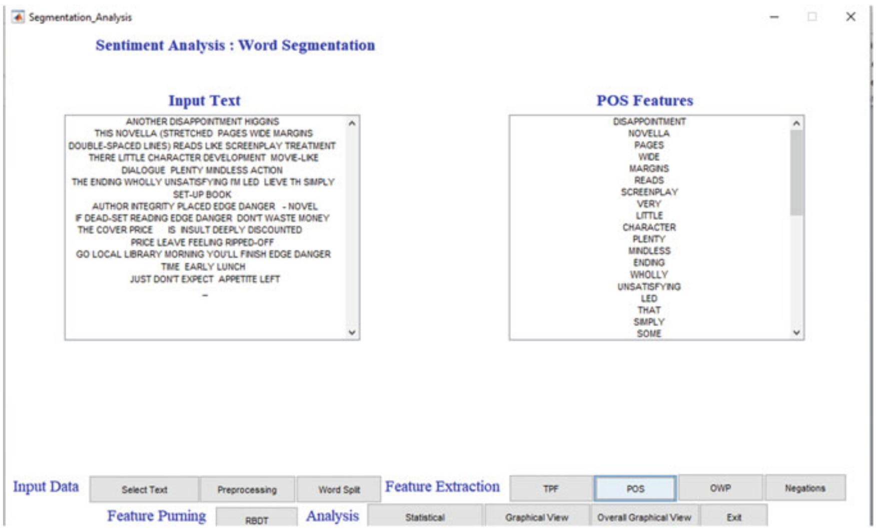


Fig. 4 Output of POS feature

OWP: Words and phrases that communicate good or negative feelings are known as opinion words (or sentiment words). Good, wonderful, outstanding, excellent and brilliant, for example, have a positive connotation, whereas terrible, dull, slow, worst and poor have a negative connotation. Though adjectives and adverbs make up the majority of opinion words, nouns and verbs can also express an opinion. In certain texts, words like garbage (noun), hate (verb) and like (verb) can express a point of view. Regular opinions might be expressed in a direct or indirect manner. For example, quality of the book is excellent. It directly refers to a feature of "book quality" in this case, implying a direct opinion. "My skin absolutely broke out after using the lotion". In this case, the entity "cream" is used to indicate a negative judgement of the attribute "skin". It's easier to digest direct opinions. To determine the polarity of indirect opinions, one must be familiar with the data source domain. Unlike regular opinions, they may express different opinions for same entity. There are two types of comparative opinions: explicit and implicit. Explicit comparisons are easy to analyse since they offer a single, positive or negative judgement. For example, Intel 5 has a faster CPU than Intel 3. The element of CPU speed is specifically contrasted here, and Intel 5 comes out on top. The term "implicit comparison" refers to an objective statement in which opinions are expressed in an indirect and oblique manner. For example, programme x takes longer to execute than programme y. In this case, a longer execution time indicates that programme x's performance is inferior to programme y's. As a result, an unfavourable view about programme x is conveyed. This feature collects all of the opinion words that are either explicitly or implicitly known in this suggested approach, and weights are assigned to these terms. The words that were counted based on weights are then tabulated for the trimming procedure. In this phase, before the stop words removal, this feature has been extracted and the output is shown in Fig. 5.

Negations: Negations are crucial in linguistics because they influence the polarity of other words. Words like no, not and should not are examples of negatives. When a negation appears in a sentence, it is critical to figure out which words are affected by this phrase. The scope of negation can be confined to the next word after the negation, or it can be expanded to include further words after the negation [10]. Negation processing affects virtually every context or domain, since neglecting negations can lead to erroneous implications or false interpretations [11]. Negation scopes are difficult to detect because they are latent, unobservable and extremely subjective, even among specialists. Anecdotal data shows that even for very basic phrases, this can result in divergences. In the sentence "this mobile is not nice but it functions correctly", for example, the scope of negative is only limited to the next word after negation. The scope of negation is extended to the conclusion of the phrase in another sentence, as in "the battery does not work for a long time". These examples show how the scope of negation varies based on linguistic features like as conjunctions, punctuation marks and the negation's part of speech (POS), among others. Furthermore, the presence of a negation phrase does not guarantee that the entire sentence's polarity-carrying words will be inverted. The two forms of negation are clear negation (with obvious indicators such as not, no, etc.) and implicit negation. At the highest

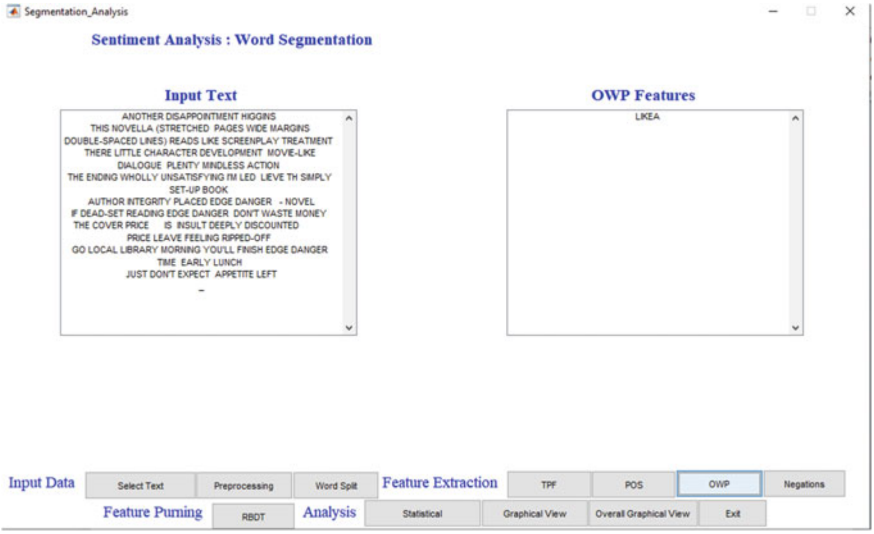


Fig. 5 Output of the OWP feature

structural level, there are two types of negations: morphological and syntactic negations. A suffix (e.g., ir-, non-, un-, etc.) or a prefix (e.g., ir-, non-, un-, etc.) modifies the root word (e.g. -less). In syntactic negations, explicit negation signals are used to flip a single word or a series of words. Negations may occur implicitly in complex settings, such as irony, without the presence of explicit words in the sentence. In this recommended technique, which analyses all types of negations, the words are gathered and tallied. The output of this phase is shown in Fig. 6.

Feature Pruning

Pruning the feature set for emotion categorization on a sentence-by-sentence basis may result in the loss of critical information for new instance classification and the omission of uncommon features (multiple occurrences). As a result, for accurate prediction, the model must be capable of identifying strong opinion sentences concerning the relationship between the features of each category, as well as adjust the weights of less frequently used features in order to expand the opinion indicators of sentiments with strong indications. For this aim, RBDT algorithm was introduced in our prior paper for efficient sentiment word segmentation.

RBDT algorithm: This algorithm has been proposed by this author which can be found here [1]. The main advantage of this algorithm is that it combines the advantages of rule-based (Apriori) and decision tree algorithms into a single algorithm rather than utilising them individually. There are no data holes since every sort of

Fig. 6 Output of the negations feature

word is considered, whether it is short or lengthy. The term is predicted with a high accuracy of 90% and minimal time complexity, and no data in the supplied input text is missing. The final phase in the ASFW framework described in this study is feature pruning, which is accomplished using this algorithm by gathering the amount of words extracted by feature extraction algorithms. With a high accuracy of 92% and a low time complexity of 0.0008 ms, the pruning is done in such a way that just sentiment words are segregated without any data loss. Further, they are being categorised into any one of the above-mentioned categories. The pseudocode for this framework has been stated below, and the sample output of this stage is shown in Fig. 7.

Experimental Analysis

Word Segmentation Accuracy

The words are gathered independently from all of the above-mentioned features like TPF, POS, OWP and negation. The algorithms then prunes these data using RBDT algorithm such that just the necessary sentiment words are identified, eliminating all other data. This procedure is carried out with great care to ensure that no data is overlooked. With a sample of 20 input text files, the experimental findings are reported in Fig. 8. Our suggested framework (ASFW) has obtained an excellent accuracies in segmenting the emotion words rather than the words obtained from

Input: Text Data (TD)

Output: Segmented Sentimental Words (SSW) are collected

Step 1: Input Text Data (TD)

Step2: RNV= Extra characters and Numerical values from the supplied text are removed.

Step3: RSW= RNV's stop words are removed.

Step4: SW= RSW's words are splitted

Step5: EF=**Proposed** Extract the Features from SW

Step6: SWE= Feature Pruning from EF based on RBDT method

Step7: SSW=Extract Segmented Sentimental words from SWE

Pseudocode of the Proposed Methodology

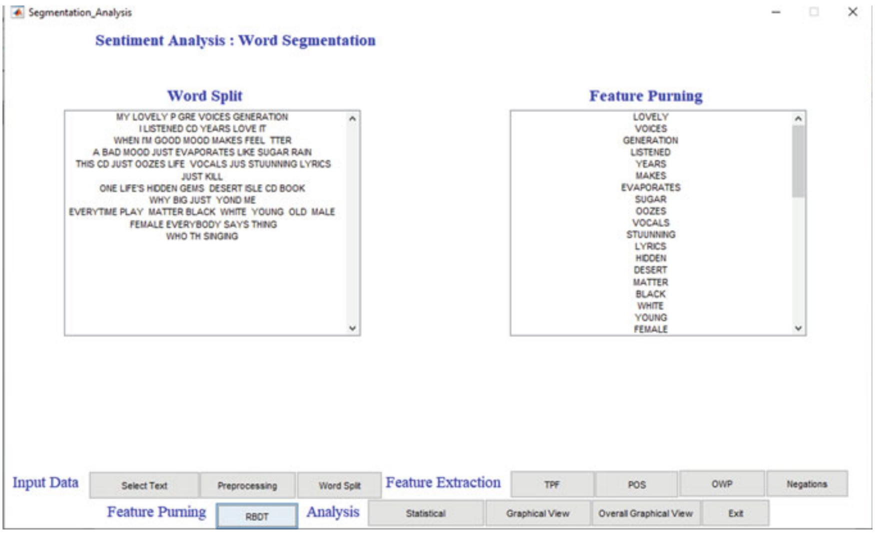


Fig. 7 Final output for the RBDT

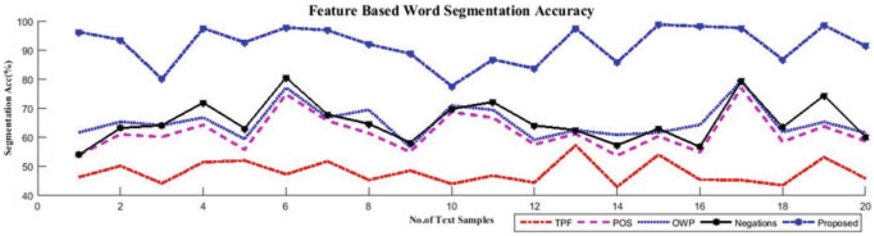


Fig. 8 Word segmentation accuracy rates in % using features

each features individually by achieving an accuracy of 92%. Finally, the segmented sentiment word is categorised into any one of the below mentioned categories.

Sentiment Word Categorization

Following the RBDT algorithm's sentiment word trimming, the sentiment words are classified into one of five categories: interested, not-interested, happy, sad or angry. Table 1 shows the % of the sentiment words extracted from each category by the features individually and through the ASFW framework which are summarised for only a sample of three input text files. But this categorization has been carried out to the entire dataset. This procedure is demonstrated for each of the above-mentioned groups. Table 1 shows that the suggested framework (ASFW) has achieved an excellent accuracies in sentiment word categorization.

Overall Performance

The overall accuracy rates in segmenting the sentiment word achieved by each features individually and through the ASFW framework with the time complexity and to achieve the word segmentation procedure for the whole dataset are given in Table 2. The findings reveal that the proposed framework ASFW segmented the sentiment word with high accuracy of 92% with a time complexity of just 0.0008 ms.

In terms of accuracy rates in percent, the overall classification of emotion words in each of the five categories by every features and through the ASFW framework is shown in Fig. 9. The results reveal that the category of sad emotion words has a high accuracy rate of 94%, whereas the category of not interested emotion words has a low accuracy rate of 89%.

Finally, the overall time complexity in categorising the sentiment words that are obtained by each features individually and through the framework of ASFW is given in Table 3. The findings reveal that the category of angry and interested has a

Table 1 Percent accuracy of sentiment word categorization

Samples	TPF	POS	OWP	Negations	ASFW
Interested					
Text 1	46.15	53.85	61.54	53.85	96.15
Text 2	51.61	65.59	66.67	67.74	96.77
Text 3	45.16	61.29	69.35	64.52	91.94
Happy					
Text 1	50.00	60.87	65.22	63.04	93.48
Text 2	51.28	64.10	66.67	71.79	97.44
Text 3	43.82	68.54	70.79	69.66	77.53
Not interested					
Text 1	44.00	60.00	64.00	64.00	80.00
Text 2	44.26	57.38	59.02	63.93	83.61
Text 3	57.14	61.04	62.34	62.34	97.40
Sad					
Text 1	51.85	55.56	59.26	62.96	92.59
Text 2	47.13	74.71	77.01	80.46	97.70
Text 3	53.85	60.26	61.54	62.82	98.72
Anger					
Text 1	46.67	66.67	69.33	72.00	86.67
Text 2	53.03	63.64	65.15	74.24	98.48
Text 3	44.62	70.77	78.46	73.85	89.23

Table 2 Overall accuracy rates in % and time complexity (ms) of sentiment word segmentation

Measurement	TPF	POS	OWP	Negations	ASFW
Word segmentation (%)	47.90	61.86	65.76	66.09	91.96
Time complexity (ms)	0.0012	0.0177	1.0793	0.0416	0.0008

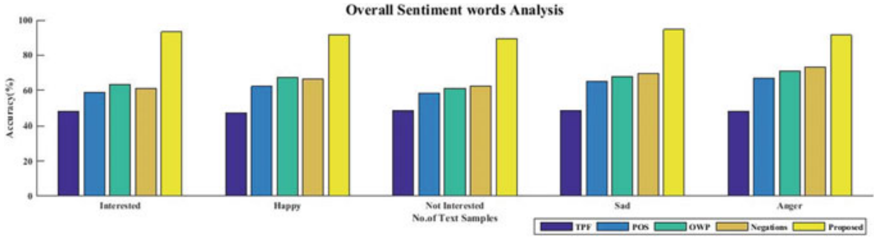


Fig. 9 Overall sentiment word accuracy rates in %

Table 3 Overall time complexity for sentiment word categorization in ms

Sentiments	TPF	POS	OWP	Negations	Proposed
Interested	0.0008	0.0156	0.8931	0.0348	*0.0006*
Happy	0.0008	0.0197	0.7396	0.0573	0.0008
Not interested	0.0009	0.0173	1.1301	0.0478	0.0010
Sad	0.0024	0.0172	1.6576	0.0326	0.0007
Anger	0.0009	0.0176	1.0549	0.0168	*0.0006*

low time complexity of 0.0006 ms, and the category of not-interested has a high time complexity of 0.0010 ms in the proposed framework.

Conclusion

The segmentation process is the difficult task in the sentiment analysis which is an open challenge for the upcoming researchers can still contribute to this research area. To overcome our research problem, a unique framework has been proposed for the purpose of both sentiment word segmentation and classifying the sentiment words according to their category. There are five categories used in this paper where the experiment results show that the proposed framework outperforms in all the aspects in segmenting the sentiment word effectively rather than using the feature extraction methods individually. Then, the classification of sentiment words is also done effectively in such a way that no data is missed in the given input text files. Experiments show that our proposed techniques are successful and efficient in segmenting the words necessary for sentiment classification without incurring data loss, with 92% accuracy and a time complexity of 0.0008 ms. Furthermore, with a time complexity of only 0.0006 ms, the categorization of sentiment words obtained excellent accuracy of 94%.

References

1. Sumathi M, Parvin SA (2021) Nuances of data pre-processing and its impact on business. In: 5th international conference on intelligent computing and control systems (ICICCS), pp 1006–1012. https://doi.org/10.1109/ICICCS51141.2021.9432376
2. Pasarate S, Shedge R (2016) Comparative study of feature extraction techniques used in sentiment analysis. In: International conference on innovation and challenges in cyber security (ICICCS-INBUSH), pp 182–186. https://doi.org/10.1109/ICICCS.2016.7542328
3. Avinash M, Sivasankar E (2019) A study of feature extraction techniques for sentiment analysis. arXiv:1906.01573v1 [cs.CL] 4 June
4. Ahuja R, Chug A, Kohli S, Gupta S, Ahuja P (2019) The impact of features extraction on the sentiment analysis. Proc Comput Sci 152:341–348. ISSN 1877-0509. https://doi.org/10.1016/j.procs.2019.05.008

5. Kumar A, Abirami S (2015) An experimental study of feature extraction techniques in opinion mining. Int J Soft Comput Artif Intel Appl (IJSCAI) 4(1)
6. Chandrasekaran G, Nguyen TN, Jude Hemanth D (2021) Multimodal sentimental analysis for social media applications: a comprehensive review. WIRE's Data Mining Knowl Discov. https://doi.org/10.1002/widm.1415
7. Tamrakar L, Shrivastava P, Ghosh SM (2021) Student sentiment analysis using classification with feature extraction techniques. arXiv:2102.05439
8. Birjali M, Kasri M, Beni-Hssane A (2021) A comprehensive survey on sentiment analysis: approaches, challenges and trends. Knowl Syst 226. ISSN 0950-7051. https://doi.org/10.1016/j.knosys.2021.107134
9. Kaushik C, Mishra A (2014) Comparative analysis of sentiment analysis techniques. ITSI Trans Electric Electron Eng (ITSI-TEEE) 2(1). ISSN (PRINT): 2320-8945
10. Farooq U, Mansoor H, Nongaillard A, Ouzrout Y, Qadir MA (2017) Negation handling in sentiment analysis at sentence level. J Comput 12(5). https://doi.org/10.17706/jcp.12.5.470-478
11. Pröllochs N, Feuerriegel S, Lutz B, Neumann D (2020) Negation scope detection for sentiment analysis: a reinforcement learning framework for replicating human interpretations. Inf Sci 536:205–221. ISSN 0020-0255. https://doi.org/10.1016/j.ins.2020.05.022

Emotion Recognition Using Deep Learning in Pandemic with Real-time Email Alert

Arnab Dey and Kousik Dasgupta

Abstract Emotion recognition is considered as one of the most vital and challenging studies in the domain of computer vision. Due to the ongoing pandemic the people's emotional and mental state is getting affected due to less physical interaction and increased virtual interaction. The motivation of the work lies in the recognition and monitoring of the emotional state of the human beings in live environment with higher accuracy rate and fast recognition time so as to keep them aware about their emotional state in the pandemic situation. Emotion recognition is achieved using the proposed deep convolution neural network (DCNN) using custom Gabor filter with 85.8% accuracy. Emotion recognition is also being investigated and compared with other deep learning models such as AlexNet, VGG-Net for testing the efficacy. When the emotion of the person will be found sad or angry then an alert sound is played and an email alert is automatically sent.

Keywords Emotion recognition · Deep learning · Gabor filter · Pandemic · Real-time alert · DCNN · Preprocessing

Introduction

Emotions have a vital role in the daily life of the people as it directly affects reasoning, decisions, attention, prosperity and the quality. People express their feelings and communicate with one another non-verbally through their facial expressions. It is found that the people rely more on non-verbal communications more than verbal ones [1]. Presently, due to the ongoing COVID-19 pandemic situation it is observed that there is a huge influence of computer devices on the life of the people and nowadays they mostly depend on it from online learning to remotely working. Due to this it is seen that there is a huge interest in improving the interaction or the communication between the computer devices and the human beings. The recognition of the emotional state is the process of identification of the facial expressions

A. Dey (✉) · K. Dasgupta
Department of Computer Science & Engineering, Kalyani Government Engineering College, Kalyani, West Bengal, India

V. Bindhu et al. (eds.), *Proceedings of Third International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 844,
https://doi.org/10.1007/978-981-16-8862-1_13

of the people. The emotional state of the people gets changed with the movement of the facial muscles which connect to the skin and fascia in the face. Researchers have found that the right part of the brain which is also called the posterior superior temporal sulcus (pSTS) allows human beings to recognize the emotions from the expressions of face. In 1971, the paper given by Ekman et al. [2] identified basic human universal emotions such as happiness, anger, sadness, fear, surprise. People express their emotions through the help of various facial expressions just by changing their facial muscles [3, 4]. It is investigated that the neural networks which are based on deep learning approach have achieved significantly good outcomes in classification and prediction [5–8]. The target of this study is to recognize the emotional state of people in pandemic through webcam accurately with real-time email alert using deep learning. Understanding people's emotion [9] mainly during the pandemic situation and suggesting them to overcome from their present state is important to keep the person happy, tension free and stress free. The emotion recognition is achieved utilizing the deep convolution neural network (DCNN). The novelty of the work is in the implementation of the proposed DCNN model with the Gabor filter and the real-time alert function. The proposed approach of emotion recognition with the DCNN mainly utilizes the Sobel filter and the Gabor filter for extracting the facial features.

Literature Survey

Computer vision is a field of study which is under artificial intelligence that teaches the computer system to understand, visualize, analyze and learn useful data from the various images or videos. It tends to achieve an automatic visual understanding [1, 10]. The earliest experiment on emotion recognition was conducted in the year 1978 by [2]. In this approach, facial action coding system (FACS) methodology for the identification and classification of human mood was proposed. From this work it was observed that each expressions of the face are a mixture of a number of action units that is in FACS. Here, the muscles movements of the face are compared to a specific emotional state based on the expression of the face. The target of the study [3] is to institute a method that is based on emotion recognition in predicting the students understanding in the entire distant learning process. Here, they have used Haar cascade based method for face detection and the method of Sobel edge detection is utilized for getting the characteristic value. In study [11], video-based mood recognition is being proposed with the convolution neural network (CNN), recurrent neural network (RNN) and C3D hybrid networks and the model classifies the emotional state with 59.02% accuracy rate without the utilization of any additional emotion labeled video clippings in the train sample. The study [12] utilizes the approach of cascading-based neural network model for identifying the emotional state of the people. The authors have found that with utilization of the multi-layered neural network model the emotional state of human beings are recognized well than the single layered network model on KDEF dataset. Arriaga et al. [13] have implemented a simple mini xception CNN model framework for classifying the facial expressions and gender of the people using the FER-2013 database. Another study proposes a modular multi-channel Expression-Net model which is dependent on CNN. The size of the neural network module gets reduced maintaining the accuracy of network when the ReduceV2 module is used [14]. The authors in the study [15] have experimented and observed that edges in the image represent discriminatory knowledge due to which their utilization helps to train the CNN neural network and improve the performance of the classifier. The proposed model has two-tower based CNN architecture and it also has an additional tower that is termed as the edge-tower, and it utilizes the edge images as inputs to the neural network. Lee et al. [16] used EmotionNet nano model. In study [7], facial expressions are classified using deep learning-based frameworks. The framework utilizes Gabor filters for extracting the features of the human face and classification is performed by convolution neural network. It is observed that the work proposed using Gabor filter requires less amount of training duration than the normal convolution-based approach. The speed of learning using the CNN has increased considerably as the Gabor filter extracts the image sub-feature more accurately than other feature extraction methods generally utilized by the CNN model. A comparative study is shown in Table 1.

Outline

The entire methodology adopted for recognizing the emotional state of the human beings using proposed DCNN model is highlighted in Fig. 1.

At first the input image is received, then image preprocessing techniques are applied, after that the face is detected using some algorithm, then the facial features are extracted and finally emotion is classified as the five emotions as shown using the trained model. The face is detected using Haar cascade classifier, and the facial features are extracted using various kinds of filters such as Sobel filter, Gabor filter using the proposed DCNN model. The main problem statement of this work lies in the prediction of the emotional state of the people in real-time and categorize into five emotional states which are neutral, happy, surprise, angry and sad accurately with real-time alert mechanism mainly during pandemic situation. It is a multi-class classification problem. The problem is a mixture of various sub-problems. The sub-problems include robust detection of the face of the people, extraction of the important facial features, predicting the emotional state of the person in real-time environment. And finally send alert to the user and authorities when the emotion will be found sad or angry.

Input: 48×48 facial image (webcam)

Output: Emotion of the person recognized into 5 (five) expressions of the face.

Table 1 Comparative study of various existing emotion recognition approaches

S. No.	Existing work/title	Key method	Classifier	Dataset	Accuracy (%)
1	Emotion recognition with deep learning [17]	–	CNN	CMU	78
2	Emotion recognition with cascade of neural net [12]	Cascade model	Neural network	KDEF	76
3	Emotion recognition model based on Haar cascade [3]	Haar cascade, sobel edge	Neural network	JAFFE	70
4	Emotion recognition and characterization	HOG, DWT	SVM	CK+	81–85
5	FER based on multi layer perceptron	Natural data division	MLP	Self made	76
6	Video-based mood recognition [11]	CNN-RNN C3D hybrid model	–	–	59.02
7	Emotion and gender classification in real-time based on CNN [13]	Gender and emotion, relu activ	Mini xception, CNN	FER2013	66
8	Facial expression recognition based on CNN [14]	ReduceV2 module, nine layers	Expression-net, CNN	FER2013	68.4
9	Mood recognition using machine learning [18]	Facial landmarks points	KNN, gradient boosting	Mixed Set	81
10	Facial expression detection using deep learning [6]	–	CNN	1. FER2013 2. JAFFE	1.70.14 2.98.65
11	Facial expression detection with less training time [7]	Gabor filter	Simple CNN	JAFFE	97
12	Predicting children's behavior using hybrid learning model [8]	Emotional features	Hybrid model-NB, DT	–	85 (approx.)

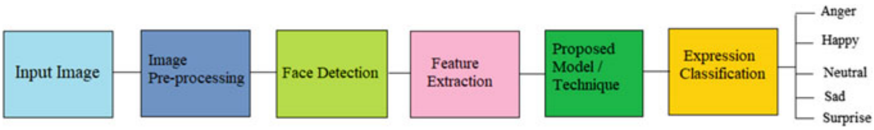


Fig. 1 Proposed methodology of emotion recognition with proposed DCNN model

The problem is based on supervised learning approach. The system will detect the face then it will understand the characteristics (i.e., the features) and predict the emotion using the proposed DCNN model within some fraction of seconds.

Model Implementation

The implementation of the model using the DCNN-based approach is discussed here. The deep learning-based approach is found as the most suitable method [19] for solving high level predictions [20] that are very complex in terms of prediction and solution and is normally utilized.

Dataset Used

The number of images samples taken for performing the emotion recognition using proposed DCNN with Gabor filter study is shown Fig. 2 with a bar graph.

In this study the selected mixture of the four datasets namely, Extended Cohn-Kanade [21], FER-2013 [22], JAFFE [23] dataset and self prepared data consisting of 16,500 images in total out of which 13,200 images are reserved for training and 3300 images are reserved for testing the proposed model. The images in the combined datasets are resized into same dimension and channel. The combined final database consists of grayscale facial images of size 48 by 48. The database taken for this study

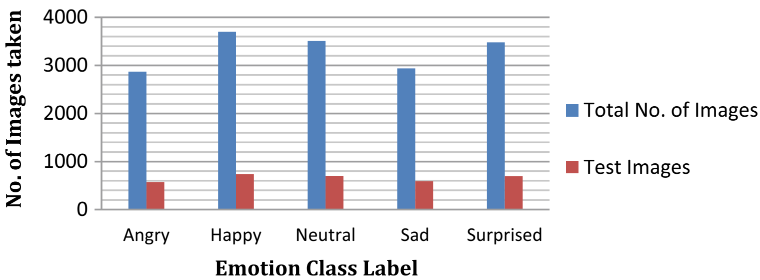


Fig. 2 Graph of total number of images taken in each expression class for this study

is being divided into 80% that is kept for training the proposed model and the rest 20% is kept for testing purpose. The 80:20 split is found to be the best and efficient split by experimental analysis.

Proposed Method

The flowchart of the emotion learning process utilizing the proposed DCNN is highlighted in Fig. 3.

It is seen from the flowchart that at first the entire dataset utilized for performing the study is divided into 80:20 ratios. Then the data augmentation technique is applied using the image data generator function. Data augmentation is a method which is utilized for creating newer train data from the existing train data artificially. This technique is utilized before feeding the images to the neural network [24]. The data augmentation parameters utilized for performing this study are rotation, zoom, brightness, rotation, horizontal flip, shift etc. Then a sequential model and the layers of the proposed DCNN model are defined. The Sobel and Gabor filter are mainly utilized for extracting the facial features such as eyebrows, mouth and eyes which are

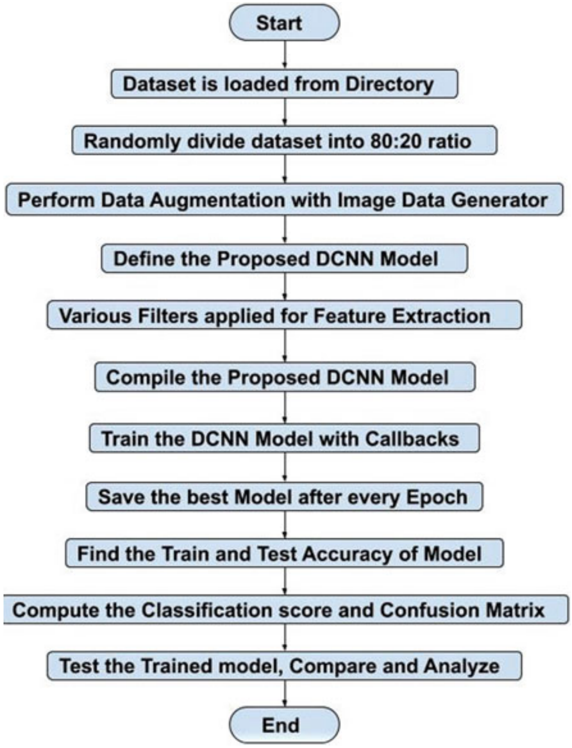


Fig. 3 Flowchart of the proposed DCNN model for learning emotions

useful for recognizing the emotions. The proposed model is trained with an initial learning rate of 0.001 and accuracy as metrics. Various callback functions have also been used in the training so that the model can decrease the learning rate when after 3–4 steps there is no improvement in accuracy, restore the best weights and save the best model after the completion of each epoch. The callback functions are used to automate the training process. The performance of the proposed DCNN model is being tested with various metrics.

In real-time implementation at first the input data comes from the webcam of the computer or laptop then the input image is converted into grayscale and the Contrast Limited Adaptive Histogram Equalization (CLAHE) [25] is being applied. Then utilizing the Haar classifier [26] the face is detected and a bounding box highlighting the facial region is formed, then the image is resized to the same size as that of the size of the images in the dataset, then this image goes to the trained DCNN model utilizing which the emotional state of the person is being predicted on the Region of Interest or ROI. Then the emotional state of the person that is matching the most is shown with percentage level. When the emotion of the person will be found sad or angry then an alert sound will be played and an email alert is sent.

Proposed DCNN Architecture

The layered architecture of the proposed deep convolution neural network (DCNN) is highlighted in Fig. 4. Firstly, a sequential DCNN model is created. Then various layers are added in the sequential model.

The proposed model comprises of six convolution layers, three max pooling layers and four dense layers. Firstly, the input image is being received in the input layer and after passing it through various hidden layers and it finally returns a vector with five emotion categories or classes as the output. The network structure is gradually made deeper and the number of filters are gradually increased which is observed to learn better in the proposed work. The ELU activation is utilized. The presence of four dense layers helps in better learning. The batch normalization and dropout techniques are utilized to make the model more powerful, generalized and free from over-fitting. Max Pooling is utilized for creating pooled feature map. The last layer is dense layer

Fig. 4 Proposed DCNN model architecture based on VGG-16 with ELU activation

which is often said as the Softmax layer which returns the probability scores of the five emotional states that is responsible for facial expression classification.

Classification

In the beginning of the CNN layers the low level features are being extracted and at the last layer high level features are being extracted. After the extraction of the useful facial features the next step is to classify the expressions based on the learned data. The classification is performed by the model using Softmax function at the last output layer of the proposed DCNN model. We generally get the probability score with the most matching emotion class having the highest score. The emotional state of the people is being classified into 5 facial expression categories which are angry, surprise, sad, happy, neutral and surprise.

Face Detection

In this study, Haar cascade classifier and OpenCV is adopted for detecting the face from the image. Haar cascade approach [26] is a method which is based on machine learning where a huge number of positive and negative images are utilized for training the classifier. The Haar cascade classifier is being trained with 7000 positive images (i.e., images which contain the human face) and 3500 negative images (i.e., images which do not contain human face) for a good number of steps due to which it is considered as a very powerful face detection method. For getting accurate results and decreasing the computational complexity the facial images were converted in grayscale. The result of the face detection method is a window which is having coordinates x, y, w and h (in which w stands for the width and h stands for the height) which highlights the presence of the face in the image and the face is being detected.

Feature Extraction

The feature extraction process normally decreases the dimensionality of the input image without losing any vital information of the image. The main facial features for the target facial expression classification are the eyebrows, eyes and the mouth [27]. The feature extraction process in deep learning is composed of successive convolution layers and pooling layers with Elu as nonlinear activation function. It is observed that at intermediate layers mainly in the second and third layer of the model the edge features of the facial image are extracted removing the other background elements and at fifth layer the proposed model extracts mouth, eyes, eyebrows, etc. In the sixth

Fig. 5 Sobel edge filter operation of neutral expression of a face sample

layer, it can detect the parts of complete objects as a whole face i.e., it extracts the high level features. Various filters that are used to extract the features are described below:

A. Sobel Filter: It is one of the most used filters for edge detection that are utilized at present. It is considered as one of the most popular approach in feature extraction because it gives much better performance most of the edge filters. The edge detection is dependent on matrix-based operations for computing the regions of varied intensities of an image. In the part of the image where the gradient magnitude is high, the Sobel filter detects the edges observed in Fig. 5. Sobel method is effective in producing good edge maps. It is considered that the edges are the most primitive features of the face [15].

The Sobel edge detection operator comprises of a pair of convolution kernels. These kernels are being prepared in such a way that it can respond maximally to edges running both horizontally and vertically relative to the pixel grid, with one kernel for each perpendicular orientation. Here, ΔG denotes the absolute magnitude of gradient is shown in Eq. 1. When Sobel edge detection approach is utilized the image is processed both in X and Y directions separately first, and then they are added together to form a new version of the image that constitutes the summation of the X and Y edges of the image. It is used for extracting the edge features from the facial images.

$$\Delta G = \sqrt{{G_x}^2 + {G_y}^2} \tag{1}$$

B. Gabor Filter: It extracts the sub-features of the image and gives it to the proposed deep neural network. In this way, the proposed DCNN model receives a good number of sub-features and improves in extracting the facial features and thus improves the recognition rate of the facial expressions. The Gabor filter bank is used in this study as it is easy to derive various filters from one Gabor mother filter just by utilizing varied orientations and scales. The approach of feature extraction with Gabor filter is a good way to extract the subtle changes in facial expressions, since mammalian visual cortex simple cells can be simulated using Gabor function, which means that

Gabor transform is similar to perception of human visual system [7]. The complex equation of Gabor filter is shown in Eq. 2. The Gabor filter has a real and imaginary part as shown in Eq. 3 and 4 which represents the orthogonal directions. The two components taken together is a complex number. The resultant value of the computation is normally the weight of the filter in the (x, y) location.

Complex Part:

$$g(x, y; \lambda, \theta, \psi, \sigma, \gamma) = \exp\left(-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}\right)\exp\left(i\left(2\pi\frac{x'}{\lambda} + \psi\right)\right) \tag{2}$$

In above Eq. (2) the Real part is:

$$g(x, y; \lambda, \theta, \psi, \sigma, \gamma) = \exp\left(-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}\right)\cos\left(2\pi\frac{x'}{\lambda} + \psi\right) \tag{3}$$

And the Imaginary part is:

$$g(x, y; \lambda, \theta, \psi, \sigma, \gamma) = \exp\left(-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}\right)\sin\left(2\pi\frac{x'}{\lambda} + \psi\right) \tag{4}$$

where $x' = x\cos\theta + y\sin\theta$ and $y' = -x\sin\theta + y\cos\theta$.

Here, the Gabor filter bank is being utilized which is having various orientations and scales. The vital parameters of the Gabor filter are illustrated in Table 2.

The orientation and frequency representations of the Gabor filters are identical to that of the human visual system. It is seen that the output location of the Gabor filter is mostly in the eye, eyebrow and mouth part. A image from the dataset is visualized using Gabor filter as shown in Fig. 6. The Gabor filters ($\phi = \pi$) extract sensitively the characteristics such as eyes, eyebrows and mouth portion. They are robust and are not sensitive to the intensity of illumination. It is also found that the local energy

Table 2 The vital parameters of the Gabor filter equation

S. No.	Various parameters	Description
1	K = (x, y)	The size of Gabor Kernel
2	ψ	The phase offset of sinusoidal function
3	θ	The orientation of normal to parallel strips of the Gabor
4	σ	Std. deviation of Gaussian function utilized
5	γ	The spatial aspect ratio
6	λ	The wavelength of sinusoidal factor in the equation
7	KF	Kernel frequency

Fig. 6 Visualization of the Gabor filter with the orientation chosen as π and KF as 0.2

spectrum is being distributed mostly in the regions of eyebrows, the mouth and the eyes that vary greatly to represent various emotional states.

Results and Observations

The results of the proposed work are shown and discussed in detailed manner with comparison in this section. The proposed work is being implemented using Python language in Spyder, Anaconda platform.

Performance Measurement

The training and test accuracy curve of proposed DCNN model without utilizing custom Gabor filter is shown in Fig. 7a and proposed DCNN model after utilizing additional Gabor filter is illustrated in Fig. 7b. The maximum test accuracy rate achieved is 80.29% without the utilizing additional Gabor filter in the proposed DCNN model and achieved 85.8% improved test accuracy after using Gabor filter in the proposed DCNN model. It is observed that proposed DCNN model with Gabor filter banks performs much better in terms of emotion recognition and improves the model performance by extracting more useful features.

The performance matrix aids us to find the performance of the classifier (the emotion classifier) on test data whose actual values are known. The screenshot of the output of confusion matrix of the proposed DCNN model with Gabor filter is shown in Fig. 8. The diagonal elements in confusion matrix are rightly predicted class and all the non-diagonal elements are incorrect predictions.

It is seen from the graph shown in Fig. 9 that the proposed model has excellent precision score of 0.83 and excellent recall score of 0.85 and also an excellent F1-score of 0.84. All the individual expressions of face are classified well as observed

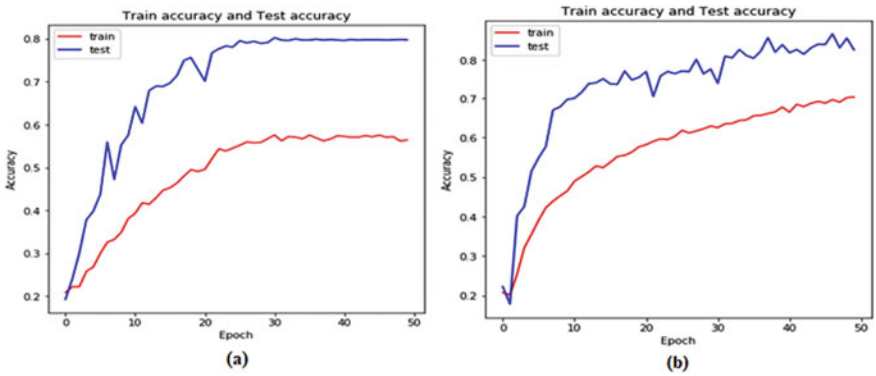


Fig. 7 Train and test accuracy of the DCNN model **a** without utilizing additional Gabor filter is shown **b** utilizing additional Gabor is shown

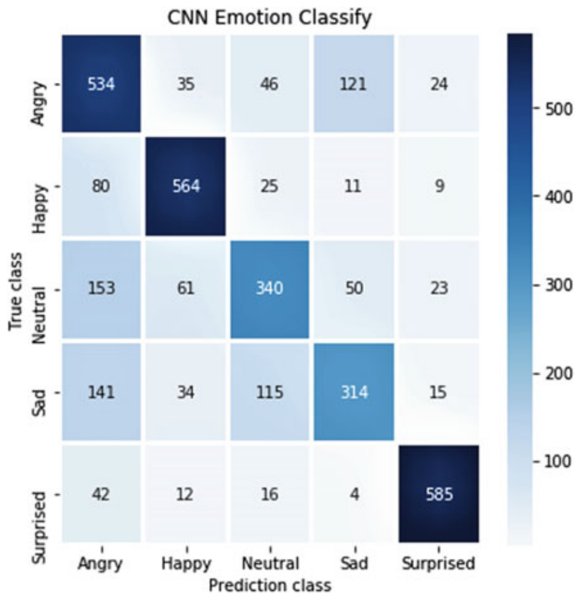


Fig. 8 Screenshot of confusion matrix of proposed DCNN model (with Gabor filter)

in from the classification score of proposed DCNN model with Gabor filter. Happy and surprise emotion class have the best classification score.

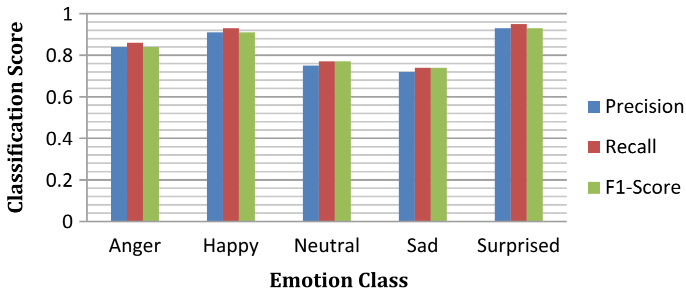


Fig. 9 Classification performance of the proposed DCNN model (with Gabor filter)

Investigation Results and Application

It is observed from Table 3 that the proposed DCNN model (with Gabor filter) and AlexNet model performed much better in terms of emotion recognition than all the other deep learning models being tested. But the proposed DCNN model with Gabor filter has proved to be more efficient and accurate in terms of prediction with test accuracy of 85.80% and has the highest Net F1 score.

It is found from Table 4 that the proposed model performed much better when the ELU activation function is being applied and it give satisfactory results. Tanh and ReLU activation function also performs well but not as accurate as ELU activation

Table 3 Compative investigation results with various deep learning models

S. No.	Model taken	Training accuracy (%)	Test accuracy (%)	F1-score (Net)
1	AlexNet	65.40	84	0.80
2	VGG-16	69	76.20	0.73
3	VGG-19	48.80	53.60	0.51
4	DNN	34	37	0.33
5	Proposed DCNN (without Gabor filter)	57.40	80.29	0.79
6	Proposed DCNN (with Gabor filter)	69.61	85.80	0.84

Table 4 Results of investigation with several activation functions in proposed DCNN

No. of epochs	Kernel/filter size	Learning rate	Activation used	Training accuracy (%)	Test accuracy (%)
50	3X3	0.001	ReLU	65	82.60
50	3X3	0.001	SeLU	57.20	78.40
50	3X3	0.001	ELU	69.61	85.80
50	3X3	0.001	Tanh	59.12	83.04

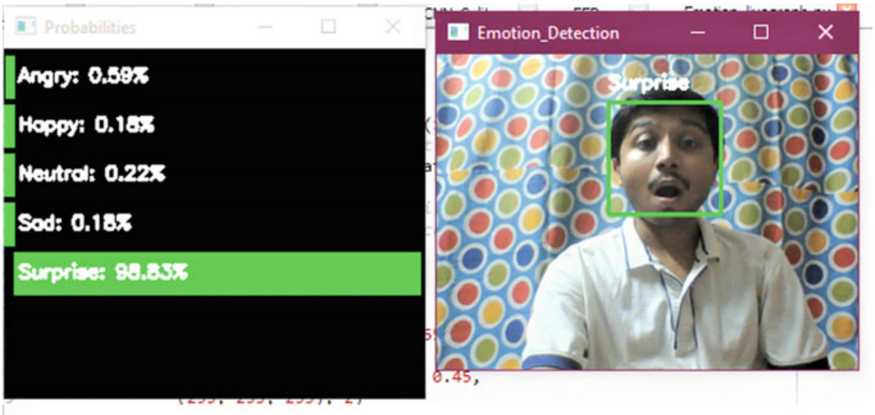


Fig. 10 Output of the emotion recognition with probability percentage is shown

function. The ELU is a powerful function which tends to converge the cost to zero faster and it has the ability to give more accurate results. Whereas, SeLU shows moderate performance.

The proposed DCNN model with the utilization of Gabor filter took 14.3 min duration to complete one epoch whereas the proposed DCNN model (without Gabor filter) took 15.8 min duration to complete one epoch. So, it is observed from experimental analysis that upon the utilization of Gabor filter the duration of each epoch takes lesser amount of time and hence learning becomes faster. It is found from various studies including the present study that increasing the number of hidden layers boosts the performance of the deep learning-based models. The deep learning-based emotion recognition approaches shows much better performance [20, 24, 28, 29] than the machine learning-based approaches. The proposed model has been used to develop a real-time application for recognizing the emotional state of the people with percentage level bar as seen in Fig. 10. Emotion recognition has numerous applications in monitoring drivers emotional state [30] for safe driving and in healthcare for monitoring the patient's mental state [31, 32].

Conclusion and Future Scope

The target of this work is to classify and predict the people's emotional state in five emotion categories utilizing deep learning approach in real-time with the best accuracy. In this study, the Contrast Limited Adaptive Histogram Equalization (CLAHE) is being utilized for detecting the emotions in real-time. The utilization of the Gabor filter in the proposed DCNN model extracts the features of the face more accurately. The notable accuracy rate of 85.8% is achieved with the proposed DCNN model with Gabor filter. It is found from the experimental results that proposed DCNN model with Gabor filter shows much better performance than the traditional models. The activation function ELU is found to be the most efficient. The proposed DCNN model with Gabor filter has excellent Net F1-score thus its classification score is high. It is also observed that the proposed model with Gabor filter is energy efficient as it takes less training time than the same model without the use of Gabor filter. When the emotion is found to be sad or angry then an alert sound is played in user end and email with details of user such as Geo-location, IP address is automatically sent. The accuracy achieved portrays that the proposed work can be utilized for recognizing the emotional state of the people in live environment mainly during the pandemic situation and inform authority on emergency. The future scope of this work will be to improve the performance of model by taking more data from the live sources. And examine the emotion classification performance utilizing transfer learning approach and also classify sentiment from the text.

References

1. Huang Y, Chen F, Lv S, Wang X (2019) Facial Expression recognition: a survey. Symmetry 11:1189. https://doi.org/10.3390/sym11101189
2. Ekman P, Friesen W (1978) Facial action coding system: a technique for the measurement of facial movement. Consulting Psychologists Press, Palo Alto
3. Yang D, Alsadoon A, Prasad PC, Singh AK, Elchouemi A (2018) An emotion recognition model based on facial recognition in virtual learning environment. Proc Comput Sci 125:2–10
4. Dols JMF, Russell JA (2017) The science of facial expression. Oxford University Press, Oxford
5. Xi Z, Niu Y, Chen J, Kan X, Liu H (2021) Facial expression recognition of industrial internet of things by parallel neural networks combining texture features. IEEE Trans Ind Inf 17(4):2784–2793
6. Jaiswal A, Krishnama Raju A, Deb S (2020) Facial Emotion detection using deep learning. In: 2020 ınternational conference for emerging technology (INCET), pp 1–5
7. Taghi Zadeh MM, Imani M, Majidi B (2019) Fast facial emotion recognition using convolutional neural networks and Gabor filters. In: 2019 5th conference on knowledge based engineering and ınnovation (KBEI), pp 577–581
8. Kumar TS (2021) Construction of hybrid deep learning model for predicting children behavior based on their emotional reaction. J Inf Technol 3(01):29–43
9. Smys S, Raj JS (2021) Analysis of deep learning techniques for early detection of depression on social media network-a comparative study. J Tends Comput Sci Smart Technol (TCSST) 3(01):24–39
10. Zhang F, Zhang T, Mao Q, Xu C (2020) Geometry guided pose-invariant facial expression recognition. IEEE Trans Image Process 29:4445–4460
11. Fan Y, Lu X, Li D, Liu Y (2016) Video-based emotion recognition using CNN-RNN and C3D hybrid networks. In: Proceedings of the 18th ACM international conference on multimodal interaction, pp 445–450
12. Kukla E, Nowak P (2015) Facial emotion recognition based on cascade of neural networks. In: Zgrzywa A, Choroś K, Siemiński A (eds) New research in multimedia and internet systems. Advances in intelligent systems and computing, vol 314. Springer, Cham
13. Arriaga O, Ploger PG, Valdenegro M (2017) Real-time convolutional neural networks for emotion and gender classification. arXiv:1710.07557v1 [cs.CV]
14. Yue Z, Yanyan F, Shangyou Z, Bing P (2019) Facial expression recognition based on convolutional neural network. In: 2019 IEEE 10th ınternational conference on software engineering and service science (ICSESS), pp 410–413

15. Bhandari A, Pal NR (2021) Can edges help convolution neural networks in emotion recognition? Neurocomputing 433:162–168
16. Lee JR, Wang L, Wong A (2021) Emotionnet nano: an efficient deep convolutional neural network design for real-time facial expression recognition. Front Artif Intell 105
17. Mehendale N (2020) Facial emotion recognition using convolutional neural networks (FERC). SN Appl Sci 2:446
18. Dey A, Dasgupta K (2021) Mood recognition in online sessions using machine learning in real-time. In: 2021 5th International conference on computer, communication and signal processing (ICCCSP), pp 1–6
19. Bursic S, Boccignone G, Ferrara A, D'Amelio A, Lanzarotti R (2020) Improving the accuracy of automatic facial expression recognition in speaking subjects with deep learning. Appl Sci 10:4002. https://doi.org/10.3390/app10114002
20. Barros P, Churamani N, Sciutti A (2020) The FaceChannel: a fast and furious deep neural network for facial expression recognition. SN Comput Sci 1:321
21. Lucey P, Cohn JF, Kanade T, Saragih J, Ambadar Z, Matthews I (2010) The extended Cohn-Kanade dataset (CK+): a complete dataset for action unit and emotion-specified expression. In: 2010 IEEE computer society conference on computer vision and pattern recognition—workshops, pp 94–101
22. FER Dataset [Online] https://datarepository.wolframcloud.com/resources/FER-2013. Accessed 30 Mar 2021
23. Lyons MJ, Kamachi M, Gyoba J (2020) Coding facial expressions with Gabor wavelets (IVC special issue). arXiv:2009.05938
24. Mungra D, Agrawal A, Sharma P et al (2020) PRATIT: a CNN-based emotion recognition system using histogram equalization and data augmentation. Multimed Tools Appl 79:2285–2307
25. Yadav G, Maheshwari S, Agarwal A (2014) Contrast limited adaptive histogram equalization based enhancement for real time video system. In: 2014 International Conference on Advances in Computing, Communications and Informatics (ICACCI), pp 2392–2397
26. Viola P, Jones M (2001) Rapid object detection using a boosted cascade of simple features. In: Proceedings of the 2001 IEEE computer society conference on computer vision and pattern recognition. CVPR 2001, pp I-I
27. Pantic M (2009) Facial expression recognition. In: Encyclopedia of biometrics (pp 400–406). Springer
28. Christou N, Kanojiya N (2019) Human facial expression recognition with convolution neural networks. In: Yang XS, Sherratt S, Dey N, Joshi A (eds) Third ınternational congress on information and communication technology. Advances in intelligent systems and computing, vol 797. Springer
29. Sajjad M, Zahir S, Ullah A et al (2020) Human behavior understanding in big multimedia data using CNN based facial expression recognition. Mobile Netw Appl 25:1611–1621
30. Jeong M, Ko BC (2018) Driver's facial expression recognition in real-time for safe driving. Sensors 18(12):4270
31. Fei Z, Yang E, Li DDU, Butler S, Ijomah W, Li X, Zhou H (2020) Deep convolution network based emotion analysis towards mental health care. Neurocomputing 388:212–227
32. Muhammad G, Alsulaiman M, Amin SU, Ghoneim A, Alhamid MF (2017) A facial-expression monitoring system for improved healthcare in smart cities. IEEE Access 5:10871–10881

A Detailed Study on State of Charge Estimation Methods

Rakhee Kallimani, Shweta Gulannavar, Krishna Pai, and Prachi Patil

Abstract Lithium-ion (Li-ion) batteries play a vital role in electric mobility, portable electronic devices, energy storage systems and power backups. In all these applications it is essential to estimate the state of charge (SOC) as it indicates the performance of the battery. Accurate estimation of the battery SOC serves many purposes like battery protection, preventing the battery from overcharging/over-discharging, improves the battery life, etc. This paper reviews the major SOC estimation methods concentrates on the advantages and identifies the drawbacks by surveying the numerous existing SOC estimation methods. In the end, the scopes and challenges of estimation methodologies are documented.

Keywords Lithium-ion battery · State of charge · Electric vehicle (EV) · Battery management system (BMS) · SOC estimation methods

Introduction

Li-ion batteries have their applications in significant sectors like electric vehicles, hybrid electric vehicles and energy storage systems, etc., therefore, BMS is required to detect the battery state and guarantee safety. Among all the algorithms within BMS, the SOC algorithm is the most complex one [1]. SOC estimation is the key issue in battery usage. The SOC of the battery is majorly influenced by battery current, voltage, temperature [2]. Other factors such as capacity, internal resistance and impedance also affect the SOC of a battery. The SOC is the indicator of battery performance. In addition to protecting the battery, the SOC estimation avoids over-charge and over-discharge and hence improves the battery life. There are several methods for SOC analysis, each method has positive and negative consequences.

R. Kallimani · S. Gulannavar · K. Pai (✉) · P. Patil
Department of Electrical and Electronics Engineering, KLE Dr. M. S. Sheshgiri College of Engineering & Technology, Belagavi, India

R. Kallimani
e-mail: rakhee.kallimani@klescet.ac.in

V. Bindhu et al. (eds.), *Proceedings of Third International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 844,
https://doi.org/10.1007/978-981-16-8862-1_14

The battery SOC is the ratio of the current capacity to the nominal capacity of the battery. Estimating the SOC accurately is a challenge for battery usage as it is the main parameter that indicates the battery performance. Many articles project a large variety of SOC estimation methods [3–6]. Based on the methodology, they are classified into 4 groups: Direct method, bookkeeping method, adaptive systems and hybrid methods as classified in [7]. Based on experimental and theoretical characteristics, SOC estimation methods are categorized into 3 groups: The traditional estimation algorithm formed on experiments, modern methods formed on control theory and other methods formed on innovative ideas as mentioned in [8]. The major five SOC estimation methods are conventional methods, adaptive filter algorithms, learning algorithms, nonlinear observers and other methods as in [9].

Direct measurements methods use physical parameters of the battery such as internal resistance and impedance as SOC indicators but these methods are not suitable in EV's [10]. The coulomb counting method is based on the integration of battery current concerning time to obtain the charge added or withdrawn from the battery. The coulomb counting method is preferred due to its simplicity and ease of implementation. The accuracy of the coulomb counting method is directly dependent on the precise measurement of battery current and accurate estimation of initial SOC [11]. This method also has problems such as initial SOC value errors and sensor errors [5]. In the open circuit voltage (OCV) method, SOC is estimated using the relationship between OCV and SOC which takes a long relaxation time for OCV measurement. Kalman filter-based and particle filter-based algorithms are used to achieve high estimation accuracy and adaptability. With the development in computer technologies, new methods like artificial neural network (ANN), support vector machine (SVM), Gaussian process regression and fuzzy logic are used for SOC estimation. Hybrid methods benefit from both data-driven and adaptive methods. For example, if OCV and neural network (NN) methods are combined, the results are more beneficial because the OCV method is used in the vehicle still state condition and the NN method is advantageous when the vehicle is in driving state. Using these two methods at different stages yields effective results [3]. Any algorithm can be used with the Kalman filter method to get fewer measurement errors during SOC estimation [3]. It is very important to choose a suitable algorithm to achieve high accuracy. The paper presents the various SOC estimation methods with their advantages and the evolution of SOC estimation methods is described.

As there is no unique method for the accurate estimation of SOC, hybrid methods (combination of two or more methods) are the most reliable methods, as they give accurate results by extracting the benefits from various methods.

For example, OCV is advantageous during vehicle still state, whereas NN is a more suitable in-vehicle driving state, therefore combining these two methods at different modes can result in more desired outputs arriving at the one reliable method is a major challenge, due to the drawbacks of each method.

The combination of OCV and CC method also give accurate results as per the requirements as OCV method takes care of the initial SOC value while the CC method gives its benefits during the driving state.

Literature Survey

The survey on the SOC estimation method primarily is divided into four categories based on the approach of estimation. In conventional methods, SOC is calculated by directly measuring the parameters such as internal resistance, OCV, current, impedance, etc. The equivalent circuit model is designed for estimating the internal resistance since it is the dominant parameter for the SOC estimation [12]. The internal resistance method is not ideal for EVs because the internal resistance of the battery does not differ much hence it is much more difficult to measure [6].

In the OCV method, SOC is estimated using the discharge curve which shows the OCV-SOC relation. This relationship is not the same for all batteries, hence the method is not feasible for many batteries mainly lithium ferro phosphate (LFP) batteries which have a flat OCV-SOC curve [6].

Coulomb counting is a simple method with less computation involved which makes the implementation easy. Here, SOC is calculated by integrating the battery current over time and the remaining capacity is estimated just by calculating the amount of capacity transferred from or to the battery [4, 6].

The enhanced coulomb counting method is simpler as internal resistance and terminal voltage are not involved [11]. The main drawback is the requirement of long time monitoring and memorizing hence requiring more memory. The estimation error of the enhanced coulomb counting method at the 6th cycle is 2.43% and it is 8.93% at the 21st cycle, therefore, we can say that estimation error increases with operating cycles. But the estimation error decreased to 1.08% at the 28th cycle after the correction of efficiency. The SOC of the battery is equal to the SOH of the battery when the battery is fully charged.

The simplified mechanistic model for $LiFePO_4$ battery can be easily expanded from single cell to battery pack [13]. As $LiFePO_4$ batteries have relatively flat OCV-SOC curves over the range of 20–80% SOC, correction of initial SOC is a challenge. Model parameter variation during battery ageing is the major challenge faced in the proposed model. After the detailed study of the article and analysing the results, it can be concluded that under different ambient temperatures the estimated SOC is accurate and cut-off voltage is predicted.

Adaptive Filter

Adaptive filter algorithms include various filters like Kalman filter (KF), dual Kalman filter (DKF), extended Kalman filter (EKF), unscented Kalman filter (UKF), etc. These filters are widely used because of their self-correcting nature. The combination of two extended Kalman filters are used in combination with the new OCV-SOC value for the estimation of SOC and capacity, respectively [14]. This method is more effective than other conventional methods as it is independent of battery conditions. The battery capacity changes for the new OCV-SOC are determined, hence capacity needs to be estimated. As per the estimation results, estimated SOC values match with the real SOC values even when the initial value error is in both smaller and larger than the real value and the results are within ±5% specifications. The estimated capacity also matches with the real capacity in both the conditions ($C_{ini} < C_{real}$ and $C_{ini} > C_{real}$).

The standard Kalman filter method is preferred in this article due to the presence of less computational effort and separation of variables is possible in this method as two filters are used [15]. This approach is demonstrated on nickel manganese cobalt (NMC) pouch cells. The estimation results indicate that the estimated SOC is more accurate with an error of less than 1% and the resistance is determined accurately.

Learning Algorithm

An estimation technique for SOC in Li-ion batteries based on Adaptive Neuro-Fuzzy Inference Systems (ANFIS) modelling of cell characteristics is presented [16]. The coulomb counting method is used for the comparison of results. The proposed method is simple and feasible as there is no requirement of measuring, and the analytical calculations are very easy. Although the methodology improves the SOC estimation accuracy with no added cost or complexity, deterioration of cell characteristics with time could be an issue. The estimation results show that SOC values estimated from this method are nearer to measured values when compared with the traditional coulomb counting method. The proposed method is more accurate than the coulomb counting method.

Hybrid Methods

Hybrid methods contain a combination of two or more methods for example data-driven coulomb counting method. The data-driven coulomb counting method is more effective as Incremental Capacity (IC) values are directly obtained without using a filtering process hence reducing the error [17]. The battery operating condition i.e. constant current-constant voltage (CC-CV) during charging mode and constant current (CC) during discharging mode employed in the suggested technique is rarely employed in practical applications, and the effect of sensor precision is overlooked in the data-driven coulomb counting technique. Results of the proposed novel method show that the SOC errors can be reduced from 25 to 0.05% and from 20 to 0.25% in discharging mode by calibrating the initial SOC. Thus online SOC estimation is done accurately in both charging and discharging mode.

State of Charge Estimation Methodology

The various SOC estimation methods of different Li-ion batteries like cobalt lithium, lithium manganate, lithium polymer, lithium iron phosphate which have their applications in electric vehicles are explored in [3]. After presenting various methods for SOC estimation with their benefits and drawbacks. The authors conclude that the combination of two or more estimation methods would give accurate results. But the precision of SOC estimation not only depends on the used method but also on the vehicle driving state, battery material, battery model, hence there is a requirement of both laboratory and field experiments.

The various SOC estimation methodologies with their advantages and disadvantages to ensure safe usage of the $LiFePO_4$ battery and reduce its average life cycle cost is presented in [5]. The existing SOC estimation methods are the Coulomb counting method (Ampere hour method), open circuit voltage method, impedance spectroscopy method, ANN-based method and model-based method. For online SOC estimation, nonlinear filters are used to compare the results. The merits of various estimation methods are: coulomb counting is an effective and simple method where SOC can be directly calculated but accurate initial SOC is needed for this method the drawback of this method is the presence of current sensor error. Computation is easier in open circuit voltage methods but OCV measurement needs more time for relaxation and temperature, age and battery types influence the OCV. The impedance spectroscopy-based method is sensitive to SOC variation but is difficult for online measurement. A model-based method is a robust and highly accurate method where diverse parameters indicate SOC but computation cost is high. An ANN-based method, implementation of hardware is easier after offline training and previous knowledge of batteries is not needed but this method is hard to generalize to various working conditions as large amounts of training samples are needed. For online SOC estimation, nonlinear filters are used to compare the results. The estimation results of SR-CDKF, SR-UKF, PF and H-infinity filters are better compared to other nonlinear filters. The absolute error graph indicates that the EKF, UKF, CDKF have larger estimation errors. This paper focuses on the online SOC estimation techniques which are divided into 5 groups and features of each method are explained.

The development of battery management systems for EVs with the SOC state estimation of different batteries like Li-ion batteries ($LiFePO_4$, $LiCoO_2$), NMC is focused on [6]. The existing SOC estimation methods are conventional method (Ampere hour counting method, open circuit voltage method, impedance and internal resistance method, electrochemical method and model-based method), adaptive filter algorithm (Kalman filter, extended Kalman filter, adaptive extended Kalman filter, fading Kalman filter, unscented Kalman filtering, sigma-point Kalman filter, particle filter and h-infinity filter), The learning algorithm (artificial neural network, support vector machine, extreme machine learning, genetic algorithm and fuzzy logic), observer (the nonlinear observer, proportional-integral observer and sliding mode observer) and hybrid algorithm method extended Kalman filter (EKF) Ampere hour

(Ah) algorithm, adaptive unscented Kalman filter (AUKF), support vector machine (SVM). Each method has its own merits and demerits like power computation cost are low in Ampere-hour method but its accuracy is not consistent. In the Kalman filter method, high variations are tolerated due to self-correcting nature but using this method it is not possible to predict the state of nonlinear systems. Artificial neural network methods are applicable in nonlinear conditions but require a huge memory. In hybrid methods, more methods are combined to get desired results. Each method has been explained in detail along with its merits and demerits. Accuracy of SOC estimation depends on the following factors: ambient temperature, self-discharge, c rate, battery ageing, cell unbalancing and battery dynamic hysteresis characteristics. This paper projects the evolution of battery SOC estimation methods.

As in Fig. 1, the voltage measurement method was invented by Curtis in 1963. SOC estimation method that involves the comparison of two batteries in which SOC of one battery is known was invented by Lemer in 1970. SOC estimation method using threshold in voltage levels was given by York in 1974. Voltage, current and temperature were measured by Brandwein in 1974. Christianson proposed an OCV method in 1975. Finger presented the coulomb counting method in 1975. Impedance was measured by Dowgiallo in 1975. OCV and voltage under load were developed by Eby in 1978. Kikuoka proposed a bookkeeping method in 1980. Finger proposed voltage relaxation time in 1981. Peled proposed a method using lookup tables that focuses on OCV and temperature computation in 1984. The impedance spectroscopy method was invented by Muramatsu in 1985. Koopmann proposed a method based on voltage, temperature and current relationship tables for the SOC determination in 1986. Bookkeeping and adaptive methods were proposed by Seyfang in 1988. Aylor presented OCV, OCV prediction and coulometric measurements in 1992. Gerard developed an artificial neural network method in 1997.

Salkind proposed the coulomb counting method, impedance spectroscopy and fuzzy logic in 1999. Bookkeeping, over potential, EMF and maximum capacity learning algorithm was developed by Bergveld in 2000. Garche proposed the Kalman filter in 2000. The Kalman filter and extended Kalman filter were proposed by Gregory in 2004. Sliding mode observer was invented by KIM in 2006. Jingyu yan developed the H-infinity filter in 2008 [6, 18].

SOC is a charge indicator of a battery in terms of its capacity. Accurate and precise estimation of SOC helps in protecting the battery from overcharging and over-discharging. Although it is not possible to measure the SOC directly, it can be estimated online or offline by direct measurement of parameters [19]. The existing SOC estimation methods are categorized into four groups. They are conventional methods, adaptive algorithms, learning algorithms and hybrid methods as shown in Fig. 2. All the methods are explained in brief below [5, 6, 9].

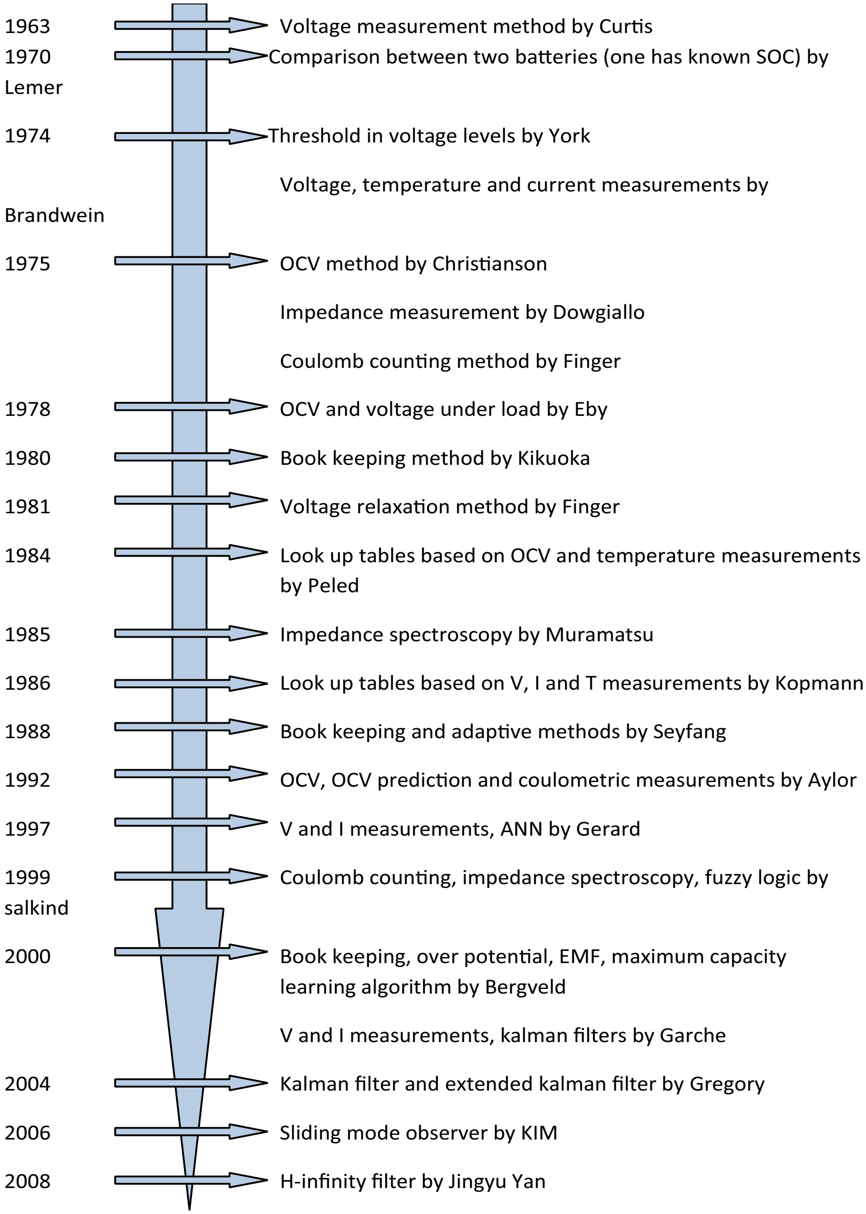


Fig. 1 Hierarchy of state of charge estimation methods

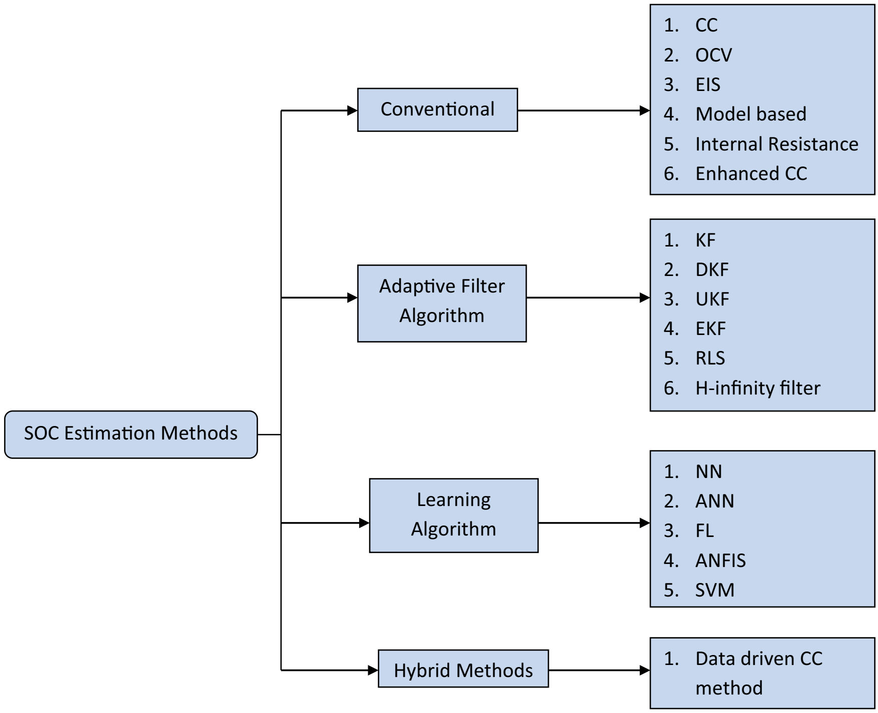


Fig. 2 Classification of state of charge estimation methods

Conventional Methods

Conventional methods consist of coulomb counting, the open circuit voltage, the electrochemical impedance spectroscopy, the model-based, internal resistance and the enhanced coulomb counting method.

Coulomb Counting Method: CC is the most common method used for the determination of battery SOC. It is a simple, direct and efficient method with less computation. Implementation of the coulomb counting method is easy as parameters are measured directly. In this method, battery current is measured using the current sensor and then measured current is integrated over time to get SOC. There is difficulty in getting the initial SOC value. In addition to that, the accuracy of the method depends on the sensor used which has measurement errors [9]. The SOC is calculated using the formula [5].

$$\mathrm{SOC}(t) = \mathrm{SOC}(t-1) \pm \left(\frac{1}{C}\right) \int i(t) * n\mathrm{d}t \tag{1}$$

where,

$\mathrm{SOC}(t-1)$	is SOC at previous time step,
n	is efficiency,
C	is capacity of a battery as $C = i * t$,
i	is charging /discharging current,
t	is charging /discharging time.

OCV Method: This method is in practice due to its comprehensive nature. The OCV method converts measured battery voltage into SOC value with the help of the OCV-SOC curve (discharge curve). The OCV-SOC curve is affected by temperature, age and battery chemistry. For example, lead-acid batteries have linear relationships, whereas LFP batteries have flat relationships. There is a requirement of a long relaxation time while measuring the OCV to attain equilibrium. The relaxation time of the Li-ion battery is approximately 10 h due to which the feasibility of the method is affected [7]. Hence this method is not applicable when the vehicles are at rest state rather than in a driving state.

Model-Based Methods: This is the most preferred method for online estimation as it is robust and accurate. The most commonly used models are electrochemical and equivalent circuit models. The function of the electrochemical model is to simulate the battery's electrical characteristics. The drawback of ECMs are some electrical characteristics cannot be represented using the circuit elements. The results of model-based methods are not dependent on initial SOC hence initial SOC error is reduced. Here accuracy of the method depends on modelling accuracy. This method is limited by its high computing cost. Figure 3 indicates the diagram of the model-based method. The measuring system measures the voltage, current and temperature of the battery which are useful to model the battery. The estimated voltage (battery voltage) is compared with the true voltage, the generated error is used by any of the model-based methods like adaptive filters, observers, etc. to estimate the SOC of the battery. The estimated SOC is compared with the predicted SOC, resulting error signal is used by the battery model.

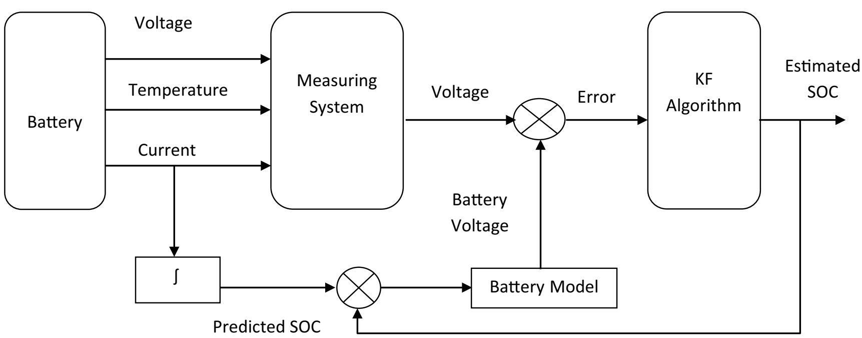


Fig. 3 Architectural diagram of a model-based method

Internal Resistance or Impedance Method: Internal resistance of the battery does not vary much hence it is difficult to observe it for SOC estimation. Because of this limitation, DC internal resistance is rarely used for SOC estimation. Online electrical impedance spectroscopy is hard to measure as it differs with battery type and experimental condition. The battery ageing and temperature are affecting the accuracy of measurement. Hence these two methods are not suitable for EV's.

Enhanced Coulomb Counting Method: To overcome the limitations of the coulomb counting method and to increase its accuracy enhanced coulomb counting method was proposed. This is a simple method. However, it requires a long time monitoring and memorizing hence requires more memory. An enhanced coulomb counting algorithm is used to estimate both the SOC and SOH of the battery. In this method, initial SOC is obtained using open circuit voltage or charging/discharging voltages. The losses that occurred can be minimized using charging and discharging efficiency. Due to its simple calculation and uncomplicated hardware, this method has its application in EVs and other portable devices.

Adaptive Filter Algorithm

Kalman Filter Method: Kalman filters are best suited for the state estimation of linear systems and it is used to estimate the dynamic state of the battery. Due to their self-correcting, high variations are tolerated during the running state. This method involves complex calculations. This method is applicable in automobiles, radar tracking, aerospace technologies and navigator tracking. The KF is limited by its complex calculations.

Extended Kalman Filter: This is the extension of Kalman filters where the linearization process is applied to estimate the SOC of nonlinear systems. But for the highly nonlinear systems, the accuracy reduces due to linearization errors. The relationship between OCV and SOC are assumed to be linear which matches with most of the batteries behaviour. This method is limited by high computation and complex hardware requirements. Figure 4 depicts the architecture of the extended Kalman filter method. The measured cell current is given to the model which outputs the predicted cell voltage and predicted system state. These two predicted values along with the measured cell voltage are used by the extended Kalman filter algorithms to give the corrected system state. The corrected system state is given as feedback to the model to enhance the prediction. With the help of the corrected system state, SOC is obtained.

Dual Kalman Filter: Dual Kalman filter facilitates the separation of variables. Its main advantage is less computational effort. The drawback of this method is slow convergence.

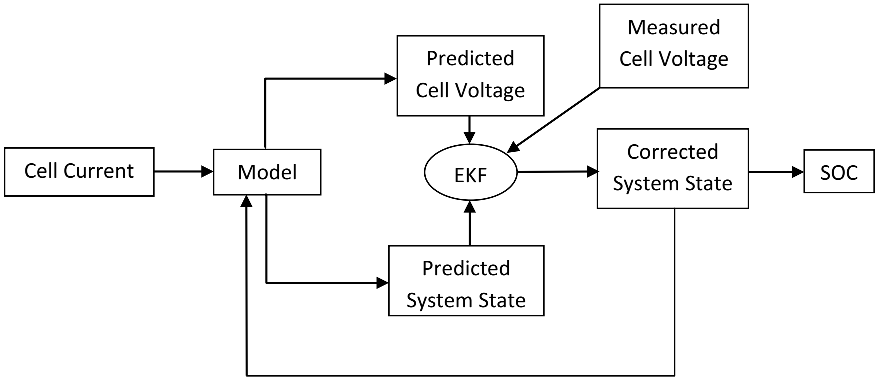


Fig. 4 Architectural diagram of extended Kalman filter

Unscented Kalman Filter: Unscented Kalman filter is the best state estimator for highly nonlinear systems. This is the improved version of EKF. Although the system is more accurate than EKF, the system is not highly robust due to the system disturbances and the uncertainty in the modelling. Jacobian matrix is not necessary.

Recursive Least Square (RLS): This method is easier to implement in online applications. The RLS algorithm is advantageous for the estimation of model parameters, which is carried out with the help of the forgetting factor. Model parameters are identified accurately here.

H-infinity Filter: This is the simple and robust approach where separate practical tests are not required for the measurement of parameters. But ageing, hysteresis and temperature are going to affect the accuracy of the H-infinity filter.

Learning Algorithm

Neural Network: Neural network is efficient for nonlinear systems when the battery is in charging/discharging condition. This method uses the nonlinear characteristics of the battery to estimate SOC. An accurate battery model is not necessary for analysis.

Artificial Neural Network (ANN): ANN can estimate the SOC without having prior knowledge about either the internal structure of the battery or the initial SOC. In this method terminal voltage, temperature, charging current/discharging current are the inputs and SOC is the output. A huge memory storage system is the need for the ANN, due to the requirement of large amounts of data.

Fuzzy Logic: SOC can be estimated without the help of an accurate model using fuzzy logic. Implementation of fuzzy logic is easier. Prior knowledge of the battery is needed. The major challenge in fuzzy logic is the requirement of more training and experimental data hence they need large storage memory. Fuzzy logic has a powerful ability to predict nonlinear models.

Adaptive Neuro-fuzzy Inference Systems (ANFIS): Adaptive Neuro-Fuzzy Inference Systems a simple, feasible and straightforward method for SOC estimation requires less computational effort. Deterioration of cell characteristics is a limitation of this method which can be solved by accommodating cell degradation mechanisms.

Support vector machine (SVM): SVM is a simple and elegant method whose performance is good in high dimension models. Parameters are adjusted using the trial and error method hence it requires more time. SVM is the method that uses the regression algorithm to transform a lower dimension model into a high dimensional linear model. This method is time-consuming.

Hybrid Methods

Hybrid methods are the combination of two or more algorithms to strengthen their advantages and overcome their limitations. Hence hybrid methods are more reliable and effective. Some of the hybrid methods used are the EKF-CC method, OCV-CC method, OCV-KF method and AUKF -SVM method.

All these methods are generally used in automotive applications such as electric vehicles and hybrid electric vehicles (Tables 1 and 2).

Scope and Challenges

The primary objective of the study is to realize the various SOC estimation methods for Li-ion batteries. SOC and SOH are the prime parameters that indicate the battery performance and in turn, it protects the battery and improves its life. This study which considers many batteries such as nickel manganese cobalt (NMC), $LiFePO_4$, lithium cobalt oxide ($LiCoO_2$), lithium nickel manganese cobalt (LiNMC), lithium iron phosphate, lithium polymer, lithium manganate has its applications in portable electronic devices, energy storage systems and automotive sector like electric vehicles, hybrid electric vehicles. Each estimation method has its extension and drawbacks because of its accuracy and feasibility.

1. The coulomb counting method is advantageous for the Li-ion batteries with high charging and discharging efficiencies. The accuracy of this method is not consistent due to the error sources such as unknown initial SOC, capacity fading, current sensor error and self-discharge rate. This method has a considerably high estimation error of about 11.4%. Reducing the error sources which in turn reduces the estimation error is the scope for further studies [6].

Table 1 Comparative study on state of charge estimation methods

No	Method	Category	Error
1	Coulomb counting method [6]	Conventional method	11.4%
2	Open circuit voltage [20]	Conventional method	(Max) 1.2%
3	Model-based method [20]	Conventional method	(Max) 4.327% Average 1.423%
4	Enhanced coulomb counting method [11]	Conventional method	2.43% (6th cycle), 8.93% (21st cycle), 1.08% (28thcycle)-after correcting with operating efficiency
5	Kalman filter method [9]	Adaptive filter algorithm	Less than 5%
6	Dual Kalman filter method [14]	Adaptive filter algorithm	<1%
7	Unscented Kalman filter [20]	Adaptive filter algorithm	(Max) <3.85%
8	Extended Kalman filter [20]	Adaptive filter algorithm	(Max) <4%
9	RLS [9]	Adaptive filter algorithm	(Max) 2.121%
10	H-infinity filter [9]	Adaptive filter algorithm	2.49%
11	Neural network [9]	Learning algorithm	(Max) <4%
12	ANN-based method [9]	Learning algorithm	(Max) <4%
13	Fuzzy logic [20]	Learning algorithm	(Max) <10%
14	Adaptive neuro-fuzzy inference systems (ANFIS) [20]	Learning algorithm	Mean <1%
15	SVM [21]	Learning algorithm	Mean <1% For dynamic SOC: (Max) +12% and −2%
16	Data-driven coulomb counting method [9]	Hybrid method	0.05% after calibration

2. OCV method is simple with high precision, however, its main drawback is the requirement of high rest time to attain equilibrium. OCV method is applicable when the vehicle is kept at parking, scope for future work is to extend this application in driving mode also. The challenge in this method is that the OCV -SOC curve is not the same for all batteries, because of which its applicability is restricted [9]. For example, lithium-ion batteries and LFP have a flat OCV -SOC curve which makes it difficult to obtain the OCV-SOC relationship.
3. Being a well-designed method, the Kalman filter method is used in many applications such as automobiles, radar tracking, aerospace technologies and navigator tracking [9]. For all linear systems, KF is the most applied method for state prediction, however, the designing and analysis of state prediction for a nonlinear system is beyond the scope of the KF method. Due to the self-correcting nature

Table 2 Attainments and limitations of various estimation methods

No.	Method	Attainments	Limitations
1	Coulomb counting method	1. Ease of computation 2. Measurement of the parameters directly makes it simple	1. Current sensor error is accumulated 2. Accurate initial SOC is needed
2	Open circuit voltage	1. Small computational burden 2. Easy to understand	1. Not applicable to the batteries with flat SOC-OCV relations 2. Long relaxation time is required
3	Model-based method	1. High accuracy and robust 2. Initial value error is reduced as this method is insensitive to initial SOC	1. Computational cost is high 2. The accuracy of the method is dependent on modelling accuracy
4	Enhanced coulomb counting method	1. Simple method	1. Long time monitoring and memorizing is required 2. More memory is needed
5	Kalman filter method	1. Self-correcting nature 2. KFs are insensitive to initial SOC error	1. Estimation errors are more 2. Not applicable to nonlinear systems
6	Dual Kalman filter method	1. Less computational effort 2. Easy variable separation	1. Slow convergence
7	Unscented Kalman filter	1. More accurate results 2. Applicable to nonlinear systems	2. Some parameters used are not chosen correctly
8	Extended KF	1. Implementation is easy	1. Linearization error may occur
9	RLS	1. Implementation is easy 2. Model parameters are identified accurately	1. Special battery tests are not needed
10	H-infinity filter	1. Separate practical applications are not required	1. Ageing, hysteresis and temperature are
11	Neural network	1. Efficient for nonlinear systems in charging/discharging conditions	1. Sensitive to amount and quality of training data
12	ANN-based method	1. It is easier to transplant hardware after offline training 2. Prior knowledge about the battery is not needed	1. A large amount of training data is needed 2. Hard to generalize to different working conditions 3. Large memory storage

(continued)

Table 2 (continued)

No.	Method	Attainments	Limitations
13	Fuzzy logic	1. Easy to implement as an accurate model of the battery is not needed	1. A large amount of training and experimental data is needed 2. Prior knowledge of the battery is needed
14	Adaptive neuro-fuzzy inference systems (ANFIS)	1. Calculations are understandable and implementation is easy 2. No extra sensors and measurements are needed	1. Deterioration of the cell characteristics
15	SVM	1. A simple and elegant method 2. Performance is good in high dimension models	1. Trial and error are needed for parameter adjustment hence this method is very time-consuming
16	Data-driven CC method	1. IC values are directly obtained without using the filters	1. The effect of sensor accuracy is ignored

of the KF method in running state their scope of work lies in running conditions but is limited to linear systems.

4. The model-based methods are relevant in different current loads and different ambient temperatures [13]. These models can be easily converted from cells to a battery pack hence these have potential application in EV's. The major drawback is its computational cost. Applying the electrical models on aged cells is the scope for future research, at present, it is applied only on fresh cells [22]. The models based on Particle Swarm Optimization (PSO) are used to identify the parameters [23]. The accuracy of model-based methods is dependent on the accuracy of their model.
5. Hybrid methods have a wide range of scopes because they are featured with the benefits of more methods which make estimation results more effective and reliable for example the combination of UKF and CC method gives results with errors within 5% [23]. Scope of hybrid methods includes the driving state of the vehicle, still state of the vehicle, linear systems and nonlinear systems, etc. Combining two to three methods is a hectic task hence they are limited by computational complexity and they require large storage memory [9].

Conclusion

The importance of battery SOC estimation is getting more attention gradually as it is not only indicating the charge left in the battery but is also used to estimate other parameters of a battery like RUL [24], hence researches are conducted to achieve high accuracy. This article provides the history of SOC estimation. A review of main SOC estimation methodologies with their features, benefits, drawbacks and estimation error is presented. As of now, there is no distinctive ideal solution for SOC estimation. Based on the working condition and the data available, the best combination must be selected. After the detailed analysis of estimation methods, it can be concluded that the most critical task in the process is to build a model that describes the exact internal structure of the battery. This article is limited by consideration of only 16 methodologies for estimating SOC for analysis and comparison. Although the various SOC estimation methods are presented along with their advantages, disadvantages, applications and estimation errors, the future work is directed towards the mathematical analysis, cost efficiency and complexity of the different approaches.

References

1. Gao DW (2015) Energy storage for sustainable microgrid. Academic Press
2. Ogura K, Kolhe ML (2017) Battery technologies for electric vehicles. In: Electric vehicles: prospects and challenges. Elsevier, pp 139–167
3. Zhou Y, Li X (2015) Overview of lithium-ion battery SOC estimation. In: 2015 IEEE international conference on information and automation. IEEE, pp 2454–2459
4. Murnane M, Ghazel A (2017) A closer look at state of charge (SOC) and state of health (SOH) estimation techniques for batteries. Analog Dev 2:426–436
5. Meng J, Ricco M, Luo G, Swierczynski M, Stroe DI, Stroe AI, Teodorescu R (2017) An overview and comparison of online implementable SOC estimation methods for lithium-ion battery. IEEE Trans Ind Appl 54(2):1583–1591
6. Zhang R, Xia B, Li B, Cao L, Lai Y, Zheng W, Wang W (2018) State of the art of lithium-ion battery SOC estimation for electrical vehicles. Energies 11(7):1820
7. Chang WY (2013) The state of charge estimating methods for battery: a review. International Scholarly Research Notices
8. Zhang M, Fan X (2020) Review on the state of charge estimation methods for electric vehicle battery. World Electric Veh J 11(1):23
9. Hannan MA, Lipu MH, Hussain A, Mohamed A (2017) A review of lithium-ion battery state of charge estimation and management system in electric vehicle applications: challenges and recommendations. Renew Sustain Energy Rev 78:834–854
10. Berecibar M, Gandiaga I, Villarreal I, Omar N, Van Mierlo J, Van den Bossche P (2016) Critical review of state of health estimation methods of Li-ion batteries for real applications. Renew Sustain Energy Rev 56:572–587
11. Ng KS, Moo CS, Chen YP, Hsieh YC (2009) Enhanced coulomb counting method for estimating state-of-charge and state-of-health of lithium-ion batteries. Appl Energy 86(9):1506–1511
12. Sun L, Li G, You F (2020) Combined internal resistance and state-of-charge estimation of lithium-ion battery based on extended state observer. Renew Sustain Energy Rev 131:109994
13. Li J, Lai Q, Wang L, Lyu C, Wang H (2016) A method for SOC estimation based on simplified mechanistic model for LiFePO4 battery. Energy 114:1266–1276
14. Lee SJ, Kim JH, Lee JM, Cho BH (2007) The state and parameter estimation of an Li-ion battery using a new OCV-SOC concept. In: 2007 IEEE power electronics specialists conference. IEEE, pp 2799–2803
15. Andre D, Appel C, Soczka-Guth T, Sauer DU (2013) Advanced mathematical methods of SOC and SOH estimation for lithium-ion batteries. J Power Sources 224:20–27
16. Awadallah MA, Venkatesh B (2016) Accuracy improvement of SOC estimation in lithium-ion batteries. J Energy Storage 6:95–104

17. Zhang S, Guo X, Dou X, Zhang X (2020) A data-driven coulomb counting method for state of charge calibration and estimation of lithium-ion battery. Sustain Energy Technol Assess 40:100752
18. Pop V, Bergveld HJ, Notten PHL, Regtien PP (2005) State-of-the-art of battery state-of-charge determination. Meas Sci Technol 16(12):R93
19. Wikipedia contributors (2021) State of charge. In: Wikipedia, The Free Encyclopedia. Retrieved on June 19, 2021, from https://en.wikipedia.org/w/index.php?title=State_of_charge&oldid=1000489172
20. Lu L, Han X, Li J, Hua J, Ouyang M (2013) A review on the key issues for lithium-ion battery management in electric vehicles. J Power Sources 226:272–288
21. Hansen T, Wang CJ (2005) Support vector based battery state of charge estimator. J Power Sources 141(2):351–358
22. Huang SC, Tseng KH, Liang JW, Chang CL, Pecht MG (2017) An online SOC and SOH estimation model for lithium-ion batteries. Energies 10(4):512
23. Vijayakumar T, Vinothkanna MR (2020) Efficient energy load distribution model using modified particle swarm optimization algorithm. J Artif Intell 2(04):226–231
24. Bhalaji N, Rimi C (2020) Remaining Useful Life (RUL) estimation of lead acid battery using bayesian approach. J Electric Eng Autom 2:25–34
25. Zhigang H, Dong C, Chaofeng P, Long C, Shaohua W (2016) State of charge estimation of power Li-ion batteries using a hybrid estimation algorithm based on UKF. Electrochim Acta 211:101–109

Understanding Partial Discharges and Its Role in Condition Monitoring of Insulators

Lekshmi Kaimal and Ramesh Kulkarni

Abstract Partial discharges (PDs), the local discharges occurring in an insulator subjected to various stresses during its operation, are an indicator of dielectric breakdown. The complex processes take place inside insulation before the occurrence of partial discharge (PD) in a dielectric. PDs cause irreversible damage to the insulator and consequently pose a danger to the life of the equipment. Their analysis requires a thorough understanding of underlying physics. A visual representation has been presented to aid the study of typical PD processes. These depictions clearly explain the atomic and molecular processes that take place in a gaseous dielectric. An attempt has been made to explain the differences in the process seen in a solid dielectric. The physio-chemical changes taking place in an insulator and the temporal changes seen in the pulses at the different stage of aging have been briefed. The parameters to be extracted for analysis, characterization, and localization of PD have been briefed. The different spectroscopy methods used for PD analysis by researchers have also been enumerated. The main contribution of this paper is the detailed explanation of PD with the help of illustrations for the ease of understanding of the complex PD phenomenon and its role in condition monitoring of insulators.

Keywords Condition monitoring of insulators · Gaseous dielectric · Solid dielectric · Partial discharge

Supported by Organization VES Institute of Technology, Mumbai, India.

L. Kaimal (✉)
VES Institute of Technology (Affiliated to University of Mumbai), Mumbai, India
e-mail: phd14.ajeshlekshmi@ves.ac.in

R. Kulkarni
Department of Electronics and Telecommunications, VES Institute of Technology (Affiliated to University of Mumbai), Mumbai, India
e-mail: ramesh.kulkarni@ves.ac.in

V. Bindhu et al. (eds.), *Proceedings of Third International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 844,
https://doi.org/10.1007/978-981-16-8862-1_15

Introduction

Most electrical failures in electrical or electronic engineering systems are caused by a breakdown in insulating materials, electrical aging, or PD [1]. Insulating polymers experience aging when subjected to various stresses like electric, dielectric, thermal, mechanical, chemical, and working environmental conditions like temperature, humidity, pressure, etc., over an extended period [2]. Aging produces an irreversible deterioration of physical, chemical, structural, dielectric, and other electrical properties. Initially, a local electric breakdown is seen at the site of the most stressed part of the insulator. This local breakdown causes more irreversible damage to the insulating material, which further generates more sites of stressed parts in the insulator. This vicious cycle of aging and breakdowns is rapidly resulting in permanent damage to the material, and finally to complete breakdown in insulation material.

It is necessary to analyze the impact of stresses on the characteristics of the insulator. Resistive and capacitive current flowing through the insulation material—both in a degraded as well as a virgin insulation—at different voltages and varying temperatures were investigated [3]. The capacitive, as well as the resistive current in an insulator, increases with applied voltage and temperature. The capacitor made up of the degraded insulation material behaves as a resistor rather than a capacitor when subjected to electrical and thermal stresses. This is because the insulation material undergoes irreversible chemical changes due to which its properties undergo changes. This phenomenon is verified even in a virgin insulator subjected to thermal stresses. It is noted that the low energy discharges can cause large damage even in a simple paper insulation in presence of external controlled stresses in a laboratory environment.

Partial discharge in electrical apparatus can be envisaged as a useful whistle-blower to the impending damages that can occur in it. The correct identification of the type of PD can help in correctly diagnosing the severity of damage and the extent and urgency of attention needed [4]. For e.g., the high amplitude corona PD in switch-gear boxes is to be treated differently by maintenance personnel, from the low amplitude surface PD in bushings or internal PD in the cable terminations. The internal PD in cable joints or termination will require immediate attention and surface PD on bushings demands only careful periodic checkup. The corona PD may not need any short-term action also. So, the identification of the type of PD and its characteristics is very important in electrical maintenance. The maintenance personnel with reasonable experience normally learn to identify the type of PD and its nature.

It thus becomes necessary to monitor the electrical equipment for any failure or breakdowns in check in real-time when subjected to external stresses. Thus, ‘condition-based monitoring’ or ‘condition maintenance’ should replace the ‘periodic maintenance’ practices applied to date [5]. These automated condition monitoring diagnostic systems can only be successful if there exists a way of capturing information that indicates the pre-breakdown phenomena. Hence, a study of PD characteristics to help in monitoring the equipment for any signs of PDs is essential and has been an active area of research [6–26]. The physical and the chemical processes of PD have been studied by several researchers [6–17]. The effects of various parameters like dielectric properties, applied voltage, and its frequency [8–13], the time lag for initiation, i.e., the statistical time lag [13, 14], cavity conditions, and cavity parameters [13–17] have been studied. The role of space charges [18–22] and environmental conditions like pressure, temperature [14, 16], etc., on PD have also been investigated. The PDs are random in nature. The PDs at different stages of insulation aging are observed to exhibit certain peculiar properties. This stochastic nature have also been investigated [23–26].

Modeling the PD requires thorough understanding the initiation, formation, and development of PD; the role of space charges; and the various stages of PD. Several researchers have implemented mathematical models and software simulation tools for better understanding and explanation of the influence of different parameters of PD, one such being [27]. The uniqueness of this simulator is that it can be used for detailed study and comprehensive analysis of PD without the use of a high-voltage (HV) setup. The simulator generates pulses of varying amplitudes, shapes, and time of occurrence of pulses observed in practice using National Instruments LabVIEW software as a base platform. The PD phenomenon is a complex physio-chemical process taking place at the atomic and molecular level inside the dielectric, not seen by a naked eye. It is difficult to visualize and understand the processes going on inside the dielectric. *Hence the objective of this paper is to present the complicated PD activity concisely. A small attempt has been made in this paper to use illustrations to the best of the ability of the author to describe and explain the PD process in a simple way. This will help the new researchers and the students working in the field of PD to understand and visualize the phenomenon effortlessly. An attempt has also been made in this paper to introduce recent research for the condition monitoring of the insulation material.*

The organization of this paper is as follows. A brief introduction to PD is included in Sect. 15.2 followed by a detailed explanation of the phenomenon and physics of PD in inside a gas-filled void in a solid dielectric in Sect. 15.3.2. The effects of PD on the insulation material, the different stages of aging, and the related physical and chemical changes are summarized. PD's effect on the pulse shapes and the related research work is covered in Sect. 15.4. The existing methods used for condition monitoring and analysis of PD and the related research papers are discussed in Sect. 15.5 followed by conclusion in Sect. 15.6.

Introduction to Partial Discharge

Partial discharge (PD), according to IEC standard 60270 [28], is defined as a localized electrical discharge that only partially bridges the insulation between conductors. In general, PDs can be classified into four types based on their discharge phenomena: internal discharges, surface discharges, corona discharges, and electrical treeing [5]. Figure 1 shows the diagrammatic representation of the types of discharges.

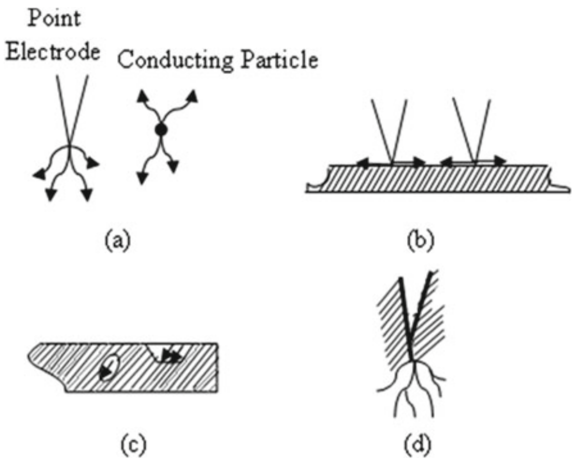


Fig. 1 Four types of partial discharges: **a** corona discharges, **b** surface discharges, **c** internal discharges, and **d** electrical treeing involving corona and internal discharges - Ref [1]

The localized discharges that take place in a gas or liquid insulation around the conductors that are away or remote from the solid insulation, in the presence of inhomogeneous fields are called corona discharges. When discharges are formed on the surface of the solid insulation not covered by the conductor or appearing at the boundary of different insulation materials, it is called surface discharges. Internal discharges are the ones that are formed in cavities or voids which lie inside the volume of the dielectric or at the edges of conducting inclusions in a solid or liquid insulating media. Electrical treeing may be considered a combination of corona and internal discharges caused due to the continuous impact of discharges in solid dielectrics forming discharge channels (treeing) [1].

The most common types of PDs seen in polymeric insulation systems are internal discharges in cavities (or voids) and the resulting electrical treeing. Such discharges are the most damaging as the solid insulation undergoes irreversible changes with each PD occurrence. This results in the insulation material losing its bulk properties such that the insulator loses its capacity to insulate in the course of its lifetime.

Discharges in cavities are almost unavoidable in the material due to imperfect manufacturing and handling processes. The voids are seen in virgin insulators as well. The process of formation of an insulation structure is a very complex process including selection and preparation of raw materials, thermal or chemical treatment, etc. The total process of fabrication of insulation material depends on man, materials, machines, and environmental conditions. Hence, it is practically not possible to manufacture perfect insulators [29]. During fabrication of the solid insulation materials, gaseous bubbles or cavities are formed in the structure [30]. Defects may also develop while the insulation material is in operation due to various stresses. The process of oxidation embrittlement which is prominent in warmer climates, in combination with mechanical and thermal stresses, lead to the formation of voids [31].

Nowadays, nano-sized inorganic particles are added to the epoxy matrix to improve its dielectric property. The breakdown strength and the resistance to PD of the epoxy matrix could be modified by the nano-sized particle with the proper type and content [32].

Understanding PD Process

As we have seen, PD is assumed to be a type of impulse which is confined to the area of defect. However, it does not lead to a complete electrical breakdown instantaneously. It is made up of electron avalanches which grow substantially to form streamer discharges. The PD phenomenon is presumed to fall into the category of gaseous discharge. Hence, the discharge growth is usually explained using the gas-phase description. The difference in the PD phenomenon seen in solid insulation is the modification of the local electric field within the void as compared to the surrounding dielectric. The influence of the secondary electron generation at the surface also affects the PD phenomenon. The charges produced by PD is also seen to impact the further PDs in a solid dielectric, giving rise to memory effects. Many researchers have studied the process that takes place during streamer [33–39]. The study of PD at an atomic and molecular level is essential, summarized in Sect. 15.3.1. Section 15.3.2 explains the process when the PD takes place in a gaseous medium surrounded by solid insulation [5, 12, 40, 41].

Process of PD at Atomic and Molecular Level

The basic processes taking place at the atomic and molecular levels are similar in the case of a gaseous as well as a solid dielectric. Here, the explanation is presented assuming the gas inside the cavity is air, and that the cavity is assumed to be as large as the dielectric for illustration purposes. Air inside void is made up of approximately 78% nitrogen and 20% oxygen as shown in Fig. 2a [40]. Gas molecules are moving constantly in a random direction. The force with which they hit the walls defines gas pressure inside the void, while twisting and stretching of molecules defines temperature. Molecular gas movement produces inter-molecular gaps. Initiation of a discharge can occur in these gaps. Initially in a virgin void, when the pressure is 65 kPa, the inter-molecular gaps are less. Thus, collisions will be more but force will be less. While in the case of an aged void having 2 kPa, molecular gaps are high. This means molecules are spread out, aiding more discharges and path for the discharge to move.

When the field is applied, electronic polarization takes place and the atom undergoes changes as shown in Fig. 2b. All the molecules align themselves in the direction of the field. It can be seen from Fig. 2c that the negative ions represented by darker portion of the atom are aligned toward the anode. A PD is said to occur when the voltage across the dielectric exceeds a certain threshold, calculated based on the pressure in the void, dimension of the void and the parameters that characterize the ionization processes of the gas in the void [27]. An additional criterion for PD occurrence is the presence of a free electron as shown in Fig. 2c. The various sources of starting electron include field emission due to electric field which starts at 100 MV/cm, temperature emission above 10,000 K, cosmic rays at 1 ionpar/cm$^3 s$, nuclear radiation like X rays and impurities. The presence of this free electron then starts ionization in

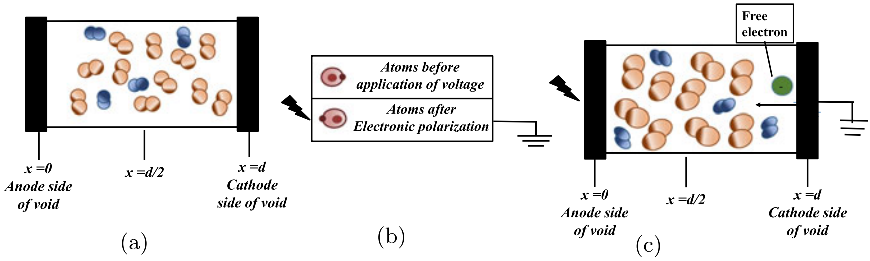


Fig. 2 Atomic representation of cross section of a solid dielectric—before and after field application, **a** gas molecules in random motion in a dielectric at rest. **b** Effect of applied voltage on atoms—electronic polarization. **c** Electronic polarized atoms and free electron release

the stressed parts of the insulator, which releases more electrons, leading to further ionization of the material.

When the free electron with charge e is accelerated in an electric field, E and x is the average inter-molecular distance when the PD occurs, energy supplied by an electric field is given by eEx. The electron having mass m acquires velocity, v under the electric field with kinetic energy (KE), $\frac{1}{2}\text{mv}^2$. The value of the kinetic energy of the electron as compared to the binding energy of the electron, W_i, is responsible for the type of reaction taking place. When the $KE > W_i$, electronic capture is said to take place, i.e., electron is trapped into atom. While when $KE = W_i$, metastable electron, i.e., electron which is excited from inner orbit to outer orbit, moves back with energy release, i.e., quantum of light in a random direction. This causes photoionization. Whereas when $KE > W_i$, successful ionization due to electronic collision takes place leading to avalanche initiation. Thus, it can be seen that the second and third case, i.e., when $KE \geq W_i$, initiation of PD can occur.

Partial Discharge in Solid Dielectrics

Figure 3 shows the field developed in an insulator at various stages of operation.

When a voltage is applied to a virgin insulator (without cavities), a constant field is created throughout the insulator. Figure 3a shows the constant field (E) and electric displacement (E) inside a perfect insulator. Once this virgin insulator is pressed into service, over a period of time, the insulator is subjected to various stresses and a cavity or void is developed inside the solid insulator. Now, the constant field E gets distorted and two fields are developed: E_{ins}—field in the solid insulator outside the void and E_{void}—field inside the void, as shown in Fig. 3b. The presence of imperfections inside the insulator distorts the field. However, the electric displacement given by Eq. 1 remains the same.

$$D = \epsilon_0 \epsilon_r E = \epsilon_0 \epsilon_{\text{air}} E_{\text{void}} = \epsilon_0 \epsilon_{\text{dielectric}} E_{\text{ins}} \tag{1}$$

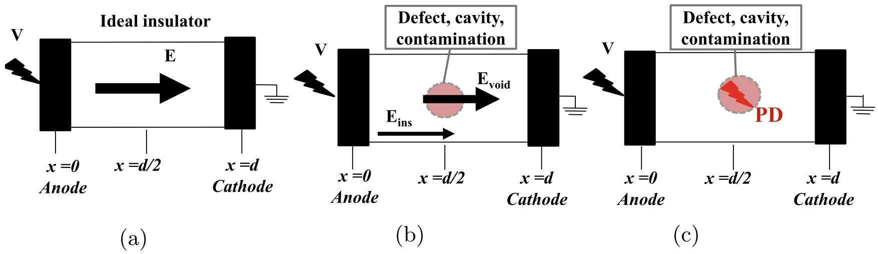


Fig. 3 PD development stages in brief **a** field applied across ideal insulator. **b** Field distortion created when the defect or cavity formed in an insulator. **c** PD produced when enhanced stress inside the defect crosses a threshold

Here, ϵ_0 represents the vacuum permittivity, while ϵ_r gives relative permittivity. The ϵ_{air} and $\epsilon_{\text{dielectric}}$ give the permittivities of air (assuming air is the gas inside the void) and the dielectric, respectively.

$$E_{\text{ins}} = \frac{\epsilon_{\text{air}} E_{\text{void}}}{\epsilon_{\text{dielectric}}} \tag{2}$$

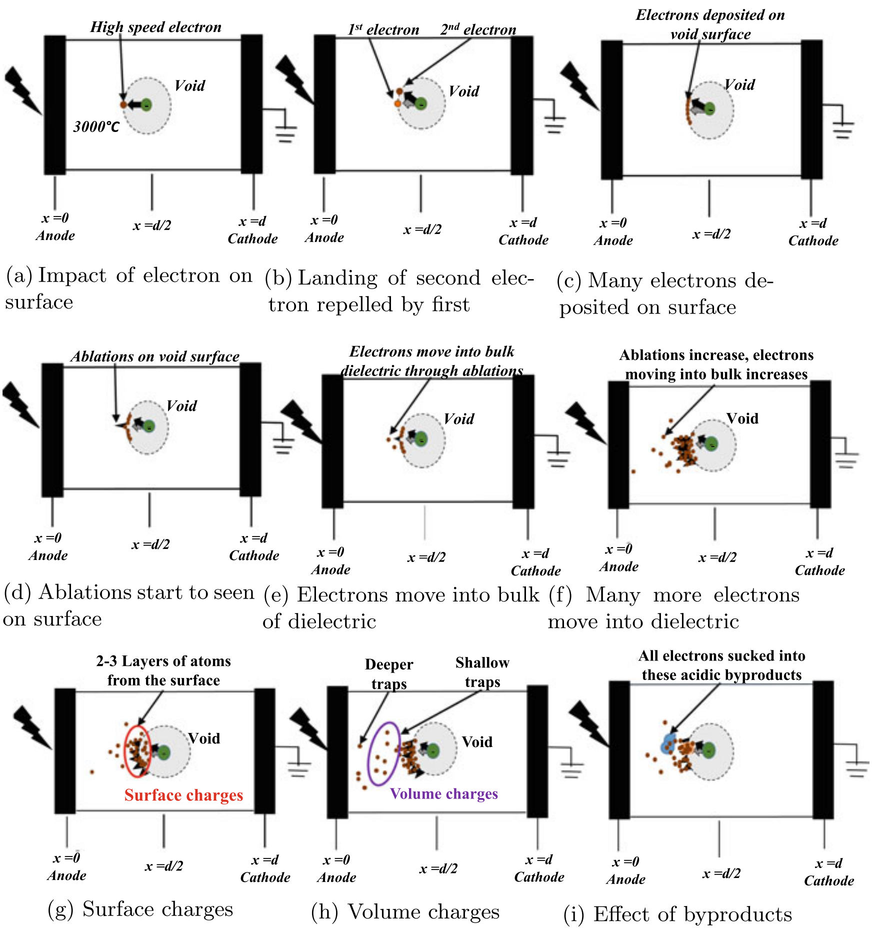
Since $\epsilon_{\text{dielectric}} \gg \epsilon_{\text{air}}$, $E_{\text{void}} \gg E_{\text{ins}}$. Thus, an enhanced field inside the void (E_{void}) and reduced field (E_{ins}) in the insulator outside the void is produced as represented in Fig. 3b.

When this enhanced stress inside the imperfections exceeds a certain threshold, local breakdown occurs, i.e., the gas inside the imperfection changes into a conductor as shown in Fig. 3c, i.e., partial discharge. It is called partial because breakdown bridges only the void as opposed to the entire insulator.

Propagation of PD in a Solid Dielectric with a Gas-Filled Cavity: Once the first PD pulse is produced in a void, it aids the generation of further discharges. This propagation phenomena taking place inside the gas-filled cavity needs to be understood. Figure 4 describes the processes involved in propagation of PD in an epoxy resin with a gas-filled cavity [40]. Here, the explanation is presented assuming the gas inside the cavity is air.

When the first electron reaches the anode side of the void, the electron, traveling with high speed, hits the surface with very high temperature resulting in the impact point to burn, evaporate, and then disappear as shown in Fig. 4a. The impact point is no longer a part of polymer. It gets pulled into the void, and the incoming electron gets deposited into this hole. The electron's mobility is reduced on surface impact. When the next electron arrives at the surface, the previous electron repels it and the impact is made at a new impact point as shown in Fig. 4b. This repulsion from the previously present electron increases their landing time [12].

This process continues, and many such electrons are deposited onto the surface as shown in Fig. 4c. The more the electrons deposited, the less is its power of impact. Avalanche impact causes mechanical and chemical damages. Due to this electron impact on the surface, ablations are also seen on the void surface as shown in Fig. 4d.



(a) Impact of electron on surface (b) Landing of second electron repelled by first (c) Many electrons deposited on surface

(d) Ablations start to seen on surface (e) Electrons move into bulk of dielectric (f) Many more electrons move into dielectric

(g) Surface charges (h) Volume charges (i) Effect of byproducts

Fig. 4 PD propagation in a solid dielectric with a gas-filled cavity

The electrons move into the bulk of the dielectric through these ablations as in Fig. 4e. They create phonons with energy ready to interact. Many more electrons then move into the dielectric as in Fig. 4f. Therefore, the ablations increase in number. The increasing number of electrons is deposited onto the surface and moves into the dielectric. Some electrons are deposited near the surface inside the bulk. The charges within 2–3 layers of atoms from the surface are called surface charges in Fig. 4g. Some electrons move more into volume of dielectric, Fig. 4h. Depending on its location from the surface, these are classified as deeper and shallow traps [40].

Energy impact causes polymer chain scission and cross-linkage, forming new polymers or by-products. Carbon, hydrogen, and other ions in addition to metastables from areas near surface dielectric react with each other in presence of humidity. The chemical by-products such as liquids composed mainly of a mixture of nitric,

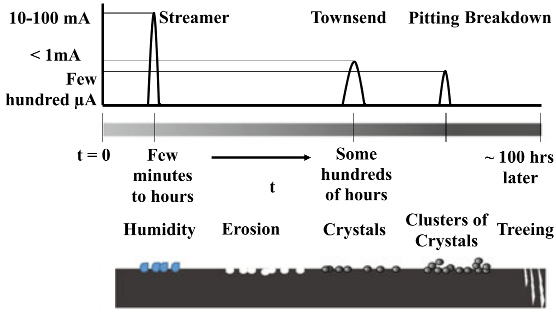


Fig. 5 Various stages of aging and the PD pulses observed

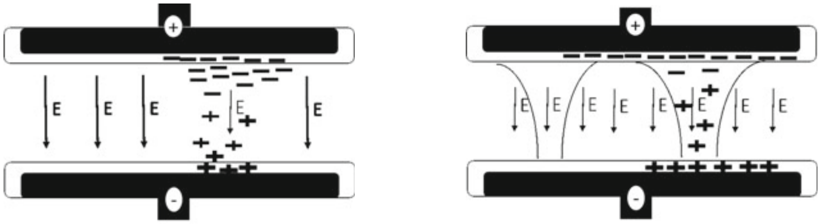
formic, glycolic, and glyoxylic acids, as well as other oxidized products such as aldehydes and alcohols [41] and acidic by-products such as alcohol, oxalic, formic, acetic, and carboxylic acids are formed in the surface resulting from the interaction among gas, temperature, polymer, and avalanche or plasma. Most of the by-products are conductive causing all the electrons to be sucked into these by-products on the surface as shown in Fig. 4i. Thus, these by-products become the preferential site of PD. By-products formation increases charge mobility and prevents overstressed conditions, causing a discharge as soon as the breakdown value is achieved, leading to slow fronted pulses with long duration. More by-product deposits lead to longer tail time. The longer duration corresponds to volume degradation conditions [40].

Impact of Aging in a Solid Dielectric on PD Pulse Shapes

When a virgin insulation material is pressed into service, it is observed to undergo several stages of aging in the course of its lifetime. The physical and chemical changes taking place at the atomic and molecular level is seen as a change in pulse shape at different stages of aging [42]. Each stage of aging has been identified by the different pulse shapes observed, namely streamer, Townsend, and pitting [42]. Figure 5 shows the changes on the cavity surface and the corresponding pulse shapes seen during the lifetime of a dielectric [41, 43–47]. The changes taking place inside a cavity in a dielectric is summarized in this section.

Aging Stage 1—Physical Changes and Pulse Characteristics

Within few hours of service, a sheet of insulator reacts with humidity and dissociation products of air [44]. Consequently, the conductivity of the cavity surface increases. The electric field throughout the insulation material is more or less uniform. The PDs occurring in this stage do not have a specific point of impact.



(a) Localized reduction in electric field due to localized space charge, [48]

(b) Reduction in electric field throughout due to spread out of space charge, [48]

Fig. 6 Atomic representation of cross section of a solid dielectric

Since the surface of the cavity mainly consists of the epoxy matrix which has relatively higher resistivity, the space charges are accumulated over a small region [48]. Thus, a discharge event does not affect the charge distribution over a large part of the cavity surface. Hence, in this stage, the discharge covers only a part of the cavity surface. The electric field reduction due to the space charge is also localized as explained in Fig. 6a [48]. The reduced electric fields inside the void cause faster field recovery after a discharge. This process takes place at each weak point of the insulator, thus giving rise to multiple PD sites.

Hence, due to the lack of free electrons, even after the local electric field has reached its breakdown value, the PD occurs only after a sufficient time lag in this stage. Thus, the PDs are observed at higher over-voltages in a virgin insulator.

The pulses observed during this stage are long, very short duration pulses (10–100 mA, 1–8 ns rise time, 0–25 ns pulse width) as shown in Fig. 5 [42, 46]. The pulse characteristics do not change with the void geometry. The pulse magnitudes are seen to be almost constant in this stage of aging [43]. The pulses seen in this stage are observed to be triangular [46]. These pulses belong to the streamer type.

The successive avalanches started in gas by high energetic photons produced as a result of ionization of high local space-charge field give rise to streamers [42]. In the streamer mechanism, the electrons and positive ions are produced throughout the gap [49]. Hence, the time taken by the charged particles to reach the respective electrode is less, giving rise to a very short rise time and duration of pulses. The front time is observed to be less than the transit time of an electron to cross the void [45, 46].

Transition to Stage 2—Physical Changes and Pulse Characteristics

On further aging, a liquid layer is observed on the surface of the cavity. Simple organic compounds, like formic, acetic, and carboxylic acids are formed as by-products. Due to this acidic layer, there is a strong increase in the conductivity of the cavity surface

(6–7 orders of magnitude) causing the space charges to spread out leading to the electric field reduction throughout as described in Fig. 6b [44].

As the time of operation of the insulator increases, the shallow traps on the surface of the cavity increases. Free electrons are available freely as compared to the virgin insulator. This causes reduced time lags, thus leading to PDs at lower overvoltages.

The overall field inside the cavity is reduced as compared to the first stage. Since the magnitude of a PD pulse depends upon this field, long pulses are seen in the streamer stage while shorter pulses in the next stage.

Aging Stage 2—Physical Changes and Pulse Characteristics

The liquid layer deposited earlier is crystallized with time to form hydrated oxalic acid crystals localized on the cavity surface [44]. The field enhancement is observed on the tips of these crystals. This leads to intensification and localization of PD activity concentrating on the tips of the crystals.

The increased surface conductivity and the consequent field reduction prevents the occurrence of multiple points. The discharge event covers the cavity surface entirely. The copious availability of free electrons, lead to reduced time lags and lower overvoltages.

Small, long duration pulses (several hundred μA, rise times more than 10 ns, and more than 15 ns wide) are the characteristics of this stage as shown in Fig. 5 [42, 46]. In this stage, pulse magnitude fluctuates as compared to the streamer pulses [43]. The rise times in this stage are also seen to have a larger scatter. The void geometry affects the pulse characteristics. These pulses belong to the Townsend type of pulses.

A self-sustained avalanche with successors originated at anode result in a Townsend pulse [42]. According to the Townsend mechanism, the charged particles are generated near the anode [49]. Hence, the time taken to reach the respective electrode is relatively more, giving rise to a larger rise time and duration of pulses as compared to the streamer. The observed front times are almost ten times more than the transit time of an electron to cross the void [45, 46].

Aging Stage 3—Physical Changes and Pulse Characteristics

Clusters of crystal are formed on the surface. PD activity is ignited on these clusters [44]. The enhanced electric field at the clusters causes sharp pit formation consequently leading to the initiation of tree growth from a point of field enhancement. The tree grows rapidly and finally leads to the breakdown of the insulation material.

This stage is characterized by very small magnitude and intermediate duration pulses (few hundred μA, rise times less than 10 ns, and 10–15 ns wide) as shown in Fig. 5 [42]. The pulse shapes and the rise times are seen in this stage which are similar to the streamer pulses [42]. A little scatter is seen in the pulse amplitudes in

Fig. 7 Typical PD spectroscopy workflow

this stage. These pulses belong to the pitting type of discharges. The repetition rate of PDs is very high (several tens in a μs).

Condition Monitoring of Insulation

Condition monitoring is the process of acquisition and processing of data relevant to different attributes of the insulation so as to predict and prevent the failure of a insulation. The condition of the insulation is expressed as a function of operating time and aging of the electrical apparatus. Based on the condition monitoring of the insulation, a health index is normally defined for the power utility. The health index gives idea about the likely residual life of the apparatus and the maintenance that can be carried out to extent the life of the apparatus [51].

It is seen from the previous section that the insulation material undergoes many changes and also exhibits its effects as PD in every stage of its aging. Hence, detecting the void in the initial stages of aging itself, monitoring it, and also analysis of PD data are crucial for the health tracking of the electrical equipment. Thus, PD prevention and detection are essential to ensure reliable, long-term operation of high-voltage equipment used by electric power utilities.

Once the PD is extracted from the field, it needs to be analyzed to estimate the aging of the insulation material and to predict the residual life of the insulator. This will help schedule the maintenance of the electrical equipment in service. The spectroscopy techniques and the parameters used for analysis are summarized in Sect. 15.5.1.

Spectroscopy for PD Analysis

Once the PD is measured by various methods, the relevant data has to be extracted from the detected PD pulses [47, 53] and analyzed. Therefore, this spectroscopy forms an integral part of condition monitoring system. A detailed description of this is already given in recent literature survey papers [54–58]. For brevity, the repetition is avoided in this paper. However, a brief summary of the spectroscopy in PD analysis is presented in the following paragraph. Figure 7 shows the typical PD spectroscopy workflow.

The first stage being PD measurement and acquisition. The raw data gathered from the field is then denoised and filtered. Three different categories of PD pulse data are gathered from the digital PD detectors during the experiments based on the

Table 1 PD pulse data patterns from field [53]

	PD pulse data patterns
Phase resolved data	$\phi \sim$ q $\sim$ n $\phi \sim$ q, q $\sim$ n $\phi \sim$ n [53]
Time resolved data	Individual pulse, q~t [53]
Data having neither phase nor time information	q~V [53]

Table 2 PD fingerprints data used in research [27, 42, 46, 47, 50, 52–63]

	PD fingerprints data
Statistical methods	Mean, variance, skewness, kurtosis, Phase symmetry, discharge symmetry [53]
Pulse characteristics	Rise time, fall time, pulse width, pulse height, and rise time distributions [27, 42, 46, 53, 62, 63]. repetition rate, average discharge current [53] pulse sequence analysis, voltage difference between consecutive pulses- $\Delta V_k(V_{k+1} - V_k)$[60, 61] time histograms, density histograms, cumulative discharge repetition rate [50] charge/pulse height distributions, pulse energy distributions, system energy distribution, pulse interval distribution [59]
Signal processing	Fourier series analysis/Fourier transform, Haar and Walsh transforms, wavelet transform [53] discrete wavelet transform (DWT) [57]
Image processing	Texture analysis algorithms, fractal features, wavelet-based image decomposition [53]
Time series	q~V, auto-regressive time series to model qV [53]

parameters discharge phase (ϕ), charge transfer (q), discharge rate (n), time (t), and test voltage, V, as summarized in Table 1. An identification set or PD fingerprint data pattern or feature vector is generated from the data acquired from the experiments. This is followed by a pattern recognition classification algorithm for defect separation and classification, severity assessment, and localization.

Different feature vectors and classification algorithms (as summarized in Tables 2 and 3) combinations have been used by different researchers [52, 54–56]. Most recently, machine learning and deep learning-based techniques have been used for PD classification [57, 58]. Table 2 lists the various PD fingerprint data pattern, while Table 3 enumerates the classifiers used by various researchers for analysis of PD to make inferences on type of PD defect, localization, and time to breakdown [27, 42, 46, 47, 50, 52–63].

Table 3 PD classifiers used in research [27, 42, 46, 47, 50, 52–63]

	Classifiers
Fuzzy logic	Fuzzy inference system (FIS) [57]
Artificial neural network-based classifiers	Back-propagation neural net, Kohonen self-organizing feature map, learning vector, quantization network, counter propagation, neural net modular and cascaded neural nets [53]
Statistical classifiers	Bayes classifier, recognition rate classifier [53] support vector machine (SVM) [57]
Distance classifiers	Minimum distance, nearest neighbor, percent score (centour score), polynomial [53] k-nearest neighbors (kNN) [57]
Conventional machine learning (ML) classifiers	Classifiers: artificial neural network support vector machine, fuzzy inference system, decision tree (DT), and random forest (RF) Dimension reduction techniques: fisher discriminant analysis (FDA) and principal component analysis (PCA), Kernel PCA (KPCA), kernel FDA (KFDA) metric multidimensional scaling (MDS) stochastic proximity embedding (SPE), isomap, stochastic neighbor embedding (SNE), and local linear embedding (LLE) [54–57]
Deep learning (DL) classifiers	Deep artificial neural network, convolutional neural network, deep belief network, stacked autoencoder, recurrent neural network [58]

Conclusion

The local discharges that take place in the most stressed part of the insulator can cause irreversible damage to the insulator. The process that leads to the discharge is discussed in detail in this paper. The diagrammatic depictions have been used to explain the complicated PD process in a gaseous dielectric as well as in a cavity inside a solid dielectric. The various features and the pulse data that can be gathered for spectroscopy for PD analysis have been listed. The use of this data for analysis, characterization, and localization of PD has been enumerated.

The PDs in a cavity in a solid medium is a complicated affair. The electron generation mechanisms, space-charge traps formations, etc., create the complexities. In the case of solid dielectric, the previous PD aids the generation of the next discharge as opposed to air medium. This causes irreversible damage to the insulation. The pulses seen in different stage of aging vary in their characteristics.

PD detection is an important diagnostic not only to identify the onset of damage caused by electric stress but also to monitor a changed and worsening condition initiated by the broader range of stress factors as also for the characterization of dielectrics. Armed with an enhanced understanding of the PD process, personnel can efficiently analyze and characterize the insulation and equipment condition. The condition-based maintenance can be scheduled based on the health of the equipment. Thus, the insights gained will certainly be useful in the field of condition monitoring of electrical equipment and will be a boon to the manufacturers and users in the industry.

Acknowledgements The authors would like to thank Dr. P. P Vaidya, HOD, R&D and Department of Instrumentation, VES Institute of Technology, Mumbai, India, and Prof. Himanshu Bahirat and Prof. S. V. Kulkarni, Department of Electrical Engg., IIT Bombay, India, for the valuable insights and technical help extended during the study related to this paper. The authors would also like to thank Prof. Vicente Rodolfo Garcia Colon Hernández, Instituto Nacional de Electricidad y Energías Limpias (INEEL), Mexico, for helping to get a better scientific understanding of the complex physics of the topic. The authors would also like to thank the editors and anonymous reviewers for providing insightful suggestions and comments to improve the quality of research paper.

References

1. Kwan Chi Kao, Dielectric Phenomena In Solids, London: Elsevier Inc, 2004
2. Das-Gupta DK (1999) Dielectric permittivity and loss, Chapter 17, dielectrics and electrical insulation. Wiley Encyclopedia of Electrical and Electronics Engineering, UK
3. Nair LM (2011) Analysis & monitoring of degradation of insulation used in high voltage dry type transformers, Master of Engineering degree thesis, University of Mumbai
4. Montanari G, Seri P ((2018)) A partial discharge based health index for rotating machine condition evaluation. In: IEEE electrical insulation magazine, vol 34, pp 17–23
5. Kuffel E, Zaengl WS, Kuffel J (2000) High voltage engineering fundamentals, 2nd edn. Newnes
6. L. Niemeyer, A generalized approach to partial discharge modeling, IEEE Transactions on Dielectrics and Electrical Insulation, Vol. 2, pp. 510–528, 1995
7. F. Gutfleisch and L. Niemeyer, Measurement and simulation of PD in epoxy voids, IEEE Transactions on Dielectrics and Electrical Insulation, Vol. 2, pp. 729–743, 1995
8. Jian LT, Bakar AHA, Illias HMH (2012) Partial discharge simulation under various applied voltage waveform. In: Proceedings of IEEE international conference on power and energy (PECon). Malaysia, pp 967–972
9. Illias HA, Chen G, Lewin PL (2008) Modeling of partial discharges from a spherical cavity within a dielectric material under variable frequency electric fields. In: Annual report of conference on electrical insulation dielectric phenomena (CIEDP), pp 447–450
10. Illias HA, Chen G, Lewin PL (2009) Partial discharge modelling in a spherical cavity within a dielectric insulation material as a function of frequency. In: IEEE electrical insulation conference. Montreal, QC, pp 55–59
11. Forssen C, Edin H (2007) Modeling partial discharges in a cavity at different applied frequencies. In: Annual report of conference on electrical insulation and dielectric phenomena (CEIDP), pp 132–135
12. Illias Hazlee, Chen George, Lewin Paul L (April 2011) Partial Discharge Behavior within a Spherical Cavity in a Solid Dielectric Material as a Function of Frequency. IEEE Transactions on Dielectrics and Electrical Insulation 18(2):432–443
13. Chen G, Baharudin F (2008) Partial discharge modelling based on a cylindrical model in solid dielectrics. In: International conference on condition monitoring and diagnosis. Beijing, pp 74–78
14. H. A. Illias, M. A. Tunio, H. Mokhlis, G. Chen, and A. H. A. Bakar, Determination of partial discharge time lag in void using physical model approach, IEEE Transactions on Dielectric and Electrical Insulation, Vol. 22, no. 1, pp. 463–471, 2015

15. Illias HA, Chen G, Lewin PL (2009) Partial discharge measurements for spherical cavities within solid dielectric materials under different stress and cavity conditions. Electrical Insul Dielectric Phenomena (CEIDP) 388–391
16. Sivaji Chakravorti and Prithwiraj Das, “Simulation of PD patterns due to a narrow void in different E-Field distribution,” Journal of Electrostatics, Vol. 68, pp. 218–226, 2010
17. C. Forssen and H. Edin, “Partial discharges in a cavity at variable applied frequency part 2: measurements and modeling,” IEEE Transactions on Dielectrics and Electrical Insulation, Vol. 15, no. 6, pp. 1610–1616, 2008
18. Seralathan KE, Mahajan A, Gupta N (2008) Modelling of electric tree progression due to space charge modified fields. J Phys D Appl Phys 41:1–9
19. S. Das and N. Gupta, Effect of Ageing on Space Charge Distribution in Homogeneous and Composite Dielectrics, IEEE Transactions on Dielectrics and Electrical Insulation, Vol. 22, No. 1, pp. 541–547, 2015
20. Pandey JC, Gupta N (2014) Thermal aging assessment of epoxy-based nanocomposites by space charge and conduction current measurements. In: Proceeding of electrical insulation conference, pp 59–63
21. Montanari GC (2010) Bringing an insulation to failure: the role of space charge. In: Annual report conference on electrical insulation and dielectic phenomena, pp 1–25
22. H. Illias, G. Chen and P. L. Lewin, Modelling of surface charge decay in a spherical cavity within a solid dielectric material using finite element analysis, *Proceedings of the 16th International Symposium on High Voltage Engineering,* Johannesburg, Paper E-18 , pp. 1-6, 2009
23. Heitz C (1998) A general stochastic approach to partial discharge processes. In: Proceedings of IEEE international conference on conduction and breakdown in solid dielectrics, pp 139–144
24. Van Brunt RJ, Cernyar EW, von Glahn P (1993) Importance of unraveling memory propagation effects in interpreting data on partial discharge statistics. IEEE Trans Electrical Insul 28:905–916
25. B. Fruth and L. Niemeyer, “The importance of statistical characteristics of partial discharge data,” IEEE Transactions on Electrical Insulation, Vol. 27, pp. 60–69, 1992
26. R. J. Van Brunt, “Stochastic properties of partial discharge phenomena,” IEEE Transactions on Electrical Insulation, Vol. 26, pp. 902–948, 1991
27. L. A. Kaimal, H. J. Bahirat, P. P. Vaidya and S. V. Kulkarni, “Design of a Universal Partial Discharge Simulator,” 2019 IEEE 4th International Conference on Condition Assessment Techniques in Electrical Systems (CATCON), Chennai, India, 2019, pp. 1-6
28. IEC60270: High-voltage test techniques—partial discharge measurements
29. Ramu TS (2010) Partial discharge based condition monitoring of high voltage equipment, New Age International Publishers
30. Tanaka T (1986) Internal partial discharge and material degradation. IEEE Trans Electrical Insul 21(6):899–905
31. Devins JC (Oct. 1984) The 1984 J. B. Whitehead Memorial Lecture the Physics of Partial Discharges in Solid Dielectrics. IEEE Transactions on Electrical Insulation 19(5):475–495
32. Li Z, Okamoto K, Ohki Y, Tanaka T (Jun. 2010) Effects of nano-filler addition on PD resistance and dielectric breakdown strength of Micro Al2O3 Epoxy composite. IEEE Trans. Dielectr. Electr. Insul. 17(3):653–661
33. L. E. Kline, “Effect of negative ions on current growth and ionizing wave propagation in air,” J. Appl. Phys., 46, pp. 1567–1574, 1975
34. Hugh Ryan, “High-Voltage Engineering and Testing,” The Institution of Engineering and Technology, London, United Kingdom, 2013
35. Hoof M, Patsch R (1997) A physical model, describing the nature of partial discharge pulse sequences. In: Proceedings of the 5th international conference on properties and applications of dielectric materials, vol 1. Seoul, pp 283–286
36. Morrow R (Sept. 1985) Theory of negative corona in oxygen. Phys. Rev. A. 32(3):1799–1809
37. Boggs, S.A., “Partial discharge. III. Cavity-induced PD in solid dielectrics,” IEEE Electrical Insulation Magazine, Vol. 6, no. 6, pp. 11–16, 1990

38. Kang WS, Park JM, Kim Y, Hong SH (2003) Numerical study on influences of barrier arrangements on dielectric barrier discharge characteristics. IEEE Trans Plasma Sci 31(4):504–510
39. A. Bondiou, I. Gallimberti, “Theoretical modelling of the development of the positive spark in long gaps,” Journal of Physics D, 27, 12521266, 1994
40. Gupta N, Rodolfo Garcia Colon Hernández V (2017) Lecture notes of GIAN course on ultra-wide band partial discharge detection in In-service high voltage equipment
41. Hudon C, Bartnikas R, Wertheimer MR (Dec 1995) Effect of physico-chemical degradation of epoxy resin on partial discharge behavior. IEEE Transactions on Dielectrics and Electrical Insulation 2(6):1083–1094
42. Morshuis PHF (1993) Partial discharge mechanisms: mechanisms leading to breakdown, analyzed by fast electrical and optical measurements. Doctoral thesis, Delft University Repository
43. P H F Morshuis and F H Kreuger, Transition from streamer to Townsend mechanisms in dielectric voids, J. Phys. D: Appl. Phys. 23, pp. 1562–1568, 1990
44. Morshuis PHF (Oct. 2005) Degradation of solid dielectrics due to internal partial discharge: some thoughts on progress made and where to go now. IEEE Transactions on Dielectrics and Electrical Insulation 12(5):905–913
45. Garcia-Colon VR (1996) Correlation between surface degradation and PD wave shape evolution for spherical voids. In: Seventh international conference on dielectric materials, measurements and applications, (Conf. Publ. No. 430), Bath, pp 317–321
46. Garcia Colon VRH (1994) Evolution of partial discharges waveshapes in spherical voids using ultra wide band techniques. Doctoral thesis, University of Manchester
47. Morshuis PHF (1995) Partial discharge mechanisms in voids related to dielectric degradation. IEE Proc Sci Meas Technol 142(I)
48. Hudon C, Bartnikas R, Wertheimer MR (Feb 1993) Spark-to-glow discharge transition due to increased surface conductivity on epoxy resin specimens. IEEE Transactions on Electrical Insulation 28(1):1–8
49. Kasten DG, Liu X, Sebo SA, Grosjean DF, Schweickart DL (April 2005) Partial discharge measurements in air and argon at low pressures with and without a dielectric barrier. IEEE Transactions on Dielectrics and Electrical Insulation 12(2):362–373
50. Morshuis PHF, Smit JJ (2005) Partial discharges at dc voltage: their mechanism, detection and analysis. IEEE Trans Dielectrics Electrical Insul 12(2):328–340
51. Montanari GC, Hebner R, Morshuis P, Seri P (Aug. 2019) An Approach to Insulation Condition Monitoring and Life Assessment in Emerging Electrical Environments. IEEE Transactions on Power Delivery 34(4):1357–1364
52. Suzuki H, Endoh T (Jun 1992) Pattern recognition of partial discharge in XLPE cables using a neural network. Electrical Insulation, IEEE Transactions on 27(3):543–549
53. Sahoo NC, Salama MMA, Bartnikas R (2005) Trends in partial discharge pattern classification : a survey. IEEE Trans Dielectrics Electrical Insul 12(2):248–263
54. J. Hamed and K. Behzad, Towards automated statistical partial discharge source classification using pattern recognition techniques, IET High Voltage, vol. 3, no. 3, pp. 162–169, 2018
55. W. J. K. Raymond, H. A. Illias, A. H. A. Bakar, H. Mokhlis, Partial discharge classifications: Review of recent progress, Measurement, Volume 68 2015, pp. 164–181
56. Wu M, Cao H, Cao J, Nguyen H-L, Gomes J, Krishnaswamy S An overview of state-of-the-art partial discharge analysis techniques for condition monitoring. IEEE Electrical Insul Mag 31: 22–35
57. Lu, Shibo & Chai, Hua & Sahoo, Animesh & Phung, Toan. (2020). Condition Monitoring Based on Partial Discharge Diagnostics Using Machine Learning Methods: A Comprehensive State-of-the-Art Review. IEEE Transactions on Dielectrics and Electrical Insulation. 27. 1861–1888
58. Barrios S, Buldain D, Comech MP, Gilbert I, Orue I. Partial Discharge Classification Using Deep Learning Methods-Survey of Recent Progress. Energies. 2019; 12(13):2485
59. Kranz HG (Dec 1993) Diagnosis of partial discharge signals using neural networks and minimum distance classification. Electrical Insulation, IEEE Transactions on 28(6):1016–1024

60. Hoof M, Patsch R (1994) Analyzing partial discharge pulse sequences—a new approach to investigate degradation phenomena. In: IEEE international symposium on Electrical insulation, 1994, conference record of the 1994, pp 327, 331
61. Hoof M, Freisleben B, Patsch R (Feb. 1997) PD source identification with novel discharge parameters using counterpropagation neural networks. IEEE Transactions on Dielectrics and Electrical Insulation 4(1):17–32. https://doi.org/10.1109/94.590861
62. H. Okubo, N. Hayakawa, A. Matsushita, "The Relationship Between partial Discharge Current Pulse Waveforms and Physical Mechanisms," IEEE Electr. Insul. Mag., Vol. 18, No. 3, pp. 38–45, 2002
63. Mazroua AA, Salama MMA, Bartnikas R (Dec 1993) PD pattern recognition with neural networks using the multilayer perceptron technique. Electrical Insulation, IEEE Transactions on 28(6):1082–1089

Analysis and Mitigation of Co-channel Interference in Downlink of Land Mobile Satellite Systems

Rachana B. Nair and S. Kirthiga

Abstract Joint Vertical-Bell Laboratories Layered Space Time architecture and multi user detection techniques implemented in MIMO based satellite communication architecture is proposed. Signal processing systems are incorporated in the satellite payload to minimize the effects introduced by the channel impairments particularly in the downlink, which accounts for the primary objective of this article. An unprecedented value of BER with 1–2 dB gain is observed for optimal maximum likelihood (ML) detection technique in comparison to other schemes^{-8} error rate is obtained for SINR of 8 dB for ML. A low value of 0.1 outage probability is observed at a SNR of 5 dB, indicating an efficient utilization of the spectrum.

Keywords VBLAST · MIMO · Co-channel interference

Introduction

There is a high demand for spectrum bandwidth due to increase in the number of users who are opting for wireless as the medium of communication [1]. Congestion in terrestrial communication system, finds it difficult to support vast majority of mobile users [2]. Opting for satellite communication system as an alternative system to provide services to the ever-increasing number of users. Satellite communication system provides services which are broadly classified as: Land Mobile Satellite Services, Maritime Mobile Satellite Services and Aeronautical Mobile Satellite Services. In this paper, Land Mobile Satellite Services providing telecommunication services to mobile users is considered as the system design model for the work under consideration. Adoption of various technologies in satellite system would boost the received signal strength thereby which the capacity and error rate improvement can be

R. B. Nair (✉) · S. Kirthiga
Department of Electronics and Communication Engineering, Amrita School of Engineering, Amrita Vishwa Vidyapeetham, Coimbatore, India
e-mail: bn_rachana@cb.students.amrita.edu

S. Kirthiga
e-mail: s_krithiga@cb.amrita.edu

V. Bindhu et al. (eds.), *Proceedings of Third International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 844,
https://doi.org/10.1007/978-981-16-8862-1_16

observed. Terrestrial communication system [3] support wide range of applications and large number of users by incorporating various technologies [4].

As we know that satellites have fixed positions on the orbit, and many countries launch satellite for supporting their users needs and demands. This places a constraint on the orbit position and frequency spectrum allocated for satellite communication system as the orbit position and frequency spectrum are resources which are scarce, so we cannot practically implement the orbital spacing of 2°–4° as per ITU-T's regulation to avoid interference [5]. As already mentioned frequency spectrum is a scarce resource, hence emphasizing the need to use the same carrier frequency for transmission/reception by another satellite communication system is a necessary requisite, resulting in frequency reuse [6]. This often leads to interference commonly called as co-channel interference.

Co-channel interference (CCI) is an issue in satellite communication systems which needs to be addressed to achieve better performance. Co-channel interference arises due to the reuse of same frequency by two or more satellite systems where the predefined orbital separation cannot be addressed [7]. A precise measurement of CCI is required to mitigate its impact on the communication system. Now a days, DVB-S2, DVB-SH standards are incorporated in the satellite systems [8]. One of the prominent technology used in DVB-SH standard is VBLAST architecture. In this paper, two satellite systems are considered. In each of satellite systems VBLAST architecture is implemented along with MIMO [9–11] technique owing to its success in terrestrial systems.

The impact of CCI on the two systems are reduced by the adoption of various prevalent detection techniques such as ML, MMSE, SIC-ZF and ZF algorithms [12–14]. ML detection technique is an optimal technique with increased complexity, whereas MMSE detection technique is a sub optimal approach involving pure nulling strategy, the success of ZF scheme lies in the accurate estimation of channel path associated with the corresponding user. In SIC-ZF algorithms, nulling and suppression of co-channel interference is done and ordering plays a phenomenal role. The performance of the system under consideration effected by CCI and various other related factors require performance analysis [15–17] and determination of received signal estimates on fading channels in the presence of CCI [18]. It is observed that temporal factors contribute to CCI [19].

Many research has been carried out to devise an efficient detection algorithm to cancel the CCI and such low complexity algorithms have been implemented. There are various ways in which CCI can be mitigated such as efficient management of radio resources, cognitive radio technology allowing frequency reuse in various dimensions such as frequency, space and time in order to reduce the limitation imposed by the spectrum and bandwidth. A secondary transmission scheme which enables the co-existence of secondary users in DTH broadcast signal reception region. Hence, we can conclude that CCI is the problem in satellite system whose influence on the received signal needs to be mitigated and various researches with respect to the detection mechanisms are carried out across the globe to result in better and efficient communication system. This paper deals with design of a system with two satellite communication systems and its associated sub systems to improve the performance metrics and the manner in which the nulling and canceling of CCI by different detection algorithms influencing the error rate performance are simulated and analyzed. A system architecture with two satellite systems implementing VBLAST structure in conjunction with MIMO and CCI suppression by detection systems are formulated and performance evaluation is done. Section 'System Design' discusses the system model and in section 'Results and Discussion' the results are presented and analyzed. Section 'Conclusion' paper is concluded.

System Design

The system model for this work (see Fig. 1). SISO ground station receives the signals from two terrestrial users which are vector encoded as a single data stream and transmitted to the satellite system. The VBLAST architecture is implemented in the satellite payload to improve the received SNR of the downlink signal transmitted from the satellite system. The multiplexed signal from the ground station is encoded by turbo encoder with code rate =1/3 and later BPSK modulated. The single data stream is precoded with channel gain values defining the downlink channel paths. The SISO encoded signal is processed by the implementation of channel encoding technique such as turbo encoder with code rate = 1/3 and the BPSK modulated signal is transmitted across the multiple antenna architecture. MIMO technique is incorporated in the satellite system under consideration. The user terminals of SATCOM system1 and 2 have single antenna and the entire system behaves as a 4 × 4 MIMO system configuration. The signal received at the earth station suffers from intra system interference and inter system interference, i.e., the signals of another satellite system

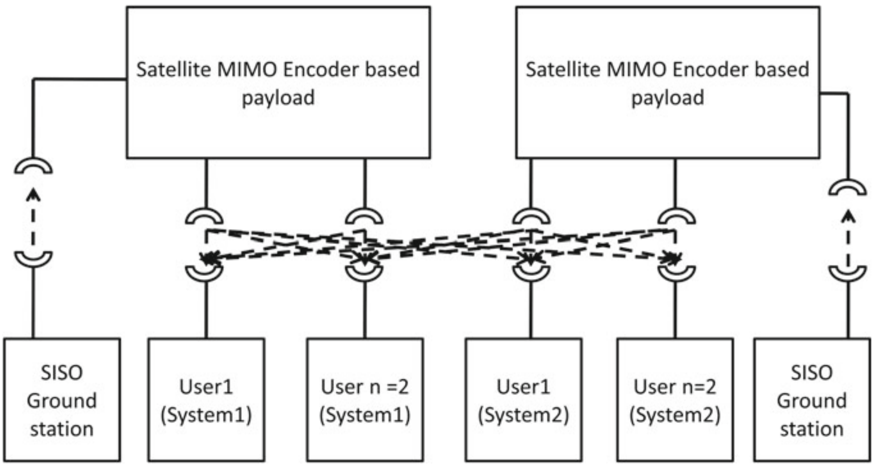


Fig. 1 VBLAST based Land mobile satellite end to end system

reusing the same frequency for its operation interferes with the signal of the desired satellite system users.

Hence, the received signal at the corresponding receiver unit is given as:

$$y_i = x_i h_i + \sum_{j \neq i}^{n} x_j h_{ij} + n_i i = 1 \ldots n \tag{1}$$

x and h are the data vector, and the channel coefficient values for n users. The two users later perform the detection mechanism to suppress the CCI from other satellite system reusing the same frequency as that of the system under consideration. Various detection mechanisms are considered for analysis purpose, namely, zero forcing, SIC-ZF, MMSE and ML detection schemes. In zero forcing detector, the mutual interference between the different antennas of the system is suppressed. The G_{ZF} is mathematically represented as:

$$G_{ZF} = \mathrm{H}^{+} = (\mathrm{H}^{*}\mathrm{H})^{-1}\mathrm{H}^{*} \tag{2}$$

which is computed by the Moore Penrose pseudo inverse of the channel transfer matrix, H. The zero forcing detector is obtained by the left multiplication of G_{ZF} on the receiver side,

$$x_{ZF} = G_{ZF} = H^{+}y = x + (H^{*}H)^{-1}H^{*}n \tag{3}$$

The channel with largest SNR, symbol is detected first on the corresponding receiver unit. MMSE minimizes the mean squared error between the output of linear detector and the actual transmitted symbol. σ_n^2 is the noise power variance.

$$G_{\mathrm{MMSE}} = (H^{*}H\sigma_n^2 I_M)^{-1} H^{*} \tag{4}$$

The detected symbol is given by:

$$x_{\mathrm{MMSE}} = G_{\mathrm{MMSE}} y = \left(H^{*}H + \sigma_n^2 I_M\right)^{-1} H^{*}(Hx + n) \tag{5}$$

H is the channel transfer matrix and σ_n^2, is the noise power variance.

$$\varphi_{\mathrm{MMSE}} = E[(x_{\mathrm{MMSE}} - x)(x_{\mathrm{MMSE}} - x)] = \sigma_n^2 (H^{*}H + \sigma_n^2 I_M)^{-1} \tag{6}$$

φ_{MMSE} is the expected value of the square of minimum error rate performance.

The MMSE output can be rewritten as:

$$x_{MMSE} = \left(H^{*}H\right)^{-1} H^{*}y \tag{7}$$

MMSE minimizes the noise power component. It is a good tradeoff between the interference suppression and noise amplification. MMSE algorithm achieves a better performance as compared to zero forcing algorithm. But a significant performance gap is observed between maximum likelihood algorithm and other detection algorithms. ML decoding minimizes the metric,

$$x = \text{argmin}_{x \in X} ||y - Hx|| \tag{8}$$

$x \in \{X\}$, x is an element of the set X which involves the exhaustive approach for the same. This technique achieves the best error probability performance in comparison to other detectors, but one major disadvantage of this technique is the large complexity involved in comparison to other algorithms. Successive cancelation algorithm is an algorithm based on the BLAST detection scheme. In this scheme, interference nulling can be considered as a feed forward filter and interference canceling works like a feedback filter. The SIC algorithm includes interference nulling, interference canceling and ordering. The algorithm works in the order of ordering, nulling and cancelation.

$$x_k = (y - \sum_{i=1, j=1}^{n} h_{i,j} x_j) h_{i,k}^{+} \tag{9}$$

where k denotes the index of the user, $h_{i,j}$ is the coefficient values of the channel matrix, MR $\times$ MT where MR is the number of antenna elements at the receiver and MT is the number of antenna elements at the transmitter. y is the received vector. The signal propagated along the channel with largest SNR value is detected at the corresponding receiver unit. nth channel has the worst statistics. Similarly, for satellite communication system 2 the same architecture as described above is applicable. The same system architecture governs the implementation of the satellite system 2 as the receiver units corresponding to satellite system 2 faces similar issue of co-channel interference which arises due to the usage of same frequency as that of the users in satellite system 1. Various metrics are analyzed with respect to the performance evaluation of the system design which is discussed in the consecutive section.

Results and Discussion

BER Performance Analysis

ML maximizes the likelihood function and minimizes the cost function, but the design complexity is high with regard to other mechanisms. MMSE determines the mean squared error and the most probable symbol is estimated to be the detected data. SIC-ZF is a nulling and suppressing algorithm which successively detects the

CCI symbols and cancels the impact of these interference from the received data at the corresponding receiver units. Zero forcing involves the application of Moore Penrose pseudo inverse of the estimated channel gain on the received data without the elimination of the interference. This leads to detected data with more errors due to interference from intra system and inter system. As per the simulation results presented (see Fig. 2). Zero forcing gives the least performance in comparison to other detection techniques.

There is 1–2 dB gain in SINR for ML in comparison to other schemes. For ML technique, a SINR of 8 dB is required to achieve a BER of 10^{-8} whereas MMSE,SIC-ZF and ZF requires 9 dB, 9.5 dB and 11 dB, respectively.

As per Table 1, the error rate performance for ZF algorithm is more than SIC-ZF, MMSE and ML algorithm. ML algorithm exhibits least error rate performance owing to the exhaustive search technique. The linear techniques exhibit sub optimal performance with higher complexity for optimal technique and reduced complexity for sub optimal approaches. SIC-ZF algorithm removes interference successively with reduced complexity.

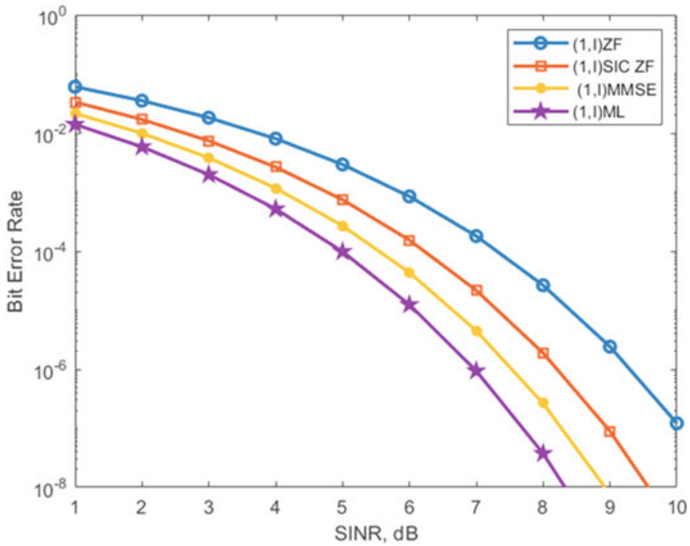


Fig. 2 BER versus SINR for different detection schemes

Table 1 Comparative analysis of different MUD algorithms

SINR	BER			
	ZF	SIC-ZF	MMSE	ML
1	10^{-1}	$10^{-1.3}$	$10^{-1.9}$	10^{-2}
5	10^{-2}	10^{-3}	$10^{-3.8}$	10^{-4}

Rate of Users of Satellite Communication System 1 and 2

The horizontal increase in rate of user1 with SNR indicates the individual capacity of user1 which is given by:

$$R_n = \log(1 + \frac{P_n}{N_o}) \quad (10)$$

where n indicates the index of user, P_n is the power associated with the nth user, R_n is the rate of the nth user and in this case $n = 1, 2$. But in actual practice, $R_1 < \log(1 + \frac{P_1}{N_o})$, user 1's rate, the constraint brings out the experimental validation that rate of individual user is lesser than the rate of a single user for a point-to-point AWGN channel. It is a valid constraint as the two users send independent signals (see Fig. 3). Rate of users in SATCOM1 and 2 exhibits behavior similar to Eq. 8. The multi user interference is observed in the sum rate region which is mathematically represented as:

$$R_1 + R_2 < \log(1 + \frac{P_1 + P_2}{N_0}) \quad (11)$$

which is lesser than individual rates of users 1 and 2 with power P_1 and P_2, respectively, in satellite system.

ML receiver achieves the best total rate among all receiver structures. The performance limit of downlink channel is characterized by the notion of capacity region. The capacity region is simulated and plotted in Fig. 3. The point-to-point capacity for user $k = 1, 2 \ldots 0.16$ is achieved by receive beamforming converting the effective channel into a SISO configuration and then decoding the data of the user. Inequality is a constraint on the sum rate that users can communicate at. The total rate that can be reached in a point-to-point channel with two users functioning as two receiver antennas with independent inputs at the transmit antennas is shown on the right. Because the receivers are optimized for reaching the total rate of a point-to-point channel with two users operating as two transmit antennas of one user, the rates for two users that this architecture can accomplish in the downlink meet inequality with equality. Moreover, if we cancel user1 first user, user 2 is only affected by the AWGN noise and the detected symbol has higher signal strength with a low probability of error. R1 and R2 represent the symmetric capacity of the downlink allowing both users to have unit spatial degrees of freedom. The received signal strength is high with reduced interference and noise owing to the processing architecture with the encoder to remove the fading effects and the noise, the modulation reduces the impact of channel impairments, with more energy per bit by the implementation of BPSK modulation.

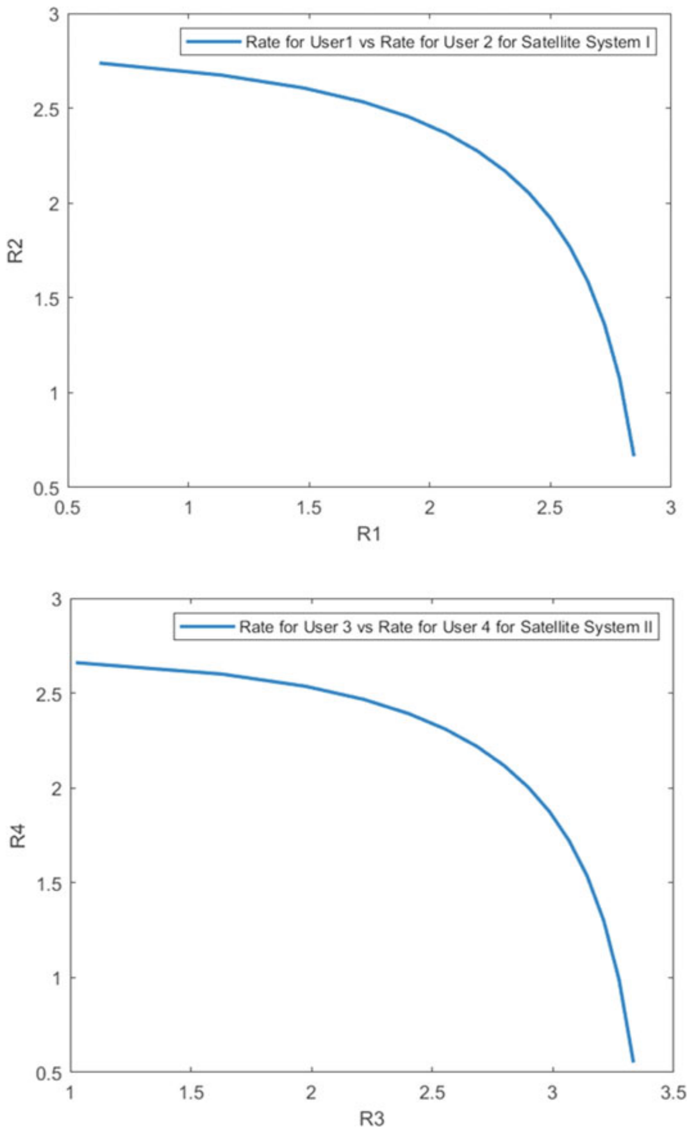


Fig. 3 Rate of users of SATCOM1 and 2

Outage Capacity

Figure 4 indicates the underutilized spectrum bandwidth allocated to users of SATCOM 1. For 7 dB SNR, a low outage capacity of 0.2 is observed, hence it indicates the transmission rate approaching the Shannon's capacity owing to the implementation of VBLAST architecture, providing more accuracy to the data transmission. It

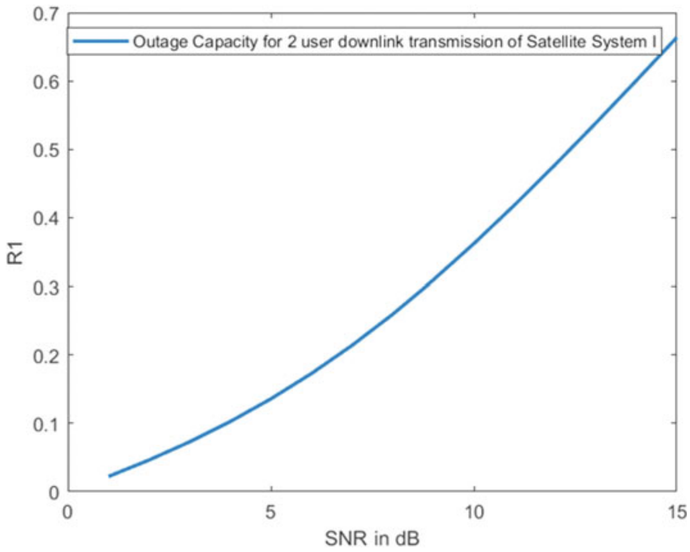


Fig. 4 Outage capacity versus SNR for SATCOM 1

is largest rate of transmission such that the outage probability is less than ϵ when we solve outage probability we would obtain

$$C_{\in} = \log\big(1 + F^{-1}(1- \in)\mathrm{SNR}\big), \tag{12}$$

where F is the complementary cumulative distribution function of channel gain power. For coded systems as per this system architecture, we are considering both high and low SNR regimes. To achieve the same rate as AWGN channel, an extra $10\ \log(1/F^{-1}(1 - \epsilon))$, of power is required and ϵ is the constant. This condition holds good independent of SNR of the environment. For the analytical purpose, we have simulated the outage capacity as a function of SNR for Loo statistical channel model. We can re frame the outage capacity equation in the following manner as $C_{\epsilon} = \log(\mathrm{SNR}) + \log(F^{-1}(1 - \epsilon)$ to analyze the outage capacity as a fraction of AWGN capacity at the same SNR. As per Fig. 4 it is evident that the impact is much more significant in the low SNR regime. At high SNR, a constant difference irrespective of SNR is observed. The reliable rate supported by Loo channel at low SNR is more sensitive than at high SNR.

Sum Rate Capacity

$$C_{\mathrm{sum}} = \log(1 + \frac{KP}{N_0}) \tag{13}$$

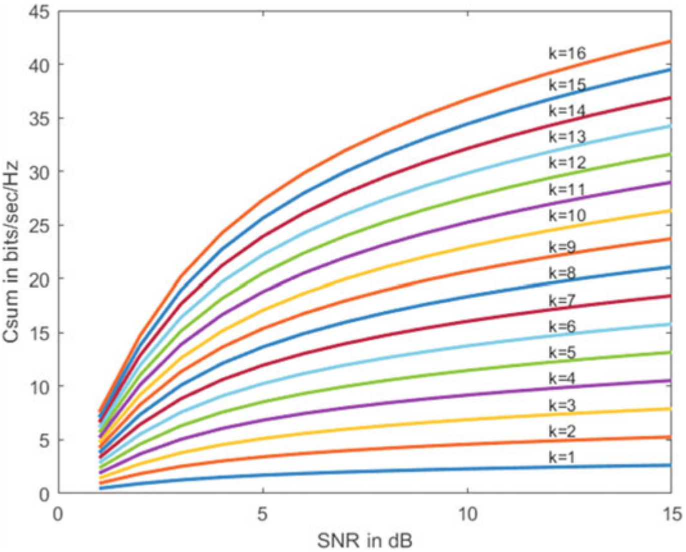


Fig. 5 Sum rate versus SNR in dB

C_{sum} is the sum rate capacity obtained for K users with individual users associated with power $= P$. It is an multiplicative increase in capacity for K users with the rate of a single user with maximum achievable spectral efficiency, depicting an ideal scenario providing a maximum achievable rate for iterative increase in the number of users from $n = 1$ to 16 with power equivalent to a single user. It is the maximum rate for each user that can be obtained if every user operates at the same rate. It is the maximum rate for each user that can be obtained if every user operates at the same rate. Sum rate is k dimensional polyhedron. The above results discussed so far dealt with 2 users for simplicity's sake, but the result plotted below illustrates an arbitrary number of users. The k user capacity region is described by $2^k - 1$ constraints. The right-hand side of the Eq. (13) corresponds to the maximum achievable sum rate by a single user with the total power of the users in the system and with no other users. The simulation plot below depicts the scenario in which the system architecture pertains to an external condition governed by only AWGN noise. But in actual practice, we know that in a multi co-channel satellite systems the performance is degraded by the fading conditions and interference which deteriorate the performance of the system resulting in deviation from the above achieved ideal AWGN conditions (Fig. 5).

Conclusion

Co-channel interference is a problem in present satellite communication system, impact of which needs to be reduced. Various detection techniques at the receiver are implemented in this paper to identify an optimal technique which can reduce CCI, improve the rate of users and also BER. From the simulations, it is known that ML scheme is an optimal technique to detect the symbols with minimum errors. It is also identified that implementation of VBLAST system in satellite payload would reduce the effect of CCI and also enhance the overall performance gain pertaining to capacity, BER as the incorporation of turbo encoder, BPSK modulation would enable error detection and correction mechanism. Precoding is performed at the transmitter earth station and also at the satellite system which would provide improvement in quality of service. Comparatively, an unprecedented BER performance is observed as per the simulations of the architecture under consideration. VBLAST processing architecture with SISO encoded data stream which is vector encoded to mitigate the effects of deep fade condition which could probably deteriorate the single stream transmission of the data vector. Moreover the multiple antenna architecture requisite for separate signal processing RF chain is not required due to SISO transmission of the multiplexed data. Hence, reducing the complexity, size of the payload system without requiring any additional structural architecture to accommodate the same. Cost is also reduced increasing the feasibility of the proposed system design with improvement in the average error rate performance and spectral efficiency. Optimization technique could be incorporated to improve the performance evaluation parameters with the existing architecture. Various meta heuristic algorithms can ensure enhanced performance with sub optimal complexity striking an appropriate balance between the non-conservative and conservative approaches.

References

1. Paulraj AJ, Gore DA, Nabar NU, Bolcskei H (2004) An overview of MIMO communications—a key to gigabit wireless, vol 92, no 2. IEEE, pp 198–218
2. Ashish N, Ramanathan R et al (2011) A novel channel estimation technique for MIMO-OFDM systems for frequency selective Rayleigh channel. In: International conferences on devices and communications
3. Chen A, Zong JL (2020) An optimized packet transfer through wireless network using age of information. IRO J Sustain Wirel Syst 2(2):76–83
4. Gopal J, Ramanarayan S, Kirthiga S et al (2018) Capacity analysis of correlated MIMO in GEOSAT downlink land mobile system. In: 7th international conference on advances in computing, communications and informatics, IEEE (ComSoC)
5. ITU-R Recommendation P.618-9: propagation data and prediction methods required for the design of Earth-space telecommunication systems, Geneva, Switzerland (2007)
6. Panagopoulos AD, Arapoglou PDM, Cottis PG (2004) Satellite communications at Ku, Ka and V bands: propagation impairments and mitigation techniques. IEEE Commun Surv Tutor
7. Arapoglou P-D et al (2011) To MIMO or not to MIMO in mobile satellite broadcasting systems. IEEE Trans Wirel Commun 10
8. Debaillie B et al (2015) RF self-interference reduction techniques for compact full duplex radios. In: IEEE 81st Vehicles Technology Conference VTC Spring, Glasgow
9. Yamuna B (2016) Reliability level List based decoding of multi level modulated block codes. Int J Inf Commun Technol 9
10. Reshma J, Ganesh HB, Barathi, Jyothi J, Gandhiraj R, Soman KP (2016) Wavelet based RTL-SDR real time denoising in GNU radio. Smart Innov Syst Technol 43

11. Mohan N, Sachin Kumar S, Soman KP (2019) Modern time -frequency techniques for signal analysis and its applications. In: Studies in computational intelligence. Springer
12. Collins GD, Treichler J (2015) Practical insights on full-duplex personal wireless communications gained from operational experience in the satellite environment. In: IEEE signal processing educational workshop (SP/SPE), Salt Lake City, UT
13. Grayver E, Keating R, Parower A (2015) Feasibility of full duplex communications for LEO satellite.IEEE Aerosp Conf
14. Beaulieu NC, Abu-Dayya AA (1995) Bandwidth efficient QPSK in co-channel interference and fading. IEEE Trans Commun 43
15. Safak A, Prasad R (1991) Effects of correlated shadowing signals on channel reuse in mobile radio systems. IEEE Trans Veh Technol 40
16. Ziqiang X, Akansu AN, Tekinay S (2000) Cochannel interference computation and asymptotic performance analysis in TDMA/FDMA systems with interference adaptive dynamic channel allocation. IEEE Trans Veh Technol 49
17. Brandao AL, Lopes LB, Mclernon DC (1994) Method for timing recovery in presence of multipath delay and cochannel interference. Electron Lett 30
18. Meyer R, Gerstacker WH, Schober R, Huber JB (2006) A single antenna interference cancellation algorithm for increased GSM capacity. IEEE Trans Wirel Commun 5
19. Salija P, Yamuna B (2017) Implementation of turbo code with early iteration termination in GNU radio. J Telecommun Electron Comput Eng 9

Comparative Analysis of Time Series Forecasting Algorithms

Deepankar Verma and Chhavi Rana

Abstract Data points that are listed according to time or the points which are sequential and equally spaced in time is known as time series. Examples are: rise and fall in stock prices, the height of the tides, flight passenger's information and medical information of the patients, etc. A lot of work has been done in the novelty discovery, and dependency detection. This kind of work exploits the fields of mathematics, statistics, machine learning and the use of information retrieval. The important key point here is the scalability of the time series data. For this paper, we have collected the patient's health attributes viz. temperature and heartbeat using sensors on IOT devices and then using this database as the input for Welles Wilder Average and Naïve forecasting algorithm. This will help us to further predict the patient's health and make a comparative analysis of the above-mentioned algorithms.

Keywords Welles Wilder · Naïve forecasting · IoT · Time series (TS) · Moving average (MA) · Forecasting

Introduction

Data mining can be defined as gleaning out information from a large volume of data or data points. In other words, we can define data mining as a procedure which can be useful for mining out knowledge from a large volume of data. The process of data mining has a wide range of application in today's world. In corporate world, data mining will help in finance planning, evaluation of assets, risk management, planning of resources. The hidden facts unveiled by the data mining techniques for instance may be used by a business organization for the purpose of increasing the sale of a particular product or for gaining the information regarding the behaviour of its customers towards their existing product. Similarly, the medical organizations may use the data mining for better diagnostics of the known and unknown diseases, forecasting the patients' health attributes which will be useful as preventive measure.

D. Verma (✉) · C. Rana
Department of Computer Science and Engineering, University Institute of Engineering and Technology, Maharshi Dayanand University, Rohtak, Haryana, India

V. Bindhu et al. (eds.), *Proceedings of Third International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 844,
https://doi.org/10.1007/978-981-16-8862-1_17

Data points that are listed according to time or the points which are sequential and equally spaced in time is known as time series. Examples are: daily opening and the closing of the stock prices, the height of the tides, flight passenger's information and medical information of the patients, etc.

Time Series Forecasting

Data points that are listed according to time or the points which are sequential and equally spaced in time is known as time series [1].

TS adds a dimension of time in between the observations. This time dimension acts both as a constraint and a structure (Fig. 1).

TS Components

1. Level (a): It is a value where demand flutters and apprehend scale of the TS. Its values are constant where no other pattern is present (Fig. 2).

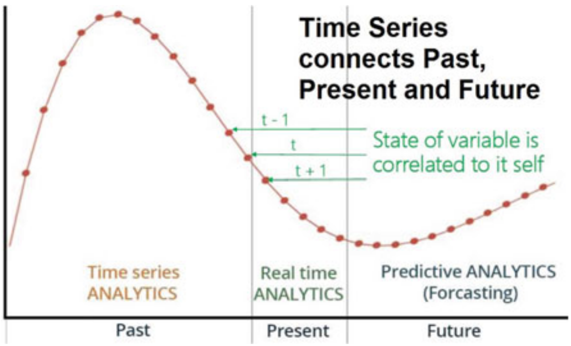


Fig. 1 Time series [2]

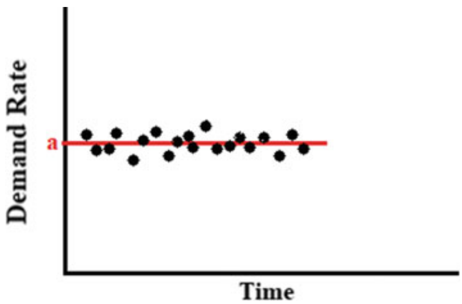


Fig. 2 Level

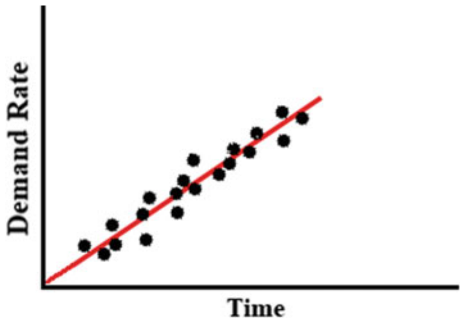
2. Trend (b): It is the rate of growth or decline. It is an incessant movement in one direction It can be an exponential pattern, a quadratic pattern but generally it a linear pattern which is optional in nature. We should note that it generally represents linearly increasing or a linearly decreasing behaviour of the series (Fig. 3).

   - Uptrend is the movement when higher highs are followed by higher lows.
   - Downtrend is the movement where lower lows follow lower highs.
   - Horizontal trend is the movement when no high highs or lower lows are produced.

     Statically, if we have one of the above trends then, it will follow the same trend [3] (Fig. 4).

3. Season Variations (F): When we find same cycle repeating again and again around a fixed point that is known to us then is called as season variation which

Fig. 3 Trends



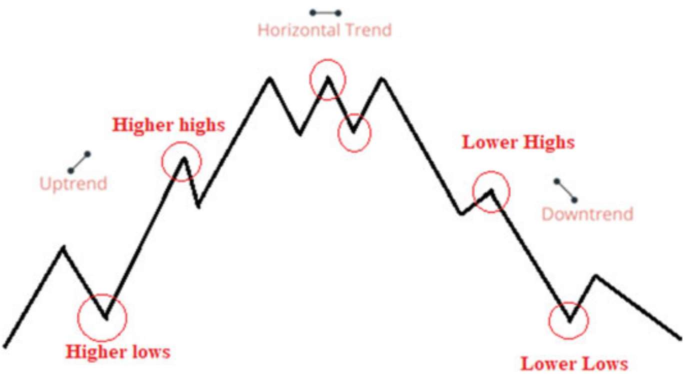


Fig. 4 Types of trends

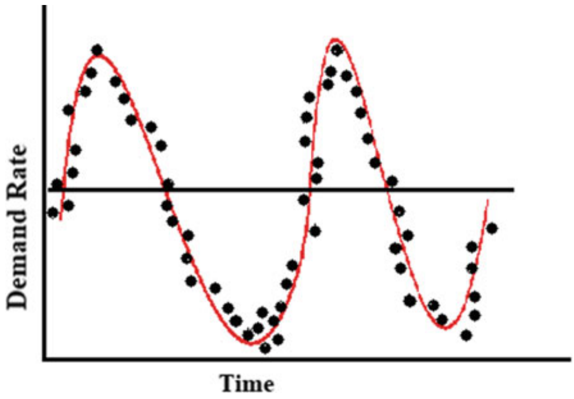


Fig. 5 Seasonality

can hour to hour, daily, weekly, etc. The causes for this can be man-made or natural. These kind of variations are also optional (Fig. 5).

4. Random Fluctuation or Noise (e or ε): It is the remainder of variability after other pieces. It can be explained as a model as they are irregular, random and erratic in nature.

These constituent components can be thought to combine in some way to provide the observed time series. For example, they may be added together to form a model as follows:

The above-mentioned components can be combines with each other in order to administer the TS. The simple example is given below.

$$y = \text{Summation}(\text{level, trend, seasonality, noise})$$

We can make assumptions about these pieces regarding their behaviour and how we combine them. This will allow us to model them using traditional statistical methods. Sometimes they are very effectual in predicting future values and when they are not, these pieces may still be useful [4].

TS Methods

See Fig. 6.

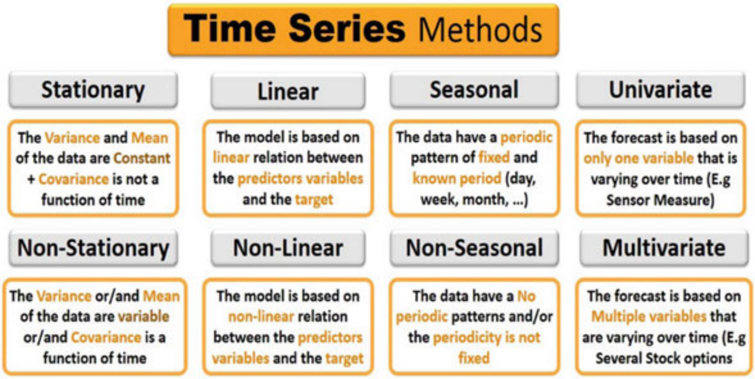


Fig. 6 Time series methods [19]

TS Forecasting Models

Autoregression (AR)

The next step in the sequence is modelled with autoregression method. Let us denote a function '*L*' which is linear in nature. It performs this function *L* of the data in previous steps of time.

Here, the specification of the model is done by its order '*p*' as a parameter to the autoregression function.

For example. AR(*p*), it is the *p*th order autoregression function.

When there is no trend and seasonal components, this is the good enough for univariate TS [5].

Moving Average (MA)

In the sequence, the next step is modelled with moving average method. Let us denote a function '*L*' which is linear in nature. It performs this function *L* of the remaining errors in previous steps of time.

Calculating the moving average of the time series is disparate from moving average model.

Here, the specification of the model is done by its order '*q*' as a parameter to the moving average function.

For example. MA(*q*), is a *q*th order moving average model.

When there is no trend and seasonal components, this is the good enough for univariate TS.

Autoregressive Moving Average (ARMA)

In the sequence, the next step is modelled with moving average method. Let us denote a function 'L' which is linear in nature. It performs this function L of the remaining errors in previous steps of time.

It is the combination of both the above methods that is autoregression and moving average.

Here, the specification of the model is done by its order AR(p) and MA(q) models which serves as the parameter to ARMA function which in term is represented as ARMA(p, q).

When there is no trend and seasonal components, this is the good enough for univariate TS.

Autoregressive Integrated Moving Average (ARIMA)

In the sequence, the next step with ARIMA method. It performs the linear function of the remaining errors as well as differenced observations at the previous steps of time.

It is the combination of both models discussed above that is AR and MA. It also performs the differencing of preprocessing steps. It is done to make a stationary sequence and is termed as integration (I).

Here, the specification of the model is done using the following terms that are AR(p), I(d) and MA(q) models which serves as the parameter to ARIMA function which in term is represented as ARIMA(p, d, q).

When there is presence of trend and there are no seasonal components, this is the good enough for univariate TS [5].

Related Scholarly Work

1. With the rise in the healthcare costs, disbursement on medications is needed to be managed [6].
2. Healthcare disbursement can be managed if future medication cost can be predicted [7].
3. Previous research has solicited TS forecasting of various healthcare costs assimilating anonymity [8, 9].
4. By assessing the predictions from the extanting forecasting methods with other methods, we can overcome the big challenge of model excerpt [10].
5. The proposal of long and short term recurrent machine learning methods [11].
6. ARIMA model is one of the known and prominent statistical TS model [12].

7. The ARIMA model is known and prominent because of its statistical properties, for example; moving averages, which are used in finding its parameters [13].
8. Comparative research and studies between ARIMA models and ANNs have been carried out previously by performing TS forecasting of the stock market data [14, 15].
9. The price of a particular product is taken as the parameter in univariate TS model which in turn influences the prognosis. The price is calculated with the use of historic sales data which is an easy approach as compared to alternative techniques [16].
10. A study had been carried out a sturdy and robust forecasting system developed for a large food distribution company. It was adopted back in November 2009 by the company [17].
11. Economical forecasting help to find the trends and pattern [18].
12. Recommender system and use of IoT sensors like smart watches, etc., helps in providing personalized preferences to the customers [19].

Basic Theory of Moving Averages

Study of Simple Moving Average

- SMA is an arithmetic moving average.
- It is calculated by adding latest data and then dividing it by the time period taken.
- The SMA is easy to construct.
- It may not be most accurate when compared with others.

Formula:

$$\text{SMA} = (D1 + D2 + \cdots + Dn)/n$$

where

Dn = data point at nth period of time

n = the number of total period.

Example:

We need to find 5 days SMA of the following data shown in the table. It is done using the above formula.

Day	1	2	3	4	5	6
Data value	16	17	17	10	17	18
5 day SMA					15.4	15.8

Study of Weighted Moving Average

- It is denoted by WMA.
- They are difficult to construct as compared to SMA.
- They are more reliable and accurate.
- A WMA less weights to old data and more weights to new and latest data.
- The weighting of the data that is how much wait should be given to each data is calculated from the sum of days.

The WMA is calculated by multiplying the given price or the input by their associated weights and then summing the values. Then the summation is divided by the sum of the days as shown is the example below.

Formula:

$$\text{WMA} = \{D1(n) + D2(n-1) + \cdots + Dn(1)\}/\{n(n+1)\}/2$$

where

$D1$ = The latest data point

Dn = The oldest data point

n = The time period.

For example: We need to calculate 5-day WMA of the given data input in the table. Sum days of 5(5 + 1)/2 = 15.

Day	1	2	3	4	5
Data value	16	17	17	10	17
Weight [n/{n(n + 1)/2}]	1/15	2/15	3/15	4/15	5/15
Weighted value (data value * weight)	1.07	2.27	3.40	2.67	5.67
5 day WMA					15.07

Study of Exponential Moving Average

- They are most reliable and accurate when compared with the primary moving averages.
- They provide an essence of weighting and with each previous day are given less and less weights with each progression.
- EMA smoothing avoids the problem that we face when we apply SMA.

- We can also determine their slopes easily because when price closes below the MA, the slope is always downwards and when the price closes above the MA, it is always in upward direction.

Exponential Moving Average Percentage

EMA% is explained as the weight that is attached to the current days value.

For example for 3 days EMA, EMA% = 50% and for 5 days EMA, EMA% = 33.3%

Formula used to calculate the EMA%

$$\text{EMA\%} = 2/(n + 1)$$

where n is the number of days.

Now, we need to calculate an EMA:

EMA(for nth day) = $D(n)$ * EMA% + EMA(for $n - 1$th day) * (1 − EMA%)

Where $D(n)$ is the input value for the Nth day.

Day	1	2	3	4	5
Data value	16	17	17	10	15
33.3% (or 1/3) EMA		16.3	16.5	14.4	14.6

Proposed Methodology

Welles Wilder Moving Average

J. Welles Wilder introduced many popular indicators in his book written in 1978 “New Concepts” in technical trading systems. These indicators are as follows:

i. Average True Range: It calculates the commitment by comparing the range for each successive day and hence is indicates the volatility. Ardor in the market is shown by contracting and expanding signals.
ii. Directional Movement System: It calculates the ability of bears and the bulls to move prices outside from the prior day’s range of the trading. It consists of 3 lines as listed below:

    - The Positive Direction Indicator (+*DI*) which is gives the upward trend movement of the market

- The Negative Direction Indicator ($-DI$) which is gives the downward trend movement of the market
- The Average Directional Movement Index (ADX) represents whether the market is ranging or trending.

iii. Relative Strength Index: When we need to compare upward movements in closing prices to the downward movements, we use this indicator. It is calculated over a selected period of time.
iv. Twiggs Money Flow: Using Wilder's moving average formula, Colin Twiggs developed this indicator which provides useful information about the trend and warn of breakouts.
1. Wilder does not use the standard exponential average formula which have compelling brunt when we are selecting convenient time periods for these indicators.

When we are using standard formula of EMA, it converts the time period using the formula EMA% $= 2/(n+1)$, where n is the given period or the number of days. This is also known as the smoothing factor of Exponential Moving Average, let us denote it with 'a'.

For example, when we take $n = 14$ that is the time period as 14 days then the EMA% for 14 days is equal to 13.3% using the above formula.

However, Wilder, uses the smoothing factor $a = 1/n$.

For example, a 3 period Wilder's smoothing is the same as a 5 period exponential moving average.

Welles Wilder's Moving Average Formula

$$\mathrm{WMA}_t = a * \mathrm{Inputt} + (1 - a)\mathrm{WMA}_{t-1}$$

where,

WMA_t = Wilder Moving Average at time t.

a is the smoothing factor and it is equal to $1/n$, where n is the period.

Input_t = input value at time t. This input can be data input as we used for our thesis work or can be one of the Welles Wilder indicators.

Naive Forecasts

1. It is the operation of statistical techniques which is used for historical observations. It is a mechanical and objective. It does not require any subjective opinion. This model produces volatile results that changes drastically because it considers that latest data points and hence it is faster and responsive.
2. It is built on a simple idea that the future events can be related to the past events.

3. This forecasting is useful as criterion for comparing more arduous methods as it is simple. It is useful in another way as its data requirement is minimum. It is also useful for the variables that are relatively stable or which changes in uniform way therefore is it applicable in forecasting of short runs.
4. The drawback is, when the variables are fluctuating too much them it tends to show really inaccurate results as it is based on the previous inputs for forecasting and high fluctuations results in high errors so it cannot forecast turning points.
5. Naïve Model: The foreseen level of the variable in the ongoing and latest period is the same as the observed level in the previous period.

$$y_{t,t+1} = y_t$$

where y_t: actual data in period t and $y_{t,t+1}$: forecast for time $t + 1$ that will be made during time t.

Dataset Preparation

Data of temperature and heartbeat of patients has been taken using IoT sensors DS18B20 and SEN-11570, respectively.

1. **Temperature Sensor**: The model DS18B20 is chosen for the collection of temperature dataset of the host. 9-bit to 12-bit Celsius temperature measurements are provided by the DS18B20 digital thermometer provides. It also has an non-volatile alarm function which is user-programmable to upper and lower trigger points. It eliminates the need for an external power supply by directly deriving power from the data line. This sensor is very easy to work with and provides with accurate data of temperature. The sensor provides data originally in the Celsius temperature scale but we can convert the scale to Fahrenheit as per our convenience. When the host touches the sensor of the device it gets saturated with his/her body temperature and shows the results aka body temperature.
2. **Pulse Sensor**: The model SEN-11574 is chosen for the collection of heartbeat dataset of the host. It is a plug-and-play heart-rate sensor. It can be used by anyone who wants to easily measure their live heart-rate data. It is an integrated optical amplifying circuit sensor. It also eliminates the noise. It is very easy to use by clipping the Pulse Sensor to the hosts earlobe or fingertip and plug it into the Arduino. This sensor can give data like Pulse Rate, Cardio Graph and Inter Beat Interval. However, for our convenience we incorporated the Pulse Rate per minute. The ear lobe or the fingertip of the host can be used on the sensor and heartbeat can be retrieved.

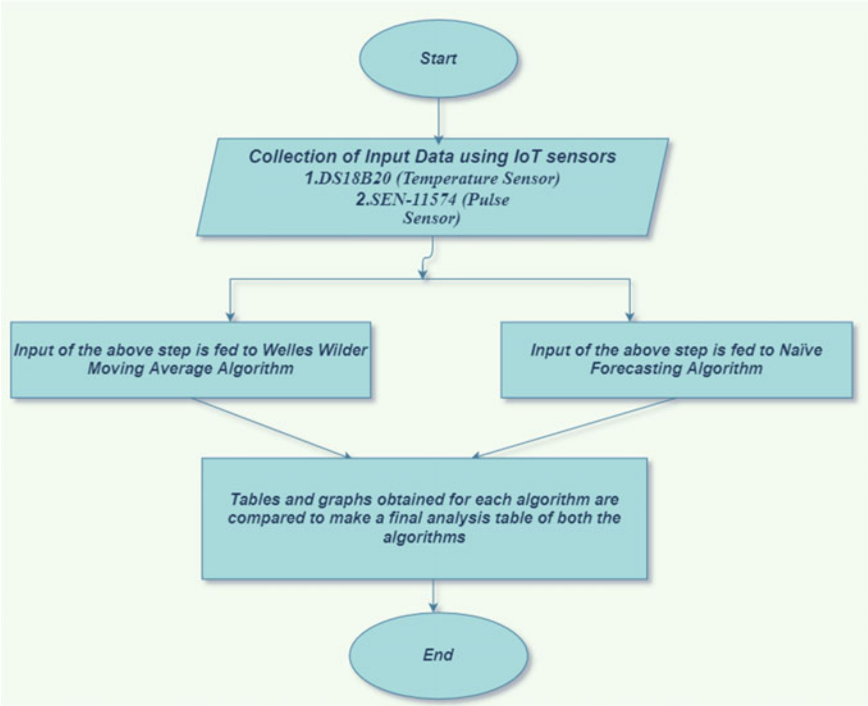


Fig. 7 Flow chart of the work

Results and Analysis

Results

Pateint A: Temperature (Figs. 8 and 9)

Pateint A: Heartbeat (Figs. 10 and 11)

Analysis of Both the Techniques

See Table 1.

Temperature With Naive

Instance	Value	Forecast	Error
1	98	98	0
2	95	98	-3
3	91	95	-4
4	93	91	2
5	93	93	0
6	98	93	5
7	99	98	1
8	98	99	-1
9	98	98	0
10	97	98	-1
11		97	

Temperature With WellesWilder

Instance	Value	Forecast	Error
1	98		
2	95		
3	91		
4	93		
5	93	94	-1
6	98	94.8	3.2
7	99	95.64	3.36
8	98	96.112	1.888
9	98	96.4896	1.5104
10	97	96.59168	0.40832

Fig. 8 Patient temperature table

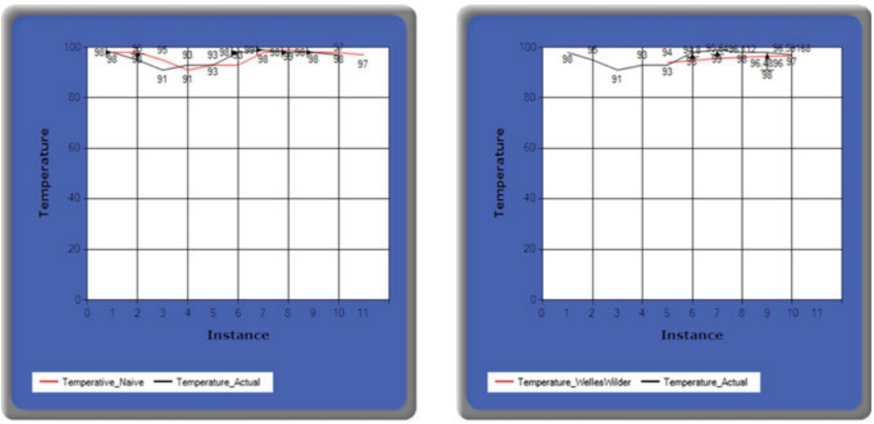


Fig. 9 Patient temperature graph

Conclusion and Future Scope

We conclude that Naïve forecast is based on the concept that what will happen in the future can be related to the past event and it is simple to use to use when data

Heart Beat with Naive

Instance	Value	Forecast	Error
1	78	78	0
2	74	78	-4
3	75	74	1
4	76	75	1
5	85	76	9
6	72	85	-13
7	75	72	3
8	78	75	3
9	85	78	7
10	90	85	5
11		90	

Heart Beat with WellesWilder

Instance	Value	Forecast	Error
1	78		
2	74		
3	75		
4	76		
5	85	77.6	7.4
6	72	76.48	-4.48
7	75	76.184	-1.184
8	78	76.5472	1.4528
9	85	78.23776	6.76224
10	90	80.590208	9.409792

Fig. 10 Patient heartbeat table

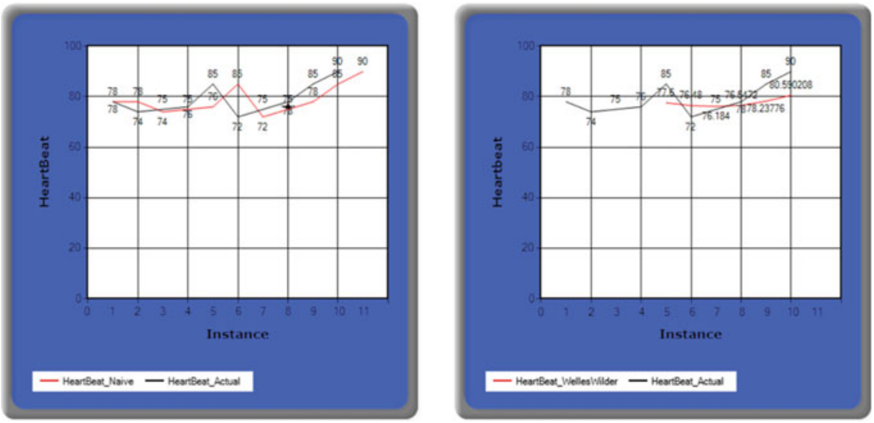


Fig. 11 Patient heartbeat graph

Table 1 Comparison of Welles Wilder and Naïve forecast

S. No.	Parameters	Welles Wilder moving average	Naïve forecasting
1	Smoothing factor	It used the concept of smoothing factor and its smoothing factor (a), $a = 1/n$ where n is the period	It does not use smoothing factor concept
2	Period	It uses the concept of periods which helps in calculating smoothing factor. More the period, more the forecasting curve sticks with the actual curve. Generally for short term, we use short time period	It does use the period concept
3	Formula used	$\mathrm{WMA}_t = a * \mathrm{Input}_t + (1 - a)\mathrm{WMA}_{t\text{-}1}$ where WMA_t = Wilder Moving Average at time t a is the smoothing factor and it is equal to $1/n$, where n is the period Input_t = input value at time t. This input can be data input as we used for our thesis work or can be one of the Welles Wilder indicators	$y_{t,t+1} = y_t$ where y_t: actual data in period t $y_{t,t+1}$: forecast for time $t + 1$ that will be made during time t
4	Forecasting at time t	It forecasts or we can say calculates the wilder moving average for time t	At time t its forecasts for time $t + 1$ and its value is the x_t where x_t: actual data in period t
5	Usage	It provides support and resistance to the actual curve It is used to predict the future trends in different fields like health, financial, etc	It is built on a simple idea that the future events can be related to the past events. This forecasting is useful as criterion for comparing more arduous methods as it is simple. It is useful in another way as its data requirement is minimum. It is also useful for the variables that are relatively stable or which changes in uniform way therefore is it applicable in forecasting of short runs

requirement is minimum. It also provides benchmarks for comparing more complicated method. On the other hand, Welles Wilder Moving Average provides support and resistance to the actual input graph and curves. It is beneficial in predicting the future trends in different fields like health: to predict the health attributes of patients and to take precautionary measures if needed, finance: to find the trends in the market and to analyse the stocks. It can further be used in predicting crimes and to reduce them, to mitigate risks, and seize new opportunities.

Future research in the area is related to find the best features and dimensions of time series data. We can also try to improve the existing forecast algorithms or to make new algorithms which will be better and more prominent to the current one's in speed, size, accuracy and requiring minimal data. So, our different sectors especially the health sector can improve in a far better way. This will not only decrease the mortality rate but will also increase the life expectancy of people. Also, large data sets can be implemented to make better forecasting and results can be shown in the forms of graphs, pie charts, bar graphs, etc.

References

1. Fenjiro Y (2019) https://medium.com/@fenjiro/time-series-for-business-ageneral-introduction-50968346e660. Accessed on 6 June 2021
2. https://medium.com/@fenjiro/time-series-for-business-a-general-introduction-50968346e660
3. https://machinelearning-blog.com/2018/01/09/time-series-forecasting
4. Brownlee J (2016) What is time series forecasting? https://machinelearningmastery.com/time-series-forecasting/. Accessed on 5 June 2021
5. Brownlee J (2016) 11 classical time series forecasting methods in Python. https://machinelearningmastery.com/time-series-forecasting-methods-in-python-cheat-sheet/. Accessed on 5 June 2021
6. Bertsimas D, Bjarnadóttir MV, Kane MA, Kryder JC, Pandey R, Vempala S et al (2008) Algorithmic prediction of health-care costs. Oper Res 56:1382–1392. https://doi.org/10.1287/opre.1080.0619
7. Zhao Y, Ellis RP, Ash AS, Calabrese D, Ayanian JZ, Slaughter JP et al (2001) Measuring population health risks using inpatient diagnoses and outpatient pharmacy data. Health Serv Res 36:180
8. Ash AS, Ellis RP, Pope GC, Ayanian JZ, Bates DW, Burstin H et al (2000) Using diagnoses to describe populations and predict costs. Health Care Finance Rev 21:7–28
9. Zhao Y, Ellis RP, Ash AS, Calabrese D, Ayanian JZ, Slaughter JP et al (2001) Measuring population health risks using inpatient diagnoses and outpatient pharmacydata. Health Serv Res 36:180
10. Makridakis S, Spiliotis E, Assimakopoulos V (2018) Statistical and machine learning forecasting methods: concerns and ways forward. PLoS ONE 13, e0194889. https://doi.org/10.1371/journal.pone.0194889
11. Gamboa JCB (2017) Deep learning for time-series analysis. arXiv:1701.01887
12. Newbold P (1983) ARIMA model building and the time series analysis approach to forecasting. J Forecast 2:23–35. https://doi.org/10.1002/for.3980020104
13. Box GE, Jenkins GM, Reinsel GC, Ljung GM (2015) Time series analysis: forecasting and control. Wiley, San Francisco, CA
14. Jain A, Kumar AM (2007) Hybrid neural network models for hydrologic time series forecasting. Appl Soft Comput 7:585–592. https://doi.org/10.1016/j.asoc.2006.03.002
15. Adebiyi AA, Adewumi AO, Ayo CK (2014) Comparison of ARIMA and artificial neural networks models for stock price prediction. J Appl Math 7:1–7. https://doi.org/10.1155/2014/614342
16. Schaidnagel M, Abele C, Lauxy F, Petrovy I (2013) Sales prediction with parameterized time series analysis. In: The fifth international conference on advances in databases, knowledge, and data applications (DBKDA)

17. Wagner N, Michalewicz Z (2011) Intelligent techniques for forecasting multiple timeseries in real world systems, IJICC
18. Jha GK, Sinha K (2013) Agricultural Price forecasting using neural network model: an innovative information delivery system. Agric Econ Res Rev 26(2):229–239
19. Joe CV, Raj JS (2021) Location-based orientation context dependent recommender system for users. J Trends Comput Sci Smart Technol (TCSST) 3(01):14–23

Binary and Multiclass Classifications Using a Deep Fusion Network

Amay Gada, Russel Lobo, Dhruv Bohara, Dhruvi Jodhawat, and Pratik Kanani

Abstract This paper describes a method for providing a unified model from multiple classification-based datasets. Dilated convolutional neural networks are used for applications that require high accuracy, speed, and real-time processing. This paper proposes a classifier that is accurate for both binary classification and multiclass classification. The model employs a number of techniques, including batch normalization, ReLU activation function, dilation blocks, and learning rate optimizations, and has provided us with a model with reasonable accuracy for classification-based datasets. The proposed model has been trained on four different classification-based datasets by including both binary classification and multiclass classification. Further, the proposed model is tested on malaria dataset, WBC dataset, driver drowsiness detection dataset, and handwritten symbol classification dataset. The experimental results provide high accuracy and elasticity.

Keywords Convolutional neural networks · Dilation rates · Batch normalization · Deep dilation network

Introduction

Various classification models are constructed based on the given application. However, this paper focuses on creating a general model that results in a reasonable accuracy for the multiple datasets obtained from different domains. The concept of varying dilation network is used here to provide the model a much better and varied view of the image it encodes. Working excellently in the domain of image-driven pattern recognition, the model can be used in various fields from detecting diseases in the health sector to identify human perception. To prove this, it has been tested on various publicly available datasets ranging from malaria-infected cells to handwritten math symbols.

A. Gada (✉) · R. Lobo · D. Bohara · D. Jodhawat · P. Kanani
Synapse-Computer Engineering Department, DJSCE, Mumbai, India

V. Bindhu et al. (eds.), *Proceedings of Third International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 844,
https://doi.org/10.1007/978-981-16-8862-1_18

Adapting to the real-time requirements of multiclass classification, the model also supports datasets with more than two classes. Out of the four datasets used here, two are binary classification problems, and two are a multiclass classification problem. The binary classed malaria dataset uses thin blood smears to identify if a cell is infected or not. The drowsiness dataset similarly uses a series of zoomed in images of the human eye in order to classify them into two classes: open and close. To test the model's inference on multiclass datasets, white blood cell images and hand written math symbols are used as datasets. The white blood cell classifies images into eight classes, while the handwritten math symbol dataset has 16 classes.

The model comprises four mini-models that ensemble to give a final model. Each model begins with a 256 $\times$ 256 sized original image which then branches into different convolution layer chains to give varying results. A special maximum dilation block is used to encapsulate the chain of convolution operations that involve varying the dilation rate. Finally, these models are spatially concatenated to form the final model (refer methodology for more information about the model).

A. Literature review

Before evaluating how the model performs on various datasets, the working of other models on these similar datasets was looked upon. An end-to-end deep learning-based model is proposed in [1], for the diagnosis of malaria using microscopic images of blood samples from a thin slide. This model makes use of dilation nets. It varies the dilation rates of convolutions to include characteristics from a large spectrum and extract features from different receptive areas in the image. After resampling, a DeepFusionNet is used to carry out a joint optimization of these spatial features, and the most relevant representation of the sample image is obtained. Classification done by this architecture is highly accurate. [2, 3] relied on physical factors for extracting the features, and further, the support vector machines and principal component analysis (PCA) techniques are applied for classification. Modern techniques relying on deep learning will outperform these models. Hence, the accuracy of these models is very low. Moreover, the model uses over 19.6 billion floating point operations, which makes it very difficult to run on the power constrained devices. Whereas, [4] uses a ten-fold cross-validation layer of convolutional neural networks. The fine-tuned CNN has provided a high accuracy rate when compared to the other CNN models. [5] uses AlexNet and VGGNet frameworks to classify blood smear images into two classes. By comparing the CNNs of different internal depths, it was found that the VGGNet is better at extracting when compared to AlexNet.

Several attempts have been made in past years of handwritten mathematical symbol recognition by utilizing deep learning. In 60s, Anderson made an attempt for performing handwritten mathematical expression recognition. In [6], researchers have used CROHME 2013 database and trained the model with the combination of bidirectional recurrent neural networks (RNNs) and LSTM, which leads to a BLSTM-RNN classifier. They used PRHLT, RWTH, and POLAR features for offline handwritten features with 93% recognition rate, but faced a problem in discrimination of symbols with very similar shape. [7] used MNIST dataset of offline handwritten digits ranging from 0 to 9. To extract feature of handwritten image, they used BWEULER function in MATLAB, which returns an Euler number for binary image. They demonstrated the accuracies based on various methods in which decision tree had the least accuracy and KNN with the most, but they faced difficulties due to errors in raw dataset. In [8], researchers have used CNN by integrating pooling and softmax with CROHME 2014 dataset and trained it with different input image size, whereas still, 12.28% was misclassified by this model structure, and it was more time-consuming. From [9], it is observed that the proposed support vector machine model [SVM] has obtained 83.61% accuracy, while [10] used generalized linear vector quantization (GLVQ) classifier and achieved 76.24% accuracy, but both have used CROHME 2014 dataset. In [11], the proposed model has achieved 79.11% accuracy by using fuzzy shape context and artificial neural neuron (ANN) with CROHME 2014 dataset. In [12], the researchers have included the symbol segmentation and accurate classification for over 300 classes by segmenting with simple linear iterative clustering (SLIC) with LNet and SqueezeNet neural network methods. They obtained 90% accuracy by transfer learning of pretrained SqueezeNet. They used MNIST dataset with 8000 images that includes images from inkML of the CROHME dataset, and a dataset was published by the Computer Vision Group of the University of Sao Paulo.

Classification of blood cells is done to illustrate the count of different blood cells in one's body. Excess or deficiency of any blood cell helps us to detect the disease. Various techniques have provided reasonable results for classification. In [13], various architectures like AlexNet, VGG 16, GoogLeNet, ResNet CNN have been trained and tested with both full and transfer learning. Best result is shown by the Resnet50 model, which is trained by using transfer learning. In [14], researchers have proposed a segmentation algorithm, which finds a discriminating region of white blood cell tones in the HSI color space. Otsu's thresholding method was used to segment the WBCs in [15]. Elements having no resemblance with white blood cells were eliminated by using mathematical morphological operations. The features were then extracted from the cell nucleus and trained by using Naive Bayes classifiers. Then, in [16], researchers have used k-means clustering for segmenting white blood cells from the acquired images and using principal component analysis (PCA), performed feature extraction and selection for classification. In [17], researchers have proposed the fast relevance vector machine (F-RVM) for the segmentation and classification of white blood cells. Fast relevance vector machine (F-RVM) is better than extreme learning machine (ELM) and standard RVM as it is very easy to train, and it consumes less time to make predictions. Although the promising results were obtained from the abovementioned works, a small testing and training dataset was used. Many of the works only relied on 50–60 images per class for training and testing the proposed models. The diversity in the acquired images of white blood cells in real life is a lot more and can often render these models trained on small datasets as ineffective.

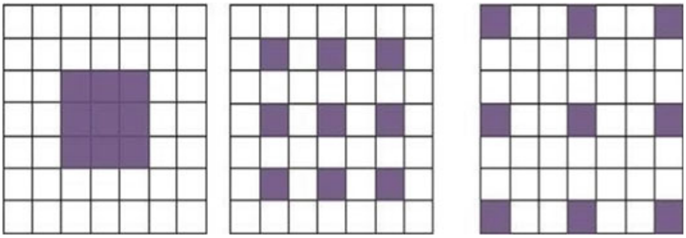
The model can detect driver drowsiness by the behavior of the driver that majorly includes eye closeness detection. In [18], the proposed model has used CNN and FD-NN, VGG16 and VGG19 for eye status classification with ZJU dataset and their extended dataset with a good accuracy, but faced issues due to the lack of a good dataset in the area of detection of eye closure. It shows how superior FD-NN is over other proposed networks for quickness of real-time tasks. Convolutional neural networks, being able to only learn features from a still image of the driver, results in unnecessary and repetitive information being captured. Hence, the facial behavior information, which is very important, is not captured efficiently. In [19], the researchers have used maximum closure duration (MCD), eyelid closure (PERCLOS), blink frequency (BF) features with SVM methods and reported an appreciable accuracy. [20] used a neural network with Gabor wavelet transformation for extraction of the lighting and gloss on face and eyeglasses of the driver can be a problem while detecting drowsiness. Hence, in [21], the proposed research work has implemented a hardware system which relied on infrared light to tackle this problem. They proposed yawning as another drowsiness symptom. In [22], the proposed model has used percent of the time eyelids are closed (PERCLOS) metrics to measure eye drowsiness. This method has provided them with a good accuracy. In [23], researchers have used KNN algorithm and recorded 134 h of data during three driving simulator studies for testing and ended up with 84.2 and 70.0% accuracy in the binary and multiclass classification. The depth of CNN network has been increased to convolve more the input data and achieve more accurate results by adding layers in traditional methods.

The problem in conventional use of dilated convolution was pointed out by the research work carried out in [24]. To improve that, they used a local feature extraction module above a dilated module to deal with smaller and crowded objects in remote sensing imagery with decreasing the dilation factor. They acquired good accuracy with testing on three different remote sensing datasets. In [25], a CNN module is specifically designed for performing dense prediction. They proposed that dilated convolutions support exponential expansion of the receptive field without any loss in resolution. They examined image classification adaptation on dense prediction and received a good accuracy. [26] has used dilated convolutions to perform multi-scale context aggregation. This was done without down sampling the resolution of the image. Hence, dilation networks are used in semantic segmentation methods due to their ability to capture large context while preserving fine details of the image.

Dilation Rate

A convolutional layer is defined by three major properties such as the kernel, stride, and padding. The kernel is a matrix that is used to perform convolutional operations over the image. The stride defines the number of steps the kernel takes while performing these convolutions. A convolutional operation will reduce the size of the image matrix if the padding is not set to "same." Padding is used in order to maintain a certain size of the matrix. In this paper, the padding is set as "same" and has used strides to reduce matrix dimensions.

One of the optional parameters used here is the dilation rate. A dilation rate of 1 is like having a normal convolutional layer without any dilation effects. A dilation



(1) (2) (3)

Fig. 1 shows the effect on convolution operation for dilation rates 1, 2, and 3, respectively

rate greater than 1 performs convolutions in such a way that it obtains a wider view of the image while using a lesser number of parameters. A dilation rate of 2 on a 3 × 3 kernel would perform a convolution similar to a 5 × 5 kernel, while using only nine operations. This makes it helpful when a wider field of view is needed, but it cannot afford a lot of computation that comes along with larger convolution kernels (Fig. 1).

$$(F_{lk}^{*})(\boldsymbol{p}) = \sum_{s+lt=\boldsymbol{p}} F(s)k(t) \tag{1}$$

Methodology

The architecture proposed in this paper is quite extensive. Before moving into the details, this section discusses about the preprocessing stages before training them.

A. *Preprocessing*:

All pixel values are normalized by dividing each of them by 255; batches of 16 images are created for batch processing; all images are resized to 256 × 256 × 3, and finally, the prefetch of the Tensorflow framework is leveraged to obtain a quicker access to the next batch.

B. *Batch normalization*:

Mini-batch gradient descent is often used while training the models in order to speed up the learning process. It is recommended that the batch size be in powers of two. Along with the mini-batch gradient descent, batch normalization is often used. Usually, only the inputs are normalized while training the model, but when using batch normalization, each output layer is also normalized before passing them through the activation function [27].

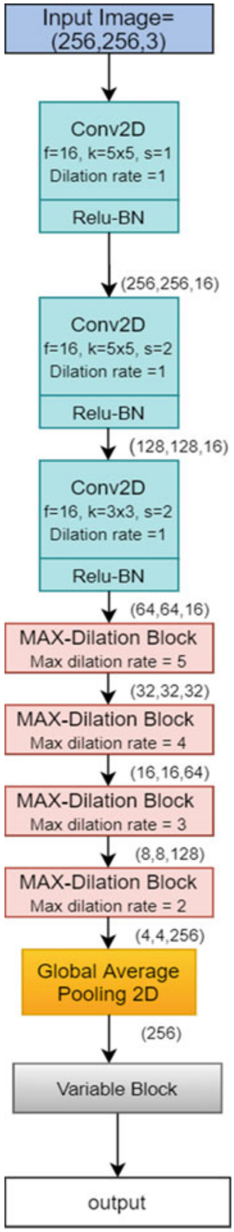
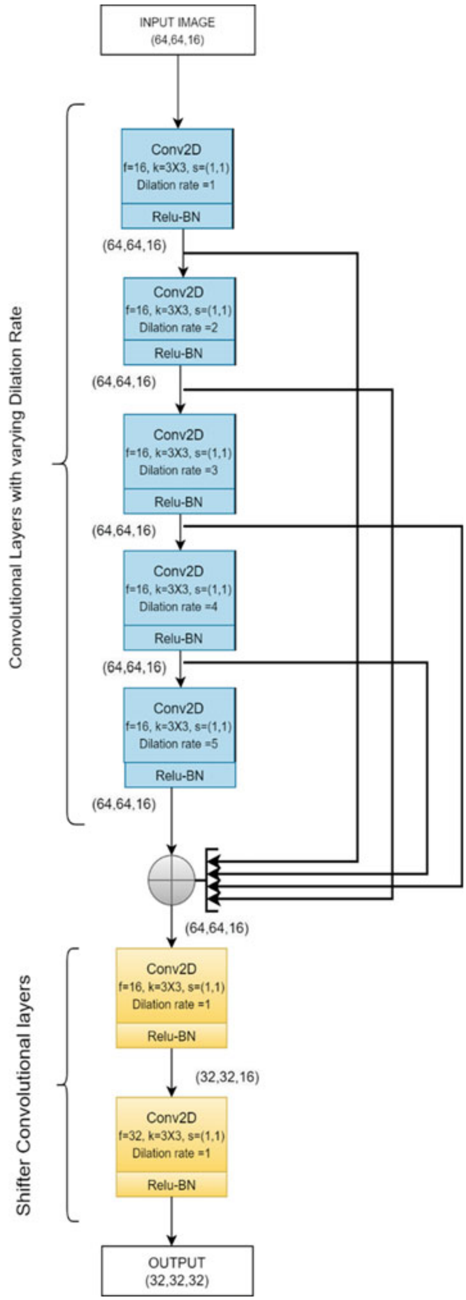


Fig. 2 Model1-architecture

Fig. 3 Figure describing a maximum dilation block for a max dilation rate of 5



$$\overline{x_I} = \frac{x_I - \mu_B}{\sqrt{\sigma_B{}^2 + \epsilon}} \tag{2}$$

$$y_I = \sqrt{x_I} + \beta \tag{3}$$

$\overline{x_I}$ represents the normalized x_1.

y_I represents the new output layer value before passing it through the activation function.

γ and β are both trainable parameters.

Neural networks suffer with the problem of covariant shift. Let's assume that the model is trained on an input set X, the distribution of X will suddenly change. So, there would raise a need to retrain the model as the model will not perform well on this changed input X. This change in distribution is called covariant shift. Similarly, in hidden layers, when the weights get updated, it can be seen as a change in distribution of the weights which will be used for further layers. This brings uncertainty in the network, and hence, batch normalization is used. Batch normalization normalizes the activations which in turn reduces how much that activation's distribution shifts around. This can help speed up learning [27] (Figs. 2 and 3).

In this paper, batch normalization has been done after each layer before being passed through a ReLU activation function.

C. *Activation functions*:

Activation functions decide whether a neuron should be activated or not. They even help in introducing non-linearity in the output.

In this paper, ReLU activation function has been made use of extensively for all the hidden layers.

$$f(x) = x^+ = \max(0, x) \tag{4}$$

For the output layers, the model is concatenated with sigmoid and softmax activation functions depending on whether it needs to do a binary classification or a softmax classification.

Sigmoid:

$$\sigma(x) = \frac{1}{1 + e^{-x}} \tag{5}$$

Softmax:

$$\sigma(z)_i = \frac{e^{z_i}}{\sum_{j=1}^{k} e^{z_j}} \tag{6}$$

$$\text{for} i = 1, \ldots, K \text{and} z = (z_1, \ldots, z_K) \in R^K$$

D. *Modeling*:

The 256 × 256 image size gives a very clear and sharp view of the dataset to the human eye. It is even a sweet spot considering the convolution computations that come with a larger image size. The model is branched out to four different models starting from a 256 × 256 × 3 image. Convolution operations are applied on the 256 × 256 × 3 to bring the size to 128 × 128 × 16 before branching out four different models.

This is done as the 256 × 256 image being of a better resolution helps the initial few layers to extract important information. The model architecture diagram (Fig. 3) is self-explanatory. The four models are trained individually with varying blocks composed of layers described in the result section of this paper. The dense layers are then clipped off these individual trained models and are concatenated to form a single one-dimension vector. The remnants of the clipped models are not trained anymore and only the variable block which is composed of a series of dense layers as described in the result section of this paper is trained to give a final concatenated and better model.

The important thing to discuss here is the maximum dilation block that can be seen in the model. Figure 3 shows how the maximum dilation block works given the maximum dilation rate is 5. Considering the same diagram, it can be seen that the tensor is passed through five different convolutions each with an increasing dilation rate till the maximum dilation rate is reached. Moreover, the output corresponding to each dilation rate is added element wise to get a final 3-dimensional tensor. This step is particularly necessary as information learned by each of these layers is considered. A layer which barely learned anything will not contribute during addition, and a layer which learned something very important will contribute heavily during addition. Over and above this, using dilations allow us to create a comparatively deeper and a stronger network as they give us a wider view of the pixel matrix, while not having to increase the number of computations. Finally, the maximum dilation block also plays a role in reducing the image dimension as the network grows in size. It must be noted that the maximum dilation block is highly inspired by a recent paper by Mahmud et al. [1].

The proposed model is tested on various datasets for binary and multiclass classifications. The model gives a good response when trained on them. Detailed explanations have been given in the results section of this paper (Fig. 4).

Results

A. *Experimental Setup*

All models have been trained by using a 2 GB NVIDIA GeForce MX13 GPU. All of the modeling and dataset segmentation along with preprocessing is done on a DELL latitude laptop with 8th generation i7 Intel processor and 16 GB RAM.

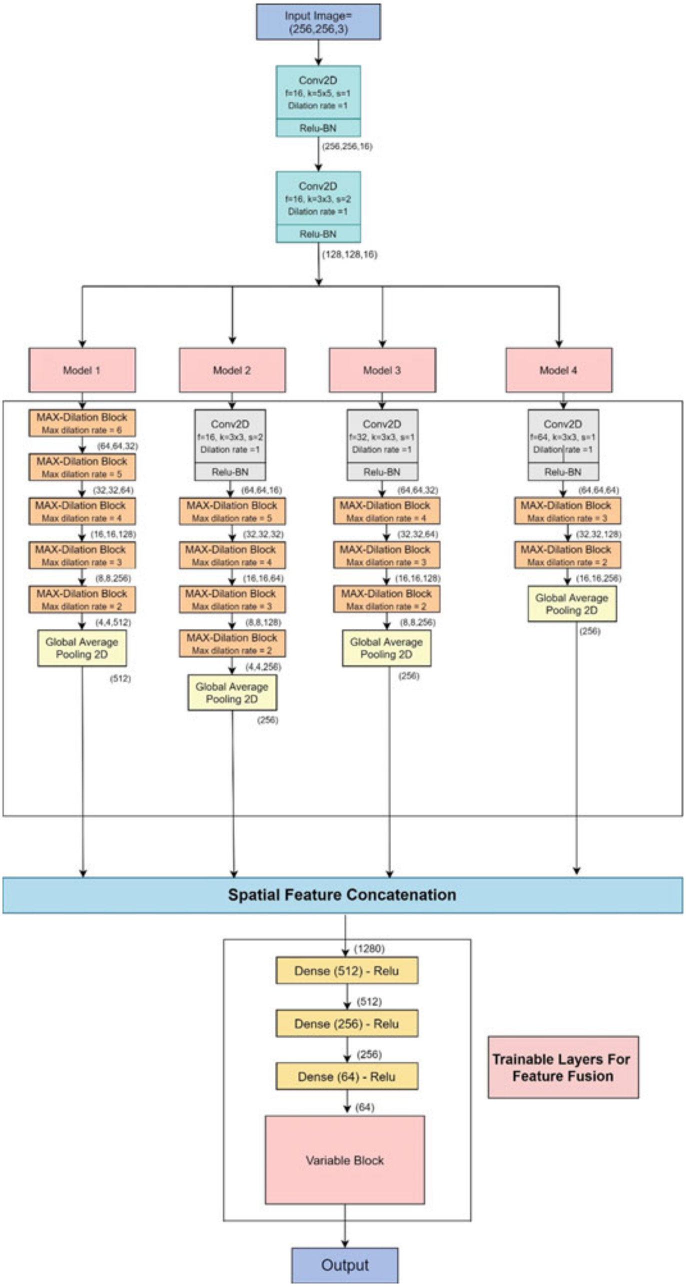


Fig. 4 Figure describing a maximum dilation block for a max dilation rate of 5

B. *Performance Analysis*

For each dataset that the model has been trained on, the results have been organized in the given format. Firstly, the dataset is explored by providing the details about the size, number of classes and dataset splitting information. After that any possible hardships caused by the dataset are discussed along with proposed solutions. Finally, the prediction results are discussed. The prediction results include the results from all the five models of the entire deep fusion network. Model 1, model 2, model 3, and model 4 are the branched mini-models, while model 5 is the spatial concatenation of all the abovementioned models. The results are ended after mentioning the prediction time for the given test set.

Binary classification

Malaria Dataset

The proposed dataset involves 27,560 images of extracted red blood cells, which are further divided into two classes.

The two classes include 13,780 images of infected cells and 13,780 images of uninfected cells.

Dataset split has been done such that the train data size is 18,000; validation data size is 4000, and test data size is 5560.

Prediction Results:

Considering the binary classification, the result is presented in a confusion matrix. The variable block used here after the 256-dimension layer is a 128-dimension layer, followed by a 64-dimension layer and a single dimension prediction node with the sigmoid activation function (Table 1).

Therefore, concatenating all the smaller models will provide a stronger classifier, which is called as the final model.

The final model can classify 5560 images in 222 s.

Drowsiness Dataset

This dataset involves 48,000 images of the human eye, which is further divided into two classes.

Table 1 Results corresponding to the malaria dataset in percentage

	Model4	Model3	Model2	Model1	Final
Accuracy	95.5	94.9	95.6	95.8	96.1
Precision	93.2	93.6	95.6	94.7	94.9
Recall	98.1	96.2	95.6	97.1	97.4
F1 score	95.6	94.9	95.6	95.9	96.1

Note that ReLU is used everywhere else as an activation function.
All results are mentioned in percentage

Table 2 Results corresponding to the drowsiness dataset in percentage

	model4	model3	model2	model1	final
Accuracy	97.9	98.7	99.1	98.9	99.6
Precision	99.8	98.7	99.1	99.8	99.7
Recall	96.1	98.8	99.1	97.9	99.6
F1 score	97.8	98.7	99.1	98.8	99.6

Note that ReLU is used everywhere else as an activation function.
All results are in percentage

The two classes include 24,000 images of open eyes and 24,000 images of closed eyes.

Dataset split has been done such that the train data size is 18,000; validation data size is 4000, and test data size is 5560.

Prediction results:

Considering the binary classification, the result is presented in a confusion matrix.

The variable block used here after the 256-dimension layer is a 128-dimension layer, followed by a 64-dimension layer with a single dimension prediction node and the sigmoid activation function (Table 2).

The final model leverages an outstanding performance for the drowsiness detection dataset. It barely overfits by giving a very good model with an accuracy of over 99%.

The final model is capable of classifying 8000 images in 336 s.

Multiclass Classification

WBC Dataset

This dataset has been slightly troublesome to work with. Each type of white blood cell was surrounded by different kinds of smaller cells and blood components. This has confused the very sensitive dilation networks as they work in such a way that they consider the tiniest and the largest detail depending on the dilation rate.

Initially, a maximum accuracy of 85% was achieved, and hence, it was decided to give the dilation net only what it required. Fortunately, the cells were centered in the image by making it possible to write a generalized cropping function, which would focus on the white blood cell type more than the background. This has improved the results. Let's look at them in more detail:

Dataset involves 17,092 images of different white blood cells, which are divided into eight classes.

Dataset split has been done such that the train data size is 1000; validation data size is 100, and the rest is test data. Symmetry and fairness were ensured while deciding the training and validation dataset size.

Table 3 Results corresponding to the WBC dataset in percentage

	Model4	Model3	Model2	Model1	Final
Accuracy	97.7	97.1	94.8	96.3	97.5

Note that ReLU is used everywhere else as an activation function.
All results are in percentage

Prediction results:

The variable block used here after the 256-dimension layer is a 128-dimension layer, followed by a 64-dimension layer which is followed by an 8-dimension prediction layer with the softmax activation function (Table 3).

The concatenation of the smaller models does not improve the final accuracy. Instead, it gets reduced. Trying a variety of learning rates doesn't seem to help. Adding a 128-dimension layer increases the accuracy slightly, but it doesn't improve the model completely. Hence, the model 4, which is used here, provides the best results.

The model 4 takes 76 s to predict 1037 white blood cell images.

Math Symbol Dataset

This dataset is comparatively easier as it includes symbols on a white surface.

Dataset involves 9000 images of different mathematical symbols divided into 16 classes.

Dataset split has been done such that the train data size is 650; validation data size is 30, and the rest is test data. Symmetry and fairness were ensured while deciding the training and validation dataset size.

Prediction results:

The variable block used here after the 256-dimension layer is a 128-dimension layer, followed by a 64-dimension layer which is followed by a 16-dimension prediction layer with the softmax activation function (Table 4).

The final model doesn't show any improvements over concatenation. Hence, the model 4, which is used here provides the best results.

The model 4 takes 23 s to predict 286 math symbol images (Fig. 5).

Table 4 Results corresponding to the math symbol dataset in percentage

	Model 4	Model 3	Model2	Model 1	Final
Accuracy	99.5	99.3	99.1	99.1	99.2

Note that ReLU is used everywhere else as an activation function.
All results are in percentage

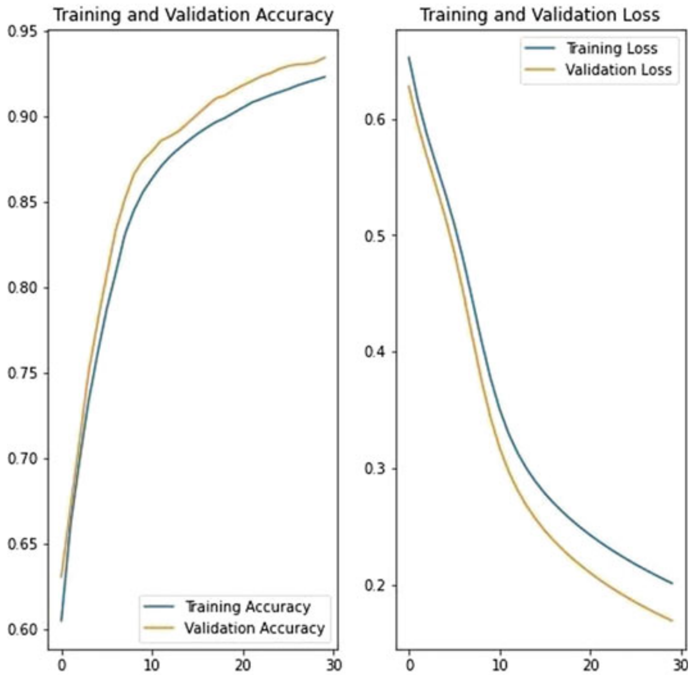


Fig. 5 Image of training and validation accuracy graph while training one of the malaria models

Conclusion

To conclude, a lot of information can be deduced from the proposed model, such as the use of the dilation net in the model, which is very important and serves as the foundation for most of the major conclusions. When dilation rates are varied in descending order, the model achieves a better global minimum, which improves accuracy. As previously discussed, depending on the dilation rate, a dilation network can see a single image pixel matrix in a variety of ways. This may necessitate the removal of unnecessary backgrounds from images in order to focus on the part of the image that needs to be classified. This is especially noteworthy while training networks for multiclass classifications. It is further noticed that the final concatenation proves to be helpful only in the cases of binary classification. In the case of multiclass classification, it fails to increase the overall accuracy. Hence, one must train the smaller four models and choose the best out of them to avoid redundant computations. During training, it is observed that reducing the learning rate after the model trains for certain epochs helps it find the minimum more accurately. This technique is in fact what brought the needed accuracy boost in the results. The model, being very deep, takes a comparatively larger time to predict over test data. However, the change is noticeable only over a large batch of testing data.

References

1. Fattah S, Mahmud T (2020) Automatic diagnosis of malaria from thin blood smear images using deep convolutional neural network with multi-resolution feature fusion. arXiv:abs/2012.05350
2. Linder N, Walliander M, Turkki R, Mårtensson A, Rahtu E, Pietikäinen M, Diwan V, Lundin JM, Lundin A (2014) Malaria diagnostic tool based on computer vision screening and visualization of Plasmodium Falciparum candidate areas in digitized BloodSmears. PLoS ONE 9:e104855
3. Eghan MJ, Opoku-Ansah J, Boampong JN, Anderson B (2014) Wavelength markers for malaria (plasmodium falciparum) infected and uninfected RBCS for ring and trophozoite stages. Appl Phys Res 6:47
4. Shekar G, Revathy S, Goud EK (2020) Malaria detection using deep learning. In: 2020 4th international conference on trends in electronics and informatics (ICOEI)(48184), pp 746–750. https://doi.org/10.1109/ICOEI48184.2020.9143023
5. Powell A, Liang Z, Hossain MA, Ersoy I, Palaniappan K, Silamut K, Sameer A, Guo P, Maude RJ et al (2016) CNN-based image analysis for Malaria diagnosis. In: 2016 IEEE international conference on bioinformatics and biomedicine (BIBM). IEEE
6. Alvaro F, Sánchez J-A, Benedi J-M (2014) Offline features for classifying handwritten math symbols with recurrent neural networks. In: Proceedings—international conference on pattern recognition, 2944–2949. https://doi.org/10.1109/ICPR.2014.507
7. Wells L, Chen S, Almamlook R, Gu Y (2018) Offline handwritten digits recognition using machine learning
8. Ramadhan I, Purnama B, Faraby SA (2016) Convolutional neural networks applied to handwritten mathematical symbols classification. In: 2016 4th international conference on information and communication technology (ICoICT), pp 1–4. https://doi.org/10.1109/ICoICT.2016.7571941
9. Ludi S, Davila K, Zanibbi R (2014) Using off-line features and Synthetic data for on-line handwritten math symbol Recognition. In: ICFHR
10. Le D, Nakagawa, Phan TV (2014) A system for recognizing online handwritten mathematical expressions and improvement of structure analysis. In: 11th IAPR workshop on document analysis systems
11. Hirata NST, Julca Aguilar F, Viard Gaudin C, Medjkoune S, Mouchre H Mathematical symbol hypothesis recognition with rejection option
12. Tavakolian N, Nazemi A, Fitzpatrick D, Suen C, Fernando C (2019) Offline handwritten mathematical symbol recognition utilising deeplearning. arxiv:abs/1910.07395: n. Pag
13. Kutlu H, Avci E, Özyurt F (2019) White blood cells detection and classification based on regional convolutional neural networks. Med Hypotheses. https://doi.org/10.1016/j.mehy.2019.109472
14. Cheng C-Y, Mu-Chun S, Wang P-C (2014) A neural-network-based approach to WBC classification. Sci World J, ArticleID 796371
15. Raman B, Gautam A, Singh P, Bhadauria H Automatic classification of leukocytes using morphological features and naıve Bayes classifier," in Proceedings of the Region10 conference (TENCON), IEEE, Singapore, November 2016.
16. Singh A, Rawat J, Virmani J, Bhadauria HS, Devgun JS (2018) Application of ensemble artificial neural network for the classification of wbcs using microscopic blood images. Int J Comput Syst Eng
17. Ravikumar S (2016) Image segmentation and classification of WBCS with the extreme learning machine and the fast relevance vectormachine. Artif Cells Nanomed Biotechnol 44(3):985–989
18. Mirrashid M, Hashemi A, Shirazi B (2020) A driver safety development: real-time driver drowsiness detection system based on CNN. SN Comput Sci 1:289. https://doi.org/10.1007/s42979-020-00306-9
19. Bamidele A et al (2019) Non-intrusive driver drowsiness detection based on face and eye tracking. Int J Adv Comput Sci Appl 10

20. Petropoulakis L, Ibrahim MM, Soraghan JJ (2014) Eye-state analysis using an interdependence and adaptive scale mean shift (IASMS) algorithm. Biomed Signal Process Control 11:53–62
21. Rahmati M, Assari MA (2011) Driver drowsiness detection using face expression recognition. In: IEEE international conference on signal and image processing applications (ICSIPA). https://doi.org/10.1109/icsipa.2011.6144162
22. Singh H, Bhatia JS, Kaur J (2010) Eye tracking based driver fatigue monitoring and warning system. In: Power electronics (IICPE), India international conference. IEEE, pp 1–6
23. Fujita A, Hamaguchi R, Nemoto H, Hikosaka S, Imaizumi T (2018) Effective use of dilated convolutions for segmenting small object instances in remote sensing imagery. In: IEEE winter conference on applications of CV (WACV), pp 1442–1450
24. Schaeck T, Dreissig M, Kasneci E, Baccour M (2020) Driver drowsiness classification based on eye blink and head movement features using the k-NN algorithm. In: IEEE symposium series on computational intelligence (SSCI), pp 889–896
25. Yu F, Koltun V (2016) Multi-scale context aggregation by dilated convolutions. CoRR abs/1511.07122: n. Pag
26. Koltun V, Yu F (2016) Multi-scale context aggregation by dilated convolutions, San Juan, Puerto Rico, May 2–4, 2016, conference track proceedings. In: Bengio Y, LeCun Y (eds) 4th international conference on learning representations, ICLR 2016, San Juan, Puerto Rico 2016, conference track proceedings
27. Yuan X, Norton M, Feng Z, Li X (2019) Generalized batch normalization: towards accelerating deep neural networks. In: Proceedings of the AAAI conference on artificial intelligence, vol 33, pp 1682–1689. https://doi.org/10.1609/aaai.v33i01.3301168

Integrating Mobile Applications ICT and Digital Tasks Through Online English Language Classrooms

J. Karthikeyan, Seng Tong Chong, and Binoy Barman

Abstract The successful integration of Mobile Applications and Information and Communication Technology depends on teachers' awareness with reference to digital teaching resources, digital equipment and ability to integrate the same effectively and creatively to the learning content so as to promote critical thinking. The former researches have proved that social network models are strong motivators in the path of English language learning among digital natives. The concept of integration of technology in both classroom context and personal context will enhance better learning environment and reduce the anxiety among learners of English. M-Learning and ICT can transform English language learning classroom as a forum for language learning, practice critical thinking and problem solving skills, and sharpen esthetic sense thus to attain cognitive development. This paper emphasizes on the importance of Digital pedagogy and reveals various Mobile applications that will be resourceful for the technocrats to enrich language skills.

Keywords Mobile application · Digital pedagogy · ICT · M-learning · Digital tasks

J. Karthikeyan (✉)
School of Social Sciences and Languages, Vellore Institute of Technology, Vellore 632 014, India
e-mail: jkarthikeyan@vit.ac.in

S. T. Chong
College of Energy Economics and Social Sciences, Universiti Tenaga Nasional (UNITEN), 43000 Kajang, Malaysia
e-mail: stchong@uniten.edu.my

B. Barman
Faculty of Humanities & Social Science, Daffodil International University, Dhaka 1207, Bangladesh
e-mail: drbinoy@daffodilvarsity.edu.bd

V. Bindhu et al. (eds.), *Proceedings of Third International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 844,
https://doi.org/10.1007/978-981-16-8862-1_19

Introduction

The urge for technology integration in education is in great demand than ever before due to lockdown across the globe. Twenty-first century is a digital era and its population is called digital natives [1]. The millennium children have a deep gulf with their former generation due to the intervention of technology which has become a part of their daily life in recent situation. It made them grow beyond the geographical barriers and enables them to imbibe the culture and ways of global citizenship. It shortens the distance between the discriminations and differences.

Though these millennium kids are technocrats and technology is an integral part of their life, the pandemic condition forces them to jam with technology. Therefore, bringing the technology to the classrooms is the need of the hour. Equipping a classroom with technological infrastructure will not bring any change. Incorporating ICT with the classroom teaching learning process via this technological infrastructure is called integrating technology with education. This widespread advancement of technology is carried out in the vehicle called ‘English language.’ The inventions of twenty-first century and advancement of international markets lead to a boosting of migration all over the world in search of job. The birth of international markets and establishment of multinational companies all over the world demands the proficiency in English language speaking skills. Therefore, the excellence in this global language becomes a decisive factor of a country’s growth. It paved the way for teaching English as a foreign language, second language, specific purpose, business purpose etc.

English is one of the subjects in the school and college curricula of most of the third world, developing countries after the colonial invasions. The twenty-first century witnessed the fastest growth of internet and the accessibility of internet to the personal space of individuals. This spread and advancement opens new ways to explore the knowledge with a drastic shift from the traditional ways to digital tech. This shift had a great impact on twenty-first century classrooms. The definition of a twentieth century classroom is a place to impart to knowledge and owed much to the behaviorist and structuralist principles. The entry of Steve Jobs and his inventions like I tune, I Pixar, I pad and I phone redefines the definition of education. Among these inventions, I phone was the most decisive one that made an information boom in society. The information carried by an internet gadget is much higher than the knowledge amassed by a person with their whole life. It released the knowledge from the walls of library and the books were transfigured as microchips or as e-version of knowledge stored in virtual clouds. The traditional library buildings were transformed into reading and browsing places with the introduction of cloud library. This information boom activates ‘learn yourself’ mode because the desired information is accessible at the fingertip. It creates a greater impact in education field and rewrites the purpose of education. The major change was a shift from teaching centered education methods to learning centered methods and formal education to informal education. In a traditional classroom, a teacher is meant for a few enrolled students but ICT elevates the teacher of a small classroom as a global teacher. A traditional teacher’s duty is to transform the four walls of a classroom as the world for students. But a modern teacher is the one who brings the world into the four walls of a classroom. The video recorded classrooms and its telecast breaks the limits of teacher availability. It made the learning possible by sitting in the home and made the school going learning insignificant. The 'countrywide classroom' DVD series of UGC of India is a brilliant example for the same. The smart phones and the various apps functioning in electronic gadgets made the teaching learning process more possible in a more attractive way than a traditional classroom can do. The Learning Management System like 'Easyclass', 'Google Classroom','Edmodo' and 'Schoology' allows the blending of learner, parents and teachers into the learning process together. Its features allow the teachers to customize their needs that include virtual meetings, assessments, off-class discussions, sharing e-resources and updating the student progress to their parents. The discussion forums, sharing options for materials and teaching videos brings the classroom to the personal space of the learner. Therefore, there is a need to redefine the classroom functions and impart changes according to the necessities of digital natives. The twenty-first century demands not only the proficiency in English language speaking skill but also the ability to use it both TBLT and CLT in virtual and real world. The negotiation with the tasks enables the learner to learn language with the support of peers and teacher. Digital tasks are the tasks that operate through or with the support of digital learning gadgets and facilities.

This paper attempts to analyze the best possible ways to explore how technology can rejuvenate an English classroom in a collaborative classroom environment as illustrated in (Fig. 1) that integrate

1. Finest Mobile application, chosen on the basis of user reviews and is freely accessible.
2. ICT adhered to digital pedagogy terminology.
3. Digital Tasks derived from digital materials in audio and video formats.

This collaborative model will yield incessant positive outcome from the learners with reference to language acquisition.

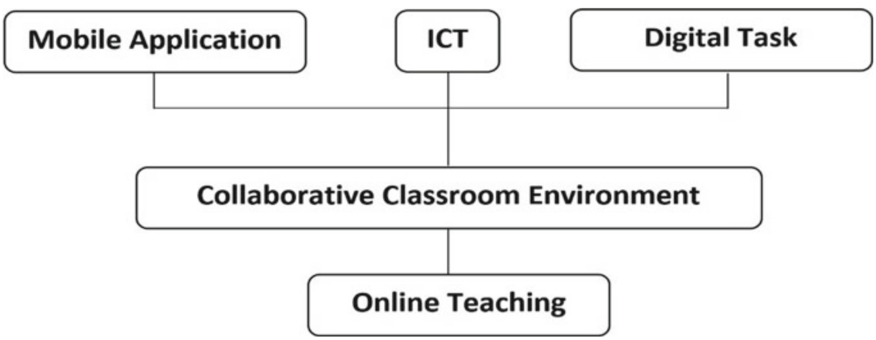


Fig. 1 Collaborative digital classroom learning environment

Novelty of This Study

- Discloses feasible user friendly mobile apps that suits the technocrats and Gen 'Z' learners.
- Reveals various digital tasks that will create conducive learning environment for better teaching and learning.
- Exemplified the research conducted over the years in the use of ICT in education and in English classroom.
- Shows path to future research enhancement at various standardized proficiency levels.

Literature Review

Many studies were carried out in different parts of the world to assess the effectiveness of ICT integration in education particularly in English classrooms. Google Translator is an effective tool to hone English writing skill [2]. This study was carried out in Turkey among the tenth grade English learners. The result of this experimental study shows that the less confident learners and low grade students are using Google Translator so often and found successful in completing the tasks. The motivational effect of ICT in English language learning is confirmed by the study that revealed how the familiarity with ICT gadgets significantly contributes to the motivation level of graduation students [3]. The study of conducted among the school students of UAE also proved that ICT is a powerful motivator to enhance better English language speaking skills [4].

Hussaini Iliyasu conducted a study in Nigeria and found that technology integrated task-based language teaching can improve the four skills of English language learning than non-ICT integrated teaching methods [5]. The experimental study recommends TITBLT among primary school children for the better learning. The study of Huang and Hong at Taiwan revealed that the flipped English classrooms provide more information and hone English reading skills [6]. The improvement of the reading skill and knowledge level is confirmed with interviews and post-tests. Daniela Matić conducted a study among the computer science students of Croatia and proved that ICT made English as a familiar language and lessens the alienation [7]. It found that students preferred the English ICT terms than the Croatian version from the given set of terms. They received the English terms as a natural one and the alternate words for the same are considered as unfamiliar and funny. Jarinabanu T and Prabhakaran K. empirical study that was conducted among medical sciences students proved that 95.92% of the students preferred the class where ICT was integrated rather than the classes where teachers use chalk and teach method [8]. The study of Raman and Yamat at Malaysia revealed that the unsuccessful ICT integration in classrooms are due to teachers' hesitancy, excess of workload, insufficient time, teaching experience, age and lack of ICT skills [9]. The study recommends training for teachers to integrate ICT with their classroom process, motivation and pleasant environment to implement the same creatively in a classroom. Nawaz A and Siddique M. insist on the need for continuous technical support in Higher Education Institutions as education has gone 24 × 7. Teaching and Learning process goes high online without any barriers with reference to time [10].

'ICT in the Algerian EFL Classrooms: An Innovative Means to Enhance Learners' Autonomy' is the study of Guerza [11] which explores the motivational effect of ICT in EFL context. The experimental group was instructed with ICT integrated teaching methods, whereas the control group with adopted non-ICT lecture method. The results show that experimental group scored more marks in all four skills in comparison with control group. The ICT tools were definitely a motivating factor for the students that encouraged them to participate more actively in the classroom process. The study of Viju among the higher secondary students of Kerala proves that ICT integrated classes are more effective in learning English language than traditional lecture method teaching English [12]. The assessment tests shows higher marks to the experimental group students than the control group students in LSRW skills. The experimental study executed by Držić, among university students found that technology-assisted online learning environment can enhance English learning better than the traditional book oriented learning environment [13]. This study progressed by allowing the experimental group to learn through Moodle–wiki online learning platform and control group in a non-ICT book oriented method. The performance of user group of Moodle–wiki online tool was higher than offline learners in oral presentation. The online learning environment significantly contributed to their vocabulary enrichment, presentation skills, fluency pronunciation and accuracy. The study of Hamid discusses the integration process of online social networking with education process [14] as well as with the digital natives which will enhance education better. A study of Arani among the medical students who pursue English for special purpose showed that internet based articles can bring better results than the traditional method of book based study [15]. The learners who used internet based lessons scored higher than the traditional book based learners in this quasi experimental study.

Assimilation of internet with ICT proved as a successful method to enhance communication skill in English based on the study of Young [16]. It found that this integration reduce the stress and create positive environment for language acquisition. Apart from that it develops critical thinking and problem solving skill.

ICT Integrated Education

Online classes have exerted tremendous impact on the behavior of student community locally and worldwide. As the doors of physical classroom have been closed, obviously for pandemic situation, the virtual doors have opened in front of them. This of course is making them more efficient in using technology needed for synchronous and asynchronous classes. They largely seem to adapt themselves to the new mode of online teaching–learning. Education is one of the fields that sort the assistance of ICT to impart quality education. Information and communication technology is a broad name which indicates the gadgets and devices that facilitates the transmission, transformation and amassing of knowledge with or without using internet. According to the report of the National Policy of Education [17], 'Education can no longer be confined to what is in the text books; internet has removed all barriers to learning and made available sources of knowledge not accessible so far.' NPE committee points out that ICT integration is highly relevant and helpful as aid to the teacher in the classroom teaching, to implement successful remedial education, and as learning tools in a classroom teaching learning process. The major areas of ICT integration in education sector can be seen in classroom process, evaluation process and classroom management. The common techniques adopted in classroom process are ICT assisted teaching, ICT teaching materials and ICT assisted learning. The video lessons are the most common techniques used by most of the teachers. The online lecture platforms offered by massive online courses and YouTube videos are another successful integration of ICT with learning process. Carrington A. The Pedagogy Wheel [18] with reference to Blooms Taxonomy allows a learner to the experience the open World with ample resources.

Other ICT integrated techniques include the use of slides, pictures, small videos, etc., as an aid in teaching process. The introduction of Artificial Intelligence in classrooms is the latest advancement of ICT in teaching learning process wherein robot-teacher are being tested to answer the questions and clear the doubts of students in a classroom using speech recognition devices. Various learning Apps helps the learners to pursue learning and prepare for competitive exams. Therefore, ICT has a Mediating role between teacher and students and students and knowledge resources. The ICT competency of a teacher lies in their ability to use it as an aid to organize and create learning material, and its successful implementation in classroom to achieve the learning objective [19]. The role of the teacher is promoting critical thinking on the basis of presented knowledge. The psychological impact of ICT in a teaching learning process is the development of this higher order thinking and its practizing.

ICT Integration in English Language Classroom

The purpose of English language classroom is to enable the learner to use the four skills of English and make them fit for the demands of a global world. Therefore, the classroom teaching learning and teaching methodologies have a greater impact in the learning process. The backbone of the learner cantered curriculum of twenty-first century is constructivist theory where the learners create knowledge through collaborative learning strategy. The Task based language teaching method was originated from the constructivist theory and it use collaborative learning strategy as a major technique. The constructivist idea of the co-operative learning fits well for a digital learning environment (Fig. 2).

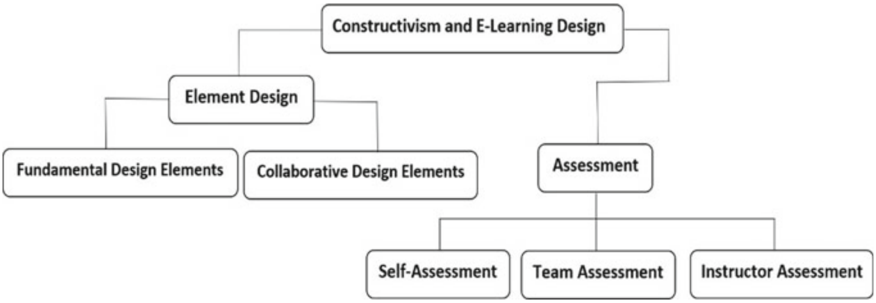


Fig. 2 Constructivist e-learning design

The traditional methods under behaviorist and structuralist theory use the explicit or deductive teaching of grammar, whereas the constructivist theory advocates inductive teaching of grammar. ICT integration to this inductive teaching method made a significant effect in teaching and learning process. There are a number of mobile and software applications to enhance the listening skills, grammar learning and enrich the vocabulary. The word journey, word mind, word up and 12 tenses are some of the vocabulary building games. The BBC developed various software and mobile applications to learn as well as to practice all four skills [20]. The teachers can make use of these apps in a classroom setting. Assigning the learning tasks as per the syllabus via chat groups or mails will warrant the learning process. The level of comprehension tasks, followed by the reading paragraphs demands the application of critical thinking skills among learners. English launch pad is an app specially designed for teachers to teach English grammar, vocabulary, pronunciation and sentence structure. It has the facility to monitor the task and plan the task as per the saved lesson plan. The options for graded quizzes ensure the learning content. CLIx or Connected Learning Initiative [21] is an initiative developed by TISS involving MIT, USA and other Indian State and Central governments to enhancing the speaking skills in English among high school children especially in rural Indian schools. The major technique employed here to develop speaking skills is through storytelling.

Digital Tasks in English Language Classroom

Technology integrated task-based language teaching is a new mode under TBLT method. According to Edgar Dale 50% the retention of the learned content is happening through what the learners see and hear [22]. In a collaborative leaning environment the learning happens 70% and when they do it raises to 90%. According to David Nunan, a task is a piece of classroom work that involves learners in comprehending, manipulating, producing or interacting in the target language, while their attention is focused on mobilizing their grammatical knowledge in order to express meaning, and in which the intention is to convey meaning rather than to manipulate form [23].

The digital tasks are the powerful tool that fastens the unraveling of the meaning in a task. A task can integrate either the four skills or focus on single skill. Digitalisation of the tasks, add color to the creative spirit of the learners in a collaborative learning environment. Tasks mentioned in (Table 1) are extracted from successful classroom practices from across the globe. But when these tasks are integrated with ICT and especially during this pandemic situation where classes are conducted online, it gives better output and makes the class environment filled with fun and more meaningful.

Table 1 Digital tasks for second language English classroom

S. No.	Task description	Materials	Process	Learning outcome
1	Dubbing task	Video clips from popular movies among students	The teacher screen the clip followed by a group discussion with reference to the clip The same clip is played again without audio and the students make their own dialogs in English	Creativity, situational dialogs & enriching fluency
2	Who Am I?	A power point slide with an image of a renowned personality. The picture is hided with blocks which will be removed step by step on each click	Student individually will collect 15 clues about a famous personality for the 15 blocks that covered the picture. The clues are arranged in an order of difficult to easy from internet sources Answering each clue will be remove the blocks and unmask the picture	Frame simple sentences Enables higher order thinking skills
3	One minute presentation	A smart phone with WhatsApp	Each student will record a video on the given topic and share it via WhatsApp The topics given for presentation should be from other subjects prescribed for the student in that term	Reduce anxiety Improves confidence Personal feedback by the teacher will improve the skills

(continued)

Table 1 (continued)

S. No.	Task description	Materials	Process	Learning outcome
4	Word mind [24]	Word mind app It is a word puzzle game to strengthen vocabulary	The teacher will ask one student to open one word listed in the App and insist them to show the work to the class without seeing it The student will identify the listed word from the clues given by his classmates For example, if the word is Lemon, the group members can say it is a yellow fruit, its citric family….	Improves Creativity and Imagination Improves explanation skills
5	12 Tenses [25]	12 Tense app MCQ Rearranging Jumbled Sentences	Teacher can assign this activity after teaching them Tenses Activities are assigned to groups. The group who complete the maximum tasks with correct answers are assigned Winner batch	Motivation Collaborative learning Error Analysis
6	Story telling-virtual reality	Virtual reality headset Smart phone Animated short stories	Animated short stories are viewed by the student and they try to convey the visuals they see to the class then and there	Improves comprehending and Narration skills
7	Learn with news	Video clips of important news collected from news channels	Students react to the news clip that is played during the class They are instructed to assume how will they react if they are in that particular scene	Skills related to Giving Suggestion, Sharing Opinion and explanation are improved

(continued)

Table 1 (continued)

S. No.	Task description	Materials	Process	Learning outcome
8	Picture description-virtual Reality	Virtual reality headset Smart phone Reality picture with adjectives	The picture is shared with the student after they wear the VR headset The student will make short sentence with an adjective For example The girl in the picture seems *angry*	Use of different types of adjectives Acquire fluency and explanatory skills
9	Birthday shoot	Smart phone Video editing app	Each student will make a video of any one of their classmate by collecting the pictures and personal info. They will edit it and add their voice sharing info about them and praising them on their birthday	Learn to greet and compliment others Use of anecdotes
10	Chat box	A laptop or smart phone with video chat facility	Student will watch a short talk on national importance Late they use the chat option to post their opinion about the video Later they will have a debate on the short talk	Functional vocabulary Spontaneous creation of short sentence while speaking

Apart from these tasks, teachers can make use of digital web tools that can enhance English language learning in an online classroom setting. Introducing blog writing will enhance critical thinking along with writing skill.

Digitalisation Through Mobile Learning

Though COVID-19 made us to stay within four walls, technology opens the doors of self-learning. If the learner is intrinsically motivated, then learning is unstoppable. For enriching one's own language skills apps and open source software are freely available in English. These resources have varied content that suits all kind of learners and segregated them by the common European framework levels. These tools are interactive to some extent as a teacher, train the learners, conduct assessments and grade them. The teacher can suggest some best personal apps that suit their students and motivate them to learn language. Some useful applications are listed below.

Buddy Talk [26]: This app is a great platform to practice speaking skills in English. This app let the users to speak to people all over the world by connecting them through voice or video call. The greatest advantage of using this app is that it will give a real life experience to the users without being identified. There is an option called buddy, through which one can add a person to buddy list and talk to them again. The app has filter option to choose speakers from desired level and country. This app will be helpful for the middle and advanced learners.

English Conversation Practice [27]: It is a best app for novice of English. It provides 200 conversation practice lessons along with instructional video. The designers gave importance to listening skill and understand English content to make a firm ground for developing good communication skill. The quizzes and tests ensure enrichment in speaking and listening skill. It has the option to save our conversation practice and track the improvement.

English Speaking Practice [28]: This is best suited for beginners. This app allows individuals to practice the pre-recorded dialogs that exist inbuilt. The short conversations are broadly categorized as beginner conversations and business conversations. Each of these topics again classified into small conversation like a surprising conversation, meeting an old friend, etc., and followed by a quiz and practice session. The learner can speak according to the scripts given in the page and record the same. These recordings will help the learner to track their progress. The app provides a model for English speaking and familiarize the conversation expressions.

ENGVARTA [29]**:** This app is beneficial for beginners and advanced learners for practizing listening, reading, speaking skills and it emphasize on grammar and vocabulary. The installation process of this app starts with a test on vocabulary and grammar to decide the level of learner as per the international standards. Based on the score, the training directs them either to the beginner level or to the advanced level.

Cake-Learn English for Free [30]: This app gives prominence to the speaking skills. There are a number of video clippings that are shorter than a minute with special significance to the key expressions. The subtitles of the video conversation will be given in a box below to the video. The video extracts are from movies, documentaries, ads, serials, YouTube etc. which offer an audio-visual treat with fun to enhance speaking skill. The key expressions will repeat 3 times, and there is an option for the learner to practice the same with the video or audio. Speech recognition technology grades the learner and helps them to identify the faults instantly by replaying the audio and corrects simultaneously.

Utter Learn English on Chat [31]: The lessons are arranged at three levels namely beginner, intermediate and advanced. The lessons focus on the grammar, vocabulary and speaking skills. The learning process proceeds through chatting. Initially, the chat proceeds through the model sentences and later there is an option to chat with an expert in real life context. The expert will rectify the mistakes of the learner and helps to learn. It has also the option to record the practice sentences as a voice clip. The lessons are arranged in a way to get a good command on business communication and official communication. It is very good for the learners who have high level of Foreign Language Anxiety and fear of negative evaluation.

ELSA Speak—Learn English Pronunciation [32]: Elsa app focuses on speaking skills and pronunciation. It records the sentences of the speaker and corrects the faults instantly. The app allows the learner to practice the same word until the learner becomes perfect in pronunciation and ensures the correct usages to move further. The topics are arranged as per the expected talks in various occasions like interview, at airport, at the restaurant…It has an inbuilt dictionary that suggests the learner with alternate expression for words.

These suggested apps enable the learner to learn and practice English off the class. The greatest advantage of this sort of learning is the convenience of time and choice of modules as per the interest and need. But such self-learning lacks motivation.

Need for ICT Skills in a Classroom

ICT integration demands certain skills from the teacher for a successful teaching learning process. Tony Bates, defines certain skills for teachers to teach the digital natives [33]. They are

(a) **Media Communication Skill**: Unlike the traditional classroom, teaching process address a global community through ICT tools. In this situation, teachers should learn to talk according to the needs of a vast learning community. The in depth knowledge about the topic with proper explanation is an essential quality to address the learning community. The online digital lectures cannot answer the doubts of the learners instantly; therefore the teachers are expected to explain the concept with clarity. The millennials have short span of attention unlike old generation, therefore content delivery should be short and apt. The online presentations should include the use of multimedia effects to avoid monotonous talk. The skill to incorporate new ideas and giving proper post class clarification via feedback is a decisive factor in digital lecturing.
(b) **Skill to Learn Independently**: The nature and range of knowledge is changing too rapidly in twenty-first century. Here, the continuous learning is inevitable for teachers. They should acquire digital skills to transfer the knowledge to teach the digital natives.
(c) **Ethics and Responsibility**: The commotion of knowledge should not fail to acknowledge the source of it. Give respect to the knowledge sources and share the knowledge which is the etiquette in digital teaching and learning process.

(d) **Teamwork and Flexibility**: Teaching should not be an independent and isolated work. Collaborative working strategy among the teachers of the same field will make the process easier and more effective in this pandemic situation.
(e) **Thinking Skills**: The learning content should be organized in a manner which kindles higher order thinking skills as per Bloom's taxonomy [34]. The instruction should pave way for the critical thinking, problem solving and expression of creativity along with the skill to collect, organize, synthesize and present knowledge.
(f) **Digital Skills**: It is the most wanted skill in digital classroom. It includes the use of digital tools as an aid in teaching learning process and a means to create knowledge. The creative integration of ICT tools with learning content has higher impact in teaching learning process.
(g) **Knowledge Management**: Information is just a click away hence; collection of knowledge should result in keeping the right information and sharing it to the learning community.

Advantages of ICT Integration in Classroom

The major advantages of ICT integration with English classroom are the precision, objectivity and accuracy in the teaching–learning process. It aids the teacher to produce the accurate pronunciation, intonation and meaning that helps the learners to learn the language as per international standards. Technology assistance in class lessens the effort of the teacher in explaining and transacting the knowledge content. An apt ICT integrated teaching content can transact much more than the verbal attempts of the teacher in the process of teaching learning.

Challenges

The technological advancement demands new skills from the teachers. The teachers of new digital era is teaching in a traditional classroom set up; therefore, it tempts the teachers to go back to the comforts of teacher oriented methods. The less technology experienced teachers are teaching the technically advanced students. In other words, the digital immigrants are teaching digital natives today. Many of these digital immigrants possess poor digital skills and computer literacy. It takes the digital natives to the old book centered era. Therefore, technology chases teachers if they are not effective in handling the subject. The technology demands the precision and accuracy from the teachers in the process of teaching. It is important to use technology as an aid in the classroom instead of teacher became the supporter of technology. Unfortunately, most of the subject teachers, especially English teachers have less knowledge to handle as well as fix the issues related to digital gadgets. Solving the issues related to software and hardware, or technical errors are a challenge for novice users of technology. The poor internet connectivity during online classes scheduled during this pandemic is a big challenge for the educational institutions. The cost of certain LMS and online teaching software packages are beyond the capacity of institutions.

Limitations of the Study

This study is limited to tertiary level learners in all branches of higher education, who are well exposed to digital gadgets, acquaintance to access mobile apps from play store and who are wide-open to digital environments. In addition to this the efficacy of teachers plays a viable role in the process of selecting and mixing the right resources that suits the classroom environment depending upon the ability of their learners.

Future Enhancement of This Work

The teaching resources in the form of digital contents shared in this study are carefully chosen to assist ESL learners in higher education institutions. The curriculum is set for learners whose language proficiency is between B1 and B2 as per the Common European Framework of Reference for Languages (CEFR) scale of assessment. Researchers can consider identifying suitable digital resources for A1 and A2 learners who are in schools and C1 and C2 learners who intend to acquire global standards same as the native speakers proficiency.

Conclusion

In current COVID situation, teaching and learning become a biggest challenge for students as well as the teachers. Parents are much worried about their wards learning process and expect the best possible ways to get connected to the education institutions. Institutions and Teachers cannot hide from using technology at its best to serve the learning community. They have to overcome any issues related to gadgets, internet connectivity, potential delivery and learning. Quality, creativity and effective use of technological tools will ensure learning at all levels of education. The anxiety disorders among learners will gradually vanish depending on the involvement of the teaching community.

References

1. Prensky M (2001) Nativos digitais, imigrantes digitais. On Horizon 9(5):1–6
2. Selcuk H, Jones J, Vonkova H The use of Google translate as an ICT tool in web-based collaborative writing: self-reported accounts of EFL learners
3. Shiny KG, Karthikeyan J (2019) An empirical study on ICT integrated task based language teaching to promote English language speaking skills. J Adv Res Dyn Control Syst 11(12):744–754
4. Amin Elkhayyat AM The effectiveness of using ICT to enhance the EFL speaking skill
5. Hussaini I The effect of using technology task based approach on English language achievement among primary four school students in Nigeria (Doctoral dissertation, Universiti Tun Hussein Onn Malaysia)
6. Huang YN, Hong ZR (2016) The effects of a flipped English classroom intervention on students' information and communication technology and English reading comprehension. Educ Tech Res Dev 64(2):175–193
7. Matić D (2014) Attitudes of computer science students to the English element in Croatian ICT magazines. ESP Today. J English Spec Purp Tertiary Level 2(2):174–198
8. Jarinabanu T, Prabhakaran K (2015) Effectiveness of computer aided teaching, as compared to traditional board teaching, among medical undergraduates of udaipur. Int J Current Res Rev 7(8):53
9. Raman K, Yamat H (2014) Barriers teachers face in integrating ICT during English lessons: a case study. Malaysian Online J Educ Technol 2(3):11–19
10. Nawaz A, Siddique M (2012) Continuous technical support for the effective working of e-learning in higher education. Int J Cur Res Rev 4(23)
11. Guerza R (2015) ICT in the Algerian EFL classrooms: an innovative means to enhance learners' autonomy. Int J Infonom (IJI) 8(1):979–985
12. Viju MJ. Application of ICT Information Technology and Communication in Communicative Language Learning.
13. Držić BF, Seljan S, Djigunović JM, Lasić-Lazić J, Stančić H (2011) Teaching English for special purposes aided by e-learning platform. Int J Excel eLearn 4(1):20–35
14. Hamid S, Chang S, Kurnia S (2009) Identifying the use of online social networking in higher education. InAscilite, pp 6–9
15. Askari AJ The effect of ict-based teaching method on medical students
16. Young SS (2003) Integrating ICT into second language education in a vocational high school. J Comput Assist Learn 19(4):447–461
17. National Policy on Education report (2016) Ministry of Human Resource Development, Government of India
18. Carrington A (2012) The Padagogy wheel… it'sa Bloomin' better way to teach. Designing outcomes
19. Valencia-Molina T, Serna-Collazos A, Ochoa-Angrino S, Caicedo-Tamayo AM, Montes-González JA, Chávez-Vescance JD (2016) ICT standards and competencies from the pedagogical dimension: a perspective from levels of ICT adoption in teachers' education practice. Pontificia Universidad Javeriana
20. Learning English (2020) https://www.bbc.co.uk/learningenglish/
21. CLIx or Connected Learning Initiative (2018) https://clix.tiss.edu/
22. Dale E (2006) Effective learning. From Chapter 1 in audio-visual methods in teaching. Dryden Press (Holt, Rinehart, and Winston), New York. Retrieved April. 1946, 3
23. Nunan D (2004) Task-based language teaching. Ernst Klett Sprachen
24. BitMango (2020) Word mind: crossword puzzle, https://play.google.com/store/apps/
25. Tình LV (2017) 12 tenses, https://play.google.com/store/apps/
26. Interaxis Networks, Buddy Talk (2019) https://play.google.com/store/apps/
27. TalkEnglish, English conversation practise (2020) https://play.google.com/store/apps/
28. TalkEnglish, English Speaking Practise (2020) https://play.google.com/store/apps/

29. ENGVARTA (2020) https://play.google.com/store/apps/
30. Playlist Corporation, Cake-Learn English for Free (2020) https://play.google.com/store/apps/
31. EML PL, Utter Learn English on Chat (2020) https://play.google.com/store/apps/
32. ELSA Speak—Learn English Pronunciation (2020) https://play.google.com/store/apps/
33. Bates AT (2018) Teaching in a digital age: guidelines for designing teaching and learning
34. Bloom BS, Krathwohl DR, Masia BB (1984) Bloom taxonomy of educational objectives. InAllyn and Bacon. Pearson Education

Energy-Efficient Hybrid Secured Routing for 5G Vehicular Ad Hoc Network (VANET)

K. Suresh Kumar Reddy, V. Tamizhazhagan, V. Senthil Murugan, and D. Rajaveerappa

Abstract These days, road accidents and traffic jams have increased significantly on the route due to the higher number of vehicles. Vehicular Ad hoc Networks (VANETs) are effective intelligent transportation system in decreasing accidents and traffic congestions. Vehicular communications also involved user-specific information such as their route and time critical information. But, it was deal previously with information about traffic and road conditions. However, the information transmission is vulnerable to eavesdropping and link failure due to high mobility of the vehicles. In this research, analysis of secured routing to provide information to the legitimate nodes in the presence of unwanted eavesdroppers is done. The hybrid energy-efficient routing algorithm is used for better stability of the connected vehicle. The beamforming technique is applied for concentrating on known users for security purposes using an array of antennas. A combined detailed analytical model of hybrid MIN NORM & ROOT MUSIC DoA, LMS adaptive beamforming, hybrid route based on Dijkstra and Bellman Ford algorithm showed a greater secrecy rate. The simulation is carried out in MATLAB 2014a MEX file package software. The result obtained such as energy efficiency and secrecy rate are convincing compared to the previous result content in the literature.

Keywords Vehicular ad hock network (VANET) · Secured routing · LMS · Min norm · Music

K. Suresh Kumar Reddy (✉)
Department of CSE, Annamalai University, Chidambaram, Tamil Nadu, India

V. Tamizhazhagan
Department of Information Technology, Annamalai University, Chidambaram, Tamil Nadu, India

V. Senthil Murugan
Department of CSE, Viswam Engineering College, Madanapalle, A.P., India

D. Rajaveerappa
Department of ECE, Adama Science & Technology University, Adama, Ethiopia

V. Bindhu et al. (eds.), *Proceedings of Third International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 844,
https://doi.org/10.1007/978-981-16-8862-1_20

Introduction

These days most people use vehicle as a transportation purpose. Due to this, it leads to traffic congestion on roads. Traffic congestion and driving problems in turn raised street accidents and deaths due to lack of information about the street. Y knowing the pre knowledge of environment of neighbors and knows the street traffic conditions with the help of on board units and road side equipment's.

VANET is one of the promising technologies for ITS. Its superior class of mobile ad hoc networks (MANETs) in which Cars act like moving nodes [1]. In VANET, vehicles are having on board units and Road side equipment's for exchanging information between vehicles as well as intelligent device and radio interface roadside units (RSUs). It allows to send messages to other vehicles or infrastructures instantly. These networks are adapted to support the highly dynamic nature of the vehicles as network nodes, fulfill their mobility requirements [2]. It enables a vehicle that to detects a risky situation to inform to other drivers on the road. Using this communication, drivers can switch time critical information such as road or weather conditions and crashes, breakdowns, or emergency vehicle on-site warning.

VANET communication take place through wireless medium. This information must be delivered to other neighboring vehicles timely within limited time. Due to rapid topology of vehicle mobility, there is high link failure [3]. Efficient routing is necessary to enable the vehicles in making the best routing decision.

VANET wireless media is also vulnerable to user information eavesdropping, interference and jamming [4, 5]. Multi antenna is one of the powerful techniques for secure vehicular communication.

The technology advancement in 5G wireless network [6] that uses massive Multiple-Input-Output (MIMO) antenna array system can be used for VANET security issue due to its spectral efficiency and low latency. Relay operation is used for no line-of-sight blocked sender and receiver vehicle link due to the mm-wave short wave length loss from penetration and reflection. To improve the performance of a VANET communication, with the help of multiple array antenna we received signals and combined to achieve this [7]. Different technique such as Direction of Arrival (DoA), and an adaptive beamforming is very essential in the signal processing [8, 9]. In beamforming technique, signal directional transmission for enhanced gain and enlarged coverage can be achieved. The bearing algorithm also calculates all incoming signals directivity.

In this paper, [10–12] a high data rate and reduced latency of massive MIMO 5G technology are used for vehicular secure routing. Beam steering antennas form narrow directional beams with reducing the error signal and increases the wanted signal by dynamically steering the activity at different vehicular nodes. In order to continues connects the vehicles, where will high velocity nodes, we use smart beam steering even in real-time which helps to collect vehicles network topology. [13, 14] The beam steering technique is done to change the direction of the radiation pattern of the antenna to get the antenna output in the desired vehicle direction. Antenna with enhanced gain and steerable radiation patterns also used to reduce intrusion, saves power, directivity and gain. This paper is organizing as, in Sect. 2, explains about literature review, in Sect. 3, related work, in Sect. 4, explains proposed method, in Sect. 5, simulations and discussions and in Sect. 6, concluded.

Literature Review

Basics of ITS

The amount of traffic accidents is in rapid growth with increase in vehicular. In a view of WHO, road accidents were annually increases approximately to 1.2 million deaths worldwide [15]. This is one of the main incentives for the development of ITS. VANET emerged as an ITS technology to help in communicating vehicles themselves without any central controller [16]. Ad hoc network is preferable for its comfort of deployment, velocity of deployment, and decreased dependence on infrastructure considering VANET dynamic topology (Table 1).

In VANET, nodes of vehicles and discovering a route between them is very critical in such scenarios. To discover the and route maintains of a vehicle is difficult to maintain. The moving cars used as nodes in a network so each vehicle can receive and transmit to others message through the wireless ad hoc network [17]. The communication between vehicular nodes can take place in two different manners either (V2V) communication or vehicle to RSU communication [18]. The VANET communications help in transmitting information about the road condition such as location, speed data, and other alerts of traffic condition to another vehicle.

VANET Component

VANET comprise vehicles equipped with a wireless device communicating with each other and RSU. The network infrastructure such as the base station or an access

Table 1 Comparison of MANET and VANET

Parameters	MANET	VANET
Energy limitation	Low	High
Vary in network topology	Poor	Frequent and very fat
Node mobility	Low	High
Network size	Medium	Large
Range	Up to 100 m	Up to 600 m
Location dependency	Low	High
Moving pattern	Regular	Random

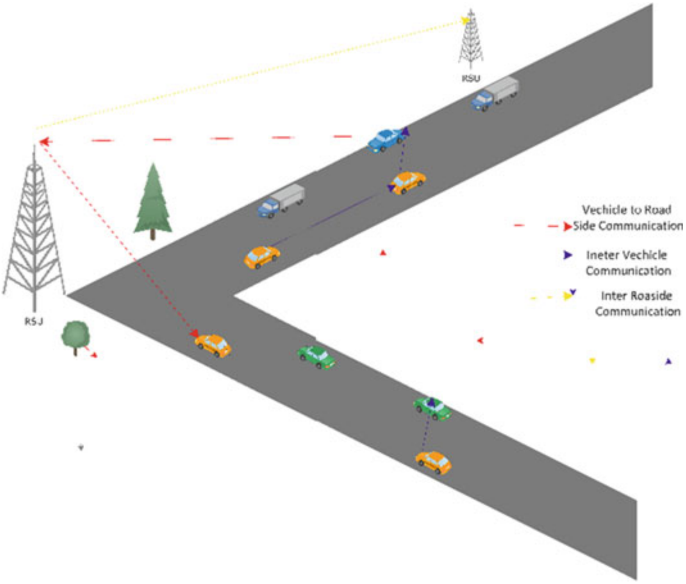


Fig. 1 VANET communication

point is not required to create VANET, although fixed nodes can be used in VANET in the form of RSU.

Onboard Unit (OBU): this equipment is transceiver its installed in every vehicle for transmit data from vehicle to vehicle and for the road side equipment. Every OBU contains abundant power with inbuilt batteries. OBU contains the sensors, with help of that receives information like location and speed of a vehicles with it surroundings. After having this, it forwards to concern vehicle which is shown in Fig. 1.

Roadside Unit: As the all nodes were highly mobile, vehicle interaction should need to communication and transmits the information. To achieve this we need to set up an equipment where it interact with on board unit to send an information like traffic awareness, breakdown vehicles details etc.

RSU can be connected to the backbone of the internet through wireless or wired and provides the services such as periodic re-transmission of a message in sparsely populated roads and advertisement services. The node mobility in VANET varies with directions and velocities [6, 19]. All nodes with in few seconds the topological structure were used to change every time.

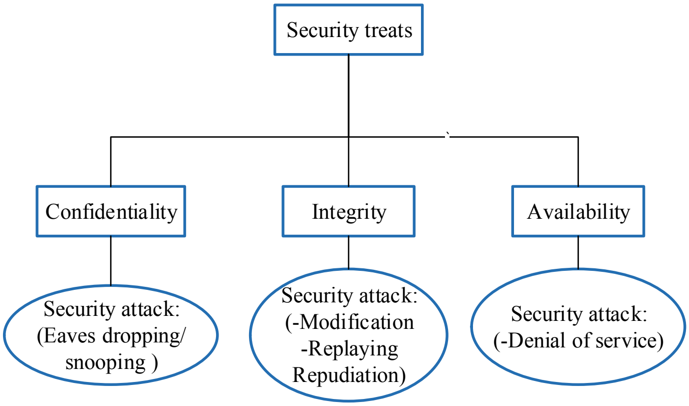


Fig. 2 Taxonomy of security goals and attacks threatening

VANET Security and Routing Protocol

Security Issue in VANET

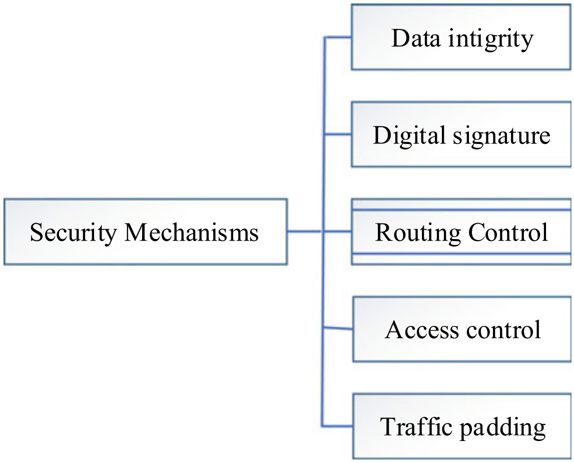
In VANET communication, security plays a key role, especially in information transformation. So that we can control to access by intruders and attackers [20, 21. For dynamically changing topology and malicious attacks, selecting best routing protocols are very important. There are various security concerns to be achieved in wireless system [22] which is shown in Fig. 2.

Snooping or Eavesdropping: This refers to unauthorized access to or interception of data. Intruder intension is to get the confidence information, in order to perform such attacks, they may be present in car or by placing the false RSU. Due to some obstacles, there will be no direct communication link exists between source and receiver, so it forwarded by relay. If sender transmits a data to different receiving nodes. Then data transmitted by one node should keep the data as secret from opponent one. Here, one is destination receiver and the other node is intruder. To overcome the security issue, there is different mechanisms which is shown in Fig. 3.

VANETs Routing Algorithm

In order to achieve the best performance a route in dynamic environment, routing plays a key role, to provide best path form one node to another node with various dynamic algorithms [23]. Such protocols must deal with the typical limitations of the Ad hoc network like high power consumption, low bandwidth, and error rates. The communication between two vehicular nodes can be a single-hop or multi-hop and the routing protocol decides the route from source to destination for the multi-hop communication. For efficient routing protocols, reliable and low latency date

Fig. 3 Security mechanisms



were needed. Also computing the shortest path between two vehicular nodes in road networks is a challenging task in vehicle routing. The routing algorithms must be able to handle scenarios where links may fail and new links may be formed dynamically which is shown in Fig. 4.

Proactive Routing: This routing protocols is mostly based on shortest path algorithms. It is table-driven protocols and they maintain information of all connected stations in the form of a table. To reduce the burden over the network, this routing works on only active nodes of a vehicle. In proactive protocols, each network node stores information on routes to all possible destinations at all times. The advantage

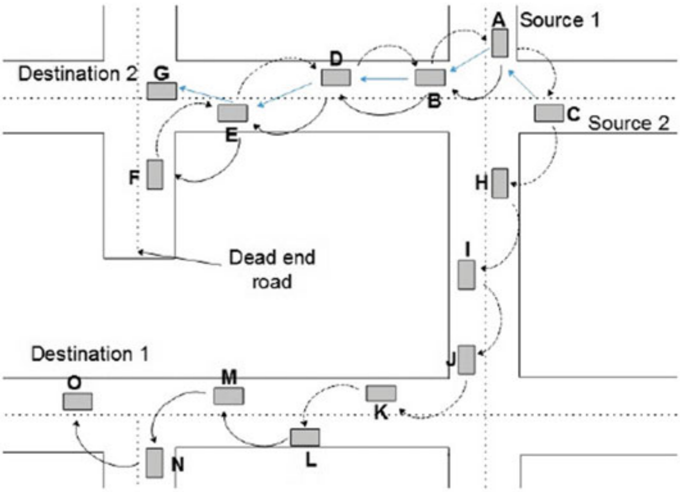


Fig. 4 Routing scenario

of this approach is that when a routing request is needed, it can be initiated with minimal delay. The disadvantage is the overhead in the form of control messages that are required to establish and maintain all routes subject to frequent failure.

These protocols sends a packets, after establishing the route only. If route not found, these packets will has to wait in queue until establish route found. Based on this routes can calculations and transmission time periodic routes will found. These can take over place in proactive protocols. Those types of protocols are difficult to maintain in a highly dynamic nature of VANET.

Reactive (On-Demand) Routing: In this case, routes are established on as-needed basis. This approach requires less control information, particularly when the network topology is rapidly changing or when the data burst or directed to only a small subset of the network nodes that are unknown in advance. The disadvantage of such an approach is that route requests can experience a large delay before being propagated from the start node, as a route to the target node has to first be determined.

Finding route is used when a font node needs to forward a packet to a particular terminus, but not have a corresponding route stored in its cache. In this case, the source node floods a query packet through the network. Intermediary nodes append their own identity to the query and forward it unless they are either the terminus node or take a route for the destination stored in their cache. Nodes that do not forward the query send a reply message back to the source. When the reply arrives back to the source, the newly discovered route is stored in the source's cache and routing can commence.

Hybrid Routing: Routing algorithms in this class incorporate elements of both proactive and reactive protocols. Two immediate strategies for routing in wireless networks are minimum energy routing, in which the link metric is the energy required to send a wireless transmission over the link, and minimum hop routing, in which the link metric is constant for all links.

Minimum hop routing sometimes minimizes the end-to-end routing delay in a network. In reality, such approach tends to use the same subset of links to transmit the network's traffic. This leads to congested routes which in turn lead to relatively long routing delays which is shown in Fig. 5.

Related Works

As the number of vehicles were increasing in vehicular ad hoc networks communication. Providing security, with the data encryption techniques in this scenario is not a great advantage, there may be an increasing losing of keys. And also key management of high complexity work is difficult to maintain. As of traditional methods were concentrated on top layers of OSI, as compared to the physical layer.

For secure transmission of data, considered a physical characteristic of radio based of sha theory. It has been widely utilized with single antenna, for getting massive

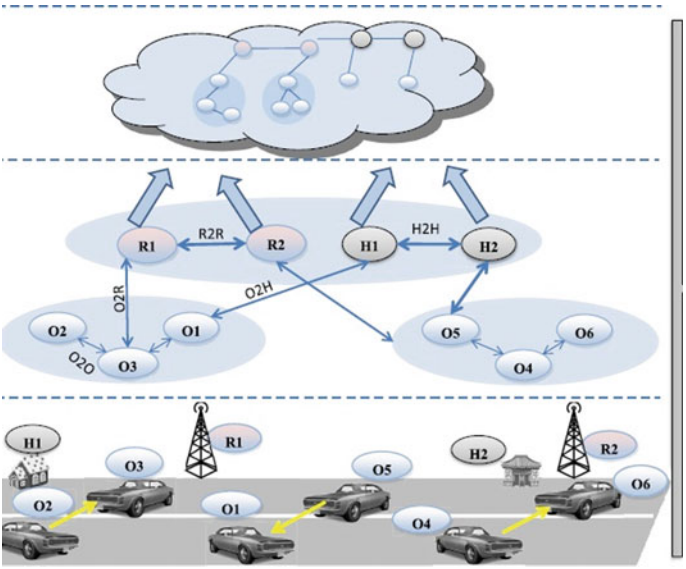


Fig. 5 Social interconnection of vehicles

free multi-in/out [12]. However, the best routing of information is not described in the article.

In [13], a system that combine information on congestion from sensors installed on the road and from inter-vehicle communication through rumor is proposed. A cost function that considers dynamic metrics on segments along with expiry, delay, and uncertainty was also proposed. The main limitation of this approach is how much congestion information on what roads are to be exchanged between 'roadside units and vehicles' and between the vehicles was not clear.

In [14] from adjacent nodes, in order to achieve the routing information we need to continuously refresh the information table on proactive routing protocols. As increasing the mobility of a vehicle's, expenses will be huge for maintain routing information.

In by using reactive routing protocols, after we found a route from source to destination only we can query floods into the network. In this case latency is high, however, it avoids necessary of required low bandwidth unnecessary control messages. The position-based routing protocols utilize the geographic location of the vehicular nodes to decide the next hop during data transmission [14].

In, authors proposed, for how to calculate node traveling time and broadcast the data to other vehicles where RSUs were placed. And also to information center. A bidirectional method were established in order to find the travel time for a road segment. In [13], the author analyzed of on-field quantities of various different antenna types (monopole antenna, patched antenna, a dipole antenna, and Yagi antenna) for vehicular networks. The radiation patterns show that the correlation of the received signals is dependent on the mobility of vehicles concerning a particular RSU.

Proposed Method

Antenna Beam Forming Technique for Secure Routing

Low directivity can be achieved with wide activity of radiation pattern of an antenna. To cover an area of specific application we need to direct the antenna for long distance. This can be done with increasing the number of antennas radiating elements. desired direction and interfere with destructively canceling each other giving rise to nulls in the undesired direction with the fields from the elements interfere constructively adding up each other giving rise to radiation pattern.

As dynamically vehicles, were moving in VANETs scenario, selection of smart antenna plays a key role, to achieve this we have two techniques were available, one is fixed beam, but for this scenario, it won't work and not must suited, second is adaptive smart antenna, with help of this we can steered in any direction and also generate the suppress of interference. It happens only when we combine the weights of complex and adaptive.

In order to achieve the minimizing an error between wanted signal and array output, adaptive beam form plays a key role. With multiple antennas, we can the control direction of the signal, by adjusting the weights of the signal phase amplitudes.

Minimum Variance Distortion Less Response (MVDR) Beam Former

This algorithm ensures that the signal from desired direction remain undistorted by the eavesdropper. The formed beam pattern concentrates or targeted to the DoA of interest.

$$\tilde{\omega} = \frac{\overset{\frown}{R}^{-1} \bar{\alpha}(\theta)}{\bar{\alpha}^H(\theta)\overset{\frown}{R}_{xx}^{-1} \bar{\alpha}(\theta)} \tag{1}$$

$$P(\theta) = \frac{1}{\bar{\alpha}^H(\theta)\overset{\frown}{R}_{xx}^{-1} \bar{\alpha}(\theta)} \tag{2}$$

Min Norm DoA Estimation

For linear arrays if the min norm represented by MN

$$\mathrm{AN} = \frac{(\widehat{u}^{T}\bar{E}_N\bar{E}_N^H\widehat{u}_1)^2}{\left|\overline{a}(\theta)^H\bar{E}_N\bar{E}_N^H\widehat{u}_1\right|^2} \tag{3}$$

where $\overline{a}$ = array steering vector

$\bar{E}_N$ = subspace of M-S eigen vectors

A = no. of antenna elements

S = No. of incoming signals

$\widehat{u}_1$ = Cartes basis vectors.

LMS Adaptive Beamforming Algorithm

To generate the least error of a signal by subtracting the signals of desired and original signal, it can be done will LMS.

To reduce the mean error, it based on the current time of error, for making a corrections on weight looping procedure is used, for the reverse gradient. By doing this, we can increase the SNR, with minimize error, of subtract with final signal and refer signal. If $d(c)$ is a testimonial signal sampled at the output instant k and correlated with the wanted signal $\ddot{x}(c)$ then the error signal is $\boldsymbol{\varepsilon}(\boldsymbol{c})$ **is**

$$\boldsymbol{\varepsilon}(\boldsymbol{c}) = g(c) - \tilde{\omega}^{Q}(c)\,\mathrm{x}(c) \tag{4}$$

formed fault is as

$$D = E[|d|]^2 \tag{5}$$

$$\tilde{\omega}_{\mathrm{opt}} = \widehat{R}_{xx}^{-1}\,\widehat{\mathrm{r}} \tag{6}$$

$$\widehat{r}(c) \approx d^*(c)\ddot{x}(c) \tag{7}$$

$$\tilde{\omega}(c+1) = \tilde{\omega}(c) - \frac{1}{2}\mu\Delta_{\tilde{\omega}}(J(\tilde{\omega})) \tag{8}$$

$$\tilde{\omega}(c+1) = \tilde{\omega}(c) - \mu\left[\widehat{R}_{xx}\tilde{\omega} - \widehat{r}\right] \tag{9}$$

$$= \tilde{\omega}(c) + \mu e^*(c)\ddot{x}(c) \tag{10}$$

$$e(c) = g(c) - \tilde{\omega}^{H}(c)\ddot{x}(c) = \text{error signal} \tag{11}$$

where in $\tilde{\omega}^H Q$ denotes complex conjugate transpose, and μ is the step size.

For uniform linear antenna array with an inter element spacing of '$d\alpha\theta$'.

$$AF = \sum_{n=0}^{N=1} A_n{}^{jn\frac{2\pi d}{\lambda}\cos\theta + \alpha} \tag{12}$$

The progressive shift will get from the wanted beam direction and from wavelength toward the known signal.

$$\alpha = \frac{-2\pi d}{\lambda}\cos\theta^0 \tag{13}$$

System Block Diagram

d is the final destination Node that receives information and the other one is called listener (*e*). Source node (*s*) transmits information to relay node. Here, the with *M* antennas elements equipped with relays. it may ask them to feedback channel state information of Its need *s* to be forwarded by relay, By contain the all CSIs, the signal sent to all receivers by adopt. The noise at any node is assumed additive white complex GN with Null-mean and variance in this consideration. To identify the intruder direction, when need to convert them into energy, with the confidential information. Some of relay antenna receives, while other one receives signals which is shown in Fig. 6.

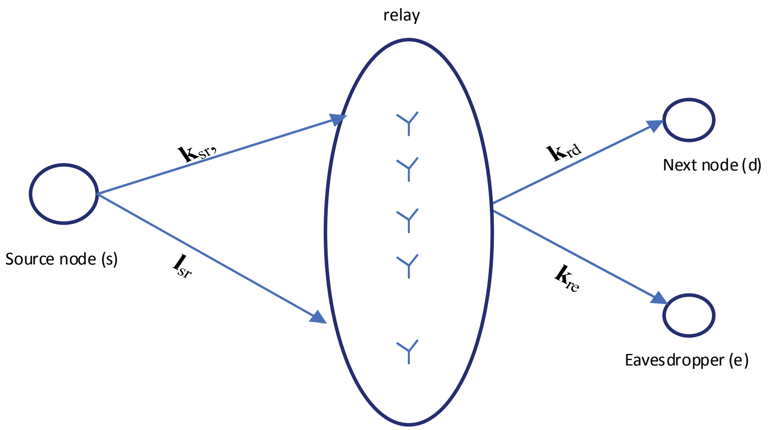


Fig. 6 System model for hybrid secured routing

Hybrid Routing Information Secrecy Analysis

Let say the relay antenna has K represents group of antennas. The group divide into two subsets of (**M, N**), and satisfies **K = M** $\boldsymbol{U}$ $\boldsymbol{N}$, and the antenna in the subset of $|M| = \boldsymbol{K} - \mathrm{N}$ used for information relay.

The established signal of the relay as stated as:

$$y_r = \sqrt{Ps} k_{sr} x + n_r \tag{14}$$

where $nr \sim CN(0, \sigma^2 I_N)$ are received antenna AN of **N**. Meanwhile, through the antennas in the subset $\boldsymbol{K}$. with relay energy. The conventional signal of the ith$(i = 1$ to$M - n)$ antenna in the subset $\boldsymbol{N}$ can be written as:

$$r_i = \sqrt{P_s} I_{sr} x + n_r \tag{15}$$

In the second stage the total power of forwarding signals is:

$$P_{r'} = P_r + P_r \alpha \sum_{i=1}^{M-N} |r_i|^2 \tag{16}$$

$$= P_r + \alpha \sum_{i=1}^{M-N} (P_s |I_{sr}|^2 + \sigma^2) \tag{17}$$

$$= P_r + \alpha P_s \|I_{sr}\|^2 + \alpha (M - N) \sigma^2 \tag{18}$$

In the second stage, by the antennas it transmit forwards information in the subset N. The conveyed signal of the relay is denoted by, $x_r = W_{yr}$ where $\boldsymbol{W}$ is a sloping matrix composed of a BF Vector

$\omega = \left[\omega_{1,} \omega_1 \ldots, \omega_N\right]^T$ i.e., $W = \mathrm{diag}(\omega)$, i.e., $\mathbf{x}_\mathrm{r}$ can be rewritten as:

$$x_r = W_{yr} = \sqrt{P_s} W k_{sr} x + W_{nr} = \sqrt{P_s} \mathrm{diag}(k_{sr}) \omega x + \mathrm{diag}(\omega) n_r \tag{19}$$

The known signals and the intruder are:

$$y_d = {k_{rd}}^T x_r + n_d = \sqrt{P_S} {k_{rd}}^T \mathrm{diag}(k_{sr}) \omega x + {n_r}^T \mathrm{diag}(h_{rd}) \omega + n_{d.} \tag{20}$$

$$y_e = {k_{re}}^T x_r + n_e = \sqrt{P_s} {k_{re}}^T \mathrm{diag}(k_{sr}) \omega x + {n_r}^T \mathrm{diag}(k_{re}) \omega + n_e \tag{21}$$

For intruder and known signals, signal to noise is calculated as below

$$\gamma_{rd} = \frac{P_s \omega^H \mathrm{diag}({k_{sr}}^*) {k_{rd}}^* {k_{rd}}^T \mathrm{diag}(k_{sr}) \omega}{\sigma^2 \omega^H \mathrm{diag}({k_{rd}}^*) \mathrm{diag}(k_{rd}) \omega + \sigma^2} \tag{22}$$

$$\gamma_{re} = \frac{P_s \omega^H \mathrm{diag}(k_{sr}{}^*) k_{re}{}^* k_{re}{}^T \mathrm{diag}(k_{sr})\omega}{\sigma^2 \omega^2 \mathrm{diag}(k_{re}{}^*) \mathrm{diag}(k_{re})\omega + \sigma^2} \tag{23}$$

The secrecy rate is defined as a max, $\{R_d - R_e, 0\}$ where R_d and R_e are, respectively, the intruder and known rates are,

$$R_d = \frac{1}{2}\log(1 + \gamma_{rd}) \text{ and } R_e = \frac{1}{2}\log(1 + \gamma_{re}) \tag{24}$$

Thus, the secrecy rate is (Fig. 7):

$$R_s = \frac{1}{2}\log(1 + \gamma_{rd}) - \frac{1}{2}\log(1 + \gamma_{re}) \tag{25}$$

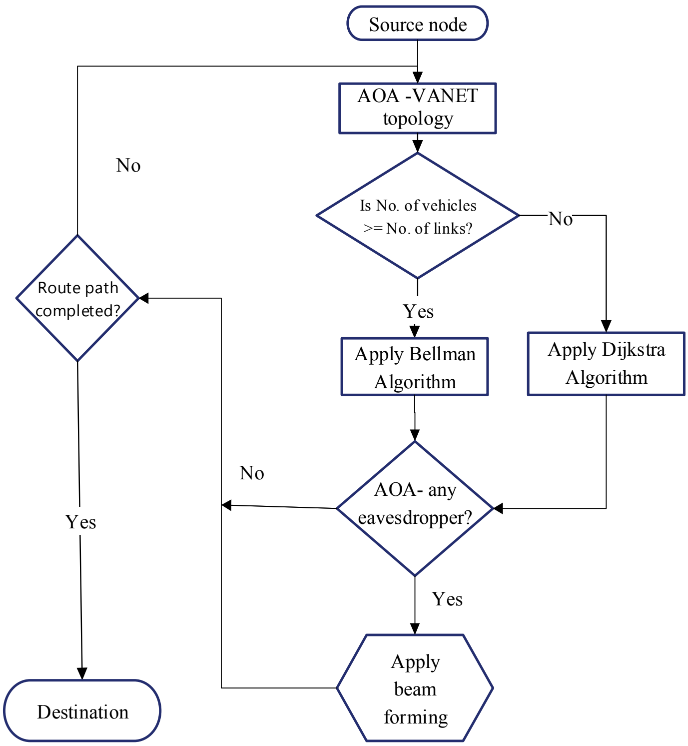


Fig. 7 Flow chart of the overall system

Simulation and Discussions

The efficient results were carried out with MATLAB, varying the number of antenna number for different beamforming algorithm considering total 64 number of antennas used in the relay. The simulation parameters are given for each result obtained with brief discussions.

By beam steering or scanning of signal around 360° the dynamic network topology of VANET identified that there is arrival of signal at angle which is shown in Fig. 8.

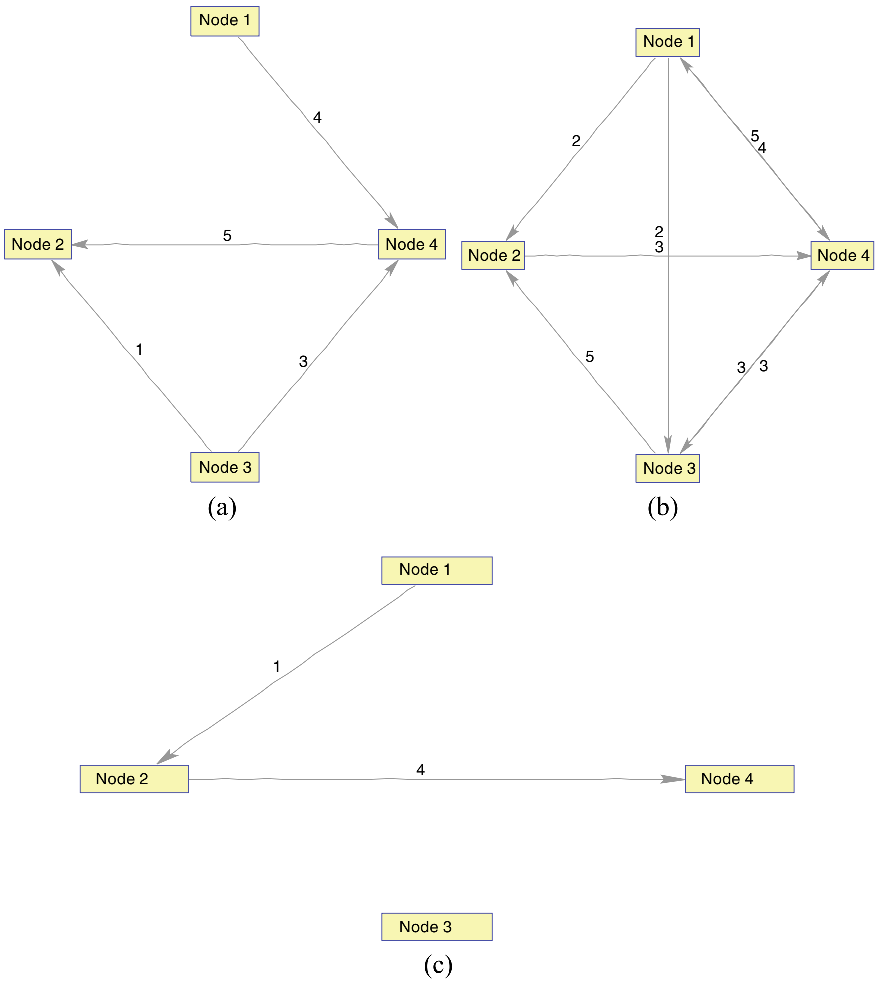


Fig. 8 Number of links for four vehicles: **a** 4-link, **b** 8-link, **c** 2-link

Table 2 Cost value for Bellman and Dijkstra algorithm for four vehicle nodes

a = Bellman				b = Dijkstra				c = Bellman			
0	1	∞	5	0	2	2	5	0	1	∞	5
∞	0	∞	∞	7	0	6	3	∞	0	∞	4
∞	1	0	3	7	5	0	3	∞	∞	0	∞
∞	5	∞	0	4	6	0	3	∞	∞	∞	0

Table 3 Simulation parameter for DoA of eavedropper

No. of antenna essentials	16
Element separation	0.5
No of intended signals 2	1
Angle of arrival of signal of interest	5
Angle of arrival for eaves dropper	−20
Noise variance	0.1

Table 4 Simulation parameter for DoA of eavesdropper

No of antenna essentials	8–64
Element parting	0.5
No of signals	2
Angle of arrival of signal of interest	0
Angle of arrival for eaves dropper	−15
Channel variance	0.1

The link connects the nodes varies instantly with the node movements. Based on the number of link and node appropriate Bellman and Dijkstra applied. The cost value based on distance and time taken is given in Table 2.

Table 5 Comparison of desired and array output signal

Iteration number	Desired signal output	Array output	Error signal
5	1	0	1
10	0.8	0.6	0.2
20	0.2	0.25	0.05
30	−0.04	−0.45	0.05
40	−0.08	−0.06	0.02
50	−1	−1	0
60	−0.8	−0.8	0
70	−0.04	−0.04	0

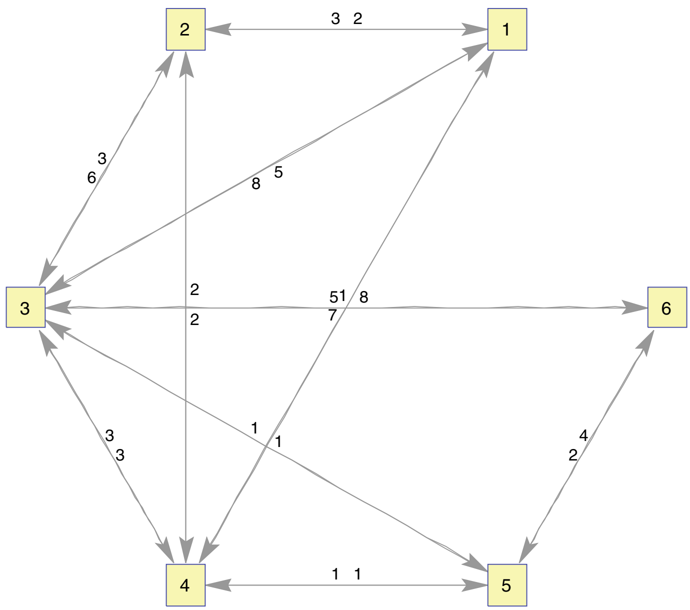


Fig. 9 Routing scenario

The cost value shows Bellman Ford algorithm is efficient for link numbers <= nodes, while Dijkstra algorithm is best, while link available is greater than the number of nodes.

Min norm and Root MUSIC algorithm identify the direction of incoming signal of intended user and eaves dropper which is shown in Fig. 9 (Table 3).

The power received from two direction is different. Signal of interest is identified at 5° and eaves dropper is −20°. The min norm algorithm estimated the direction of arrival of signal based on the signal strength which is shown in Fig. 10.

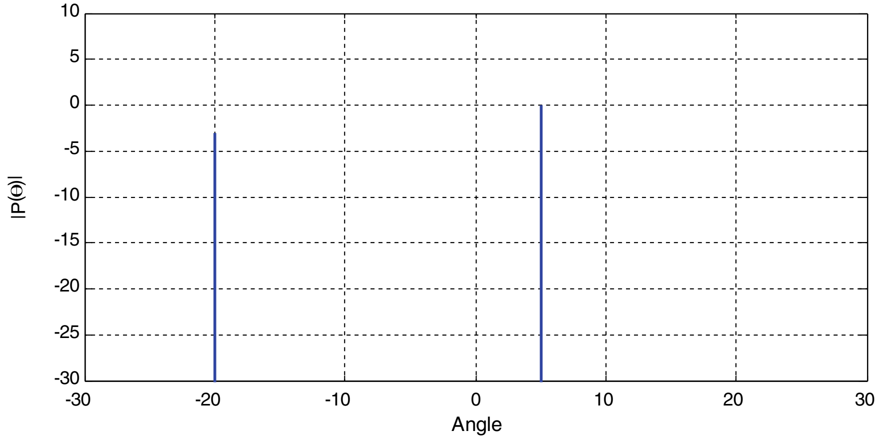
Bellman Ford algorithm for complex node consume more energy. As shown above the Dijkstra algorithm is effective fore complex node but not effective for a smaller number of vehicles. However, the combination of the algorithm saves more energy for the same node in a given time which is shown in Fig. 11.

Using adaptive beamforming technique, the beam pattern steered or shifted to the identified signal of interest direction identified, while null is put at eaves dropper incoming signal direction.

As shown in Fig. 12, the SNR increase with increase in antenna array. For 8 number of antenna array is 28.9955 dB and as the number of antenna increased to 32, the value also increased to −4262 dB.

The secrecy from interception or overhear of the information further increased by using optimal number of antennas which is shown in Fig. 12 (Fig. 13 and Table 4).

Now the error in transmitting the information is calculated using the LMSE algorithm.



(a)

Fig. 10 Power received versus angle of incoming

The simulated LMS error indicates the error of wanted signal and eavesdropping trying to retrieve information after applying n number of iterations. As the no. of antenna elements were increasing the error is decreases.

Now, based on the received signal strength the secrecy rate and signal-to noise ratio of the signal is maximized by increasing number of array elements.

The information is then routed with improved signal quality using adaptive beamforming LMS algorithm which is shown in Fig. 14.

By arranging the array weight, the error of the array output and desired signal is adjusted. This array weight vs iteration number above shows that after a number iteration the array weight becomes constant.

The quality of information transmission improves after a number of iterations, as shown in Fig. 15, the array output fit the desired signal after 55 iteration is carried out which is shown in Fig. 15 (Table 5).

Conclusions

The frequent link failure due to dynamic nature VANET topology requires suitable routing. The combined shortest path of Bellman and Dijkstra algorithm result showed better energy efficiency. During routing of information from source to next hop or destination the wireless information transmission is vulnerable to eaves dropper security attack. To improve the information secrecy array antenna beamforming technique is used. The beam steering technique employed to scan the angle of arrival of available signal around the source node and direction of incoming eavesdropper. The min norm and ROOT MUSIC algorithm beam former simulation shows the

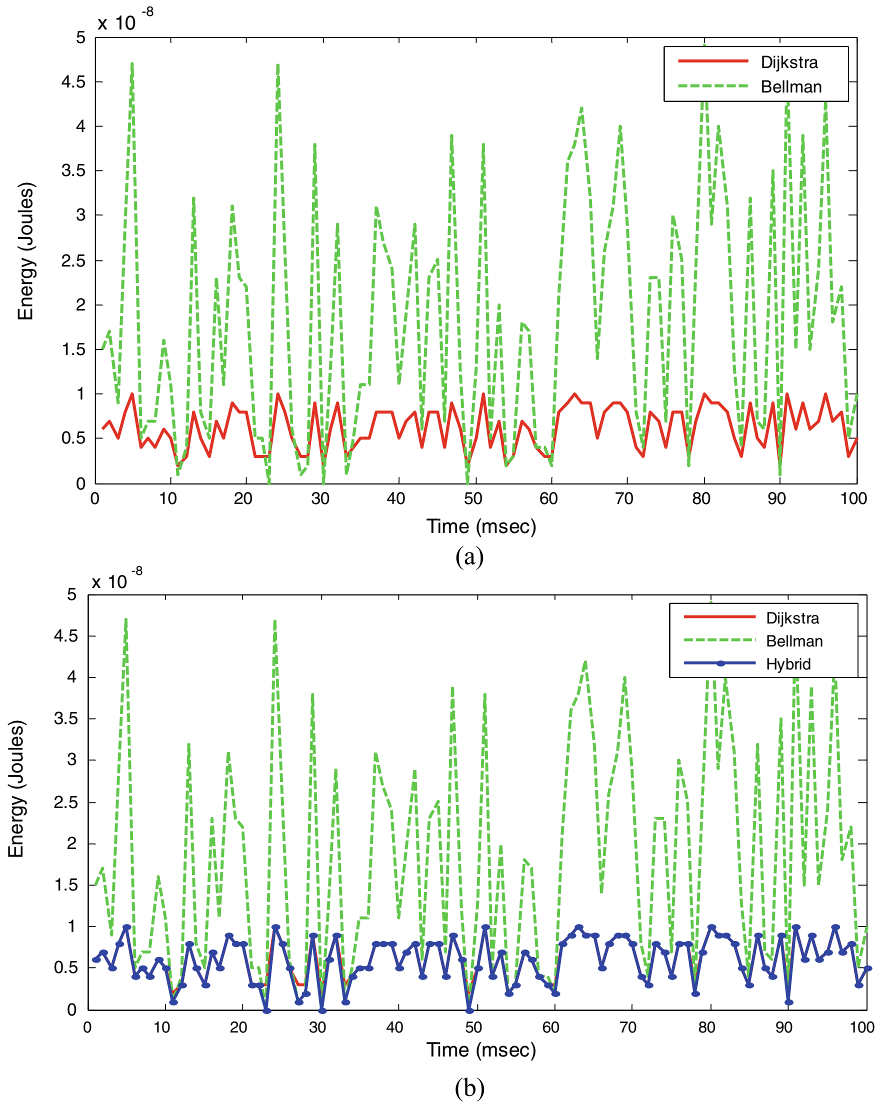


Fig. 11 Energy efficiency. **a** Bellman Ford versus Dijkstra. **b** Hybrid algorithm

main beam directed to the intended node direction, while null is put in the direction of incoming eaves dropper is simulated. Finally, LMS algorithm result indicated maximum SNR the high quality of transmitted information. Also, the analytical model compared with simulation result shows greater secrecy improvement.

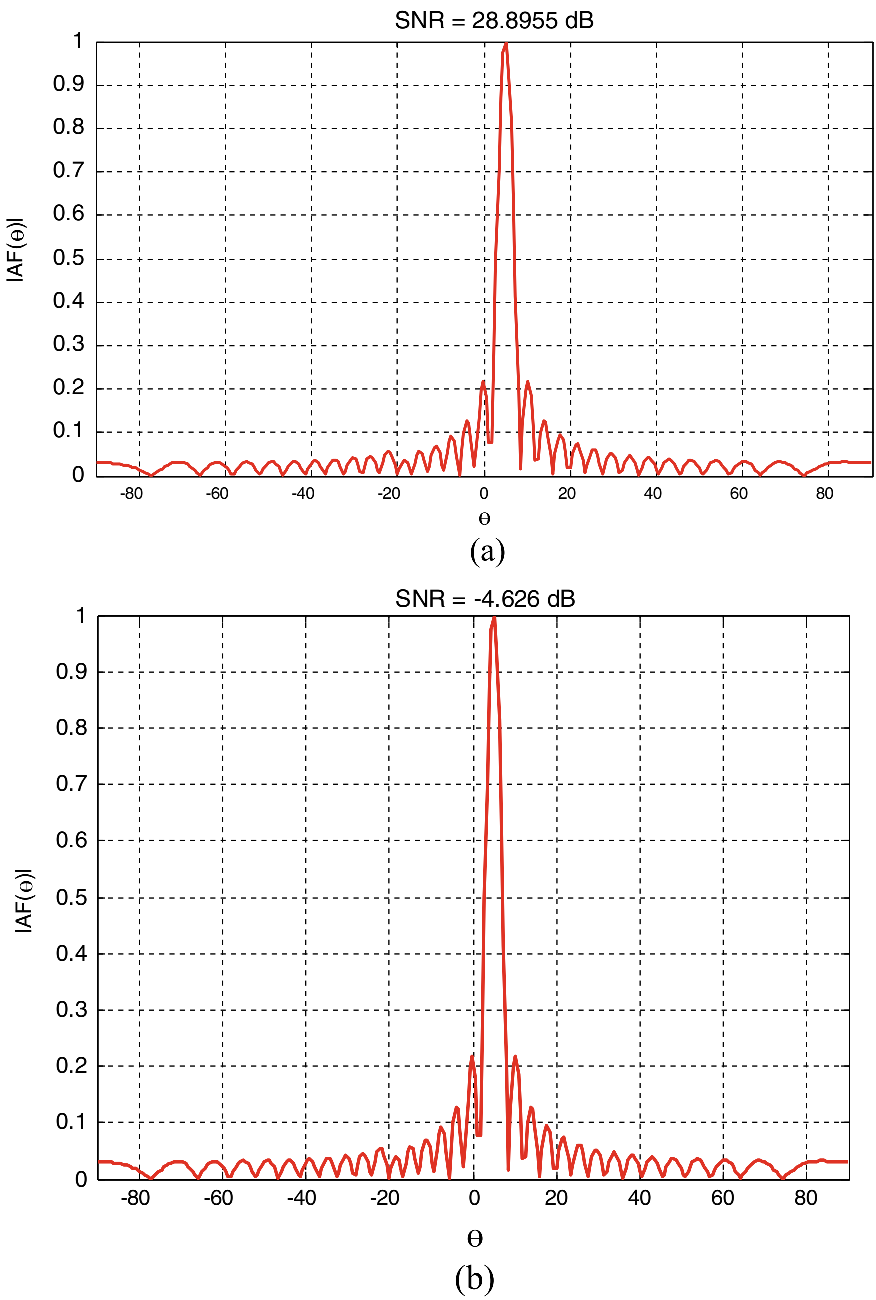
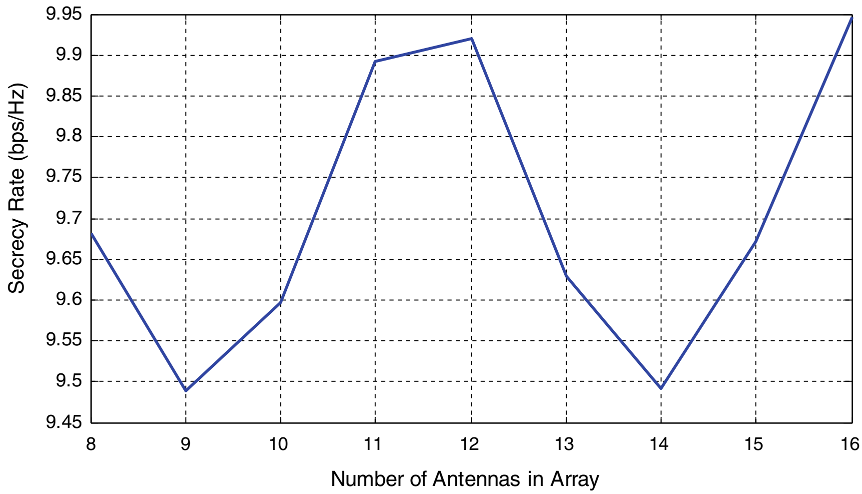
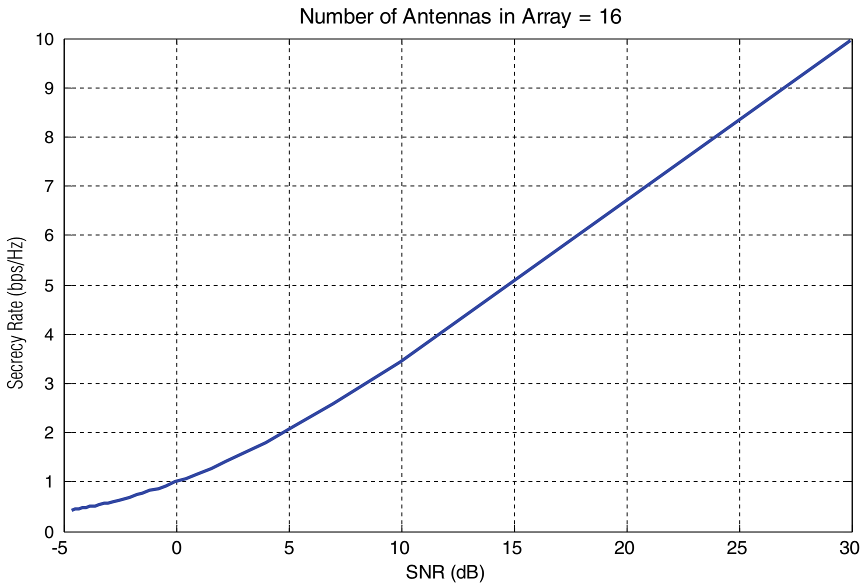


Fig. 12 Array factor versus SNR for **a** 8 antenna, **b** 32 antenna



(a)



(b)

Fig. 13 SNR versus secrecy rate for **a** 8 antenna, **b** 32 antenna

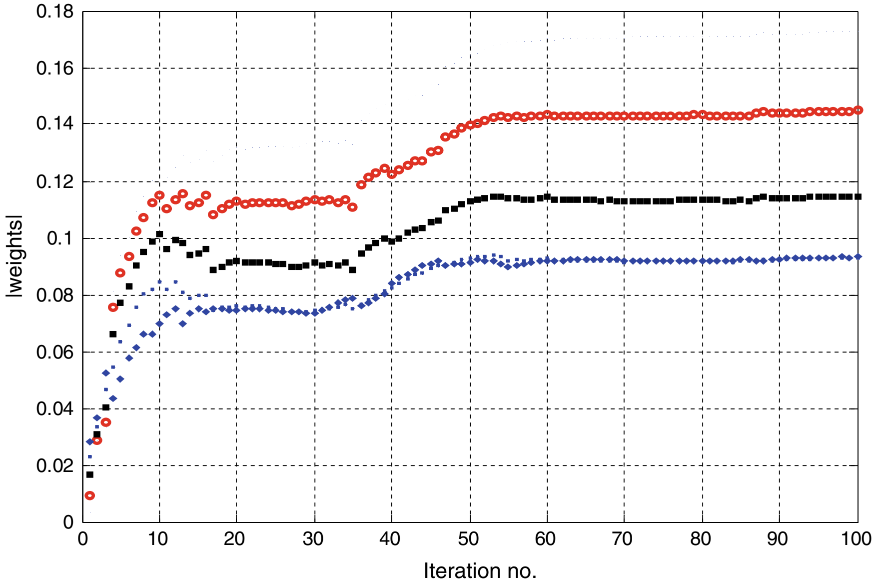


Fig. 14 Array weight

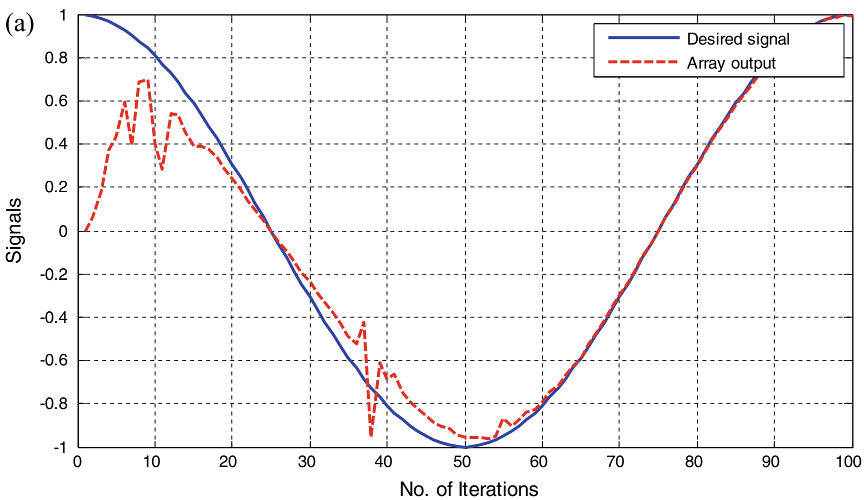


Fig. 15 **a** Array out put vs desired signal. **b** LMSE error

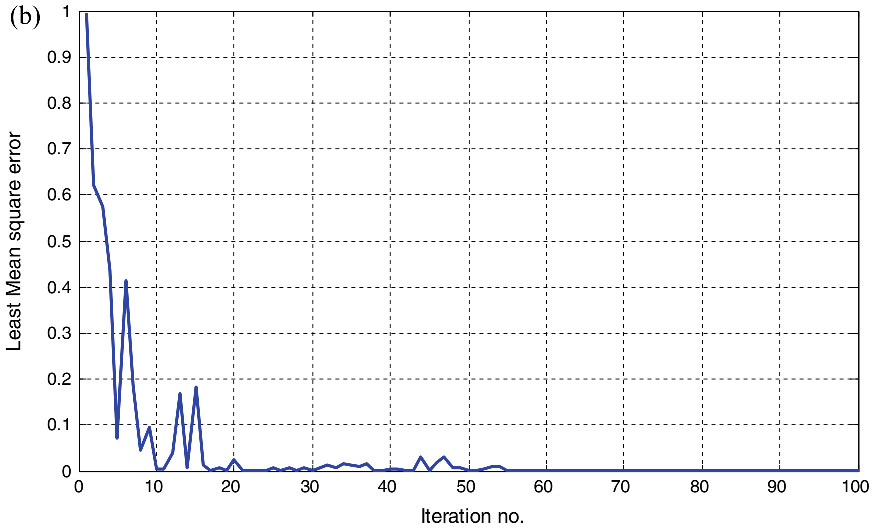


Fig. 15 (continued)

References

1. Godara L (1997) Application of antenna arrays to mobile communications, part II: beam-forming and direction-of-arrival considerations. Proc IEEE 85(8):1195–1245
2. Johnson D (1982) The application of spectral estimation methods to bearing estimation problems. Proc IEEE 70(9):1018–1028
3. Van Trees H (2002) Optimum array processing: part IV of detection, estimation, and modulation theory. Wiley Inter Science, New York
4. Shan T-J, Wax M, Kailath M (1985) Spatial smoothing for direction-of-arrival estimation of coherent signals. IEEE Trans Acoust Speech Signal Process ASSP-33(4):806–811
5. Minasyan G (1994) Application of high resolution methods to underwater data processing. PhD dissertation, N.N. Andreyev Acoustics Institute, Moscow (In Russian)
6. Liberti J, Rappaport T (1999) Smart antennas for wireless communications: IS-95 and third generation CDMA applications. Prentice Hall, New York
7. Schmidt R (1986) Multiple emitter location and signal parameter estimation. IEEE Trans Antenna Propog AP-34(2):276–280
8. Applebaum S (1966) Adaptive arrays, Syracuse University Research Corporation, Rep. SPL TR66-1
9. Widrow B, Mantey P, Griffiths L et al (1967) Adaptive antenna systems. In: Proceedings of the IEEE, vol 55
10. Suresh Kumar Reddy K, Tamizhazhagan V, Senthil Murugan V (2021) 5G VANETs: a details performance analysis of fusion beam forming techniques for vehicular environment. Turkish J Comput Math Educ 12(11):5518–5533
11. Khosroshahi AH, Keshavarzi P, KoozehKanani ZD, Sobhi J (2011) Acquiring real time traffic information using VANET and dynamic route guidance. In: 2011 IEEE 2nd international conference on computing, control and industrial engineering, Wuhan
12. Pandey K, Raina S (2015) Hop count analysis of location aided multihop routing protocols for VANET
13. Vasudeva Reddy Y, Nagendra M, Vijaya Kumar G (2010) Current research work on routing protocols for MANET: a literature survey, vol 2, no 3, pp 706–713

14. Kurien A, Diana A (2013) Survey on various position based routing protocol in vehicular ad hoc network, vol 2
15. Gabriel W (1983) Adaptive processing antenna systems. IEEE Antenna Propag Newslett 5–11
16. Reddy KSK, Tamizhazhagan V (2020) An innovative selection of road side unit emplacement in VANETs. In: Karrupusamy P, Chen J, Shi Y (eds) Sustainable communication networks and application. ICSCN 2019. Lecture notes on data engineering and communications technologies, vol 39. Springer, Cham. https://doi.org/10.1007/978-3-030-34515-0_39
17. Agee B (1986) The least-squares CMA: a new technique for rapid correction of constant modulus signals. In: IEEE international conference on ICASSP '86, vol 11, pp 953–956
18. Sorenson H (1970) Least-squares estimation: from Gauss to Kalman. IEEE Spectr 7:63–68
19. Rajaveerappa D, Belgassem FH (2010) A two ray propagation model for finding mobile location over microcellular environment. In: International conference on computer and communication technology 2010 (ICCCT2010), (IEEE Explore) IEEE Computer Society, MNNIT, Allahabad
20. Rajaveerappa D, Siddappa Naidu K (2011) Diffraction propagation model for finding mobile location over rough surface adhoc microcellular environment. In: International conference on smart technologies for materials, communication, controls, computing & energy 2011 (ICST2011), VELTECH University, Chennai
21. Suresh Kumar Reddy K, Rajaveerappa D (2017) Design and development of an efficient location based bandwidth allocation method for wireless mobile networks. Int J Control Theory Appl (ISSN: 0974-5572) (Scopus) 10(13):141–151
22. Rajaveerappa D, Sreenivasan M (2002) Design and development of a hybrid SDMA/CDMA MAC protocol for wireless mobile computing networks. In: The sixth international conference/exhibition on high performance computing in Asia-Pacific region (HPC ASIA 2002), vol 2, pp 561–562
23. Suresh Kumar Reddy K, Tamizhazhagan (2019) An eloborative analysis of bearing estimation techniques for VANETs. J Adv Res Dyn Control Syst 11:1457–1466

Evaluation of Dynamic Frequency Control on an Automotive Microcontroller

Anmol Kaushik, Shivaprasad Chumbalakar, Surya Musunuri, and Anju S.Pillai

Abstract This article presents an extensive study which was performed on Infineon's Aurix2G microcontroller to evaluate the impact of overclocking and underclocking on several parameters such as throughput, current, and power. Guided by the study, both overclocking and underclocking were utilized together to develop a dynamic frequency control approach to application development. An experimental application was divided into an underclocked zone and an overclocked zone. Switching between these two zones provided benefits in multiple application scenarios as compared to various static frequency approaches. In the evaluated application scenarios, the dynamic frequency control approach provided up to 33% reduction in execution time and up to 49% reduction in overall energy consumption, in comparison with the default static approach. The proposed approach also provided a reduction in average current and power in most of the evaluated application scenarios.

Keywords Overclocking · Underclocking · DVFS · Dynamic frequency scaling · Microcontroller · Automotive · Power efficiency

A. Kaushik (✉) · A. S.Pillai
Amrita Vishwa Vidyapeetham, Coimbatore, India

A. S.Pillai
e-mail: s_anju@cb.amrita.edu

S. Chumbalakar · S. Musunuri
Infineon Technologies, Bangalore, India
e-mail: Shivaprasad.Chumbalakar@infineon.com

S. Musunuri
e-mail: Surya.Musunuri@infineon.com

V. Bindhu et al. (eds.), *Proceedings of Third International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 844,
https://doi.org/10.1007/978-981-16-8862-1_21

Introduction

As the automotive industry trends toward electrical vehicles (EVs), the requirement of getting the best possible throughput while expending minimal power and energy has become very significant. Minimization of total energy consumption is of significant importance to battery-powered vehicles. With the introduction of applications such as advanced driver assistance systems (ADAS), the need for a high processing throughput is also of great importance to modern automobiles. CMOS-based circuits (such as microcontrollers) dissipate power in three major forms: static, dynamic, and short circuit [1, 2]. Static power is dissipated due to the presence of junction leakage currents [3] and is not a function of the switching frequency of a CMOS gate [1]. There are several approaches beyond the scope of this paper that can minimize static power dissipation [4]. Short-circuit power draw is caused during the time period when both NMOS and PMOS transistors conduct, causing short-circuit current to flow from supply to ground [5].

In our work, the focus is only on the dynamic power dissipation of a CMOS circuit, i.e., the switching power dissipation caused by the charging and discharging of gate capacitances when the output changes between high and low logic [1]. Dynamic power dissipation in CMOS circuits can be formalized as [1]:

$$P(\text{Dynamic}) = \alpha \mathrm{Cf}_{\mathrm{clk}} V_{\mathrm{dd}}^{2} \tag{1}$$

where f_{clk} is the clock frequency, C is the output capacitance, V_{dd}^{2} is the supply voltage, and α is the activity factor.

It is evident from Eq. 1 that a change in operational clock frequency has a direct impact on the dynamic power dissipation of a CMOS circuit. P(Dynamic) is proportional to f_{clk}; therefore, reducing f_{clk} will cause a corresponding decrease in dynamic power [2]. But a reduction in f_{clk} usually has an adverse effect on the computational throughput.

Operating a component beyond its manufacture-specified frequency rating is known as overclocking (OC). Increasing the operational clock frequency of a device generally results in more operations performed per second by the device. Thus, overclocking a device can provide a boost to the device's computational throughput. However, this boost in throughput comes at a price. As observed from Eq. 1, an increase in operational frequency leads to an increase in the current drawn and power dissipated by the device. Underclocking (UC) on the other hand is the process of operating a device below its specified maximum operational frequency. In doing so, the current draw and power dissipation by a device are reduced but at the cost of scaled down computational speed.

In the proposed dynamic frequency control (DFC) methodology, we tried to get the best of both above-mentioned approaches (OC and UC). A lot of automotive applications can be analyzed and segregated into two phases, i.e., computationally intensive data processing and low-speed data transmission phase. In our DFC approach, overclocking is performed during computationally intensive tasks (such as CRC32 calculations and DSP algorithms) to boost throughput and UC is performed during less computationally intensive tasks (such as low-rate data transmission/reception and waiting states). Impact of OC and UC on parameters such as throughput, average current, and power has been analyzed in our experiments. In the application part, a comparison of DFC versus static frequency approach has been performed under multiple application scenarios in terms of throughput, average current, power, and total energy consumption.

The rest of the paper has been organized as follows: Sect. "Related Work" provides a summary of the literature survey that was performed to examine other instances of DFC. Section "Experimentation and Analysis" covers implementation details and result analysis of all the experiments that were performed to evaluate the impact of OC and UC. Section "Dynamic Frequency-Based Application" provides insights on the DFC application implementation and analysis of the obtained results. Section "Conclusions" provides a conclusion to the paper.

Related Work

A literature survey was performed to analyze other instances of dynamic frequency control/scaling on microcontrollers. A detailed study focusing at DFC on microcontrollers was performed by the authors in [6]. In this study, the authors suggested minor modifications on both hardware and software to implement a DFC methodology that yielded up to 10% power savings. The study initially involved estimating the power consumed by each assembly instruction. This was followed by an estimate of the energy consumed by the instructions in the PIC16F's instruction set at various frequency levels. This information guided their development of a DFC approach. On the hardware side, the authors utilized a programmable clock generator which provided the external clock input to the microcontroller. This programmable clock generator was implemented on a field-programmable gate array (FPGA) in conjunction with a soft processor which could manipulate the output clock frequency based on external factors such as throughput of a system. The programmable PLL blocks in the Aurix2G microcontroller (the target for the proposed work) provide the same programmable clock generation functionality in chip.

The authors in [6] also identified a critical concern when employing DFC on a microcontroller which is the fact that microcontroller-based applications utilize constant frequency values and do not expect a switch in the middle of execution. Switching frequencies mid-execution of an application can lead to violations of a microcontroller's sub-modules' timing constraints and can also lead to inconsistencies in processes such as baud rate calculations. Therefore, such switches must be performed with care at a software level. The authors dealt with this issue with the development of a C-based library and driver modifications which accommodated these frequency shifts and changes in baud rate calculations. With a combination of these approaches, a test application was developed where calculations were performed on incoming data and the results were displayed on a liquid crystal display (LCD).

The use of a clock frequency switching technique to reduce the energy consumed by a microcontroller was discussed in [7]. The authors utilized the PLL block of an LPC2387 to generate a high PLL output frequency during data processing (in their case JPEG compression) but a lower output PLL frequency during data transmission. The lack of waiting states in the data processing part is crucial for this approach to work. This way the increase in power consumption experienced due to the use of a higher frequency is offset by the reduction in processing time to yield energy savings. The use of a low frequency during waiting states and data transmission provides an added reduction in current and power draw.

In [8], several approaches to dynamic frequency and voltage scaling (DVFS) in biomedical embedded systems were examined by the authors. The focus of the research was to develop a biomedical system that is capable of delivering max computation power (during a task such as anomaly detection) as needed while maintaining optimal energy consumption levels (to maximize battery life) [8]. The authors developed a dynamic approach to DVFS which they termed as DIAF. The aims of this approach were to develop a dynamic, intelligent, and adaptable methodology for voltage and frequency scaling. The critical features of their approach were identification of optimal performance and energy consumption at multiple operating points, and a diagnostic loop to determine the optimal operating point for various network conditions.

Another example of DFC was discussed in [9], where the authors utilized DFC in a robotics-based application (automatic path tracking) to conserve energy. The target microcontroller was the ATMEGA 16. Static frequency and DFC were compared, and it was observed that DFC provided up to 4.7% reduction in energy consumption. Dynamic frequency scaling (DFS) and dynamic power management (DPM) are two of the most popular mechanisms for conserving power in mobile devices [10].

Instances of DFS and DPM in the sphere of traditional embedded systems are limited but provide the potential for significant gains. It was observed from the survey that experimental data over a wide frequency spectrum was not provided. Our paper aims to provide the impact of OC/UC on a variety of parameters over a wide clock frequency spectrum. Guided by the experimental data, an application was developed to showcase the effects of utilizing both OC and UC in the same application.

Experimentation and Analysis

This section details the implementation of various experiments that were performed to analyze the impact of OC/UC on a variety of parameters. Also, results from the experiments have been described and analyzed in this section.

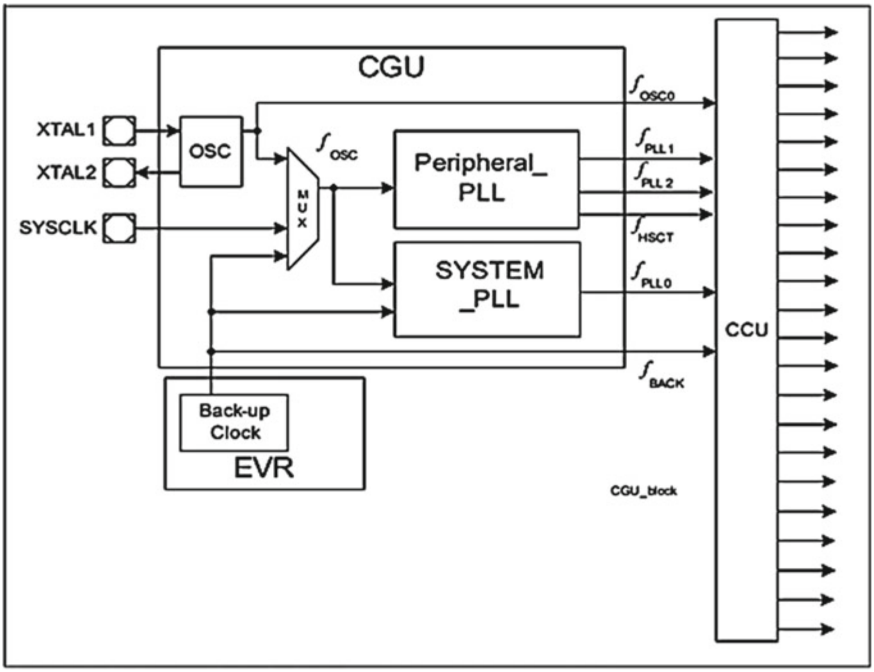


Fig. 1 Clock generation unit of A2G

Clock Generation Module of Aurix2G

The device under testing (DUT) for this project was Infineon's Aurix second-generation family (A2G) automotive microcontroller. A high-level view of A2G's clock generation unit (CGU) has been provided in Fig. 1 [11]. As depicted in Fig. 1, the A2G microcontroller consists of two internal phase-locked loop (PLL) blocks, i.e., system PLL (SYSPLL) and peripheral PLL (PERPLL). The output clocks from these two PLL blocks are fPLL0 and fPLL1, respectively.

SYSPLL Configuration

The output frequency of the SYSPLL is described by the following relation [11]:

$$\text{fPLL0} = (\text{fOSC} \times N) / (P \times K2) \tag{2}$$

As described in Eq. 2, the input frequency (*fOSC)* is scaled based on the value of scalars (*N, P,* and *K2*). These scalars are mapped to the PLL configuration registers

Table 1 Upper bounds for OC

Upper bound (MHz)	VDD (V)	% OC (%)
505	1.25	68.3
580	1.35	93.3
450–480	1.25	50–60

and can be manipulated by editing these register bit fields. By modifying these scalar values, the output frequency of the SYSPLL can be controlled.

Following the start-up procedure of the A2G microcontroller, the SYSPLL operates at a frequency of 300 MHz. The default fPLL0 value (300 MHz) was achieved with the following configuration: fOSC = 20 MHz, $N = 60$, $P = 2$, $K2 = 1$.

Determination of the Upper Bound for OC

In this experiment, the objective was to push the fPLL0 frequency as high as possible if the microcontroller could perform certain basic tasks. These basic tasks included performing CPU-based basic arithmetic and floating-point operations such as: addition, subtraction, multiplication, division, and factorial calculation. Another task was port toggling through delays generated by the system timer (STM). This was done to evaluate the upper limits of the system on chip (SoC). Additionally, this experiment provided the upper OC bound for the experiments as well as the DFC application. The results of this experiment are given in Table 1.

Here, 505 MHz and 580 MHz upper bounds are suitable for basic CPU-based computations only. The 450–480 MHz range was found to be suitable for complex applications that utilize multiple peripherals. Going past, these upper bounds led to bus errors and resulted in a trap. It is important to note that these results are applicable only at room temperature.

SYSPLL Clock Frequency Switching Latency

As displayed in Eq. 2, either of the N, P, $K2$ scalars can be varied to modify the fPLL0 value. Changing the P and N scalars can lead to a loss of lock of the SYSPLL and can result in the use of the backup clock (fBACK, which operates at 100 MHz) instead of the fPLL0. Therefore, there is a set of steps detailed in [11] that must be followed while varying these scalars. Varying the $K2$ scalar on the other hand does not lead to loss of lock of the SYSPLL, so it is faster than varying P and N scalars. A combination of all three (N, P, and $K2$) can also be used to generate more targeted output frequency values. The frequency switching latencies have been detailed in Table 2. This analysis was of importance to the application development because if

Table 2 SYSPLL frequency switching latency

Method	Latency (μs)
Varying P and N	41.4
Varying K2	1.64

the frequency switching latency is too high, the benefits of DFC approach in terms of throughput could be diminished.

Impact of OC/UC on CPU Throughput

An increase in frequency generally leads to an increase in the number of operations performed per second. Therefore, this test established the pattern of decrease/increase in processing speed as we overclock and underclock the fPLL0 output from the SYSPLL. A custom test case was developed where a few CPU-based operations were performed on a large matrix of floating-point numbers (320 samples from a signal). The test case executed the following operations in a row: calculation of mean of the elements of the matrix, calculation of variance of the matrix's elements, and the convolution operation of this matrix with another one (impulse response of a low-pass filter). The execution time of this test case was measured at various frequencies ranging from 30 to 505 MHz. The results have been depicted in Fig. 2.

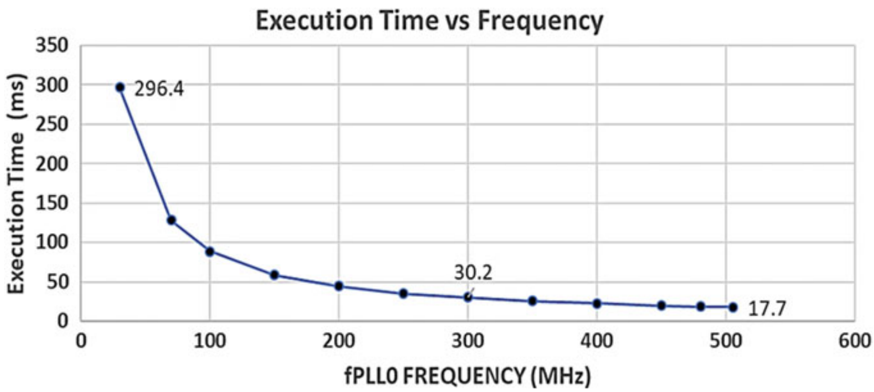


Fig. 2 Impact of OC/UC on test case execution time

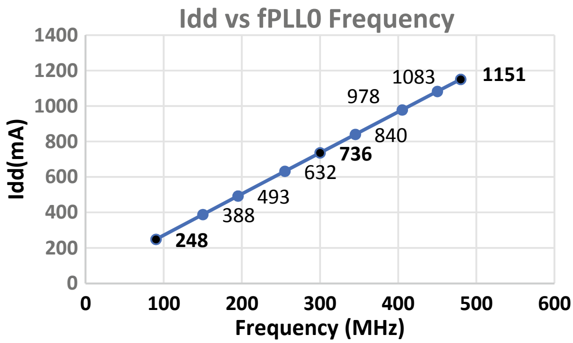


Fig. 3 Average I_{dd} versus fPLL0

Impact of OC/UC on Average Current, Power, and Core Temperature

As observed from Eq. 1, an increase in frequency causes an increase in the dynamic power dissipation of a CMOS circuit. This experiment dealt with analyzing the impact of OC/UC on current draw, power consumption, and core temperature. To obtain these readings, an Infineon-developed stress test was used to stress the various IPs of the microcontroller (such as CPU cores, ERAY, CAN). While the stress test was executing—current, power, and core temperature (core temp.), measurements were obtained, and average was calculated. The current and power readings were obtained directly from the digital power supply using a Python-based automation script. The readings have been described in Table 4.

It can be observed from Table 4 that with an increase in fPLL0 from 90 to 480 MHz, there is a corresponding increase in average current draw, power consumption as well as core temperature. This behavior is expected from the relationship described in Eq. 1. The plots of frequency versus current draw as well as frequency versus power consumption have also been provided above.

A linear relationship was observed between average current drawn and frequency at the 1.25 V voltage rail. A similar relationship was observed for power versus frequency as well (Figs. 3 and 4). The plot of core temperature versus frequency has also been presented in Fig. 5.

Conclusions from Experimentation

The experimentation part of the work resulted in very important observations that have been utilized in the application development section of the proposed work. The effect of OC/UC on CPU throughput as well as current draw and power consumption was considered while developing the DFC application. Other experiments also

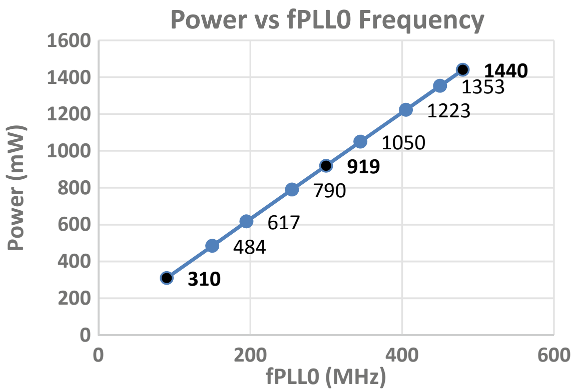


Fig. 4 Average power versus fPLL0

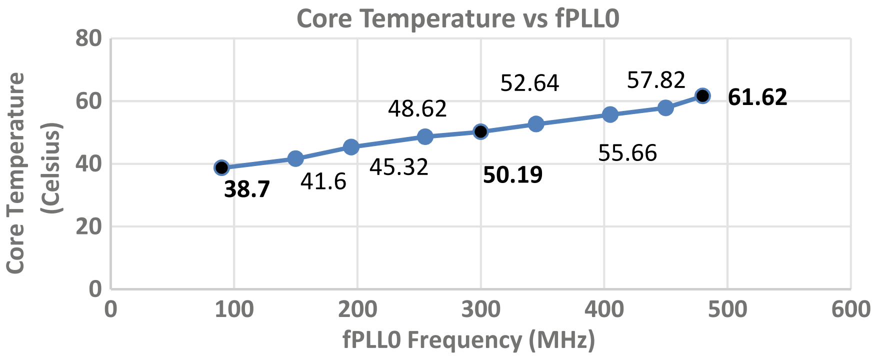


Fig. 5 Core temperature versus fPLL0

provided a lot of insights that guided the application development and helped in the identification of frequency scaling points.

Dynamic Frequency-Based Application

This section focuses on the high-level structure of the application and its implementation details. The results obtained and their analysis are also covered in this section.

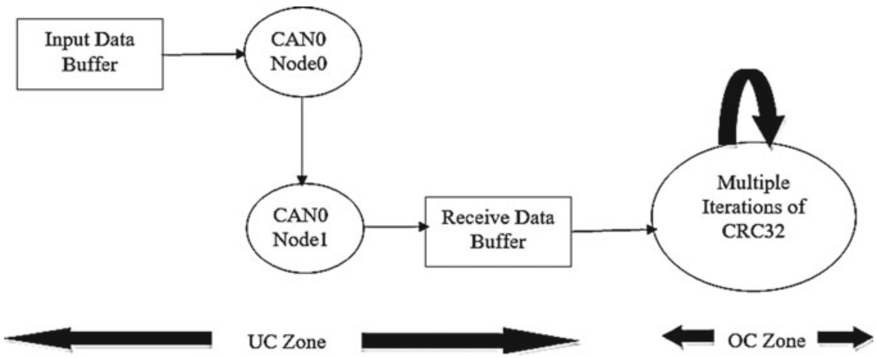


Fig. 6 High-level overview of the DFC application

Overview of the Application

A high-level overview of the application has been depicted in Fig. 6. Here, the objective was to implement an experimental application that mimics the following scenario: A sensor provides data at periodic intervals (every 5 ms) (waiting states), and this data is transmitted utilizing the controller area network (CAN) (data transmission) protocol to another node. Once whole of the data packet has been received (1000 32-bit words), multiple iterations of cyclic redundancy check (CRC32) are performed on this data (data processing). The application covers various modules of the Aurix2G microcontroller including the MCAN (Multi-CAN) module and the flexible CRC engine (FCE) [11]. The three important stages of the application are: the waiting states, data transmission, and data processing.

Description of DFC Behavior of the Application

This application utilizes a DFC approach where the application is divided into the underclocked zone and the overclocked zone. The fPLL0 is switched from a lower frequency to a higher one in between these zones. Special care was taken to ensure that the frequency switches were performed as carefully as possible in terms of maintaining MCAN module's as well as serial peripheral bus's (fSPB) timing and clock ratio constraints.

In the DFC approach, the execution starts at an underclocked frequency which is maintained during the CAN transmission/reception (Tx/Rx) and the waiting states (simulated sensor delay). Following this, a careful switch of the fPLL0 frequency is performed to overclock it. The data processing stage (multiple CRC iterations) is performed at an OC frequency. When the CRC calculations are finished, fPLL0 is underclocked again. A summary of this approach has been provided in Table 3.

Table 3 DFC application configuration

Task	fPLL0 status
Waiting states + data transmission/reception (CAN)	UC
Data processing (CRC32 calculations)	OC

Table 4 DFC application scenarios

Scenario	Application distribution	CRC calculation module	Number of CRC32 iterations
Scenario 1	tUC $\gg$ tOC	FCE	1000
Scenario 2	tUC $\cong$ tOC	FCE	50,000
Scenario 3	tUC $\ll$ tOC	CPU	1000

DFC Application Test Scenarios

The application has been tested under multiple configurations of the above-mentioned scenarios. By varying the time spent by the application in the two zones (UC and OC), we have evaluated different application scenarios. The following scenarios were considered for further analysis. Here, tUC is the time spent in the UC zone (CAN Tx/Rx + waiting states) and tOC is the time spent in the OC zone (CRC32 calculations). By increasing the CRC iterations or changing the computation method, we can make the data processing stage longer (and thus tOC). Scenarios generated by using this methodology have been described in Table 4.

Scenario I (tUC $\gg$ tOC)

This scenario tries to emulate a situation where the time spent for data transmission (CAN Tx/Rx) and in the waiting states is significantly higher than the time spent for the data processing (CRC32). The following results were observed upon the execution of this scenario as given in Table 5. Here, tPROC refers to the time taken during the data processing stage, i.e., time to complete all the CRC32 calculations.

Table 5 Results from scenario 1

Scenario	fPLL0 (MHz)	tPROC (ms)	tTotal (ms)	Avg. I_{dd} (mA)	Avg. power (mW)	E (J)
Static UC	100	187.23	2.69	141.2	176.53	0.475
Default	300	62.79	2.56	298.9	372.26	0.954
Static OC	450	41.86	2.543	420.9	526.1	1.227
DFC	100 $\geq$ 380	49.25	2.553	141.7	189.08	0.483

Table 6 Results from scenario 2

Scenario	fPLL0 (MHz)	tPROC (s)	tTotal (s)	Avg. I_{dd} (mA)	Avg. power (mA)	E (J)
Static UC	100	9.36	11.87	147.07	183.84	2.181
Default	300	3.14	5.64	313.06	390.7	1.766
Static OC	450	2.093	4.59	440.3	549.3	2.523
DFC	$100 \geq 380$	2.46	4.96	285.35	350.48	1.741

tTotal refers to the total execution time of the application including all the stages. E is the total energy consumed.

If we compare the default (no frequency scaling) and the DFC approaches, it is clear that DFC provides a substantial improvement in CRC32 calculation time (tPROC). At the same time, DFC approach results in a lower average current draw, power consumption, and total energy consumption when compared to the default approach. The static options do provide some benefits but at the price of significant setbacks in terms of throughput (in the case of static UC) or high current draw (in the case of static OC).

Scenario 2 (tUC $\cong$ tOC)

This scenario tries to emulate a situation where the time spent during data transmission (CAN Tx/Rx) and in the waiting states is lesser than or equal to the time spent for the data processing (CRC32 calculations here). The results observed from executing this scenario have been given in Table 6.

It is evident from the results that the DFC approach still provides a lower tPROC and at the same time lower average Idd and power when compared with the default approach. Same can be said about the total energy consumption. An interesting observation here is the fact that DFC's advantages in terms of current and power draw have started diminishing. This is because of the increase in the time spent in the data processing stage (OC zone). Still, in this scenario significant reductions in current, power, and energy were observed while providing better throughput.

Scenario 3 (tUC $\ll$ tOC)

This scenario emulates a scenario where time spent in the data processing stage (CRC calculations) is significantly higher than the time spent in the data transmission (CAN Rx/Tx) stage and the waiting states. On paper, this is one of the worst-case scenarios for our DFC approach in terms of power savings as a significant majority of the execution time is spent in the OC zone. An interesting thing to note here is the fact

Table 7 Results from scenario 3

Scenario	fPLL0 (MHz)	tPROC (s)	tTotal (s)	Avg. I_{dd} (mA)	Avg. power (mW)	E (J)
Static UC	100	36.08	38.59	147.2	183.98	7.06
Default	300	11.96	14.46	317.2	396.55	5.73
Static OC	450	8.02	10.52	446.6	559.11	5.88
DFC	$100 \geq 450$	8.007	10.48	379.94	474.99	4.96

that one of the CPU cores has been used to perform the CRC32 calculations here. The results observed while executing this scenario have been provided in Table 7.

In this scenario, the DFC approach provides us with a significantly improved tPROC and at the same time a much lower total energy consumption. A lower energy consumption has significant benefits especially in terms of prolonging the battery life of a device. In terms of average power consumption and current draw, DFC no longer provided benefits over the default approach as a slightly higher average Idd draw was observed. This is due to the prolonged OC zone phase because of the longer data processing phase. However, major reductions in processing time due to the OC led to significant energy savings over the default approach.

Summary of DFC versus Default

Table 8 describes a summary of the percentage differences in the various metrics between the DFC versus default approach. It can be observed that in all the three scenarios DFC leads to reduce processing times, and the same time reduced total energy consumption. In two out of the three scenarios, DFC also provides a lower average current draw and power consumption in two of the scenarios.

From the perspective of an application developer, the process of breaking down an application into waiting, data transmission, and data processing stages has the potential to be extended to workflows other than microcontroller-based ones (e.g., microprocessor-based). Potentially, if an application developer divides their application into the aforementioned phases and introduces DFC accordingly, the approach can provide the benefits that have been covered in this article. In microcontroller-based workflows, the differing factor (vs microprocessor-based) is that the application code is known in advance which aids the application analysis for DFC. Thorough

Table 8 DFC versus default approach

Scenario	tPROC (%)	Avg. I_{dd} (%)	Avg. power (%)	Total energy (%)
Scenario 1	−21.5	−52.59	−49.2	−49.4
Scenario 2	−21.6	−8.85	−10.29	−1.47
Scenario 3	−33.06	+19.77	+19.78	−13.43

knowledge of the application is instrumental in applying this DFC approach to an application. On the other hand, the vast number of peripherals situated on an automotive microcontroller with their respective timing constraints provides an additional layer of complexity when compared to microprocessor-based workflows.

Conclusions

The DFC approach to embedded application development has the potential to provide beneficial dividends if implemented with care. It is clear from what has been detailed in the paper that depending on the application, DFC can provide faster processing times and at the same time can result in savings in one or more of the following metrics: average current draw, average power consumption, and overall energy consumption. Infineon's Aurix2G automotive microcontroller was suitable for implementing the DFC approach due to the presence of two dedicated internal PLL blocks and scalable sub-module clocks. This, along with low frequency switching latencies, made the Aurix2G a suitable platform for the application. To perform DFC, an in-depth knowledge of the application is required as changing the operational frequency midway during execution carelessly can lead to violations of a microcontroller's sub-modules' timing and frequency specifications. Throughout the development of this application, special care was taken to make sure such violations do not occur especially in the case of the MCAN module. Additionally, the frequency switches should be brief and should not significantly impact the throughput of an application. With these factors taken care of, DFC can prove to be useful to application developers that seek to minimize energy and/or power consumption and at the same time improve their application's throughput.

Future Work

This work provided experimental results that can guide future works in the field of DVFS on a microcontroller. It presented the impact of frequency scaling on several parameters over an extensive frequency spectrum. Additionally, factors such as PLL frequency switching latency are reported, something which is missing from a lot of the current literature. An experimental application that showcases an implementation of DFC on a modern automotive microcontroller was implemented and analyzed in a detailed manner. The complexities that affect such an implementation have also been detailed. Based on this work, future implementations could include dynamic voltage scaling along with DFC. Further, machine learning techniques could be used to generate optimal voltage–frequency settings for reduced energy and power consumption [12]. Another useful extension of this work could be the inclusion of this approach in a real-world automotive embedded application.

References

1. Kabbani A (2008) Modeling and optimization of switching power dissipation in static CMOS circuits. In: 2008 IEEE computer society annual symposium on VLSI, Montpellier, France, pp 281–285. https://doi.org/10.1109/ISVLSI.2008.83
2. Megalingam RK, Krishnan N, Arjun Ashok V, Arunkumar M (2009) Highly power efficient, uncompromised performance cache design using dual-edged clock. In: 2009 2nd IEEE international conference on computer science and Iinformation technology, Beijing, China, pp 538–542. https://doi.org/10.1109/ICCSIT.2009.5234660
3. Shiue W-T (2001) Accurate power estimation for CMOS circuits. In: Proceedings of IEEE region 10 international conference on electrical and electronic technology. TENCON 2001 (Cat. No.01CH37239), Singapore, vol 2, pp 829–833. https://doi.org/10.1109/TENCON.2001.949709
4. Ga Priya M, Baskaran KB, Krishnaveni DC (2012) Leakage power reduction techniques in deep submicron technologies for VLSI applications. In: Procedia engineering, Coimbatore, vol 30, pp 1163–1170
5. Veendrick HJM (1984) Short-circuit dissipation of static CMOS circuitry and its impact on the design of buffer circuits. IEEE J Solid-State Circuits 19(4):468–473. https://doi.org/10.1109/JSSC.1984.1052168
6. Lueangvilai A, Robertson C, Martinez C (2012) A dynamic frequency controlling technique for power management in existing commercial microcontrollers. J Comput Sci Eng 6:72–88. https://doi.org/10.5626/JCSE.2012.6.2.79
7. Duangmanee P, Uthansakul P (2018) Clock-frequency switching technique for energy saving of microcontroller unit (MCU)-based sensor node. Energies 11:1194. https://doi.org/10.3390/en11051194
8. Raskovic D, Giessel D (2009) Dynamic voltage and frequency scaling for on-demand performance and availability of biomedical embedded systems. IEEE Trans Inf Technol Biomed 13(6):903–909. https://doi.org/10.1109/TITB.2009.2030181
9. Pillai AS, Isha TB (2014) Dynamic frequency scaling based energy consumption reduction for power-aware embedded systems—a simulation and experimental approach. Int J Electr Syst 10(1):36–47
10. Sankaran S (2016) Predictive modelling-based power estimation for embedded multicore systems, pp 370–375. https://doi.org/10.1145/2903150.2911714
11. Infineon AG (2020) Aurix TC3xx user manual part I. Accessed 3 Aug 2020. [Online]. Available: https://www.infineon.com/dgdl/Infineon-AURIX_TC3xx_Part1-UserManual-v01_00-EN.pdf?fileId=5546d462712ef9b701717d3605221d96
12. Gupta M, Bhargava L, Indu S (2021) Dynamic workload aware DVFS for multicore systems using machine learning. Computing 103:1747–1769

Remaining Useful Life Prediction Using Machine Learning Algorithms

Malcolm Andrew Madeira, Teslin Jacob, and Sai Reddy

Abstract Machinery's Remaining Useful Life (RUL) is an effective instrument for maintenance and performance. As a consequence, expenses are reduced, safety is enhanced, and operations are improved. A comparison of available Machine Learning (ML) methods to anticipate the RUL is presented in this research. The ML models were built and tested using datasets from NASA's Prognostics Data Repository for turbo fan engine data. The obtained results were then compared to the real outcomes in order to assess the accuracy. To compare prediction accuracy, eleven ML methods were chosen. The various methods were evaluated in order to find the prediction model that best predicted the RUL in terms of number of cycles and also classified using binary classification and multi-class classification. These models are Linear Regression, LASSO Regression, Ridge Regression, Polynomial Regression, Decision Trees, Random Forest (RF), Logistic Regression, K-Nearest Neighbours (KNN), Support Vector Machines (SVM), Gaussian Naïve Bayes (NB) and neural net Multilayer Perceptron (MLP).

Keywords Remaining useful life · Machine learning · PHM · Predictive maintenance · Data science

Introduction

The current industrial technological development trend is industry 4.0. It combines automation, the Internet of Things (IoT), cloud computing and AI to increase the efficiency and quality of factory output [6]. "Prognostic and Health Management (PHM)" has recently become increasingly significant in a variety of industry activities [22]. PHM is made up of three main technologies: improved diagnostics, prog-

M. A. Madeira (✉) · T. Jacob
Goa College of Engineering, Farmagudi Ponda-Goa, India
e-mail: teslinjacob@gec.ac.in

S. Reddy
Research and Development Department, Energy Automation Siemens Ltd, Verna-Goa, India

V. Bindhu et al. (eds.), *Proceedings of Third International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 844,
https://doi.org/10.1007/978-981-16-8862-1_22

nostics and health management. The objective of prognostics is to provide valuable information regarding system degradation, such as the remaining useful life (RUL); RUL prediction and health management are the two primary components of PHM, extending the lifespan of equipment and scheduling scientific health care measures can both be accomplished by properly predicting the RUL of the equipment [22]. RUL is the amount of time that a running device or system can continue to function normally for a limited amount of time [7].

There is need for an efficient model and existing machine learning algorithms to be analysed for estimating the RUL for application in industry 4.0.

Background and Related Work

Four types of maintenance policies are identified in [36], namely run-to-failure (RTF), preventive maintenance (PM), condition-based maintenance (CBM) and predictive maintenance (PdM). RTF maintenance is when repairs or restoration-related operations are carried out after a failure, PM maintenance is normal service that is carried out on a regular basis with the goal of preventing process failures. CBM maintenance: when actions on the process are done after one or more conditions have been verified. (PdM) maintenance: similar to CBM, maintenance activities are only undertaken when they are essential. The RUL prediction techniques were classified into four groups in a review article [2]: Physics-based, Experimental-based, Data-driven and Hybrid-based. Support Vector Machine (SVM) is a data-driven method for solving nonlinear problems [11]. Auto-CNNLSTM, a data-driven RUL prediction technique [31]. ANN technique is used to learn from a large number of data points and improve prediction accuracy [9]. The maximum likelihood estimation (MLE) technique was utilised to estimate the RUL's PDF [17]. To estimate RUL and forecast reliability, Bayesian theory and a novel grid-sampling techniques are proposed [37]. Approach to RUL estimation it was proposed to use EMD and CNN [38]. An ensemble hybrid model with outlier detection to forecast the remaining usable life (RUL) of a lithium-ion battery is based on an UKF [23]. BCT method is used to optimise the data-driven RUL model [14].

As it is evident from the preceding literature, there are several data-driven models; due to dearth of literature on data-driven model comparison, this study conducts an experiment data-driven models, analyses their characteristics and provides a reference for selecting data-driven models for RUL prediction (Table 1).

Methodology

The general flow diagram of the methodology is depicted in Fig. 1.

Table 1 Comparative study of recent research papers

Paper	Objectives	Algorithms used	Results	Future scope and limitations
[3]	RUL Estimation of Lead Acid Battery using Bayesian Approach	The use of the MCMC sampling method to do Bayesian inference of linear regression	Gradient enhanced regression and a tree-based algorithm technique was the most successful	The suggested approach takes longer to execute, it can be enhanced by using advanced machine learning methods
[39]	On-board diagnostics (OBD) using DNN for RUL estimation	Deep Neural Network (DNN)	The results are shown on the website using flask and Azure framework after the learned model is trained in Azure ML Studio	Limitation is that it focused mostly on only one fault condition
[19]	RUL Prediction of Equipment in Production Lines	Multilayer Perceptron neural network (MLP) ML algorithm	With a single failure working mode dataset, the ML technique worked better	Doesn't operate on datasets with many faults. Advanced ML approaches can enhance the performance of the prediction model
[24]	An ensemble-based method to modelling and forecasting the RUL for an exponential deterioration process	RFs, CART, RNN, AR model, ANFIS, RVM, and EN	The results demonstrate that a prognostic method accurately forecast the RUL of an aviation engine by 89%	This has shortfalls, as different base learners were used for RUL prediction and can be improved using Fusion method
[8]	RUL predictions for turbofan engine degradation using "semi-supervised deep architecture"	CNN, DBN or LSTM algorithms with Genetic Algorithm (GA)	The GA method only utilised 20 people and three evolutions to optimise hyper-parameters for each subgroup in the dataset, whereas combining all the hyper-parameters resulted in a total of 8,748,000 combinations	The quality of the produced run-to-failure training data labels has a significant impact on the accuracy of RUL forecasts based on data-driven methods. In different fault circumstances, the accuracy suffers
[30]	Adversarial Transfer Learning for Machine RUL Prediction	Novel ADARUL, built on top LSTM network	During the test phase, the developed end-to-end architecture based on a deep LSTM network and adversarial domain adaptation that had less error	Needs to be tested on real-Time data

(continued)

Table 1 (continued)

Paper	Objectives	Algorithms used	Results	Future scope and limitations
[25]	"Predicting future capacities and RUL with uncertainty quantification"	EMD method, IMFs functions, LSTM, GPR	The RMSE and the maximum error are within 0.0048 and 0.038, indicating that a high accuracy is attained	Effective feature extractions after collecting valuable data for trend of Li-ion batteries, and improvement of data-driven approach for the battery "knee point" prediction
[35]	"RUL Estimation of BLDC Motor Considering Voltage Degradation and Attention-Based Neural Network"	RNN and LSTM are employed for constructing the model	Predicts the RUL as well as estimate the state-of-health of motor during operation	Using a real-time updating technique, the suggested model with attention mechanism may be further improved to collect fault-related information
[33]	Data-Driven Prognostics for RTF Data Employing ML Models	Linear regression and Extra Tree regression	The predictions were compared with RMSE and R2 measures and significant inferences on the outliers and feature selection was observed	Since there are less failure modes and degradations recorded, ML model might perform less efficient
[5]	RUL Estimation with Sobolev Training	DCNN	The model was able to learn the deterioration trend over time, which can be translated to the present RUL, with fewer parameters	Utilising real-time datasets to assess the effectiveness of training networks on various models and industrial applications

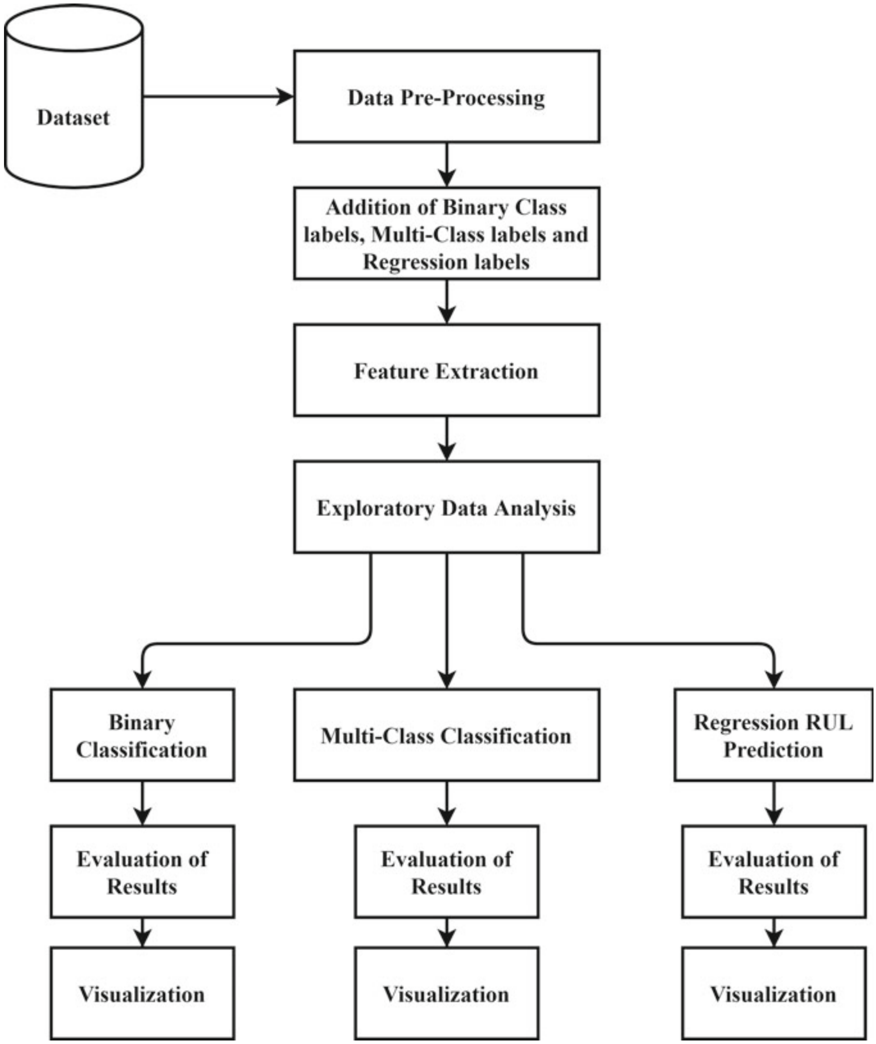


Fig. 1 General flow diagram

Dataset Description

The data was obtained from NASA's data archive. The RTF sensor readings from deteriorating turbofan engines are included in the dataset. The dataset schema is shown in Table 2.

The dataset is composed of 100 training records, 100 test records and a file that provides the actual RUL of 100 test records. After each observation period of each record, 21 sensor outputs were gathered, which are labelled as S1, S2, …, S21. [34] offers a comprehensive overview of the dataset.

Table 2 Dataset schema

S. No.	Columns	Data types	Descriptions
1	ID	Integer	Aircraft engine identifier
2	Cycle	Integer	Running cycle count
3	Setting 1	Double	Operational setting 1
4	Setting 2	Double	Operational setting 2
5	Setting 3	Double	Operational setting 3
6	S1 …S21	Double	Sensor measurement 1…21
7	TTF	Integer	Time to failure

	id	cycle	setting1	setting2	setting3	s1	s2	s3	s4	s5	...	sd15	sd16	sd17	sd18	sd19	sd20	sd21	ttf
0	1	1	-0.0007	-0.0004	100.0	518.67	641.82	1589.70	1400.60	14.62	...	0.000000	0.0	0.000000	0.0	0.0	0.000000	0.000000	191
1	1	2	0.0019	-0.0003	100.0	518.67	642.15	1591.82	1403.14	14.62	...	0.008697	0.0	0.000000	0.0	0.0	0.042426	0.003253	190
2	1	3	-0.0043	0.0003	100.0	518.67	642.35	1587.99	1404.20	14.62	...	0.007640	0.0	1.154701	0.0	0.0	0.055076	0.044573	189
3	1	4	0.0007	0.0000	100.0	518.67	642.35	1582.79	1401.87	14.62	...	0.028117	0.0	1.000000	0.0	0.0	0.076322	0.037977	188
4	1	5	-0.0019	-0.0002	100.0	518.67	642.37	1582.85	1406.22	14.62	...	0.025953	0.0	1.095445	0.0	0.0	0.073621	0.033498	187

Fig. 2 Loaded dataset along with the TTF label

The Engine RTF events, operating settings 1, 2 and 3 and sensor readings spanning from S1 to S21 are all included in the dataset. The training data file comprises 20,000+ cycle records for 100 engines, as well as run-to-failure data. Engine running data is included in the test data file, but no failure occurrences are recorded. A separate data file contains the remaining cycles for each engine; the real remaining cycles for each engine in the test data, indicated as TTF, are stored in the ground truth data file.

A snapshot of the loaded dataset along with the TTF label is shown in Fig. 2.

Data Pre-processing

For training data, the following regression and classification labels were created:

- Regression: For each cycle/engine, the "Time-To-Failure" (TTF) (number of remaining cycles before failure).
- Binary Classification: If the remaining cycles are less than a specific number of cycles (e.g. period = 35), the engine will fail in this time; otherwise, the engine is fine.
- Multiclass Classification: by subdividing the TTF into cycle bands (for example, 0–14, 15–35, 35+), we can determine when the engine will fail. TTF is given in a separate data file for test data. These two files were combined, and test data categorisation labels were produced in the same manner as stated previously.

Feature Extraction

The training and test data are subjected to feature extraction by adding two new columns for the 21 sensor columns: rolling mean and rolling standard deviation [12, 40]. Some ML methods perform better when the sensor values are smoothed over time. The final training dataset and test dataset were saved to csv files in preparation for further analysis.

"The unweighted mean of the preceding M data points forms the Simple Moving Average (SMA). The degree of smoothing sought determines M (sliding window) choices, as increasing M increases smoothing at the cost of accuracy" [40].

The SMA is calculated as follows in Eq. (1) for a sequence of data at time period t:

$$SMA_t = \frac{x_t + x_{t-1} + x_{t-2} + \cdots + x_{m-(t-1)}}{M} \quad (1)$$

where, SMA at time period t

The simple moving average was calculated by using the pandas. Series. rolling method. This method provides rolling windows over the data. On the resulting windows, we have performed calculations using a statistical functions as in our case Mean and Standard Deviation. The parameter window specifies the window's size (number of periods) [40].

The output of feature extraction using rolling mean and standard deviation is shown in Fig. 3 with window size as 5.

```
# adding extracted features to training data
df_train_fx = add_features(df_train_raw, 5)
test=df_train_fx.head(10)
test
```

	id	cycle	setting1	setting2	setting3	s1	s2	s3	s4	s5	...	sd12	sd13	sd14	sd15	sd16	sd17	sd18	sd19	
0	1	1	-0.0007	-0.0004	100.0	518.67	641.82	1589.70	1400.60	14.62	...	0.000000	0.000000	0.000000	0.000000	0.0	0.000000	0.0	0.0	0.00
1	1	2	0.0019	-0.0003	100.0	518.67	642.15	1591.82	1403.14	14.62	...	0.438406	0.035355	5.041671	0.008697	0.0	0.000000	0.0	0.0	0.04
2	1	3	-0.0043	0.0003	100.0	518.67	642.35	1587.99	1404.20	14.62	...	0.404475	0.026458	3.717450	0.007640	0.0	1.154701	0.0	0.0	0.05
3	1	4	0.0007	0.0000	100.0	518.67	642.35	1582.79	1401.87	14.62	...	0.495950	0.029439	3.050906	0.028117	0.0	1.000000	0.0	0.0	0.07
4	1	5	-0.0019	-0.0002	100.0	518.67	642.37	1582.85	1406.22	14.62	...	0.432574	0.025884	2.651326	0.025953	0.0	1.095445	0.0	0.0	0.07
5	1	6	-0.0043	-0.0001	100.0	518.67	642.10	1584.47	1398.37	14.62	...	0.425417	0.023452	0.958697	0.025727	0.0	1.140175	0.0	0.0	0.05
6	1	7	0.0010	0.0001	100.0	518.67	642.48	1592.32	1397.77	14.62	...	0.425652	0.021679	0.643141	0.023476	0.0	1.140175	0.0	0.0	0.08
7	1	8	-0.0034	0.0003	100.0	518.67	642.56	1582.96	1400.97	14.62	...	0.429919	0.021679	1.149274	0.022477	0.0	0.836660	0.0	0.0	0.08
8	1	9	0.0008	0.0001	100.0	518.67	642.12	1590.98	1394.80	14.62	...	0.341101	0.008944	3.205438	0.020740	0.0	0.836660	0.0	0.0	0.07
9	1	10	-0.0033	0.0001	100.0	518.67	641.71	1591.24	1400.46	14.62	...	0.358260	0.014142	2.883881	0.020493	0.0	0.836660	0.0	0.0	0.06

10 rows × 68 columns

Fig. 3 Output of rolling mean and standard deviation

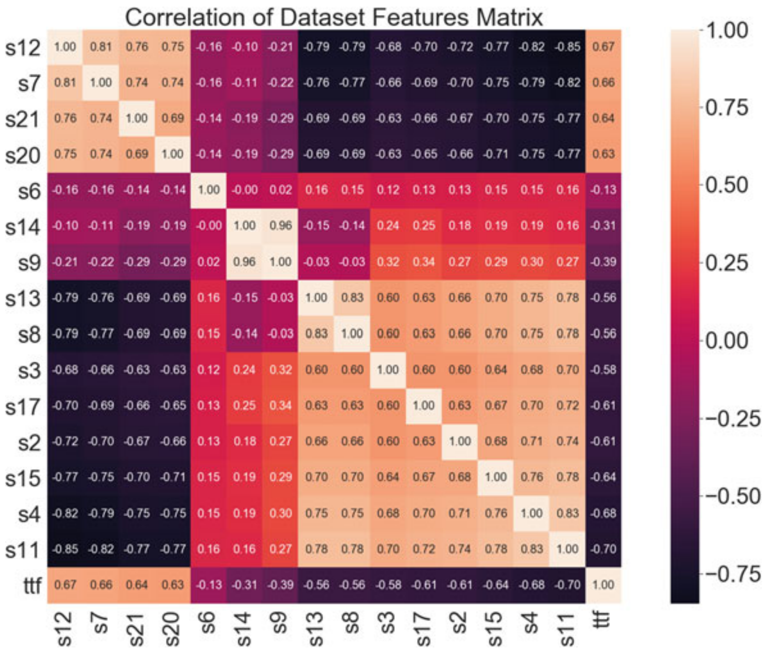


Fig. 4 Correlation among dataset features and regression label.

Exploratory Data Analysis

To identify underlying structure and retrieve significant factors, researchers looked at feature variability, distribution and correlation. Characteristics having high variability were compared to other features and regression label (TTF).

As seen from Fig. 4 following pairs of independent variables have a very high correlation i.e. greater than 0.8 as seen in the feature's correlation heatmap: (s14, s9), (s11, s4), (s11, s7), (s11, s12), (s4, s12), (s8, s13) (s7, s12) Some machine learning algorithms may suffer as a result of this multicollinearity[21]. As a result, during the modelling phase, these characteristics will be targeted for removal during feature selection.

The dependent variable (TTF) has a positive correlation with features, s20, s21, s7 and s12 in the range of 0.63 to 0.67 and features, s17, s2, s15, s4, s11 exhibit a negative correlation in the range of 0.61 to 0.70 with the dependent variable (TTF).

The distribution and correlation of characteristics were also checked using a scatter matrix. It was concluded that the majority of the variables have a normal distribution, which helps machine learning algorithms. Because the majority of features have a nonlinear relationship with the regression label TTF, polynomial models may provide superior results.

Machine Learning Algorithms Evaluated

The following machine learning algorithms were used for predicting the RUL.

1. Random Forests (RF): RF is an ensemble erudition approach that works by training a set of decision trees and then predicting based on mean of the individual trees, which combines Bagging and random feature selection [18].
2. Linear Regression (LR): LR is the most basic kind of predictive analytic method. It's a representation of the link between one or even more independent variables and the dependent variable [1].
3. Logistic Regression (LOR): LOR makes it possible to build a multivariate regression model. A discontinuous variable is used as the dependent variable in LR. The correlation between a dependent variable and numerous independent variables is modelled by LOR, although the independent variables do not require to have normal distributions [13].
4. Ridge Regression (RR): RR approach reduces the sensitivity of outcomes to training dataset by adding a bias. This can help to reduce variance and enhance the accuracy of making predictions [20].
5. LASSO Regression: The "least absolute shrinkage, and selection operator" (LASSO) is an LR method that additionally regularises functionality. The total of regression coefficients' absolute values is taken into account [27].
6. Polynomial Regression (PL): PL may be used to fit a polynomial line with the lowest possible error or cost function. PL is a kind of LR in which the connection between the independent and dependent variables is represented by an *n*th order polynomial in statistics [15].
7. Decision Tree (DT): DT is a classification approach that builds a prediction model using a tree structure. Each node serves as an attribute, the branch representing the attribute's value and the leaves representing the class [32].
8. Support Vector Machines (SVM): SVM creates an "n"-dimensional space out of the numeric input variables in the data that are in various columns. A hyper-plane is chosen in SVM to optimally partition points in the input variable space by class. Implemented using a kernel [10].
9. K-Nearest Neighbours (KNN): KNN works by looking for a group of K items in the training data that are similar and classifies a new case based on the similarity to the stored classes [28].
10. Gaussian Naive Bayes This technique estimates the distribution of each class label using just the training instances of the associated class. It is assumed that the distribution is Gaussian [16].
11. Multilayer Perceptron Neural Networks (MLP): MLP is a feedforward neural network having multiple layers between input and output layers. This network is trained using the backpropagation learning algorithm (BP) [26].

Predictive Results and Model Comparison

Experimental Setup

The models were built and tested using the Anaconda Jupiter notebook Python programming environment, using python3 programming language on Windows 10 operating system with intel i5 eight generation processor with 8 GB of RAM. The models were developed using Scikit-learn (Sklearn), the most powerful machine learning toolkit [29].

The grid search analysis was used to obtain the efficient hyper-parameters for a model. Finding hyper-parameters in training data, unlike parameters, is impossible. As an outcome, in order to find the optimum hyper-parameters, For each set of hyper-parameters, we created a model.

As a result, grid search is regarded as a fairly conventional hyper-parameter optimisation approach, as we are essentially "brute-forcing" all conceivable combinations. The models are then evaluated using cross-validation. Naturally, the model with the highest accuracy is regarded as the best.

The ML algorithms were analysed on the data sets produced during the pre-processing steps. Four data sets were used to train and analyse the ML models, each of which had data gathered from 100–250 engines. In Table 2, you'll find the data set's schema. The graphs depicting the relationship between predicted RUL and actual RUL for each dataset are included in the findings, as detailed in Sect. 22.3.

Binary Classification

During the pre-processing step, binary classification labels were produced. The classification label statistics were created to evaluate if the dataset was balanced. As shown in the output in Fig. 5, there were 17,531 negative samples and 3100 positive samples, resulting in an 85% positive sample rate and 15% negative sample rate.

Because this dataset is clearly imbalanced, we can't use classification accuracy as a model performance indicator. Instead, we may utilise AUC ROC [4].

Machine learning algorithms tried included Random Forests (RF), Logistic Regression (LOR), Decision Trees (DT), Support Vector Machines(SVM), K-Nearest Neighbours (KNN) and Gaussian Naïve Bayes (NB).

Fig. 5 Binary classification dataset statistics

```
0     17531
1      3100
Name: label_bnc, dtype: int64
Negaitve samples =  85%
Posiitve samples =  15%
```

Table 3 Performance of machine learning models using binary classification

	Random Forest (RF) **B1**	Random Forest (RF) B2	Logistic Regression (LOR) B1	Logistic Regression (LOR) B2	Decision Tree (DT) B1	Decision Tree (DT) B2	SVC B1	SVC B2	KNN B1	KNN B2	NB Gaussian B1	NB Gaussian B2
ROC AUC	0.980	0.982	0.980	0.981	0.945	0.962	0.891	0.930	0.935	0.963	0.9877	0.980
Recall	0.680	0.680	0.560	0.680	0.560	0.720	0.680	0.720	0.680	0.720	0.9600	0.960
Accuracy	0.910	0.910	0.880	0.920	0.880	0.920	0.910	0.920	0.910	0.920	0.9400	0.940
Precision	0.944	0.944	0.933	1.000	0.933	0.947	0.944	0.947	0.944	0.947	0.8275	0.827
F1 Score	0.790	0.790	0.700	0.809	0.700	0.818	0.7906	0.818	0.790	0.818	0.888	0.888

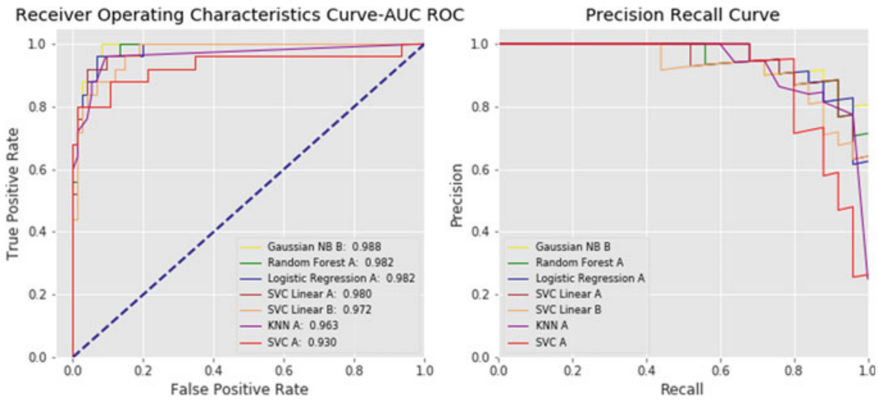


Fig. 6 AUC ROC and precision–recall curves for the ML models

Area under "Receiver Operating Characteristics Curve" (AUC ROC), Recall, Precision, F1 Score and Accuracy are some of the key binary classification performance measures calculated. Grid Search hyper-parameter tuning was done using the AUC ROC score. Several models were evaluated using the test dataset, as shown in Table 3. Where B1 refers for the original features set (before features extraction) and B2 represents for the modified feature set (After feature extraction).

AUC ROC was better for Naïve Bayes and Random Forests. It was also discovered that feature extraction enhanced the performance of most models. Figure 6 shows the Precision–Recall and AUC for ROC curves for the best models.

Figure 7, depicts the number of engines that must be maintained every period, i.e. maintenance capacity. Maximum business gain may be obtained depending on business capacity, recall and precision objectives with the option to threshold at multiple levels. For example, the RF model used a threshold of 0.6 to give 18 engines a recall of 68% and a precision of 94.4%. If maintenance capacity is not a problem, we may set the model's threshold to 0.4, resulting in 72% recall and 94.7% precision for 19 engines, obtaining superior outcomes.

Multiclass Classification

Similarly, Multiclass classification labels were created during the pre-processing stage. The statistics on the classification labels were generated to see if the dataset is balanced.

The statistics result, as shown in the output Fig. 8, confirms that, 17,531 class 0 samples, 1500 class 1 samples and 1600 2 samples, i.e. class 0 has, i.e. 85% samples and class 1 had, i.e. 7% samples and class 2 had, i.e. 8% samples. Similarly, it is observed here also its an unbalanced dataset; therefore, accuracy of the classification system cannot be used as a model performance metric. We can use AUC ROC instead.

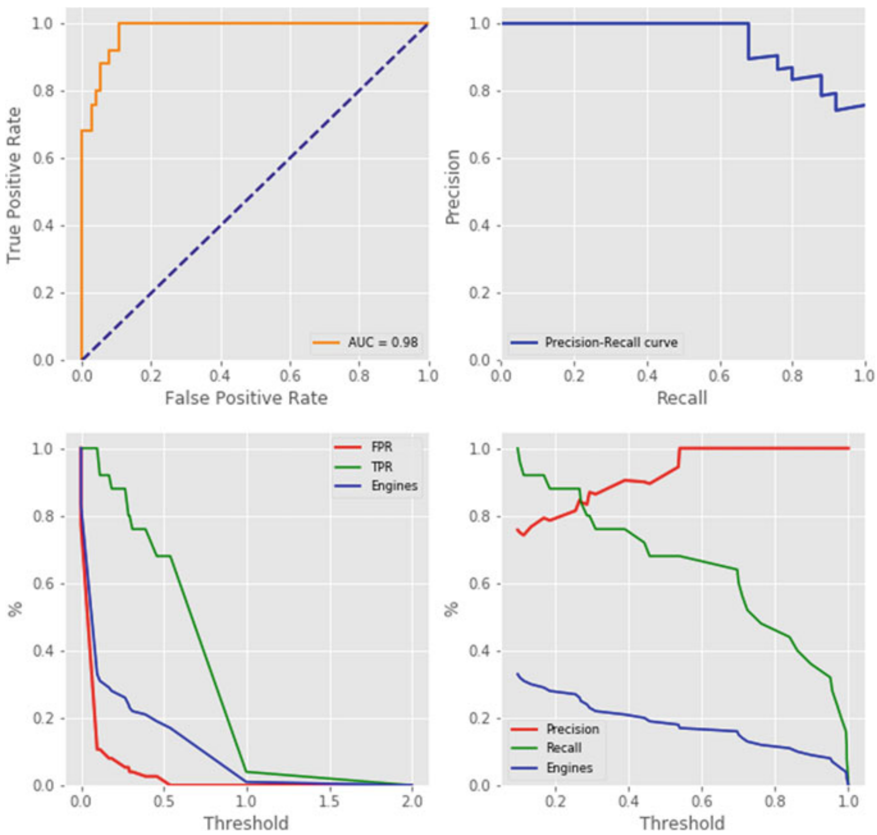


Fig. 7 Evaluation metrices for Binary Classification Random Forest (RF)

Fig. 8 Multi-class classification dataset statistics

```
0     17531
2      1600
1      1500
Name: label_mcc, dtype: int64
Class 0 samples =  85%
Class 1 samples =  7%
Class 2 samples =  8%
```

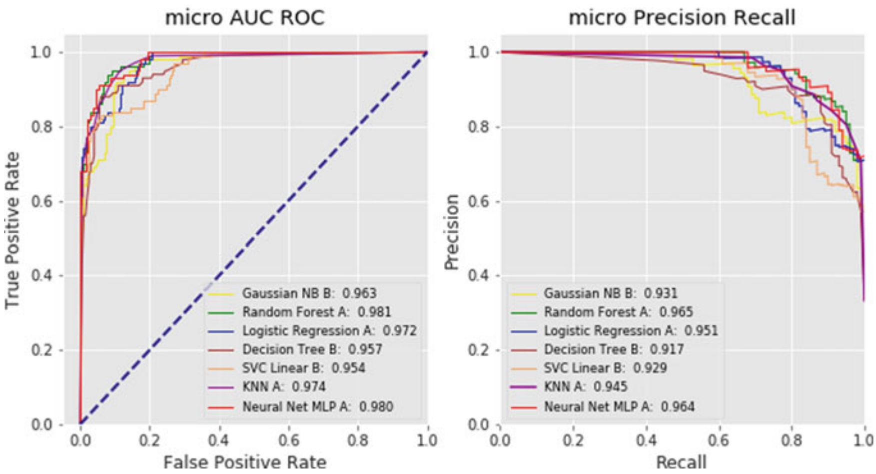


Fig. 9 AUC ROC and precision–recall curves for the ML models

In addition to Accuracy, micro and macro averages of AUC ROC, Recall, Precision and F1 were generated for model assessment. The test dataset's values for these measures are shown in Table 4, where B1 refers for the original features set (before features extraction) and B2 represents for the modified feature set (After feature extraction).

Neural Net MLP and Random Forests scored best AUC ROC in Multi-Class classification. It also noticed that feature extraction has improved most model's performance metrics. Precision-Recall curves and AUC for ROC were plotted for best models as shown in Fig. 9.

Detailed performance results of the Neural Net MLP classifier are shown in Fig. 10.

Root Mean Squared Error (RMSE), R-squared (R^2), Mean Absolute Error and among the key regression analysis metrics evaluated. The test dataset results result is shown in Table 5.

Nonlinear regression models like Polynomial and Random Forest performed better than linear regression models like LR, LASSO and Ridge regression, according to our data exploratory study. Random Forest surpassed other models with an RMSE of ±28.63 cycles, indicating that the model predicts RUL within a ±28.63 cycle average error range.

As indicated in Fig. 11, key features were also investigated using feature selection approaches such as "Recursive Feature Elimination" (RFE) and parameters supplied by models such as Random Forest feature prominence. Figure 12 shows the visualisation of RUL prediction covering both predicted and actual values using ML Regression Model. It shows the traces of actual RUL compared with the predicted RUL using regression techniques.

Table 4 Performance of machine learning models using multi-class classification

	Random Forest (RF) B1	Random Forest (RF) B2	Logistic Regression (LOR) B1	Logistic Regression (LOR) B2	Decision Tree (DT) B1	Decision Tree (DT) B2	SVC B1	SVC B2	KNN B1	KNN B2	NB Gaussian B1	NB Gaussian B2	MLP B1	MLP B2
ROC AUC Micro	0.978	0.981	0.971	0.972	0.957	0.973	0.954	0.917	0.954	0.974	0.963	0.942	0.984	0.981
ROC AUC Macro	0.964	0.967	0.945	0.942	0.906	0.949	0.935	0.943	0.905	0.949	0.951	0.945	0.972	0.972
Recall Micro	0.823	0.850	0.810	0.810	0.840	0.841	0.682	0.920	0.839	0.860	0.956	0.931	0.851	0.892
Recall Macro	0.575	0.662	0.563	0.533	0.668	0.652	0.418	0.733	0.596	0.684	0.977	0.933	0.668	0.762
Precision Micro	0.893	0.885	0.884	0.901	0.880	0.875	0.972	0.478	0.883	0.887	0.772	0.781	0.905	0.908
Precision Macro	0.773	0.802	0.557	0.587	0.819	0.852	0.594	0.661	0.801	0.793	0.655	0.664	0.873	0.890
F1 Micro	0.854	0.867	0.845	0.853	0.862	0.857	0.801	0.631	0.855	0.873	0.852	0.849	0.876	0.899
F1 Macro	0.612	0.708	0.557	0.552	0.685	0.607	0.482	0.501	0.641	0.709	0.757	0.756	0.732	0.798
Accuracy	0.820	0.851	0.812	0.813	0.841	0.840	0.681	0.021	0.831	0.869	0.741	0.740	0.850	0.881

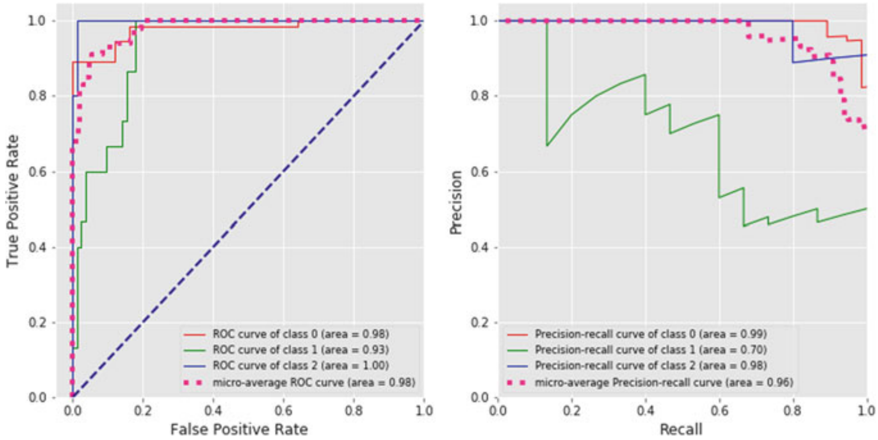


Fig. 10 Performance results of the Neural Net MLP classifier

Table 5 Performance of Machine Learning Models using Regression techniques

	Root mean squared error	Mean absolute error	R^2
Random forest regression	28.634	23.167	0.525
Linear regression	32.041	25.591	0.405
Ridge regression	31.965	25.544	0.408
Decision Tree regression (DT)	32.095	24.319	0.403
Polynomial regression	31.261	23.728	0.434

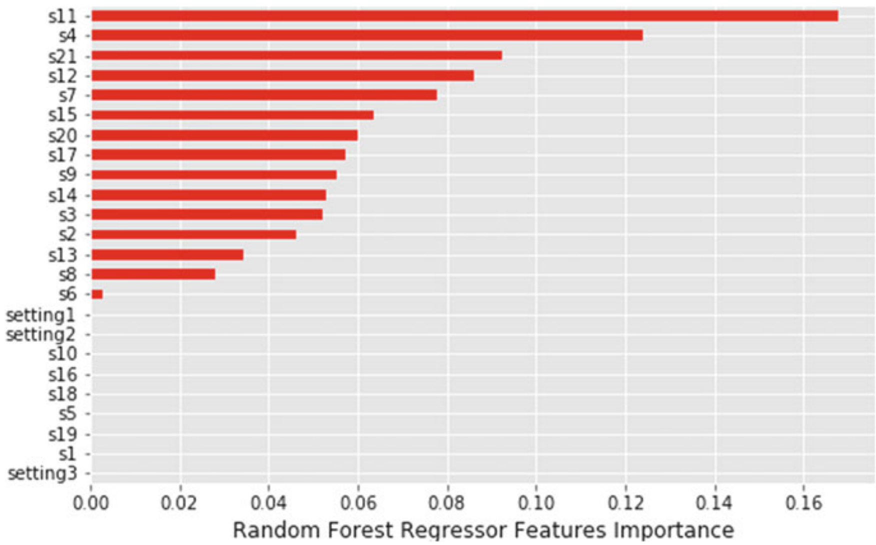


Fig. 11 Importance of features using Random Forest Regressor

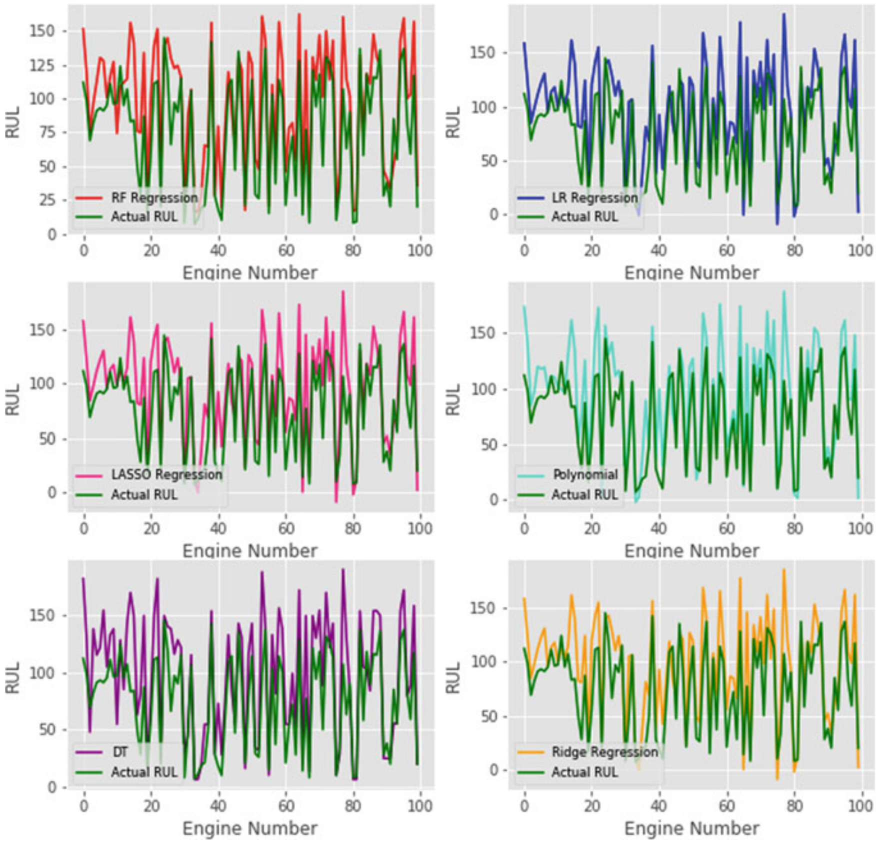


Fig. 12 Visualisation of machine ID versus using Random ML Regression Model

Conclusion and Future Scope

For a consistent comparison of outcomes, the different algorithms were tested on the same dataset. The fundamental goal of RUL prediction is to decrease the difference between the actual RUL and the predicted RUL.

In Binary classification and Multiclass classification, since the dataset was unbalanced, the models were evaluated using AUC ROC.

In Binary classification, AUC ROC was better for Naive Bayes and Random Forests. It was also discovered that feature extraction enhanced the performance of most models.

In Multiclass classification, Neural Net MLP and Random Forests scored best AUC ROC in. It can be concluded that feature extraction has improved the model's performance metrics.

Nonlinear regression models like Polynomial and RF outperformed linear regression models like LR, LASSO and Ridge regression. With an RMSE of ±28.63 cycles, RF outperformed other models, suggesting that the model accurately predicts RUL within a ±28.63 cycle average error range.

The best results were produced using the RF method, according to the RMSE findings, the RF algorithm captures the variance of several input variables at once and allows a large number of observations to participate in the prediction, which is further enhanced by using only the key features and investigated using feature selection approaches such as "Recursive Feature Elimination" (RFE) and parameters supplied by models like Random Forest feature prominence.

Predictive maintenance's major goal is to predict equipment failure. The RUL forecast was carried out in order to plan the turbofan engine's maintenance requirements. We analysed the performance of 11 ML algorithms in our research. In future, the algorithms can be tested by automating the hyper-parameter tuning for the Random Forest Regressor, LASSO and Ridge models which was done manually in this paper.

Acknowledgements I would like to take this opportunity to express my profound gratitude and deep regard to my Prof. Teslin Jacob, for his guidance and valuable feedback and constant encouragement.

References

1. Acharya MS, Armaan A, Antony AS (2019) A comparison of regression models for prediction of graduate admissions. In: 2019 International conference on computational intelligence in data science (ICCIDS), pp 1–5 (2019). https://doi.org/10.1109/ICCIDS.2019.8862140
2. Ahmadzadeh F, Lundberg J (2014) Remaining useful life estimation. Int J Syst Assur Eng Manage 5(4):461–474
3. Bhalaji N, Rimi C (2020) Remaining useful life (rul) estimation of lead acid battery using Bayesian approach. J Electr Eng Autom 2:25–34
4. Branco P, Torgo L, Ribeiro R (2015) A survey of predictive modelling under imbalanced distributions. arXiv preprint arXiv:1505.01658
5. Chadha GS, Nazmus Sakeib SM, Schwung A (2020) Remaining useful lifetime estimation with sobolev training. In: 2020 25th IEEE international conference on emerging technologies and factory automation (ETFA), vol 1, pp 1447–1450. https://doi.org/10.1109/ETFA46521.2020.9212061
6. Chen KC, Gao ZJ (2020) Integrated group-based valuable sensor selection approach for remaining machinery life estimation in the future industry 4.0 era. In: 2020 International symposium on VLSI design, automation and test (VLSI-DAT), pp 1–4. https://doi.org/10.1109/VLSI-DAT49148.2020.9196260
7. Chen Z, Yang C, Peng T, Dan H, Li C, Gui W (2019) A cumulative canonical correlation analysis-based sensor precision degradation detection method. IEEE Trans Ind Electron 66(8):6321–6330. https://doi.org/10.1109/TIE.2018.2873100
8. Ellefsen AL, Bjørlykhaug E, Æsøy V, Ushakov S, Zhang H (2019) Remaining useful life predictions for turbofan engine degradation using semi-supervised deep architecture. Reliab Eng Syst Saf 183:240–251
9. Fauzi MFAM, Aziz IA, Amiruddin A (2019) The prediction of remaining useful life (rul) in oil and gas industry using artificial neural network (ann) algorithm. In: 2019 IEEE conference on big data and analytics (ICBDA), pp 7–11. https://doi.org/10.1109/ICBDA47563.2019.8987015

10. Ghosh AK, Afroge S (2017) A comparison between support vector machine (svm) and bootstrap aggregating technique for recognizing bangla handwritten characters. In: 2017 20th International conference of computer and information technology (ICCIT), pp 1–5. https://doi.org/10.1109/ICCITECHN.2017.8281810
11. Guo B, Wang X, Wang Y, Su H, Chao S (2019) Application of support vector regression to predict the remaining useful life of polymerized styrene butadiene rubber of cable insulation. In: 2019 Prognostics and system health management conference (PHM-Qingdao), pp 1–5. https://doi.org/10.1109/PHM-Qingdao46334.2019.8942888
12. Hällman L (2017) The rolling window method: Precisions of financial forecasting
13. Hu X, Yang Y, Zhu S, Chen L (2020) Research on a hybrid prediction model for purchase behavior based on logistic regression and support vector machine. In: 2020 3rd International conference on artificial intelligence and big data (ICAIBD), pp 200–204. https://doi.org/10.1109/ICAIBD49809.2020.9137484
14. Huang Z, Zhou X, Gao D, Zhang X, Jiang F, Chen B, Yang Y, Wu M, Peng J (2019) A novel label-free supervision learning method for lithium-ion battery RUL prediction. In: 2019 IEEE energy conversion congress and exposition (ECCE), pp 3150–3156. https://doi.org/10.1109/ECCE.2019.8912697
15. Isaac OA, Adedeji AA, Ismail IR (2012) Polynomial regression model of making cost prediction in mixed cost analysis. Int J Math Theor Model 2(2):14–23
16. Jahromi AH, Taheri M (2017) A non-parametric mixture of gaussian Naive Bayes classifiers based on local independent features. In: 2017 Artificial intelligence and signal processing conference (AISP), pp 209–212. https://doi.org/10.1109/AISP.2017.8324083
17. Jia C, Cao Y (2019) Remaining useful life prediction for equipment with periodic maintenance. In: 2019 4th International conference on mechanical, control and computer engineering (ICMCCE), pp 227–2273. https://doi.org/10.1109/ICMCCE48743.2019.00059
18. Kabiraj S, Raihan M, Alvi N, Afrin M, Akter L, Sohagi SA, Podder E (2020) Breast cancer risk prediction using xgboost and random forest algorithm. In: 2020 11th International conference on computing, communication and networking technologies (ICCCNT), pp 1–4. https://doi.org/10.1109/ICCCNT49239.2020.9225451
19. Kang Z, Catal C, Tekinerdogan B (2021) Remaining useful life (RUL) prediction of equipment in production lines using artificial neural networks. Sensors 21(3):932
20. Kumar A, Das PK, Mallick RK, Nayak P (2020) Islanding detection of micro-grid using ridge regression. In: 2020 International conference on computational intelligence for smart power system and sustainable energy (CISPSSE), pp 1–5. https://doi.org/10.1109/CISPSSE49931.2020.9212236
21. Kumar S, Chong I (2018) Correlation analysis to identify the effective data in machine learning: prediction of depressive disorder and emotion states. Int J Environ Res Pub Health 15(12):2907
22. Li T, Pei H, Pang Z, Si X, Zheng J (2020) A sequential Bayesian updated wiener process model for remaining useful life prediction. IEEE Access 8:5471–5480. https://doi.org/10.1109/ACCESS.2019.2962502
23. Li Z, Fang H, Yan Y (2019) An ensemble hybrid model with outlier detection for prediction of lithium-ion battery remaining useful life. In: 2019 Chinese control and decision conference (CCDC), pp 2630–2635. https://doi.org/10.1109/CCDC.2019.8832623
24. Li Z, Goebel K, Wu D (2019) Degradation modeling and remaining useful life prediction of aircraft engines using ensemble learning. J Eng Gas Turb Power 141(4)
25. Liu K, Shang Y, Ouyang Q, Widanage WD (2021) A data-driven approach with uncertainty quantification for predicting future capacities and remaining useful life of lithium-ion battery. IEEE Trans Ind Electron 68(4):3170–3180. https://doi.org/10.1109/TIE.2020.2973876
26. Madalgi JB, Kumar SA (2017) Congestion detection in wireless sensor networks using mlp and classification by regression. In: 2017 3rd International conference on applied and theoretical computing and communication technology (iCATccT), pp 226–231. https://doi.org/10.1109/ICATCCT.2017.8389138
27. Manasa J, Gupta R, Narahari NS (2020) Machine learning based predicting house prices using regression techniques. In: 2020 2nd International conference on innovative mechanisms

for industry applications (ICIMIA), pp 624–630. https://doi.org/10.1109/ICIMIA48430.2020.9074952
28. Okfalisa, Gazalba I, Mustakim, Reza NGI (2017) Comparative analysis of k-nearest neighbor and modified k-nearest neighbor algorithm for data classification. In: 2017 2nd International conferences on information technology, information systems and electrical engineering (ICITISEE), pp 294–298. https://doi.org/10.1109/ICITISEE.2017.8285514
29. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, Vanderplas J, Passos A, Cournapeau D, Brucher M, Perrot M, Duchesnay E (2011) Scikit-learn: machine learning in Python. J Mach Learn Res 12:2825–2830
30. Ragab M, Chen Z, Wu M, Kwoh CK, Li X (2020) Adversarial transfer learning for machine remaining useful life prediction. In: 2020 IEEE International conference on prognostics and health management (ICPHM). IEEE, pp 1–7
31. Ren L, Dong J, Wang X, Meng Z, Zhao L, Deen MJ (2021) A data-driven auto-cnn-lstm prediction model for lithium-ion battery remaining useful life. IEEE Trans Ind Inf 17(5):3478–3487. https://doi.org/10.1109/TII.2020.3008223
32. Rochmawati N, Hidayati HB, Yamasari Y, Yustanti W, Rakhmawati L, Tjahyaningtijas HPA, Anistyasari Y (2020) Covid symptom severity using decision tree. In: 2020 Third international conference on vocational education and electrical engineering (ICVEE), pp 1–5. https://doi.org/10.1109/ICVEE50212.2020.9243246
33. Saranya E, Sivakumar PB (2020) Data-driven prognostics for run-to-failure data employing machine learning models. In: 2020 International conference on inventive computation technologies (ICICT), pp 528–533. https://doi.org/10.1109/ICICT48043.2020.9112411
34. Saxena A, Goebel K, Simon D, Eklund N (2008) Damage propagation modeling for aircraft engine run-to-failure simulation. In: 2008 International conference on prognostics and health management, pp 1–9. https://doi.org/10.1109/PHM.2008.4711414
35. Shifat TA, Jang-Wook H (2020) Remaining useful life estimation of BLDC motor considering voltage degradation and attention-based neural network. IEEE Access 8:168414–168428. https://doi.org/10.1109/ACCESS.2020.3023335
36. Susto GA, Beghi A, De Luca C (2012) A predictive maintenance system for epitaxy processes based on filtering and prediction techniques. IEEE Trans Semiconductor Manuf 25(4):638–649. https://doi.org/10.1109/TSM.2012.2209131
37. Wang T, Liu Z, Zhao X, Liao M, Mrad N (2020) Bayesian-based method for the remaining useful life and reliability prediction of steel structure. In: 2020 Asia-Pacific international symposium on advanced reliability and maintenance modeling (APARM), pp 1–6. https://doi.org/10.1109/APARM49247.2020.9209574
38. Yao Q, Yang T, Liu Z, Zheng Z (2019) Remaining useful life estimation by empirical mode decomposition and ensemble deep convolution neural networks. In: 2019 IEEE international conference on prognostics and health management (ICPHM), pp 1–6. https://doi.org/10.1109/ICPHM.2019.8819373
39. Yu W, Sun Y, Kadam A, Chen Y (2020) Deep learning (dl) based on-board indirect useful life prediction. In: 2020 IEEE International conference on consumer electronics (ICCE). IEEE, pp 1–5
40. Zivot E, Wang J (2003) Rolling analysis of time series. In: Modeling financial time series with S-Plus®. Springer, pp 299–346

Anomaly-Based Intrusion Detection System Through Deep Neural Network

Sabeeha Afzal and Anjna Jayant Deen

Abstract Intrusion detection has a crucial responsibility in securing our network, and also the main aim is to precisely identify each type of attacks. In this paper, efforts have been made to study about the datasets KDD-CUP'99, NSL-KDD properly; further these datasets are used after pre-processing for training machine learning models such as SVM, DT, RF and then for testing and calculating the accuracy and other metrics of machine learning models. Apart from these, focus was to build a Deep Learning approach for Intrusion Detection using Deep Neural Network on the CIC-IDS 2017 Benchmark dataset. In this model, functions made use of are Adam optimizers, Rectified Linear Unit (ReLU), sigmoid classification function for binary classification and softmax classification for multi-class classification. Overall, the intention of making this paper was to do an exhaustive study so in future we can use the best of the experimental results to enhance and detect real-time attacks. Then it has come to my notice that handling lots of datasets is easier for deep learning models by giving a little knowledge about the important features and that there is improvement in the detection rate.

Keywords Network security · Machine learning · Support Vector Machine (SVM) · Decision Tree (DT) · Random Forest (RF) · Normalization · Standardization · Deep Learning · Deep Neural Network (DNN) · Softmax · Sigmoid · ReLU · One hot encoder · Confusion matrix

Introduction

An Intrusion can be contemplated as any kind of behavior that deviates from the normal working of any system [1]. Intrusion actions are indeed set to trade-off the CIA triad (Confidentiality, Integrity and Availability), which is a model created to mentor policies, for information security within an organization [2]. Intrusion Detection is basically a system of security dedicated to consistent monitoring of the environments

S. Afzal (✉) · A. J. Deen
Department of Computer Science and Engineering, University Institute of Technology-Rajiv Gandhi Proudyogiki Vishwavidyalaya, Bhopal, Madhya Pradesh 462033, India

V. Bindhu et al. (eds.), *Proceedings of Third International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 844,
https://doi.org/10.1007/978-981-16-8862-1_23

of computing and information technology and connected devices such as networks, hosts, applications, databases, etc. Researchers currently working on IDS [3] lag nowadays because of increasing size of network data, due to the increase in use of internet and rising technology and a variety of attacks being created. So, proposing a perfect IDS is still a major challenge.

In this paper, works have been done to do an exhaustive study on various machine learning and a deep learning-based IDS. The problem Intrusion Detection System [4] is generally a classification problem; and this classification is of two types, binary classification and multi-class classification. In this the foremost one will result in anomalous or attack, but in the later one we categories all the attack in 4 types as: Denial-of-Service (DoS), Probing (Probe), User to Root (U2R), Root to Local (R2L). This result of classification can be better predicted by doing pre-processing on the dataset, to avoid unnecessary features which may lead to creating confusion for the model. There are various pre-processing methods also, which also we will be studying in this paper.

The IDS can be classified as:

1. Misuse-based IDS/Signature-based IDS
2. Anomaly-based IDS [5]
3. Hybrid-based IDS.

Misuse-based is also acknowledged as signature-based IDS because they compare the signature of known attacks with those of the testing data and then detect the data as anomalous or normal traffic data. Hence, they are unable to detect attack of unknown signature, i.e., Zero-day attack cannot be found. But anomaly-based IDS has this benefit of detecting Zero-day attack because, it's trained using Normal/Benign data and if the traffic data deviates from the normal data, it's considered anomalous. Hybrid-based IDS has the benefit of using both misuse-based as well as anomaly-based IDS. Further, one of the key aims in this detection is to obtain lower false alarm rates, so in misuse since the model is trained with known signatures and its detects attack having the same signatures only, so in this there is lower chance of getting false alarms. But in anomalous case, any single deviance from the standard traffic data will be considered as an attack, so there is a lot of chances of getting high false alarm rates. So, our second aim after calculating better accuracy is getting lower false alarm rates (Fig. 1).

So, in this paper efforts have been made to understand the datasets, then apply it in machine learning algorithms and then use the above results and compare it with Deep Learning model, Deep Neural Network (DNN). For making this exhaustive study more reliable and the models more adaptive, we have used three datasets as: KDD-CUP'99, NSL-KDD and CIC-IDS 2017. In Sect. "Related Works", we will be studying the interrelated works, Sect. "Dataset Description" will be the datasets study and the pre-processing methods used, Sect. "Proposed Work" will contain proposed model, the machine learning methods used, and Deep Learning methods, and Sect. "Result Analysis" has the Evaluation & results analysis, and finally we will have the Conclusion.

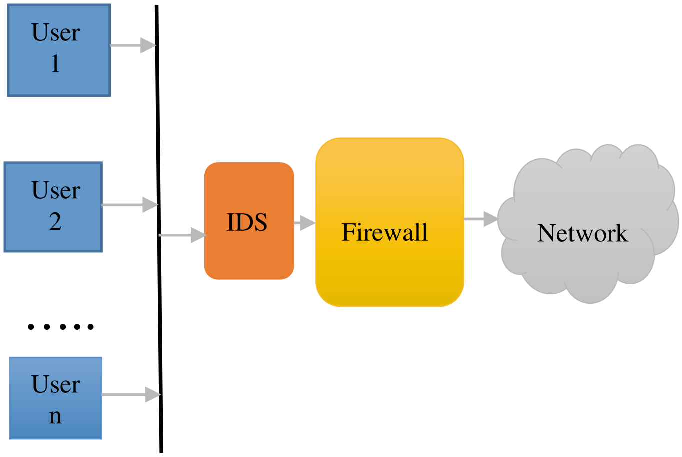


Fig. 1 Block diagram of IDS

Related Works

The chapter/article made by Chih-Fong et al. [6] contains a total of 55 related reviews in the period between 2000 and 2007 focused on researching and developing single, hybrid and ensemble-based classifiers. In this paper, they have mainly discussed about pattern classifications, single classifiers (K-nearest neighbor, SVM, Self-Organizing Maps and Naïve Bayes, decision tree, etc.) and Hybrid & Ensemble Classifiers. Further, they have compared these methods and came across a conclusion that the using only one classifier for comparing the model and evaluation the same may no longer prove to be a useful as the base classifier. Then they have talked about multiple classifiers without comparison of each other, such that all the classifiers can be used as an ensemble, and thereby gaining more accuracy be possible. Chibuzor et al. [7] proposed 'An Intrusion Detection System using Machine Learning Algorithms'. Here they have done multi-class classification by classifying the attacks of test dataset into various attack classes, and this is done on various machine learning algorithms such as (Bayes network, J48, Random Forest and Random Tree). Experimental results have shown that Random Forest and Random Tree seem to be the utmost well-organized in determining accuracy. Mohammad Almseidin et al. [8] did their research on 'Evaluation of Machine Learning Algorithms for Intrusion Detection System'. In this paper, they have evaluated the performance of KDD intrusion dataset on various machine learning algorithms. Their main emphasis was on the metrics false negative (FN) and false positive (FP) in order to increase the accuracy. The experimental results have proved that Decision Tree classifier attained the least value of false negative and the Random Forest classifier have pulled off the maximum average accuracy rate.

Riyazahmed et al. [9] proposed 'Network Intrusion Detection System Using Machine Learning'. In this author has used label encoder to encode the categorical datasets and to select the best features from among the data, they have used recursive-feature elimination (RFE). Then they have performed the said classification work using Decision Tree classifier. They have calculated the metrics on True positive rate (TPR) and false positive rate (FPR). Jayshree et al. [10] proposed 'Intrusion Detection System using Support Vector Machine'. SVM is used mainly because of its clever generalization nature and due to the ability to stun the curse of dimensionality. They have used NSL-KDD datasets which is graded using IGR and then feature subdivision selection is done using K-mean algorithm. They have also analyzed the limitation of SVM. Experimental results have shown that the reduced dataset has increased the performance with reduced set of features by dropping off unnecessary features. By now we have come to know that the handling of dataset is mainly done from human end which in turn could make a lot of mistakes plus handling such a huge size with increasing network traffic is quite difficult using machine learning models, so they switched to Deep Learning methodologies [11].

Sasanka Potluri et al. [12] proposed 'Accelerated Deep Neural Networks for Enhanced Intrusion Detection System' in this paper they are using anomaly-based network intrusion detection, and any abnormality from the usual traffic data is are detected. Further, they have used NSL-KDD dataset for training and then testing, respectively. DNN requires a lot of computation in the training phase. The whole training phase comprises the forward pass and the backward pass. The DNN used here is feed forward, and for increasing the training of multiple hidden layers which is very complex, one layer is trained at a time, so here they used the auto-encoder. Experimental results have shown better accuracy. Hassan et al. [13] proposed 'NNIDS: Neural Network-based Intrusion Detection System'. In this paper they have used the approach which comprises two smaller dimensional DNN systems which resulted in a downfall in computational power. Then they have compared the performance of Dual DNN to that of single DNN and with other machine learning algorithms. Experimental results have proved lower computational power and better accuracy. Sidharth et al. [14] proposed 'Deep Neural Network Architecture for Anomaly-Based Intrusion Detection System'. In this paper they have made effort to classify the network traffic data. At the last fully connected layer class score is computed using the dropout mechanism. NSL-KDD dataset has been used here. Further they have used a K-fold cross-validation to authenticate the model, and the cross-fold K used here is 10. Experimental results have shown better accuracy.

Dataset Description

For exhaustively studying and predicting our model, we are using three datasets in our proposed model namely: KDDCUP'99, NSL-KDD and CIC-IDS 2017. First we will be studying the dataset. Then we will apply pre-processing algorithms, for better training of model.

KDD CUP'99 Dataset [15]

- One of the datasets achieved and arranged by MIT Lincoln Labs is 1998 DARPA Intrusion Detection Assessment Program. The 1999 dataset named KDD CUP'99 dataset uses a type of this dataset. The raw training data has about 4 gigabytes of compacted binary TCP dump data from 7 weeks of network traffic. The attacks in this dataset fall into 4 main groups:

    1. **Denial-of-Service (Dos)**: syn flood
    2. **User to Root (U2R)**: unauthorized access from a remote machine, e.g., guessing password
    3. **Root to Local (R2L)**: unauthorized it can be said as an unauthorized access from a remote machine, e.g., guessing the password.
    4. **Probing**: this could be said as the surveillance and other probing, e.g., port scanning (Fig. 2).

  The features of KDD-CUP'99 can be classified into 3 groups:

1. Basic feature: It consists of all the TCP/IP extracted connection features. From among these, some of the features leads to internal interruption in detection rates.
2. Traffic features: It comprises features which are calculated with respect to window intervals as 'same host' and 'same service' features.
3. Content features: The R2L and U2R features are mostly attached in the data part of the packets, e.g., the number of unsuccessful login efforts is a content feature.

This dataset encompasses of a total of 41 features; some of them are mentioned in Table 1.

The training attack types are 22 in number, and there is a different set of 14 attack types in the test set.

The training attack types are (Table 2).

Almost all our studies have been on KDD CUP'99 dataset; however, there are lots of redundant records and mainly imbalance problem, so we shift toward NSL-KDD Dataset.

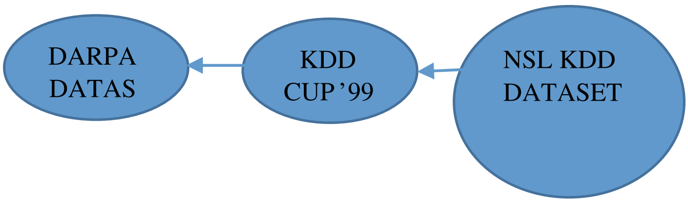


Fig. 2 Dataset

Table 1 Basic feature of individual TCP connections

Feature name	Description	Type
Duration	The span (number of seconds) of the connection	Continuous
protocol_type	Kind of the protocols, e.g. tcp, udp, etc	Discrete
Service	It signifies the network service in the destination, e.g., http, telnet, etc.	Discrete
src_bytes	Number of data bytes from source to destination	Continuous
dst_bytes	Number of data bytes from destination to source	Continuous
Flag	Normal or error position of the connection	Discrete
Land	1 if connection is from/to the same host/port; 0 otherwise	Discrete
wrong_fragment	Number of 'wrong' fragments	Continuous

Table 2 Training attack classes/labels and subcategories/sub-labels

Attack sub-category	Attack class
back	dos
buffer_overflow	u2r
ftp_write	r2l
guess_passwd	r2l
imap	r2l
ipsweep	probe
land	dos
loadmodule	u2r
multihop	r2l
neptune	dos
nmap	probe
perl	u2r
phf	r2l
pod	dos
portsweep	probe
rootkit	u2r
satan	probe
smurf	dos
Spy	r2l
teardrop	dos
warezclient	r2l
warezmaster	r2l

NSL-KDD Dataset [13]

Since KDD CUP'99 had many duplicate records, redundant records and records were not reasonable. So this NSL dataset is a proper and unique subset of the KDD. Or we can that it's a refined form of KDD which covers all the limitations of KDD. In this dataset, we have a total of 43 features out of which 41 are traffic input features and the remaining two are attack label and severity score, respectively. Figure 3 shows the importance of each feature in NSL-KDD.

There are four categories of attack types and one normal/benign label classes in NSL-KDD dataset such as: DoS, R2L, U2R, Probe, Normal/Benign (Table 3; Fig. 4).

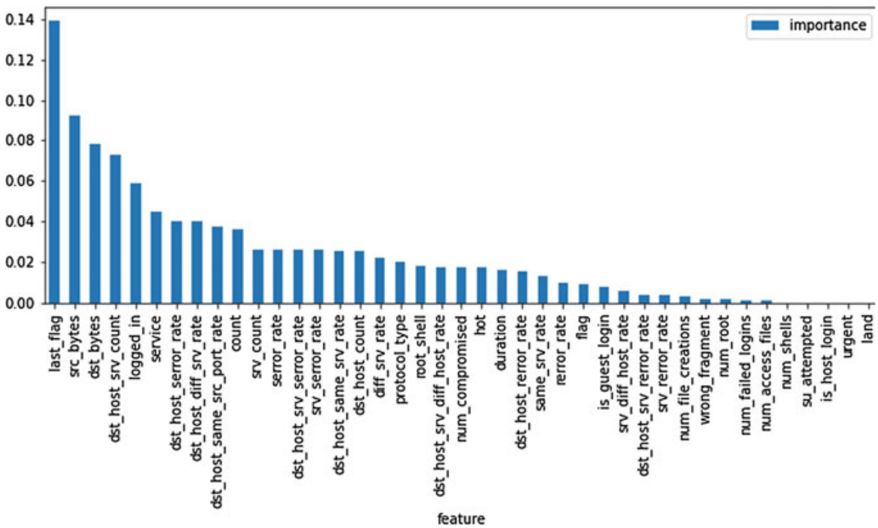


Fig. 3 Feature importance in NSL-KDD dataset

Table 3 Dataset description of NSL-KDD dataset

Attack labels	Sub-class attack labels	No. of records
Normal	Benign	2,359,087
Botnet ARES	Bot	1966
Brute Force	FTP-Patator, SSH-Patator	13,835
DoS/DDoS	DDoS, DoS, GoldenEye, DoS Hulk, DoS Slow-httptest, DoS Slowloris, Heartbleed	294,506
Infiltration	Infiltration	36
Portscan	Portscan	158,930
Web Attack	Web Attack-BruteForce, Web Attack-SQL Injection, Web Attack-XSS	2180

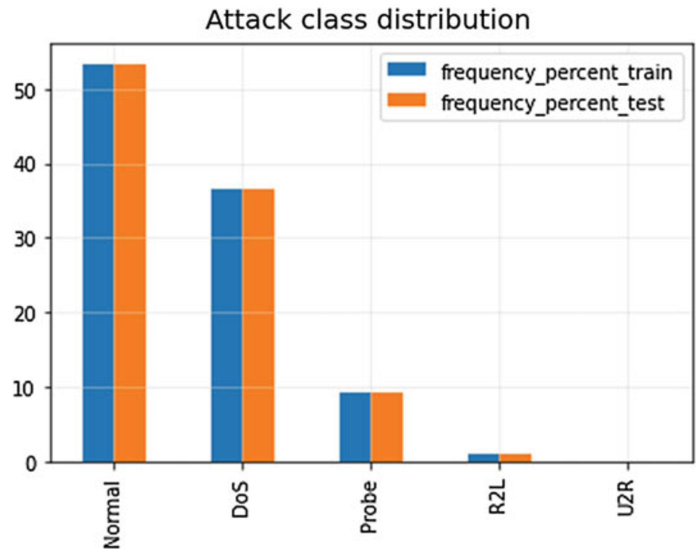


Fig. 4 Attack class distribution

CIC-IDS 2017 Dataset [16]

This dataset was recorded by the Canadian Institute for Cyber security for about one week in 2017. It contains the most common attack as well as the benign attacks. CIC-Flow meter was used to extract the eighty features from the dataset and then for using it, was stored in 10 csv files. This dataset provides a total of 14 attacks classes/types. The whole dataset comprised of a total of 2,830,743 records and 79 features as traffic features and the last one as label features determining whether it's benign/attack (Fig. 5).

Further when we got to know about the benign records that they are 2,273,097 in number, hence we calculated the attacked records and they are 557,646 in number.

We have mapped the attack types/labels (Fig. 6).

Pre-processing: The proposed method is trained and tested on the above-mentioned datasets: KDD CUP'99, NSL-KDD and CIC-IDS 2017 which includes a combination of modern attacks as well as the traditional attacks. All the datasets are pre-processed before splitting them into training and testing sets. For pre-processing

```
print ('BENIGN types: ', len(df.loc[df['Label']=='BENIGN']))
print ('Fraction of BENIGN types:', float(len(df.loc[df['Label']=='BENIGN'])/df.shape[0]))

BENIGN types:  2273097
Fraction of BENIGN types: 0.8030036637024273
```

Fig. 5 Number of benign records in the dataset

```
df_grouped=df.groupby(by='Label')
print(df_grouped.Label.count())
```

```
Label
0     2272688
1        1966
2      128027
3       10293
4      230124
5        5499
6        5796
7        7938
8          11
9          36
10     158930
11       5897
12       1507
13         21
14        652
Name: Label, dtype: int64
```

Fig. 6 Number of attack types/labels and their records

we first have to go through all the datasets we are having. I have studied the dataset thoroughly by noting down the attributes/features and then the data type of the features like numerical attributes or categorical attributes. Further I have noted down all these features separately for future processing. Sometimes these numerical features are having like huge numbers making further difficulty in processing, so these numerical features go separate processing activities which I will be introducing below. The categorical features are difficult for Machine Learning/deep learning to handle, so they too undergo processing. Further many columns have no values or infinity, removing/dropping these rows or putting some suitable values according to some standard methodology is also an important work in processing. And at times when we have a lot of features, understanding which feature makes a huge difference in increasing the detection is also the main goal of pre-processing. So, I will be mentioning the different technique/ methodologies of pre-processing below.

- CIC-IDS 2017 dataset contains some unnecessary features which do not take part in detecting but results in over-fitting, these features are like the source and destination IP addresses and the host port numbers. So, it's essential to remove these features. KDD CUP'99 and NSL-KDD do not contain such features, so this part is skipped here.

- When doing multi-class classification its necessary to remove white spaces in place of multi-class dataset, because they give will be representing unique classes.
- Then we have checked for null values columns, since none of our datasets has null values, so we have skipped this step. Had there been any null values, I would have been deleting the particular row, or may be filling in it mean value by calculating the mean.
- In this step, we will be knowing the categorical features. Like in KDD CUP'99 we have four categorical features: protocol_type, flag, service and Attack type. Here attack type is our Label class. So, in Attack_type feature when I am doing binary classification, I change the attack classes into anomalous and Normal and then map accordingly as Normal being '0' and anomalous being '1'. But in multi-class classification, all the sub-category attacks are labeled in the main 4 kind of attacks (DoS, U2R,R2L,Probe) then these 4 plus one normal is mapped as normal being '0' and DoS: '1', U2R: '2', R2L: '3',Probe: '4', similarly in the protocol_type feature we have mapped this also as 'icmp':0, 'tcp':1, 'udp':2, and the same is done flag feature as well 'SF':0, 'S0':1, 'REJ':2, 'RSTR':3, 'RSTO':4, 'SH':5, 'S1':6, 'S2':7, 'RSTOS0':8, 'S3':9, 'OTH':10, and the service feature is dropped because I found that was not contributing in the detection process. And the same thing is done in the NSL-KDD dataset. We can further use One Hot encoding in case we want dummy encoding. Label encoder method is generally used to encode the labels. In CIC-IDS 2017 dataset here, we checked on if any column is having infinity value, if any column is having it, we basically put maximum value in that particular place. Then for labels we mapped the 14 labels in this dataset.
- Then the actual dataset is not having fixed range, so during the training process it makes quite difficult for the classifier to fit for these variations in the dataset. Normalization is a technique which helps us modify the values in the numerical columns in a fixed range that is between 0 and 1. So this normalization provides the classifier with homogeneous values to the classifier.
- CIC-IDS 2017 dataset has a total of 2867 tuples which contains either missing or infinity values. In place of infinity, we have put maximum value and in place of missing value we have put mean value.

Performance Metrics: Performance metrics are usually used to evaluate different machine learning algorithms; various metrics are as: Detection rate (DR), False Alarm Rates (FAR) and Accuracy (ACC), True Positive (TP), True Negative (TN), False Positive (FP), False Negative (FN).

$$\text{Detection Rate (DR)} = \text{TP}/\text{FN} + \text{TP} \tag{1}$$

$$\text{False Alarm Rate (FAR)} = \text{FP}/\text{FP} + \text{TN} \tag{2}$$

$$\text{Accuracy (ACC)} = (\text{TP} + \text{TN})/(\text{TP} + \text{TN} + \text{FP} + \text{FN}) \tag{3}$$

Table 4 Confusion matrix

	Predicted class	
	Positive class	Negative class
Class		
Positive class	(TP) True Positive	(FN) False Negative
Negative class	(FP) False Positive	(TN) True Negative

True Positive (TP): From the dataset those attack data parts that are accurately categorized as an attack.

False Positive (FP): From the dataset those normal data parts that are imprecisely classified as an attack.

True Negative (TN): From the dataset those normal data parts that are correctly classified as Normal.

False Negative (FN): From the dataset those attack data parts that are erroneously classified as Normal (Table 4).

After doing all the pre-processing, 33 features out of 42 features were used from KDD-CUP'99 dataset in training the models and 42 features out of 43 features from NSL-KDD dataset were used in training the model. For Training Model using CIC-IDS dataset, we have used 62 features out of 79 features.

Proposed Model:

See Fig. 7.

Proposed Work

Machine Learning:

Machine Learning methodologies [17–21] have often proved to be giving excellent detection rates. We are going to study the models (SVM, DT, RF), there performance and everything.

Support Vector Machine:

SVM [22, 23] is an unsupervised learning algorithm of machine learning, it's greatly used for classification purpose; however, we can use it for regression problems as well. Its goal is to build a line or a deciding boundary to distinguish between classes and to put the correct data in the right category. And this deciding boundary or line is called a hyper plane. This algorithm basically chooses specific points at the farthest location from some center points, and those specific points are called support vectors (Fig. 8).

SVMs are generally of two kinds: linear and nonlinear. Linear meaning that we are building linearly separable data that can be distinguishable into two separate classes by a single straight line. And nonlinear SVM could be said that the data which cannot be separated linearly by a single straight line.

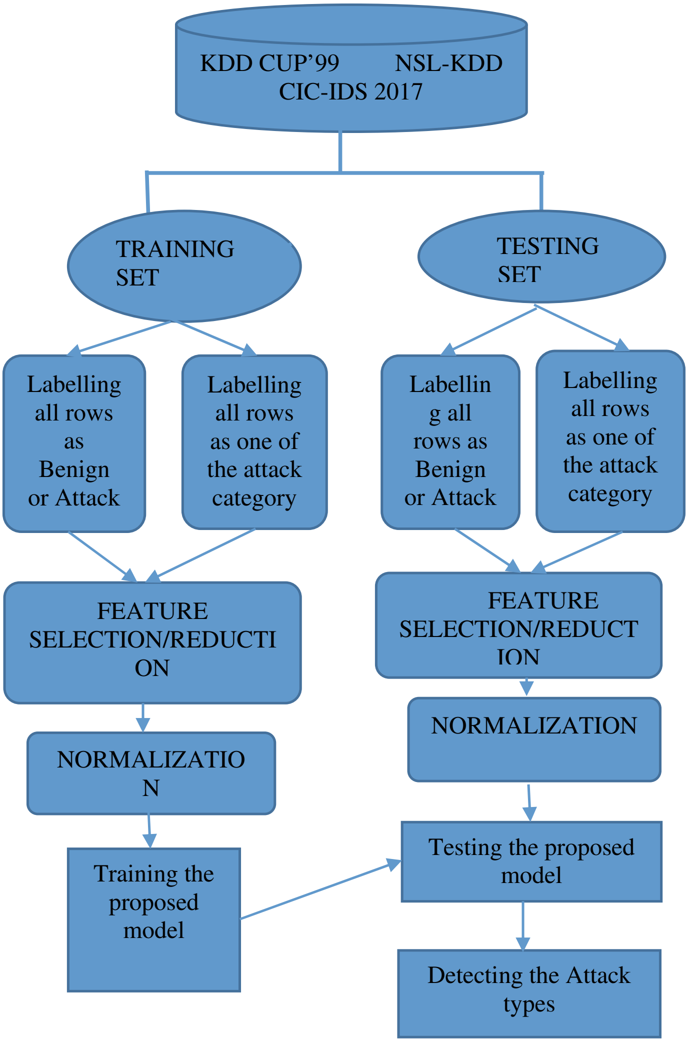
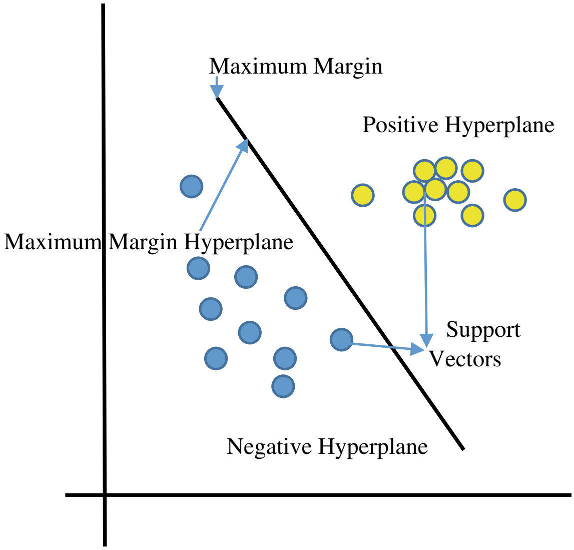


Fig. 7 Proposed methodology for binary and multi-class classification

So, for binary classification I am using linear SVM. So, in the training set, it will import SVC class from sklearn.svm library. And in the SVC classifier the parameter kernel used here is 'linear'. After the training of the model is done, we predict the model using test set. For this, we create a predict vector and call a function of SVM classifier for prediction. We have calculated the training time and testing time and that the prediction score of our model.

Decision Tree [22, 24]: This is basically an administered machine learning algorithm which is used for the purpose of both classification and regression. It's a kind of tree structured classifier, in this the internal nodes denotes the features of the dataset,

Fig. 8 SVM diagram



branches of the tree represent some decision rules and the leaves mainly represent the outcome of decision taken on the basis of the rules given.

In the decision tree, we have used a maximum depth of four, and Entropy is used to measure the quality of split. We have given these two values as parameters in our classifier function call. Then we have trained the model using the training dataset and further calculated the training time and then tested the model using the test dataset (Fig. 9).

Random Forest [4, 25]: This is basically a popular machine learning algorithm, which falls under ensemble learning, meaning thereby that it is a combination of multiple Decision Tree classifiers. And because of it being an ensemble of classifiers, its main purpose is to upsurge the accuracy/performance of the model. As its often said the increase in Decision Tree classifier leads to an increase in the accuracy of the model and further avoids the problem of over-fitting (Fig. 10).

Here in Random Forest classifier parameter, we have used n_estimator = 30, n_estimators generally tell us about the number of trees in the Random Forest, so here we are using 30 trees and Entropy is used to measure the quality of split. We have given these two values as parameters in our classifier function call. Then we have trained the model using the training dataset and further calculated the training time and then tested the model using the test dataset.

Deep Learning [26]:

Deep Neural Network:

Deep neural network or DNN [27] is a basic neural network structure, which is used in our model building. This comprises basically three components: input layer, hidden layer and the last output layer. A DNN consists of at least one hidden layer. We still

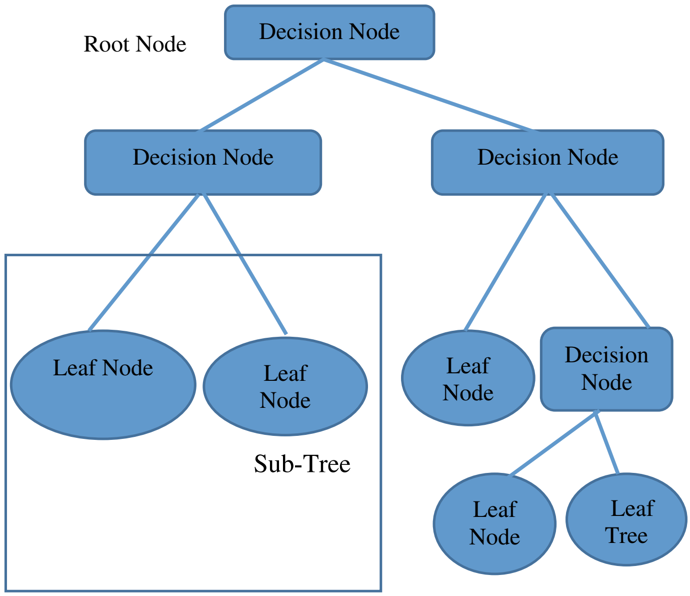


Fig. 9 Decision tree diagram

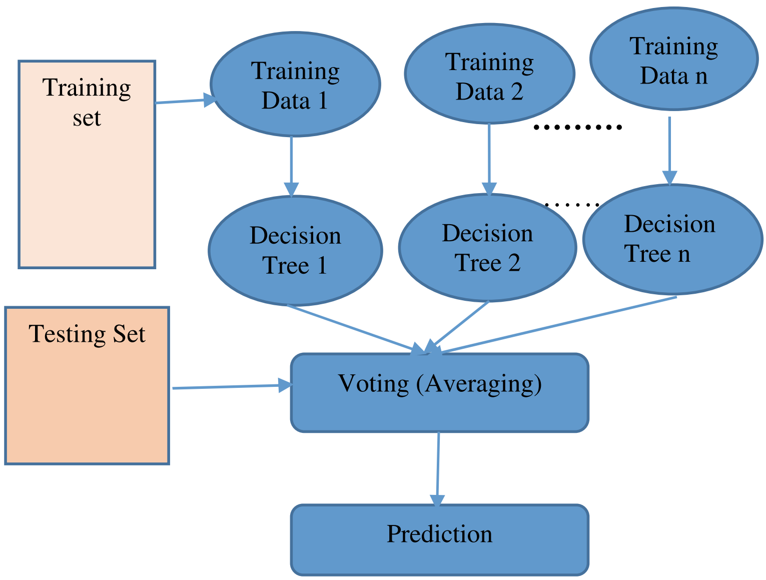


Fig. 10 Random Forest working

Table 5 Hyperparameters used in our models

Hyperparameter	Values
Optimizer	Adam
Number of epochs	150
No. of hidden layers	5
Batch size	300
Activation function	ReLU
Classification function for binary classification	Sigmoid
Classification function for multi-class classification	Softmax

can have more extra layers like in our proposed model we have used 5 hidden layers, which would provide us higher level of abstract features for better performance of the model. So, first we have defined the Keras model and then we have compiled the keras model (Table 5).

Above table shows the hyperparameters used in our model. In this we have used Adam optimizer; it is generally a replacement of stochastic gradient descent which is used for training deep learning models. ReLU activation function is used for the hidden layers as it's nonlinear and also because gradient vanishing problem can be resolved with this. Then for Binary classification, we have used Sigmoid function, and for multi-class classification we have used softmax function.

Adam Optimizer:

The optimization of neural networks is usually connected to searching such parameter that minimizes the cost function of the particular model. Adam is an example of such an optimization algorithm which optimizes the network weights recursively. This is a computationally productive optimizer, and further it requires little memory. Adam optimizers are effortless to carry out and are generally used when the dataset is too huge like our CIC-IDS 2017 dataset.

ReLU Activation Function:

Each artificial neuron or node is associated with an activation function; therefore, the output of every neuron is stipulated by it. Meaning thereby it converts the input of a node to a specific output. There are two types of activation functions, namely linear activation function and nonlinear activation function. The linear one is also said to be a simple activation function. A rectified linear unit is a kind of simple activation function which gives output 0 for input value 0 or less than 0, else its output is the input value given directly. That is it turns out as a linear increasing function.

$$z = \max(0,\ x) \tag{4}$$

Sigmoid Function:

This function is used as a classification function for solving Binary problems. For most cases, they are used at the output layer of the neural network. These are applied to each element of the raw output independently and the return value awaited for each Sigmoid function lies mostly in the range of values between 0 and 1 or −1 and 1.

$$F(x) = \text{sigmoid}(x) = 1/\left(1 + \mathrm{e}^{-x}\right) \quad (5)$$

Softmax Function:

This softmax function is used when we have to deal with a classification function for multi-class problems. In the output layer of neural networks, we sometimes use sigmoid function, which has its own importance. That is in order to get the sum of all outputs as 1; this function transforms the output of all neurons present between 0 and 1. Hence, it outputs the probability distribution of each classes.

$$\text{Softmax }(Zj) = \mathrm{e}^{z}j/\left(\sum_{k=1}^{k} \mathrm{e}^{z}k\right)\text{for } j = 1, \ldots, k \quad (6)$$

Result Analysis

Binary Classification:

The overall performance of the models is evaluated and later on compared for SVM, training time and testing time and the prediction score is calculated for the model for both Binary classification and multi-class classification. The accuracy of SVM is found out to be 99.88% for KDD CUP'99 dataset, 99.88% for NSL-KDD dataset for binary-class classification. For DT model, the accuracy is 99.18% for KDD-CUP'99 dataset, and 99.92% for NSL-KDD dataset in binary-class classification. For RF model, the accuracy is about 99.96% for KDD-CUP'99 dataset, and 99.98% for NSL-KDD dataset in multi-class classification. For Deep Neural Network in binary classification, we have got an accuracy of 89% on CIC-IDS 2017 dataset (Table 6; Figs. 11, 12, 13, 14, 15, 16, 17, 18, 19 and 20).

Table 6 Accuracy table for binary classification

Our model	KDD CUP'99	NSL-KDD
SVM	99.88	99.98
RF	99.96	99.88
DT	99.18	99.92

```
from sklearn.metrics import confusion_matrix
cm= confusion_matrix(y_test, y_test_pred)
```

```
cm
```

```
array([[130188,    672],
       [   652,  31515]], dtype=int64)
```

Fig. 11 Confusion matrix of binary classification from Decision Tree Model using KDD-CUP'99 dataset

```
from sklearn.metrics import confusion_matrix
cm= confusion_matrix(y_test, y_test_pred)
```

```
cm
```

```
array([[130826,     34],
       [    16,  32151]], dtype=int64)
```

Fig. 12 Confusion matrix of binary classification from Random Forest Model using KDD-CUP'99 dataset

```
from sklearn.metrics import confusion_matrix
cm= confusion_matrix(y_test, y_test_pred)
```

```
cm
```

```
array([[130759,    101],
       [    93,  32074]], dtype=int64)
```

Fig. 13 Confusion matrix of binary classification from Support Vector Machine Model using KDD-CUP'99 dataset

Multi-Class Classification: The accuracy of SVM is found out to be 99.879% for KDD CUP'99 dataset, 99.88% for NSL-KDD dataset for multi-class classification. For DT model, the accuracy is 99.05% for KDD-CUP'99 dataset, and 99.99% for NSL-KDD dataset in multi-class classification. For RF model, the accuracy is about 99.66% for KDD-CUP'99 dataset, and 99.99% for NSL-KDD dataset in multi-class classification. Lastly for Deep Neural Network model in multi-class classification,

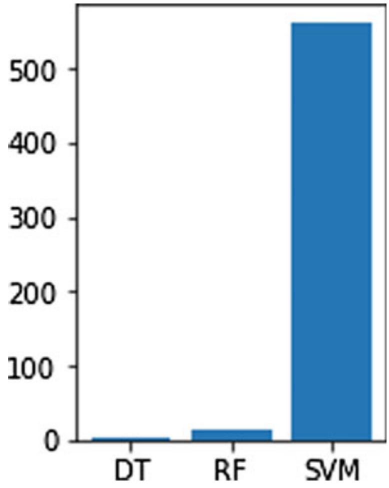


Fig. 14 Bar graph of training time for binary classification models using KDD-CUP'99 dataset

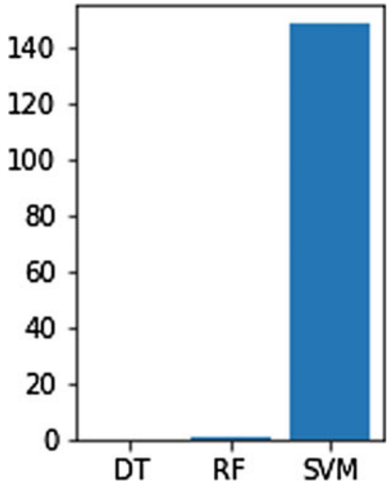


Fig. 15 Bar graph of testing time for binary classification models using KDD-CUP'99 dataset

an accuracy of about 83.33% is obtained on CIC-IDS 2017 dataset (Tables 7 and 8; Figs. 21, 22, 23, 24, 25, 26 and 27).

Conclusion

Summarizing the observations, a lot have learned about the dataset in brief, then we have trained SVM, RF, DT models using the training dataset and predicted the model using the testing dataset, for multi-class classification. Overall, our findings have confirmed that our machine learning models outstand others as mentioned above.

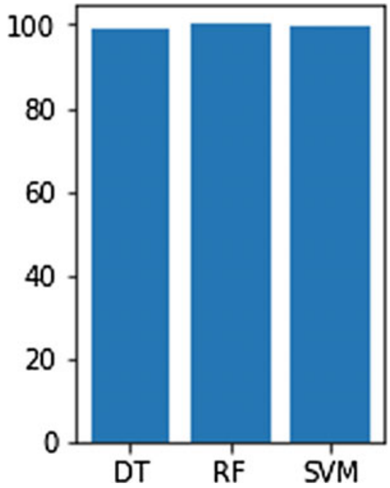


Fig. 16 Bar graph of test score for binary classification models using KDD-CUP'99 dataset

```
Cross Validation Mean Score:
 0.9992936664726596

Model Accuracy:
 1.0

Confusion matrix:
 [[13449     0]
 [    0 13449]]

Classification report:
               precision    recall  f1-score   support

         0.0       1.00      1.00      1.00     13449
         1.0       1.00      1.00      1.00     13449

    accuracy                           1.00     26898
   macro avg       1.00      1.00      1.00     26898
weighted avg       1.00      1.00      1.00     26898
```

Fig. 17 Confusion matrix for binary classification of decision tree model using NSL-KDD

Then efforts have been made to put forward a deep learning model based on a type of deep neural network for multi-class classification. It has come to my knowledge that even though my model has obtained best accuracy using Random Forest algorithm, handling this huge dataset in machine learning is quite difficult. Getting best accuracy using Deep learning is still needed; therefore, our future work would include building

```
Confusion matrix:
 [[ 9234     0]
 [    0 13449]]

Classification report:
              precision    recall  f1-score   support

         0.0       1.00      1.00      1.00      9234
         1.0       1.00      1.00      1.00     13449

    accuracy                           1.00     22683
   macro avg       1.00      1.00      1.00     22683
weighted avg       1.00      1.00      1.00     22683
```

Fig. 18 Confusion matrix for binary classification of Random Forest model using NSL-KDD

Fig. 19 confusion matrix for Binary classification of Support Vector Machine model using NSL-KDD

```
Cross Validation Mean Score:
 0.9988475145194313

Model Accuracy:
 0.9989218529258681

Confusion matrix:
 [[13440     9]
 [   20 13429]]
```

Fig. 20 Bar Graph of accuracy for Binary classification of model using NSL-KDD

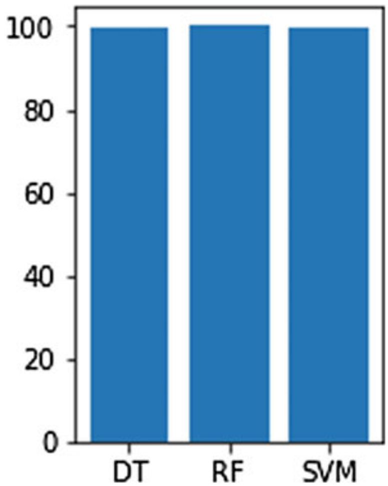


Table 7 Accuracy table for multi-classification

Our model	KDD CUP'99	NSL-KDD
SVM	99.879	99.88
RF	99.66	99.96
DT	99.05	99.98

Table 8 Comparison of our model with other models

Models	Accuracy
SVM [22]	99.66
Proposed SVM	99.88
RF [28]	94.70
Proposed RF	99.92
DT [22]	99.64
Proposed DT	99.98
Proposed binary classification DNN	89
Proposed multi-class classification DNN	83.33

```
In [20]: from sklearn.metrics import confusion_matrix
         cm= confusion_matrix(y_test, y_test_pred)

In [21]: cm

Out[21]: array([[128775,    261,     70,      0,      0],
                [     5,  31510,    652,      0,      0],
                [     8,    143,   1197,      0,      0],
                [     0,    157,    230,      0,      0],
                [     0,     18,      1,      0,      0]], dtype=int64)
```

Fig. 21 Confusion matrix for multi-class classification of decision tree model using KDD-cup'99

```
from sklearn.metrics import confusion_matrix
cm= confusion_matrix(y_test, y_test_pred1)

cm

array([[129105,      1,      0,      0,      0],
       [     1,  32158,      5,      3,      0],
       [     0,     18,   1330,      0,      0],
       [     1,     14,      0,    371,      1],
       [     0,      6,      0,      0,     13]]
```

Fig. 22 Confusion matrix for multi-class classification of Random Forest model using KDD-cup'99

Intrusion Detection System using Ensemble learning approach and getting better real-time detection rate. For detaching the bias and variance, an ensemble [29] method is used which merges all the set of predictions to obtain a final result with enhanced accuracy.

```
from sklearn.metrics import confusion_matrix
cm1= confusion_matrix(y_test, y_test_pred)
```

```
cm1
```

```
array([[129097,      8,      0,      1,      0],
       [    27,  32087,     10,     41,      2],
       [     1,     51,   1296,      0,      0],
       [     1,     47,      1,    337,      1],
       [     0,      3,      0,      4,     12]], dtype=int64)
```

Fig. 23 Confusion matrix for multi-class classification of Support Vector Machine model using KDD-cup'99

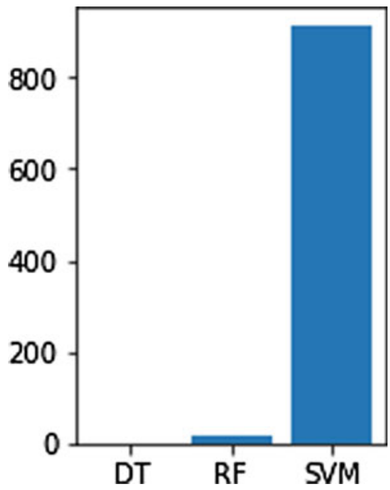


Fig. 24 Bar graph for training time of multi-class classification of models using KDD-cup'99

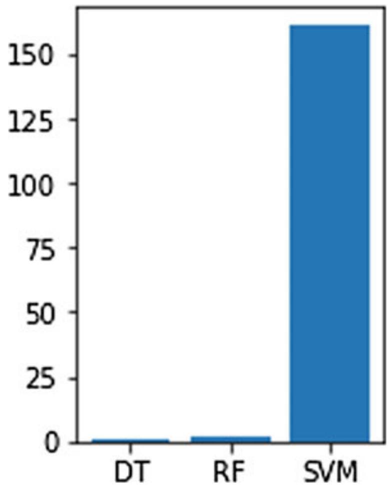


Fig. 25 Bar Graph for testing time of multi-class classification of models using KDD-cup'99

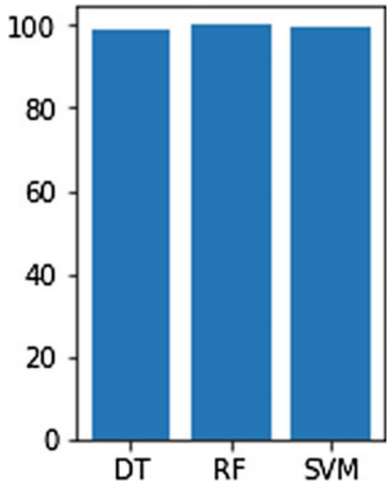


Fig. 26 Bar Graph for accuracy of multi-class classification of models using KDD-cup'99

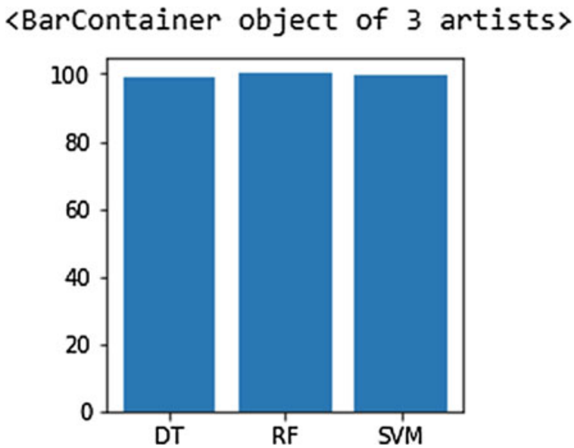


Fig. 27 Multi-class bar graph for NSL-KDD dataset on DT, RF and SVM models

References

1. Abraham A, Jain R, Soft computing models for network intrusion detection systems. Computer (Long. Beach. Calif)
2. "43474201-MIT.pdf"
3. Tang TA, Mhamdi L, McLernon D, Zaidi SAR, Ghogho M (2018) Deep recurrent neural network for intrusion detection in SDN-based networks. In: 2018 4th IEEE conference on network softwarization work. NetSoft 2018, no NetSoft, pp 462–469. https://doi.org/10.1109/NETSOFT.2018.8460090
4. Zhang J, Zulkernine M (2006) Anomaly based network intrusion detection with unsupervised outlier detection. In: IEEE international conference on communications, vol 5, no c, pp 2388–2393. https://doi.org/10.1109/ICC.2006.255127
5. Prasad NR, Almanza-Garcia S, Lu TT (2009) Anomaly detection. Comput Mater Contin 14(1):1–22. https://doi.org/10.1145/1541880.1541882

6. Tsai CF, Hsu YF, Lin CY, Lin WY (2009) Intrusion detection by machine learning: a review. Expert Syst Appl 36(10):11994–12000. https://doi.org/10.1016/j.eswa.2009.05.029
7. Ugochukwu CJ, Bennett EO, Harcourt P (2018) An intrusion detection system using machine learning algorithm, vol 4, no 1, pp 39–47
8. Alsahli MS, Almasri MM, Al-Akhras M, Al-Issa AI, Alawairdhi M (2021) Evaluation of machine learning algorithms for intrusion detection system in WSN. Int J Adv Comput Sci Appl 12(5):617–626. https://doi.org/10.14569/IJACSA.2021.0120574
9. Jamadar RA (2018) Network intrusion detection system using machine learning. Indian J Sci Technol 11(48):1–6. https://doi.org/10.17485/ijst/2018/v11i48/139802
10. Mulay SA, Devale PR, Garje GV (2010) Intrusion detection system using support vector machine and Decision Tree. Int J Comput Appl 3(3):40–43. https://doi.org/10.5120/758-993
11. Atefinia R, Ahmadi M (2021) Network intrusion detection using multi-architectural modular deep neural network. J Supercomput 77(4):3571–3593. https://doi.org/10.1007/s11227-020-03410-y
12. Potluri S, Diedrich C (2016) Accelerated deep neural networks for enhanced intrusion detection system. In: International conference on emerging technologies and factory automation ETFA, vol 2016. https://doi.org/10.1109/ETFA.2016.7733515
13. Al-Maksousy HH, Weigle MC, Wang C (2018) NIDS: neural network based intrusion detection system. In: IEEE international symposium on technologies for homeland security HST 2018, pp 1–6. https://doi.org/10.1109/THS.2018.8574174
14. Behera S, Pradhan A, Dash R (2018) Deep neural network architecture for anomaly based intrusion detection system. In: 2018 5th International conference on signal processing integrated networks, SPIN 2018, pp 270–274. https://doi.org/10.1109/SPIN.2018.8474162
15. Tavallaee M, Bagheri E, Lu W, Ghorbani AA (2009) A detailed analysis of the KDD CUP 99 data set. In: IEEE symposium on computational intelligence in security and defense applications, no Cisda, pp 1–6. https://doi.org/10.1109/CISDA.2009.5356528
16. Sharafaldin I, Habibi Lashkari A, Ghorbani AA (2019) A detailed analysis of the CICIDS2017 data set, vol 977, no Cic. Springer International Publishing
17. Kumar Singh Gautam R, Doegar EA (2018) An ensemble approach for intrusion detection system using machine learning algorithms. In: Proceedings of the 8th international conference confluence 2018 on cloud computing, data science and engineering, pp 61–64. https://doi.org/10.1109/CONFLUENCE.2018.8442693
18. Halimaa A, Sundarakantham K (2019) Machine learning based intrusion. In: 2019 3rd International conference on trends electronics informatics, no Icoei, pp 916–920
19. Xiaofeng Z, Xiaohong H (2018) Research on intrusion detection based on improved combination of K-means and multi-level SVM. In: International conference on communication technology proceedings, ICCT, vol 2017-Octob, pp 2042–2045. https://doi.org/10.1109/ICCT.2017.8359987
20. Waskle S, Parashar L, Singh U (2020) Intrusion detection system using PCA with random forest approach. In: Proceedings of international conference on electronics and sustainable communication systems ICESC 2020, no Icesc, pp 803–808. https://doi.org/10.1109/ICESC48915.2020.9155656
21. Ingre B, Yadav A, Soni AK (2018) Decision tree based intrusion detection system for NSL-KDD dataset. In: Smart innovation system technology, vol 84, no Ictis 2017, pp 207–218. https://doi.org/10.1007/978-3-319-63645-0_23
22. Peddabachigari S, Abraham A, Grosan C, Thomas J (2007) Modeling intrusion detection system using hybrid intelligent systems. J Netw Comput Appl 30(1):114–132. https://doi.org/10.1016/j.jnca.2005.06.003
23. Patel KK, Buddhadev BV (2014) Machine learning based research for network intrusion detection: a state-of-the-art. Int J Inf Netw Secur 3(3):1–20
24. Ben Amor N, Benferhat S, Elouedi Z (2004) Naive Bayes vs decision trees in intrusion detection systems. In: Proceedings of ACM symposium applied computing, vol 1, pp 420–424. https://doi.org/10.1145/967900.967989

25. Resende PAA, Drummond AC (2018) A survey of random forest based methods for intrusion detection systems. ACM Comput Surv 51(3). https://doi.org/10.1145/3178582
26. Moraboena S, Ketepalli G, Ragam P (2020) A deep learning approach to network intrusion detection using deep autoencoder. Rev d'Intelligence Artif 34(4):457–463. https://doi.org/10.18280/ria.340410
27. Gao N, Gao L, Gao Q, Wang H (2015) An intrusion detection model based on deep belief networks. In: Proceedings of 2014 2nd international conference on advanced cloud big data, CBD 2014, pp 247–252. https://doi.org/10.1109/CBD.2014.41
28. Zhang J, Zulkernine M, Haque A (2008) Random-Forests-based network intrusion. MAN Cybern 38(5):649–659
29. Folino G, Sabatino P (2016) Ensemble based collaborative and distributed intrusion detection systems: a survey. J Netw Comput Appl 66:1–16. https://doi.org/10.1016/j.jnca.2016.03.011
30. Ahmad Z, Shahid Khan A, Wai Shiang C, Abdullah J, Ahmad F (2021) Network intrusion detection system: a systematic study of machine learning and deep learning approaches. Trans Emerg Telecommun Technol 32(1):1–29. https://doi.org/10.1002/ett.4150
31. Pastrana S, Mitrokotsa A, Orfila A, Peris-Lopez P (2012) Evaluation of classification algorithms for intrusion detection in MANETs. Knowledge-Based Syst 36:217–225. https://doi.org/10.1016/j.knosys.2012.06.016

A Data Analysis Pipeline to Explore Demographic Information for Identifying COPD Patients

B. Jyothi, V. S. S. Harika Koundinya, Israr Ahmed Khan, and Narapongu Sravan

Abstract Chronic obstructive pulmonary disease (COPD) is a life-threatening lung disease and a major cause of morbidity and mortality worldwide. Monitoring of biomarkers that reflect the disease progression plays a pivotal role in the effective management of COPD. Hence, the accurate examination of respiratory tract fluids like saliva is a promising approach for staging disease and predicting its upcoming exacerbation in a point-of-care (PoC) environment. Due to the lack of anticipation, the rate of deaths in COPD patients has been increasing tremendously; hence, the proposed model envisages the disease. Here, we propose a data analysis pipeline to analyze the saliva-metric and demographic data and use it for identification of COPD patients. The pipeline includes two phase data analyses, i.e., descriptive and predictive analyses. Descriptive analysis focuses on statistically describing and visualizing the dataset. It helps in selecting the model for prediction that is useful for doctors.

Keywords COPD · Lung disease · Secondhand smoker · Machine learning · Prediction · Descriptive analysis · Statistical analysis · Visualization · Breathing problems

Introduction

Chronic obstructive pulmonary disease (COPD) is a pulmonary disorder that causes respiratory problems in patients by restricting airflow to the lungs. It is a disorder that worsens over time. It develops slowly over time, and the symptoms often worsen. The chronic disorder is a leading cause of mortality worldwide, affecting many people and putting a tremendous financial strain on the healthcare system. The main reason for the pulmonary disorder is the prolonged exposure to cigarette smoke (active secondhand smokers) or other respiratory irritants such as toxins, industrial dust, or chemical pollutants. Alpha-1 antitrypsin deficiency is a genetic condition which can also lead to lung damage and subsequently COPD in rare cases. The most common

B. Jyothi (✉) · V. S. S. H. Koundinya · I. A. Khan · N. Sravan
Anurag University, Ghatkesar, Hyderabad, India
e-mail: bjyothicse@anurag.edu.in

V. Bindhu et al. (eds.), *Proceedings of Third International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 844,
https://doi.org/10.1007/978-981-16-8862-1_24

sign of COPD is shortness of breath, persistent cough, wheezing, chest tension, and abnormal mucus development. While lung disease is irreversible, early intervention has been proven to be critical in the successful management of COPD.

Purpose

As there is no proper system that diagnoses lung disease in initial stages, many people are battling with their lives. This creates tension among the family and disturbs people mentally which affects the overall degradation of the lives of the people.

Spirometry tests and the other physical diagnosis systems are only effective when the disease has affected the person up to 65–70% and treating the disease in that scenario is vital and tough. The results of the person being cured are unpredictable, and to minimize this tension and to give an intuition to the person about his/her health being affected through COPD is the prime purpose of the project. To minimalize the deaths occurred due to this disease by ideating the disease in the initial stages is the foremost motivation.

Existing Approaches

COPD is a lung discomfort in which the ventilation to the lungs is limited, making it impossible for patients to breathe. It is a slow-moving condition with signs that sometimes escalate over time. COPD is one of the leading causes of death globally, afflicting many individuals and putting a tremendous financial strain on healthcare systems.

Long-term exposure to tobacco smoke or other respiratory irritants like toxins, chemical vapors, or factory dust is the most common cause of COPD. Shortness of breath, persistent cough, wheezing, chest tension, and abnormal mucus development are the most common signs of COPD. Early detection can help in successful control of COPD (Fig. 1).

Spirometry lung function tests are the most simple and standardized means of diagnosing COPD, out of all the screening and diagnostic tools available. Spirometry is a procedure that tests the patient's lung volume. However, as the sensitivity of the spirometry test is limited to 64.5–79.9% COPD is not diagnosed in several vehicles. These methods have shown results in collecting traffic data (Fig. 2).

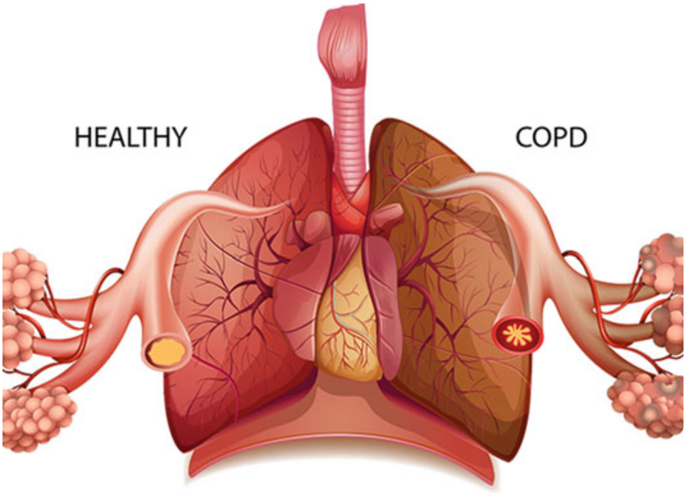


Fig. 1 Healthy versus COPD lungs

Fig. 2 Spirometry test depiction

Challenges in Existing Approaches

Since the spirometer's intensity ranges from 64.5 to 79.9%, it is often misguided. The major drawback is that the lung discomfort disease, COPD, cannot be diagnosed in the initial phases.

The Proposed System and Its Objectives

The system is divided into a sequence of steps in which the first one is to collect the data (gender, age, and saliva metrics) from patients suffering from lung diseases such as HC, asthma, and COPD. Three hundred and ninety-nine records are taken for the analysis purpose. Saliva-metric bio-sensing values play a major role in identifying if the person has COPD or not.

After collecting the data, the system works as follows:

- Data cleaning
- Descriptive analysis
- Predictive analysis.

Data Cleaning

Data cleaning includes missing value imputation. The real-world data is dirty and consists of many missing values which must be handled for further processing.

Descriptive Analysis

Descriptive analysis is used to provide:

Class-based statistical distribution of patterns.

Correlation analysis of saliva sample and demographic features with the class samples.

Exploratory analysis will focus on selection of predictive models.

And finally, **prediction analysis** will primarily focus on predicting if the user has COPD or not using saliva metrics. The workflow of the model is shown in Fig. 3.

Objectives of the Proposed System

The key aim is to use saliva-metric and demographic data to predict COPD.

To detect the disease in early stages to minimize the effect on the subjected patients.

Besides that, the complete technique can be used to know the degree of the disease caused in the patient that can be used to inform the criticality of the situation.

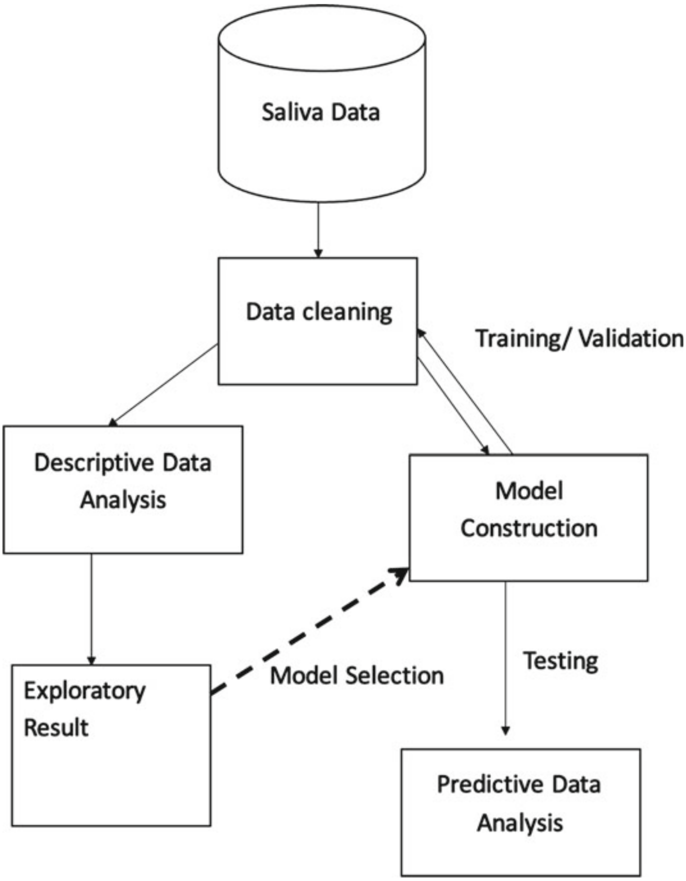


Fig. 3 Workflow of the model

Literature Survey

Diagnosis of COPD Based on Integrated Model and Knowledge Graph

Chronic obstructive pulmonary disease (COPD) is a persistent lung infection that causes a reformist decrease in respiratory capacity. Diagnosing COPD in the early treatable stages is vital also and may even save the existence of a patient. In this paper, we present a coordinated model for diagnosing COPD considering an information chart.

To start with, we build an information diagram of COPD to dissect the relationship between subsets and further find the information on inferred infections from the information. Second, we propose a calculation for arranging highlights and a versatile element subset determination calculation CMFS-η, which chooses an ideal subset of highlights from the first high-dimensional set. At long last, the DSA-SVM incorporated model is recommended to assemble the classifier for the finding and forecast of COPD.

We performed broad tests on the dataset from the emergency clinic outpatient electronic clinical record dataset. The order exactness of our strategy was 91.1%.

Detection of Chronic Obstructive Pulmonary Disease in Computer

Aided Diagnosis System with CNN Classification

Chronic obstructive pulmonary disease (COPD) is a sort of noticeable ongoing illness. The principal factors that cause COPD sickness are breathing in dust, windedness, climate contamination, weariness, and regular respiratory diseases.

COPD is represented by air shaft impediment coming about because of ongoing incendiary reactions in the aviation routes and harmful particles or gases. Air shaft obstacle is a normal part in chronic obstructive pulmonary disease (COPD). Area of impediment and its assessing is fundamental.

PC-supported analysis framework helps the experts for clarification of clinical pictures. The PC-helped analysis framework has been proposed to conclude the COPD by using CT pictures. Processed tomography (CT) pictures are by and large picked because of less mutilation, less time utilization, and negligible exertion.

The proposed work of an electronic-based analysis framework for a chronic obstructive pulmonary disease (COPD) is to analyze the sickness with assurance utilizing convolutional neural network (CNN). CNN classifier is to group the CT pictures, and it will be assessed utilizing execution measurements. The PC-helped finding framework for COPD is made of preprocessing, highlight extraction, division, and arrangement. The preprocessing is to improve the nature of the picture like eliminating commotion and confining locale of interest. The component extraction is a technique of catching visual substance of pictures for ordering and recovery. The division partitions the CT picture into various districts. The CNN classifier is to portray the divided CT pictures for upgrading the assurance of bunching under clamor.

Spirometry Pulmonary Test

This is a physical test which is usually taken in the premises of the hospital. It is a functional test that will measure the speed of the amount of the air inhaled or exhaled by the person.

The test results in the following:

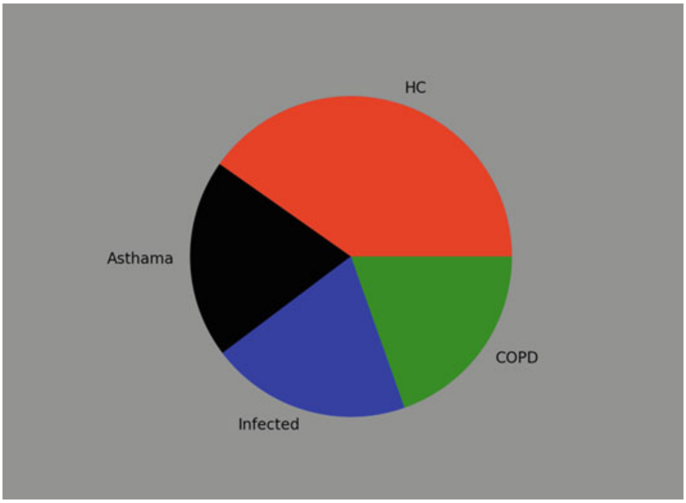


Fig. 4 Dataset distribution

Peak expiratory flow (PEF): The peak air flow during the session.

Maximal voluntary ventilation (MVV): The exhaled volume of air for which the person had to put maximum pressure.

Total lung capacity (TLC): The sum of vital capacity (volume of air breathed out after deepest inhalation) and RV (the volume of air breathed in after deepest exhalation).

Working of the Model

Dataset

Biosensor outputs for dielectric characterization of saliva samples for COPD, HC, and asthma are included in the data collection. Other details such as gender, age, and smoking status (smokers, former smokers, and non-smokers) are also included in the data collection (Fig. 4).

Data Cleaning

Data cleaning includes missing value imputation, deletion of duplicate rows, and removing the outlays from the dataset.

Models	Train	Test	Overall Accuracy
KNN	100%	88%	90%
Decision tree	100%	90%	95%
Naive Bayes	88.8%	90%	89%
SVM	87%	91%	89%

Fig. 5 Model analysis based on the accuracy

Data Imputation

Imputation is the primary task involved in our project. Imputation is done using Bayes' theorem.

Descriptive Analysis: Descriptive analysis is used to provide the following.

- Class-based statistical distribution of patterns.
- Correlation analysis of saliva sample and demographic features with the class samples.
- Exploratory analysis will focus on selection of predictive models.

Three models have been used to test with the dataset and for exploring on the dataset (decision tree, Naive Bayes, and SVM). After performing crystal clear analysis on the dataset, it is known that the decision tree works best for the dataset and to do the prediction analysis (Fig. 5).

Decision Tree Algorithm

The decision tree algorithm is based on taking stepwise decisions for every input given by the user.

- First the user is asked to give the name and age, and then, he/she will be asked if he/she is a smoker or non-smoker or secondhand smoker.
- Then, the person will be asked to take the saliva test and the corresponding imaginary_min, imaginary_average, real_min, and real_average biosensor values from the saliva sample.
- Based on these values, the details are compared with the dataset and the best approximate class is given as the result.

Prediction analysis: Prediction analysis will focus on predicting COPD patients using decision tree after gradient boosting (enhanced accuracy).

Inputs

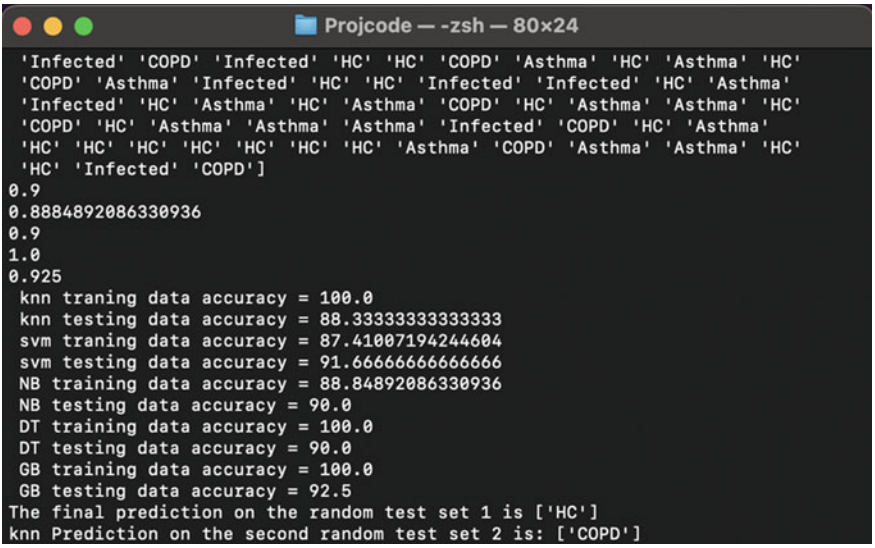
The inputs consist of the details of the patient: gender of the person, age of the person, if the person is a smoker or non-smoker or a secondhand smoker, and the saliva biosensor values of the person.

Output

The model is selected based on the accuracy, and that is used for processing the inputs given by the user for predicting whether the person has COPD or not (Fig. 6).

Testing

Testing is the process of evaluating a system or its components with the intent to find whether it satisfies the specified requirements or not. Here, we have taken real-time test samples (Fig. 7).



```
Projcode — -zsh — 80×24
 'Infected' 'COPD' 'Infected' 'HC' 'HC' 'COPD' 'Asthma' 'HC' 'Asthma' 'HC'
 'COPD' 'Asthma' 'Infected' 'HC' 'HC' 'Infected' 'Infected' 'HC' 'Asthma'
 'Infected' 'HC' 'Asthma' 'HC' 'Asthma' 'COPD' 'HC' 'Asthma' 'Asthma' 'HC'
 'COPD' 'HC' 'Asthma' 'Asthma' 'Asthma' 'Infected' 'COPD' 'HC' 'Asthma'
 'HC' 'HC' 'HC' 'HC' 'HC' 'HC' 'HC' 'Asthma' 'COPD' 'Asthma' 'Asthma' 'HC'
 'HC' 'Infected' 'COPD']
0.9
0.8884892086330936
0.9
1.0
0.925
 knn traning data accuracy = 100.0
 knn testing data accuracy = 88.33333333333333
 svm traning data accuracy = 87.41007194244604
 svm testing data accuracy = 91.66666666666666
 NB training data accuracy = 88.84892086330936
 NB testing data accuracy = 90.0
 DT training data accuracy = 100.0
 DT testing data accuracy = 90.0
 GB training data accuracy = 100.0
 GB testing data accuracy = 92.5
The final prediction on the random test set 1 is ['HC']
knn Prediction on the second random test set 2 is: ['COPD']
```

Fig. 6 Final output on random test sets

Functionality	Initial State	Input	Expected Output	Actual Output	Status
Initialize system	Pause state	"Hi"	Receive Input	Receive Input	Success
Upload Saliva metrics	Pause state	Receiving the saliva metrics	Analysis of the model	Model Analysed	Success
Saliva metrics analysis	Load the appropriate model	Receiving the data and model	Receiving the data and model.	Receiving the data and model.	Success
Prediction	Pause state	The saliva metrics and the model are taken to analyse the patients data.	The prediction is done based on the saliva metrics comparison.	The prediction is done based on the saliva metrics comparison.	Success

Fig. 7 Model stage-wise testing

Conclusion

The COPD prediction system for predicting COPD is successfully designed and implemented. The data was visualized using different plots which are computed to get a clear understanding of the dataset. For disease prediction, a machine learning model was trained and tested successfully with significantly good metrics. The model is employed to make predictions on newer data. The study thus demonstrated the testing accuracies of different classifier models on the dataset where gradient boosting model got highest accuracy of 92.5; thus, we can use this model for predicting COPD.

Future Scope

In the future of the project, we can implement other algorithms which give better performance metrics values. We can make the project available in real time by integrating this model with the biosensor. We can predict COPD instantly, the images of the saliva can be uploaded or taken instantly, and they can be used for the further processing of the prediction.

Bibliography

1. Barnes PJ (2000) Mechanisms in COPD: differences from asthma. J Chest 117(2):10S-14S
2. Mathers CD, Loncar D (2006) Projections of global mortality and burden of disease from 2002 to 2030. PLoS Med 3(11):e442
3. Csikesz NG, Gartman EJ (2014) New developments in the assessment of COPD: early diagnosis is key. Int J Chronic Obstructive Pulm Dis 9:277–286

4. Dong T, Santos S, Yang Z, Yang S, Kirkhus NE (2020) Sputum and salivary protein biomarkers and point-of-care biosensors for the management of COPD. Analyst 145(5):1583–1604
5. Mirza S, Clay RD, Koslow MA, Scanlon PD (2018) COPD guidelines: a review of the 2018 GOLD report. Mayo Clinic Proc 93(10):1488–1502
6. Chiappin S, Antonelli G, Gatti R, Elio F (2007) Saliva specimen: a new laboratory tool for diagnostic and basic investigation. Clinica Chim Acta 383(1):30–40
7. Bayes Theorem in Machine Learning: Bayes Theorem in Machine Learning: introduction, how to apply & example | upGrad blog
8. Gradient Boosting Algorithm: gradient boosting—Wikipedia
9. SVM Algorithm: SVM Algorithm | Working & Pros of Support Vector Machine Algorithm (educba.com)
10. KNN Algorithm: KNN algorithm—finding nearest neighbors—Tutorialspoint
11. Price D, Crockett A, Arne M, Garbe B, Jones R, Kaplan A, Langhammer A, Williams S, Yawn B (2009) Spirometry in primary care case-identification, diagnosis and management of COPD. Prim Care Respir J 18(3):216–223
12. Haroon S, Jordan R, Takwoingi Y, Adab P (2015) Diagnostic accuracy of screening tests for COPD: a systematic review and meta-analysis. BMJ Open 5(10), Art. no. e008133
13. Chen JI, Hengjinda P (2021) Early prediction of coronary ArteryDisease (CAD) by machine learning method—a comparative study. J Artif Intell 3(1):17–33
14. Palani U, Vasanthi D, Rabiya Begam S (2020) Enhancement of medical image fusion using image processing. J Innov Image Process (JIIP) 2(4):165–174

Deep Learning-Based Implicit Continuous Authentication of Smartphone User

Christy James Jose and M. S. Rajasree

Abstract Mobile devices especially smartphones are an unavoidable personal companion of the twenty-first-century human being. The data they are carrying and stored in the cloud through connected applications are private in nature and crucial. These are to be protected from unwanted usage to avoid both mental and financial distresses to its owner. The traditional security methods for authentication are the usage of PIN, fingerprint and face unlock. But the issue with these mechanisms is they are just an entry point authentication and once the user is authenticated it believes that the device is in the hands of its genuine user until it is locked. An intruder who has access to any of the traditional authentication credentials can steal the private data. Similarly, an unlocked device which falls at the hands of an intruder can also act as genuine owner. In addition to the entry point authentication, a continuous monitoring of the user throughout the interaction time is highly essential. Continuous authentication is the solution to have an absolute protection to the valuable data. In this paper, we are proposing an implicit continuous authentication mechanism using face recognition. Deep neural network-based classifiers were used for the model creation and subsequent binary classification of the current user of the device as genuine or intruder. Two CNN models were implemented and tested with unseen data. Model-01 gave an accuracy of 86% with specificity 0.91, whereas Model-02 returned an accuracy of 96% with specificity 1. Thus, Model-02 can be used in real time as it is robust in detecting the intruders correctly. Advantage of the proposed mechanism is that there is no need for any new sensors to be added to the smartphone. Our system would be using the inbuilt front camera to continuously monitor its owner.

Keywords Authentication · Deep learning · Face recognition · Smartphone · Implicit continuous authentication

C. J. Jose (✉)
Government Engineering College, Idukki, Kerala, India
e-mail: christy@gecidukki.ac.in

M. S. Rajasree
APJ Abdul Kalam Technological University, Thiruvananthapuram, Kerala, India

V. Bindhu et al. (eds.), *Proceedings of Third International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 844,
https://doi.org/10.1007/978-981-16-8862-1_25

Introduction

Smartphones have become inevitable in the professional and personal life of men and women [1]. Today, mobile phones are not just a two-way communication equipment. It is a multitasking device embedded with digital camera, navigation equipment, Web browser, personal assistant, multimedia players and vast number of useful applications including mobile banking and e-commerce [2]. The computing power of latest smartphones has been surpassing personal computers. As a result, they carry a huge amount of personal and sensitive information. Photographs, email, private documents and banking credentials are among them. There is a high probability for the device falling into an intruder's hand considering the portable size of them. In the year 2013 alone, 3.1 million was the reported number of stolen smartphone cases in America and that was almost double that of previous year [3]. Traditional user authentication mechanisms built with smartphones include PIN, graphical passwords, fingerprint sensors and face unlock. These mechanisms provide some amount of security as a one-time entry point authentication. Reports say that these mechanisms can be hacked and intruders can gain access to the mobile phones. Simple shoulder surfing and smudge attack is enough to bypass the PIN and graphical passwords, while sophisticated methods were used by intruders to bypass fingerprint sensor and face recognition [4–7]. In other words, the smartphone believes that the user who is successful in the entry point authentication is the genuine user. But in reality, an insider who knows the login credentials or an intruder who can bypass the traditional authentication mechanisms can pose as the genuine owner and be able to steal/misuse the private information. To address this concern, continuous authentication has been suggested. In this scheme, genuineness of the current user is verified continuously throughout the interaction time with the device. Various methodologies were suggested for implicit continuous user authentication, in which the data generated by the embedded sensors of the smartphones during the interaction of users with the device are used. Continuous authentication based on keystroke dynamics, swipe actions on touch screen, browsing history, call and SMS data, app usage, activity recognition, voice data and facial data are the most researched ones. In some of the suggestions, a fusion of two or more of the above-mentioned were used. Figure 1 lists the available entry point authentication schemes and the most researched continuous authentication schemes. A general framework of continuous authentication is illustrated in Fig. 2; this mechanism can authenticate a legitimate user throughout the session.

In this paper, we are suggesting deep learning-based face recognition for implicit continuous authentication. Face recognition is widely used as an entry point authentication modality. Here, we have used it for continuously monitoring the device user. Inbuilt front camera of the smartphone is made to work in the background without giving any alarms to the current user, to capture the images of him/her. A deep neural network classifier is used. Lighting situations and background are problems with mobile camera setup. Advancements in imaging technology and the sensor upgrades have reduced the adverse effect of lighting situations. Deep learned classifiers are popular for its accurate classification. This work relies on a dataset created in house,

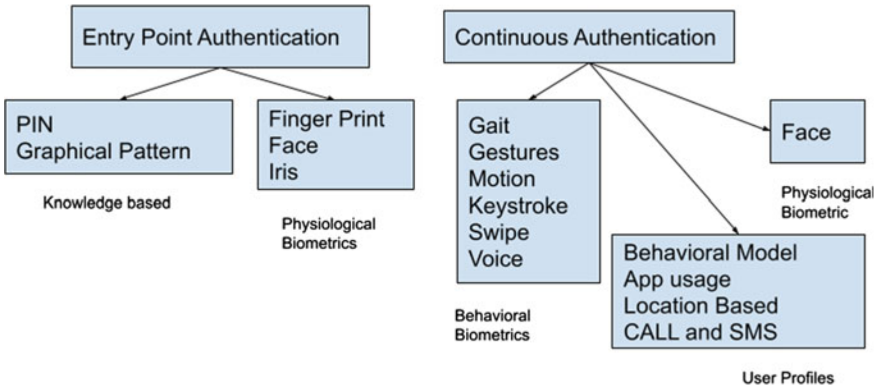


Fig. 1 Entry point and continuous authentication schemes

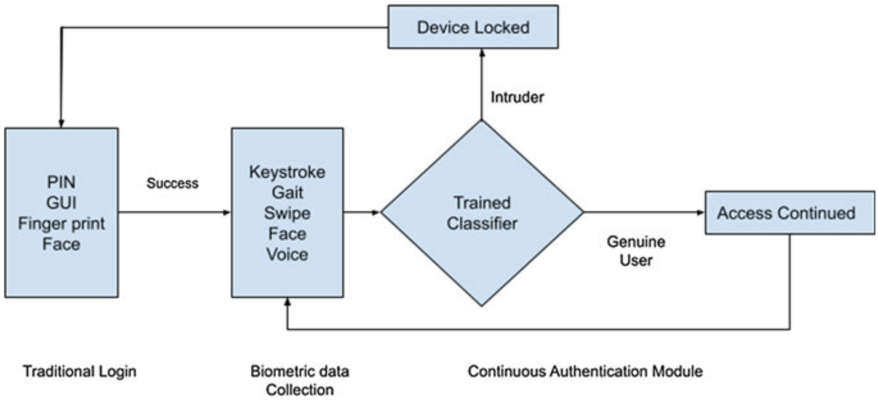


Fig. 2 Biometric continuous authentication framework

mimicking most of the situations in real-time usage of the smartphone by the genuine user/intruder. Continuous authentication using face recognition is more realistic compared to the other listed methodologies since no user can escape the front camera in normal usage.

The advantage of the face recognition system is that no additional hardware is required as almost all the smartphones are equipped with a front camera. There are challenges too, like the image quality which depends on the camera sensor, lighting situations, angle of captured face image. Further, aging-related changes, scars or tattoo on face, after the registration process, can adversely affect the authentication. Whenever an intruder detection system is suggested, it shall be balanced in terms of security and usability. The factors that need to be considered are (1) accuracy of the detection system. A higher false-positive rate is not good for security, and a higher false-negative detection degrades the usability. (2) Latency is another factor which deals with time taken for detection. More time in detection means the intruder will get enough time to steal the data. (3) Efficiency is another important factor which handles the device resources utilization. Being a program running in the background, there should be a balance as more usage of computing power will slow down other operations and battery drainage will happen at a faster rate. In this work, more focus is on accuracy of the system and other two factors are suggested as a future work. Experimental result iterates that the proposed scheme can be used in real time.

Related Works

As mentioned earlier, extensive research was found in continuous authentication with different methodologies. In continuous authentication, facial data is one among the widely used modalities. Here, we are discussing only those works where facial data is the sole/one of the inputs to the binary classifier. Normally, the stages in this type of classification system are described as follows. First stage is the face detection from the video or image captured using the front facing camera of smartphones followed by second stage of feature extraction from the detected face image, and in the third stage, these features are given as input to a classifier for verification of genuineness of the current user. If the classifier decides the current user as genuine, the access to the device is continued; otherwise, the device shall be locked. In the following paragraph, we are discussing some of the recent methods proposed in the area of continuous authentication based on facial data.

Pramuditha Perera and Vishal M Patel proposed a face recognition-based active authentication scheme for multiple users called external open set rejection which is a fusion of identification and verification method. The user count was found to have affected the performance of the system [8]. David Crouse et al. introduced a continuous authentication scheme based on a fusion of face image and inertial measurement unit utilizing the gyroscope, magnetometer and accelerometer sensor along with the mobile camera [9]. Mauricio et al. came with continuous 3D face authentication using RGB-D cameras [10]. This work requires an additional kinetic sensor, they were suggesting this mechanism for a high security environment, and the computation cost for this proposal was high. Mahbub et al. suggested partial face detection for continuous authentication. Fourteen facial section detectors are trained, and the algorithm will estimate a complete or partially detected face. They have claimed improvement in performance over existing methods using pure facial section cascade classifiers [11]. A continuous authentication scheme using face recognition and users' cloth color was proposed by Niinuma et al. in which the current face biometric data and body histogram data are evaluated and compared with enrollment data to obtain a final similarity vector. Based on the threshold of similarity, the system can decide on termination or permission to continue access to the current user [12]. Stoimenov et al. proposed face recognition in Android using feedforward back propagation neural network model designed in Java platform. They concluded that the results obtained in MATLAB simulation and on Java in mobile devices were similar [13]. Simonyan et al. in their paper very deep convolutional networks for large-scale image recognition suggested 19-layered VGG for large-scale image classification. Error percentage was less for their model when compared to other models like GoogLeNet and MSRA [14]. Smith-Creasey et al. introduced a new concept called contextual face authentication to continuously authenticate mobile phone users. Here, the users are to be verified based on factors such as face illumination, face activity, ambient light, accelerometer and gyroscope readings. A virtual user is created with this variety of biometrics, and the same was used for verification [15].

A standardized evaluation procedure seems absent with the works related to continuous authentication. Comparison of works is almost impossible due to the following facts: different datasets and metrics, dissimilar evaluation scenarios and public availability of the datasets.

Most of the research suggestions were on the basis of simulations on popular platforms like MATLAB, Google Colab, etc. A very few implementations on the device were found. The smartphone operating system limitations and the computing power were the main problems. During the last few years, computational power of the mobile phones is being updated and some of these devices are at par with the laptops and desktop in terms of hardware specifications. This would address the onboard computing limitations. Another solution is to have the computation done in a remote server, and the decision may be sent back to the device so that it can lock the device if an intruder found out to be using it.

Materials and Methods

In the proposed implicit continuous user authentication system, the front camera of the device is made to capture the face images of the current user in specified intervals. This facial data is used to verify the genuineness of the user by cross-checking the same with the already enrolled user data. If the DNN classifier confidence is above the prefixed threshold, the user is allowed to continue the usage of the phone; else, the device would be locked and goes back to the initial login stage. The algorithm is depicted in Fig. 3.

Once the initial login is successful, we need to capture the face image of the user in frequent intervals and this has to be repeated throughout the interaction time. Just after the initial login is successful, a variable STRT is set to one. A predefined

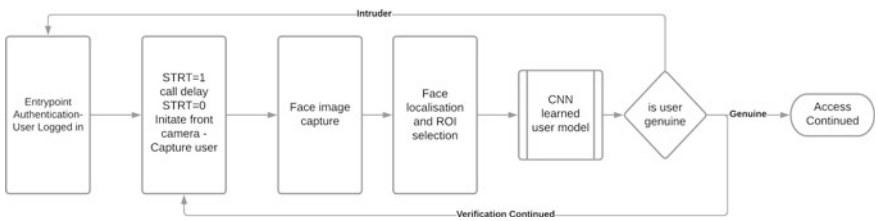


Fig. 3 Block diagram: implicit continuous authentication

Algorithm 1 Face recognition based continuous Authentication

```
Require: Traditional login
Ensure: login successful
set STRT = 1
CALL delay(ms)
initiate front camera
set STRT = 0
while capture front camera = TRUE do
    send capture front camera to ROI module
    if face is seen then
        send capture front camera to Trained classifier model
    else if face not seen then
        lock device and go to Traditional Login
        if Trained classifier output == Genuine then
            Access Continued
            Go to Step 1
        else if Trained classifier output == Intruder then
            lock device and go to Traditional Login
```

Fig. 4 Algorithm: face recognition-based implicit continuous authentication

delay time tdelay in seconds is observed, and after that, the STRT is set to zero. This initiates the front camera of the device, and user face image capture will happen in the background without any signals to the one who is currently using the device. This captured image after face localization and region of interest (ROI) selection fed as input to the already trained CNN classifier to predict the genuineness of the current user. If the model is confident that it is the owner above a specified threshold, the user will be allowed to continue the interaction with the device and the algorithm again sets STRT to one and the process would be repeated. If the prediction says an intruder, the mobile phone will be locked and goes back to entry point authentication. Figure 4 illustrates the algorithm.

Deep Learning-Based Solution

A convolutional neural network is proposed for the deep learning approach. The CNN model, Model-01 architecture (Fig. 5), contains 6 layers, and the first 4 layers have similar blocks. A block contains a set of two convolutional layers accompanied by a max pooling layer. The number of filters is increased as 32, 64, 128 and 256 for the first four blocks of the model. It is followed by fully connected layers. The 6-layer neural network uses ReLU activation for hidden layer and sigmoid output layer activation. RMSProp optimizer is used in this model. The CNN model, Model-02, is an improvement on top of Model-01. There is a change in the block structure for

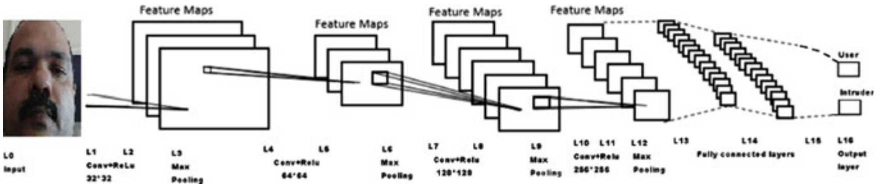


Fig. 5 CNN model overview

the first four layers. Now each block has only a single convolutional and a pooling layer. He-uniform weight initialization is used in all layers. Padding in convolutional layers is implemented. The RMSProp optimizer is replaced by SGD.

Dataset

Most of the referred datasets other than the UMDAA dataset were created in controlled environments and seen that no considerations were given to the viewing angle of the mobile phone screen when used for normal operations [16]. The face image used for the entry point face recognition system is more or less similar to selfie capturing angle. The mentioned dataset was collected by a program running in the background, and the frames were having partial images, background information and sometimes no face at all. Another fact worth mentioning is that the camera sensor of the smartphones being upgraded to have more pixels and, as a result, clarity of the face images captured by recent phones are better in terms of resolution which would help feature extraction stages. Furthermore, user adaptation is crucial in our binary classification and custom dataset will give an optimum performance. Considering the above mentioned facts, we have created our own dataset and the process is described below. We have selected 4 subjects, 1 male and 3 females, aged 30–46 years, and the data collection device, an android smartphone with the front camera resolution of 12 MP, was given to them with an instruction to open some of the apps and to use it for a small time period. Specific instructions were given to them that they have to use it keeping in mind that the usage pattern should be the same as that of a normal day-to-day usage of smartphones. A video recording program was running in the background which continuously recorded the frames through the front camera. This video file is the base of our dataset creation. From this video, frames were saved as images. From these images, background and other noises are eliminated and only the face part is saved using a suitable face detection program. A sample of the dataset is given in Fig. 6. These face images are augmented for increasing the number of samples. Face localization and ROI selection reduced the size of the images and resulted in less computational complexity for training and validation. The same process is repeated with all the subjects. Now, we have enough data with face images of 4 different persons. One person is selected as genuine, and

Fig. 6 Samples from dataset

others considered intruders. These samples were split into training and validation data. Convolutional neural networks are trained using this dataset, and a user model is created. Model so created is tested with unseen test data. Performance metrics were noted (sensitivity, specificity, precision, accuracy and F1 score). The two CNN models Model-01 and Model-02 are implemented using Keras in the TensorFlow 1.x version. All simulations are carried out in Google Colab [17].

CNN Model-01 and Model-02

The models consist of 5 hidden layers having ReLU activation function and one output layer with sigmoid function. Rectified linear unit (ReLU) activation function is commonly used in the initial layers of a neural network model [18, 19]. The ReLU function is defined as $f(x)$ =max (0, x). The function value is either x or 0. If x is positive, $f(x)$ is x itself. If x is negative, $f(x)$ is 0. ReLU activation function is useful for the initial hidden layers of CNN as it fixes the vanishing gradient problem where the error gradient tends to vanish as the error gradient propagates backward through the network from the output layer to the input layer. Sigmoid function is usually used at the final layer closer to the output layer as the function ranges between 0 and 1. At the final layer, the neuron potential converges to the limit values (0, 1) [or (−1, 1)] and, thus, provides a probabilistic solution. Here, it provides the probability of the test image being a user or intruder. Optimization algorithms form the basis on which a model is able to learn through its training. The models compute gradients and attempt to minimize the loss function. In Model-01, root mean square propagation (RMSProp) algorithm is used, whereas in Model-02, stochastic gradient descent (SGD) algorithm is used. The difference between RMSProp and SGD is on how the gradients are calculated. The SGD algorithm trains the whole dataset at one shot [12]. Its variation is stochastic gradient descent that performs the training on each data image. In the RMSProp algorithm, gradients are assembled together into an exponentially weighted mean. RMSProp dumps the history and sustains only the latest gradient information. The need for uniform weight initialization is that the random initial value cannot be too small or too large. If the weights in each neuron of the network initialized at a value are too small, then the signal shrinks as it propagates through each layer till it becomes too small to be useful. On the other hand, if the weights are initialized at a large value, then the signal grows and becomes too huge to be useful. So as an improvement to Model-01, in all the layers of Model-02, the weights of all the neurons have been initialized. In He-uniform distribution, an initial random symmetric weight with variance U(0, 1) is implemented [20]. Figure 7 describes the model summary of Model-01. Figure 8 describes the model summary of Model-02. Both the summaries describe the output shape, size and the total number of parameters in each layer.

Results and Discussion

Our main objective is to design an implicit continuous authentication of smartphone user using face recognition. In this paper, we suggested deep learning-based face recognition to indirectly authenticate the smartphone user. CNN classifiers give very good results on image classification. All the simulations were carried out in Google Colab. The accuracy graph and the loss graph of both the CNN models Model-01 and Model-02 are compared. In the accuracy graph of Model-01, Figure 9, there is a gap between the training and validation plot and they are not converging to a single point. In the accuracy graph of Model-02, Figure 10, the training and validation plot is converging at a point. That denotes minimum loss point; that is, the value of weights of each neuron in the model is optimized at that point. The two deep learned user models, Model-01 and Model-02, are tested using another unseen dataset. Here, the unseen dataset consists of 250 facial images, out of which 130 were of genuine users and 120 were of intruders. Confusion matrix for both models is given in Fig. 11. Other performance metrics are also given in Fig. 12. Simulation results clearly show that

```
Layer (type)                 Output Shape              Param #
=================================================================
conv2d_1 (Conv2D)            (None, 200, 200, 32)      896
_________________________________________________________________
conv2d_2 (Conv2D)            (None, 200, 200, 32)      9248
_________________________________________________________________
max_pooling2d_1 (MaxPooling2 (None, 100, 100, 32)      0
_________________________________________________________________
conv2d_3 (Conv2D)            (None, 100, 100, 64)      18496
_________________________________________________________________
conv2d_4 (Conv2D)            (None, 100, 100, 64)      36928
_________________________________________________________________
max_pooling2d_2 (MaxPooling2 (None, 50, 50, 64)        0
_________________________________________________________________
conv2d_5 (Conv2D)            (None, 50, 50, 128)       73856
_________________________________________________________________
conv2d_6 (Conv2D)            (None, 50, 50, 128)       147584
_________________________________________________________________
max_pooling2d_3 (MaxPooling2 (None, 25, 25, 128)       0
_________________________________________________________________
conv2d_7 (Conv2D)            (None, 25, 25, 256)       295168
_________________________________________________________________
conv2d_8 (Conv2D)            (None, 25, 25, 256)       590080
_________________________________________________________________
max_pooling2d_4 (MaxPooling2 (None, 12, 12, 256)       0
_________________________________________________________________
flatten_1 (Flatten)          (None, 36864)             0
_________________________________________________________________
dense_1 (Dense)              (None, 256)               9437440
_________________________________________________________________
dropout_1 (Dropout)          (None, 256)               0
_________________________________________________________________
dense_2 (Dense)              (None, 256)               65792
_________________________________________________________________
dropout_2 (Dropout)          (None, 256)               0
_________________________________________________________________
dense_3 (Dense)              (None, 1)                 257
_________________________________________________________________
activation_1 (Activation)    (None, 1)                 0
=================================================================
Total params: 10,675,745
Trainable params: 10,675,745
Non-trainable params: 0
_________________________________________________________________
```

Fig. 7 CNN Model-01 summary

Model-02 outperforms Model-01 in all performance metrics. Our aim is to detect the presence of an intruder and prevent data stealing/misuse. Model-02 has specificity (true-negative rate) value 1 which is highly essential for our binary classification. Misclassification of the genuine user as intruder for Model-02 is very low as the sensitivity (true-positive rate) is 0.9154. F1 score of 0.9558 assures the system would be a better one in assessing false positives and false negatives. These results indicate that the Model-02 is better in all performance metrics and highly suitable for real-time usage in a smartphone to authenticate the current user in an implicit continuous manner. In real-time implementation, we can have two methodologies onboard or offboard. Onboard means the learning as well as detection happens on the smartphone itself. In case of offboard, the learning, model creation and testing happen in the remote server and decisions will be sent back to the device. In our simulations, a time of approximate 1s is taken for decision making by the deep learned model. A better solution will be a mixture of both schemes. A remotely learned model that can be

Model: "sequential_1"

Layer (type)	Output Shape	Param #
conv2d_1 (Conv2D)	(None, 200, 200, 32)	896
max_pooling2d_1 (MaxPooling2	(None, 100, 100, 32)	0
conv2d_2 (Conv2D)	(None, 100, 100, 64)	18496
max_pooling2d_2 (MaxPooling2	(None, 50, 50, 64)	0
conv2d_3 (Conv2D)	(None, 50, 50, 128)	73856
max_pooling2d_3 (MaxPooling2	(None, 25, 25, 128)	0
conv2d_4 (Conv2D)	(None, 25, 25, 256)	295168
max_pooling2d_4 (MaxPooling2	(None, 12, 12, 256)	0
flatten_1 (Flatten)	(None, 36864)	0
dense_1 (Dense)	(None, 128)	4718720
dense_2 (Dense)	(None, 1)	129

Total params: 5,107,265
Trainable params: 5,107,265
Non-trainable params: 0

Fig. 8 CNN Model-02 summary

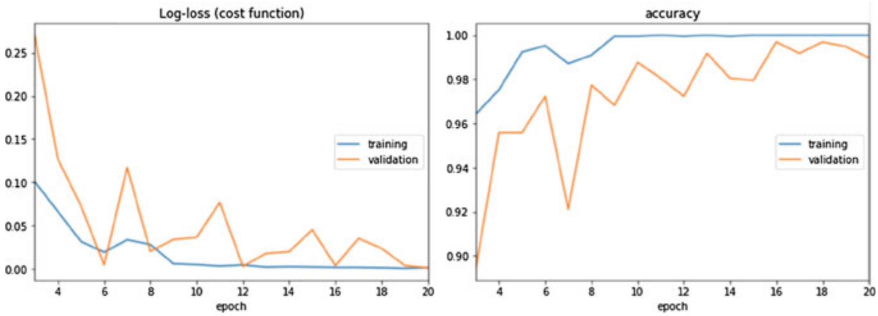


Fig. 9 CNN Model-01—loss and accuracy graph

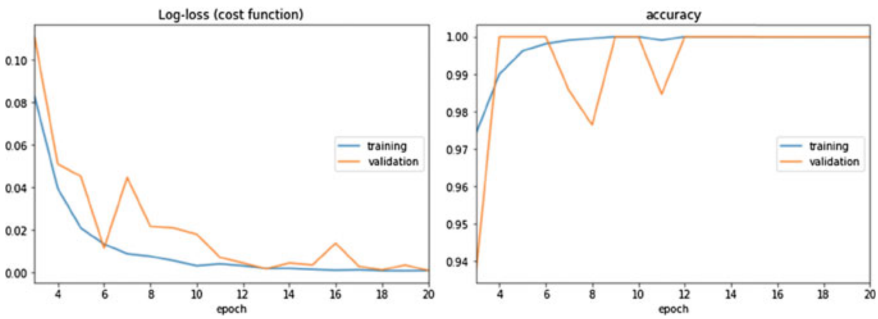


Fig. 10 CNN Model-02—loss and accuracy graph

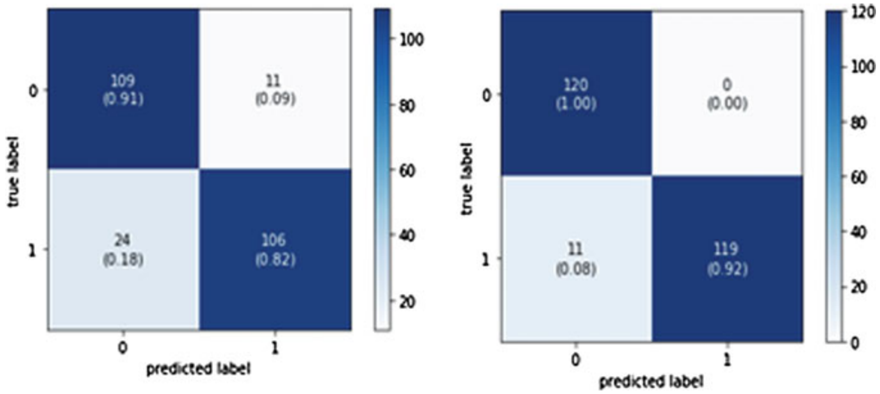


Fig. 11 Confusion matrix for Model-01 and Model-02

Test Performance				
Sl No	**Measure**	**Model -01 Value**	**Model -02 Value**	**Derivations**
1	Sensitivity	0.8154	0.9154	TPR = TP / (TP+FN)
2	Specificity	0.9083	1	SPC = TN / (FP +TN)
3	Precision	0.9060	1	PPV = TP / (TP +FP)
4	Accuracy	0.86	0.9560	ACC = (TP+TN) / (P+N)
5	F1 Score	0.8583	0.9558	F1 = 2 TP / (2TP+FP+FN)

Fig. 12 Performance of the system

updated periodically will be ported back to the device for onboard user classification. Nowadays, the smartphones are equipped with faster CPUs and dedicated GPUs. These embedded GPUs can ensure a better onboard classification process using the already ported user model.

Conclusions and Future Work

Research on continuous/active authentication on mobile devices still attracts researchers due to the fact that more sensors are being added to the new-generation mobile phones and computation power of latest devices is capable of running complex computations on the device itself. The embedded sensors of smartphones are being updated every now and then. Taking the case of the front camera, which is the prime data source in this work, resolution is upgraded with every new model. A dataset created will be outdated very early. Device-independent large-scale realistic dataset which is being updated frequently is highly essential for the research. The effect of various lighting conditions is to be carried out. From the perspective of a smartphone user, the authentication time plays an important role. It should be minimum, without giving any burden on the user. Furthermore, the proposed solutions need to be implemented on smartphones so that real-time computational analysis can be carried out. The contributions of this work can be summarized as follows. An optimized deep learning model is implemented for implicit continuous authentication of smartphone users. A face localization and ROI selection are introduced to reduce the computational complexity in user modeling which can be helpful in onboard implementation of the proposal on the smartphones. Experimental results show the proposed method can be used in real time for implicit continuous authentication of smartphone user. Compared to other continuous authentication modalities like activity recognition, voice, typing, gestures, face recognition can be considered fool proof. We cannot expect an intruder to make use of the above-mentioned modalities other than face recognition as alternate methods are available to steal data from a mobile device. But no user can fully stay away from the camera, and a single snap is enough to recognize the genuine owner.

References

1. Van Deursen AJ, Bolle CL, Hegner SM, Kommers PA (2015) Modeling habitual and addictive smartphone behavior: The role of smartphone usage types, emotional Intelligence, social stress, self-regulation, age, and gender. Comput Hum Behav 45:411–420
2. Islam N. and Want R., "Smartphones: Past, Present, and Future" in IEEE Pervasive Computing, (13) 89–92 (2014)
3. Farrell, G. "Preventing phone theft and robbery: the need for government action and international coordination". Crime Sci 4, 4 (2015)
4. Khan H, Hengartner U, Vogel D (2018) Evaluating attack and defense strategies for smartphone PIN shoulder surfing. In: Proceedings of the 2018 CHI conference on human factors in computing systems (CHI '18). Association for Computing Machinery, New York, NY, USA, vol 164, pp 1–10
5. Airowaily K, Alrubaian M (2011) Residuals security threat on smart phones. In: First international conference on robot, vision and signal processing, pp 300–302
6. Winder D (2019) Hackers claim 'any' smartphone fingerprint lock can be broken in 20 minutes. Available at https://www.forbes.com/sites/daveywinder/2019/11/02/smartphone-security-alert-as-hackers-claim-any-fingerprint-lock-broken-in-20-minutes/?sh=13f1df8e6853
7. Foltýn T (2019) Face unlock on many android smartphones falls for a photo. WeLiveSecurity. Available at https://www.welivesecurity.com/2019/01/10/face-unlock-many-android-smartphones-falls-photo/
8. Perera P, Patel VM (2019) Face-based multiple user active authentication on mobile devices. IEEE Trans Inf Forensics Secur 14(5):1240–1250
9. Crouse D, Han H, Chandra D, Barbello B, Jain AK (2015) Continuous authentication of mobile user, fusion of face image and inertial measurement unit data. In: IEEE 2015 international conference on Biometrics (ICB), Phuket, pp 135–142

10. Pamplona Segundo M, Sarkar S, Goldgof D, Silva L, Bellon O (2013) Continuous 3D face authentication using RGB-D cameras. In: IEEE conference on computer vision and pattern recognition workshops, Portland, OR, pp 64–69
11. Mahbub U, Patel VM, Chandra D, Barbello B, Chellappa R (2016) Partial face detection for continuous authentication. In: IEEE international conference on image processing (ICIP), Phoenix, AZ, pp 2991–2995
12. Niinuma K, Jain AK (2010) Continuous user authentication using temporal information. In: Proceedings of SPIE. The International Society for Optical Engineering
13. Stoimenov S, Tsenov GT, Mladenov VM (2016) Face recognition system in Android using neural networks. In: 13th Symposium on neural networks and applications (NEUREL), Belgrade, pp 1–4
14. Simonyan K, Zisserman A (2015) Very deep convolutional networks for large scale image recognition. ICLR Computer Vision and Pattern Recognition. https://arxiv.org/abs/1409.1556
15. Smith-Creasey M, Albalooshi FA, Rajarajan M (2018) Context awareness for improved continuous face authentication on mobile devices. In: IEEE 16th international conference on dependable, autonomic and secure computing, Athens, pp 644–652
16. Mahbub U, Sarkar S, Patel VM, Chellappa R (2016) Active user authentication for smartphones: a challenge data set and benchmark results. In: IEEE 8th international conference on biometrics theory, applications and systems BTAS, pp 1–8
17. Bisong E (2019) Google collaboratory. In: Building machine learning and deep learning models on Google cloud platform. Apress, Berkeley, CA
18. Kirana KC, Wibawanto S, Hidayah N, Cahyono GP, Asfani K (2019) Improved neural network using integral-RELU based prevention activation for face detection. In: International conference on electrical, electronics and information engineering (ICEEIE), Denpasar, Bali, Indonesia, pp 260–263
19. Femin A, Biju KS (2020) Accurate detection of buildings from satellite images using CNN. In: International conference on electrical, communication, and computer engineering (ICECCE), pp 1–5
20. Zaheer R, Shaziya H (2019) A study of the optimization algorithms in deep learning. In: Third international conference on inventive systems and control (ICISC), Coimbatore, India, pp 536–539

Fish Species Detection Using Deep Learning for Industrial Applications

K. Yashaswini, A. H. Srinivasa, and S. Gowrishankar

Abstract In this world, almost 50% depend on sea food which serves as a major protein source and also one of the important sources of the economic considering for countries from all over the world. Directly, the seafood can be used for many of the applications which will be processed further to a very high functional and also as well as the products of the nutraceutical food which will affect the health of the human beings. In many of the examples, skin of the fish and fins is used in order to extract collagen and gelatin. In the fishing industry, for the classification purpose it is necessary to identify the fish species is very important. Our proposed methodology is based on the CNN and faster RCNN technique for the fish species identification in the industrial applications. In this proposed work, CNN and faster RCNN almost show 95 and 98% of the accuracy.

Keywords Deep learning · CNN · RCNN · Fish species · Gelatin · Collagen

Introduction

Fishing industry is mainly concerned with processing, preserving, storing, culturing, taking, marketing, transporting and selling the fish products [1]. Resources of aquaculture various transformations of those resource products are used for sale. This is also referred as seafood industry.

Around the world, the important sources for their economics are the marine-based food industries. It's known that around 60% of the waste products is being generated around the world under this seafood processing industry, and for human consumption,

K. Yashaswini (✉) · A. H. Srinivasa · S. Gowrishankar
Department of Computer Science and Engineering, Dr.Ambedkar Institute of Technology, Bengaluru, Karnataka 560056, India

A. H. Srinivasa
e-mail: srinivasaah.cs@drait.edu.in

S. Gowrishankar
e-mail: gowrishankarnath@acm.org

V. Bindhu et al. (eds.), *Proceedings of Third International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 844,
https://doi.org/10.1007/978-981-16-8862-1_26

around only 40% is being utilized. The food and the nutraceutical industries gained popular attention because of seafood-derived by-products and their potential benefits.

There are many nonfood items, such as fish products, which include fish oil, fish glue, and are also mainly for medical purposes. The categories of these by-products which will be caught or discarded parts of the fish body under the commercial usage of fish, for example, heads, gills, skin, trimmings, fins, frames, bones, viscera, blood, and roes [2]. For example, from the fish skin and fins extract collagen, which is used for many cosmetic products.

For small-scale fisheries which is having the small volume of fish catch easy to classify the fish but for large fishing industries tonnes of fishing done at a time in sea, because of more varieties and large quantity it is difficult to classify the mixed species [3]. Due to the large quantity and many varieties of fishes in fishing industry, it is very important to classify the fish for processing and marketing. The major problem to classify the image is different environmental condition, noise and segmentation [4].

Deep learning is one of the maximum using techniques in every field [5]. It is a machine learning method where it helps to make the machine think like humans and act like humans based on good training. To classify the images, audio and video deep learning have top most techniques. This important technique gives a high level of accuracy and performance while prediction of the class [6]. The main aim of this paper is to classify the image using deep learning for the industrial applications.

In deep learning, convolutional neural networks are one of the methods used for classification and segmentation problems [7]. It has several convolution layers which help extract features from the images. Because of the multiple layers, the data representation is also improved [8].

However, there was one problem. CNNs were too slow and computationally very expensive. It was impossible to run CNNs on so many patches generated by the sliding window detector [9]. R-CNN solves this problem by using an object proposal algorithm called Selective Search, which reduces the number of bounding boxes that are fed to the classifier to close to 2000 region proposals [10].

The proposed work mainly concentrates on the food and the medicinal applications in fishing industry, for example, Otolithes ruber, Lutjanus vitta, Megalaspis cordyla; these are the fish names from our dataset used for isolate collagen which is mainly used for cosmetic purpose [11].

In order to classify the image CNN, faster R-CNN is mainly used which classifies the image easily. Finally, the comparison is made between them which method give the more accuracy. We do fish classification in several ways by analyzing the shape, texture and color of the fish. Apart from this, the main work is to classify the fish with a good accuracy.

The rest of the paper is organized as follows. In Sect. "Literature Survey", we explore the current research on this topic. The methodology used for our proposed work is explained in Sect. "Methodology". Dataset is discussed in Sect. "Dataset Analysis". Results are presented in Sect. "Results", and finally we conclude.

Literature Survey

In [12], a convolutional neural network is used which helps for classification, and R-CNN is used for localizing the images to get better accuracy. The important classification of the fish can be done by the proper dataset training at the level of high accuracy, which will be helpful by providing the improved different activation functions. At last done the comparison between two methods, which gives better result find out.

In deep learning neural network for the classification of the fish, which is labeled image automatically, by using a certain camera with no human intervention. The classification of the image is done in two steps; the first one at the instance level, and the second one is the image-level classification. The following method is proposed in [13].

In [14], a proper novel technique which can be used for the convolutional neural network under the consideration of the deep learning and also as well as image processing to get the proper high level of accuracy is discussed. There is an implementation in this paper known as Otsu's thresholding which is used for noise removal and also as well as a gray-level histogram which is supposed to be created under grayscale image processing. In the process, there will be an implementation of the consider operation known as morphological and also erosion, including the deletion process. The reason importance is given to the foreground thickness for the fish decreases. And this is the application of the convolutional neural network for the implementation to consider the classification of the species of fish.

In [15], a proper deep learning methodology for the lightweight neural network techniques where the proposed model is considered to be having the faster RCNN is proposed. There will be any change in the convolutional neural network, and also there will be an adding block such as ReLu along with the consideration of inception and also hyperspecies. For the set of images, construct a match and model where the proper development and identity of the fishes will be done automatically, and the proper output is calculated with the high level of accuracy for each fish.

A model which helps in statistical extraction of the images depends upon the behavior of the fish with and also without considering the cleaning of the dataset. İn [16], this method uses the technique of deep learning consider in the important algorithm known as clustering, where it is helpful for the removal of unwanted noises and also fast detection as well, and also it is helpful to correct the trajectories of the assignments. The final results can be observed that the higher the temperature speed of the fish is also increased.

Rekha et al. [17] proposes three phases, augmentation, detection and classification. Augmentation is used to take fish images using camera. CNN is used for detection and classification of images. Name of implementing of a system is to help the researchers and also investigators for the proper analysis of the image which is captured by the consideration of the board cameras which is, later on, helpful for the classification of the fishes irrespective of their species.

Fish species are identified automatically and separately calculated the each fish species accuracy, to accomplish this in [18], propose novel method to classification of image dataset which is used to identify many fish species with precise result. Algorithms for image dataset building, modeling help to develop the fish species recognition automatically.

Methodology

Convolutional Neural Network

In deep learning, convolutional neural networks (CNNs) is the one of the prominent methods, which is helpful for the training and tested the multiple layers in an effective manner. Considering the convolutional neural network, there will be mainly four categories of layers: CNN layer, ReLu layer, max pooling layer and fully connected layer.

Convolutional neural networks (CNNs) which are the ability to extract the proper information which is categorized based on color shape and also as well as texture. As the number of the input images and training increases, the accuracy of the output results also increases regardless of the dataset which is collected.

ReLu layer activates the node if the input is in above certain quantity, max pooling layer mainly used for shrink the image and the output of the pooling layer given to the fully connected layer where it classifies the image to the specific class.

CNN Architecture of Fish Species Detection

The appearance of the fish can be easily identified by the image, considering some of the conditions in the environment sometimes the image which will be blurry, and also it is cannot be viewed properly and some of the important initial knowledge which is must be needed for the context information under the identification of the fish. During this kind of consideration, the prediction can be done by using the methodology CNN by the utilization of the prediction probability theory. The input will be the entire image of the fish and the prediction, which will be the output.

For the proper classification of the fish images, must have good knowledge of relationship between the image features and the classes of the image which are classified for the particular class. Hence, we can classify the fish species easily by training dataset. As shown in Fig. 1 in this proposed work, there will be two convolutional layers and two pooling layers, and the next one will be the two fully connected layers and ReLu activation function is used for the classification of the fish species with 95% accuracy.

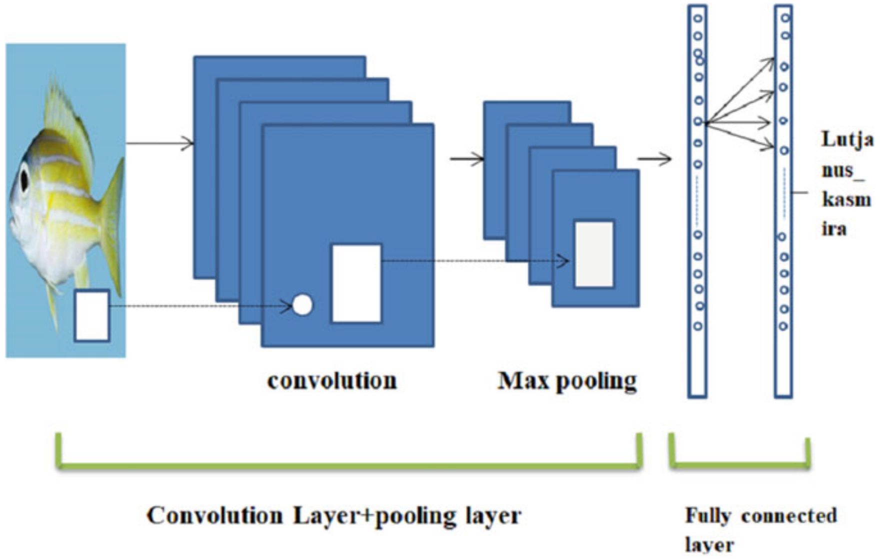


Fig. 1 CNN architecture of fish species detection [19]

Faster R-CNN

R-CNN is the term which can be explained as region-based convolutional neural network. The main method of this is included by the combination of the region proposal segmentation of the object and also considering the high capacity for the object detection CNN, the algorithm-based method the R-CNN technique, in which as described below.

Considering the selective search algorithm, the extraction for the image input is considered from the several proposals in the region which is obtained from a particular candidate. Considering this algorithm, the candidate region is a selection that is generated considering the initial sub-segmentation process.

If any 10 regions are found to be very similar and later on are combined to form the region bigger using the algorithm known as greedy, hence the final result which can be obtained for region proposals. The CNN component for proper vector output, the extraction of the features, can be done under the proper proposals.

The considered extracted features of the images are later on fed into the support vector machine, which is in short known as SVM such that for the recognition of the object under the proper interest of the proposal. The faster R-CNN training and prediction process is faster than the other methods of the region's proposals, R-CNN and fast R-CNN.

.

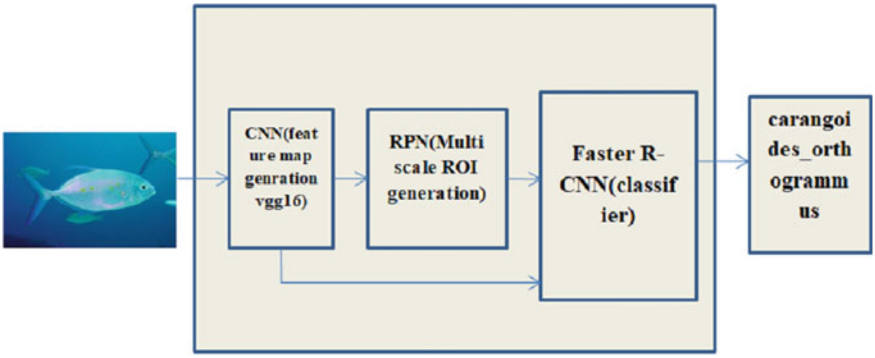


Fig. 2 Faster RCNN architecture [20]

Faster R-CNN Architecture

In faster R-CNN architecture, entire image is directly fed to the backbone of CNN and the features which can be extracted from the layer of convolutional. The output seems to have much smaller than that of the original size of the image where the feature map extracted. The obtained feature is properly divided based on the proposal windows, which is derived from that of the regional proposal algorithm**.**

Like regular pooling, ROI pooling is carried out in every channel individually which will cause in proper fixed-size features at the output. The output features from the ROI pooling layer (N × 7 × 7 × 512), where N is the number of proposals which is later on directly fed into the successive fully connected layer, where it predicts the particular class as shown in Fig. 2.

Dataset Analysis

The dataset used in the project is QUT FISH Dataset downloaded from Kaggle. In this dataset, 3,960 fish images are there which are collected from around 468 species. From this large dataset, 10 fish species are selected. The fish names used in our work are aphareus_furca, bodianus_axillaris, carangoides_orthogrammus, caranx_ignobilis, elagatis_bipinnulata, lutjanus_kasmira, lutjanus_vitta, megalaspis_cordyla, otolithes_ruber and xiphocheilus_typus.

Results

There are 1200 images used from10 different fish species, and they are divided into 80% and 20% for training and testing, respectively. In CNN, the pretrained

Table 1 Training parameters

Parameters	CNN	Faster R-CNN
Epochs	30	20
Optimizer	Adam	Adam
Batch size	50	32
Objective	Categorical cross-entropy	Categorical cross-entropy

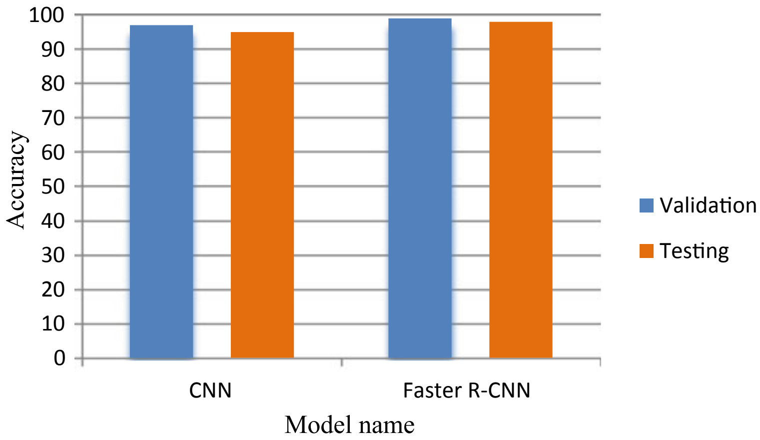


Fig. 3 CNN and faster R-CNN validation and testing accuracy

Inception_v3 architecture is used for classification of fish species, and in the faster R-CNN classifier, the VGG-16 pertained network is used on the dataset and fine-tuned features are extracted from the training data. The parameters used while training the network are listed in Table 1.

Convolution neural network gives 95% testing accuracy and 97% validation, and one of the region proposal network faster R-CNN gives the test accuracy 98% and validation accuracy 99% as shown in Fig. 3.

Conclusion

Both CNNs and faster R-CNN give the remarkable result in detection of fish species, where CNNs is very slow and computationally very expensive, so in our work faster R-CNN gives better result to compare to CNN. In fish species, detection achieved an accuracy of about 95% with CNN. Faster R-CNN also shows the better result with 98% accuracy. Both methods which will not show up the 100% accuracy since noises in background images and dataset images are less and quality of the image affects the result.

References

1. Fishing industry. https://en.wikipedia.org/wiki/Fishing_industry
2. Vidanarachchi JK, Ranadheera CS, Wijerathne TD, Udayangani RMC, Himali SMC, Pickova J (2014) Applications of seafood by-products in the food industry and human nutrition. In: Kim S-K (ed) Seafood processing by-products. Springer Science+Business Media New York, pp 463–527
3. Fische J (2013) Fish identification tools for biodiversity and fisheries assessments. In: Fischer J (ed) FAO fisheries and aquaculture technical paper, pp 107
4. Knausgard KM, Wiklund A, Sordalen TK, Halvorsen KT, Kleiven AR, Jiao L, Goodwin M (2020) Temperate fish detection and classification: a deep learning based approach, pp 1–14
5. Cui S, Zhou Y, Wang Y, Zahi L (2020) Fish detection using deep learning, pp 1–13
6. Khalifa NEM, Hamed MN, Taha, Hassanien AE (2019) Aquarium family fish species identification system using deep neural networks, pp 347–356
7. Tseng SL, Li HY (2020) Fish detection using convolutional neural networks with limited training data, pp 1–13
8. Iqbal MA, Wang Z, Ali ZA, Riaz S (2019) Automatic fish species classification using deep convolutional neural networks, pp 1–10
9. Object detection, https://cv-tricks.com/object-detection/faster-r-cnn-yolo-ssd
10. Adiwinataa Y, Sasaokab A, Bayupatia IPA, Sudanaa O (2020) Fish species recognition with faster R-CNN inception-v2 using QUT FISH dataset, pp 144–154
11. Venkatesan J, Anil S, Kim SK, Shim MS (2017) Marine fish proteins and peptides for cosmeceuticals: a review, pp 1–18
12. Varalakshmi P, Rachel J (2019) Recognition of fish categories using deep learning technique, pp 168–172
13. Chen G, Sun P, Shang Y (2017) Automatic fish classification system using deep learning, pp 24–29
14. Rathi D, Jain S, Sreedevi I (2018) Underwater fish species classification using convolution neural network and deep learning, pp 1–6
15. Li X, Tang Y, Gao T (2017) Deep but lightweight neural networks for fish detection, pp 1–5
16. Beyan C, Vasiliki-Maria K, Fisher RB (2018) Extracting statistically significant behavior from fish tracking data with and without large dataset cleaning, pp 162–170
17. Rekha BS, Srinivasan GN, Reddy SK (2020) Fish detection and classification using convolutional neural networks, pp 1221–1231
18. Shafait F, Mian A, Shortis M, Ghanem B, Culverhouse PF, Edgington D, Cline D, Ravanbakhsh M, Seager J, Harvey ES (2016) Fish identification from videos captured in uncontrolled under water environments, pp 2737–2746
19. Vani R (2017) Towards efficient intrusion detection using deep learning techniques: a review, pp 375–384
20. Bravo MG, Calero MF (2018) Pedestrian detection at night based on faster R-CNN and far infrared images, pp 48–57

Performance Analysis of Video Transmission Over OWC-PON in Weak Turbulence Regimes

Harpuneet Singh Gill, Maninder Lal Singh, Mandeep Singh, Priyanka, and Sehajpal Kaur

Abstract Optical wireless communication (OWC) is an attractive candidate for the next-generation communication networks. OWC in the form of passive optical networks, i.e., PON can provide real sense broadband connectivity from core to the end user. However, the atmospheric turbulence severely impairs the performance of such systems. In this paper, we investigate the performance of video transmission over OWC-PON network in the presence of weak turbulence. For holistic performance description, both the domains of Quality of Service (QoS) and Quality of Experience (QoE) are investigated and compared for limiting cases. Variation of metrics for both the realms is shown to vary with channel parameter, i.e., Rytov variance. Our results show that in the case of weak turbulence due to QoE visual quality assessment, an additional loss of 55 m is incurred in the system. This loss would be hidden had only been QoS performance evaluated.

Keywords Optical wireless communication (OWC) · Passive optical networks (PONs) · Quality of Service (QoS) · Quality of Experience (QoE) · Atmospheric turbulence · Structure similarity index (SSIM) · Peak signal-to-noise ratio (PSNR)

H. S. Gill (✉) · M. L. Singh · M. Singh · Priyanka · S. Kaur
Department of Electronics Technology, Guru Nanak Dev University, Amritsar, India
e-mail: harpuneetece.rsh@gndu.ac.in

M. L. Singh
e-mail: mlsingh.ece@gndu.ac.in

M. Singh
e-mail: mandeepece.rsh@gndu.ac.in

Priyanka
e-mail: priyankaece.rsh@gndu.ac.in

S. Kaur
e-mail: sehajpal.ece@gndu.ac.in

V. Bindhu et al. (eds.), *Proceedings of Third International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 844,
https://doi.org/10.1007/978-981-16-8862-1_27

Introduction

With the convergence of information communication, broadcast and broadband services [1–4], the last decade has seen an enormous amount of data generation. Enabling technologies for dissemination of this data have undergone drastic changes and are yet in this process for even new service delivery [5]. With the successful rollout of 5G in 2020, the industry and experts driven by exploration of new services such as extended immersive multimedia, extended virtual reality, holographic communication and teleportation to name a few, are constantly looking way ahead in the realm of 6G [6, 7]. With the advent of smart and more sophisticated devices, the computing has gone from cloud to edge in device so that real-time services can be delivered with low latency yet providing personalized Quality of Experience to the end user. This convergence of computing with triple play of communication, broadband and broadcast marks the dawn of era of 6G where delivery of services with high speeds, low latency, ubiquitous coverage, pervasive networks, large mobility support, high spectrum efficiency shall be inevitable feature. This would attempt to create an intelligent digital world. 6G would be a seamless network of connected intelligence.

Area traffic capacities of now (few Mbps/m^2) and devices connected per km^2 are going to increase ten folds in 6G [8]. To support such massive traffic and devices, the access networks would have to be updated. The existing access networks on RF suffer from various limitations such as spectrum congestion, low availability, licensing and management overheads, cell overlap, microcell handoffs limitations, etc. [9]. The use of optical fiber in access networks such as FTTx [10] in the form of passive optical networks (PONs) [11] has provided an inexpensive yet effective solution to last mile bottleneck. However, there are certain situations where possibility of a wired link is out of question. The geographical barriers would not allow a physical wired links in tough terrains. In these scenarios, to provide a huge potential bandwidth, optical wireless communication (OWC) [9, 12] networks come in handy. In OWC-PON, optical wireless carrier bearing information is launched into atmosphere from OLT and received at ONU where user devices are connected. An all optical splitter splits the carrier into N nodes. The downlink is obviously a broadcast link. The situation is shown in Fig. 1.

However due to interaction with atmosphere, certain phenomenon arises such as atmospheric turbulence which impairs the system performance [13]. In multimedia service, dissemination Quality of Service (QoS) alone cannot portray the performance of the system [14]. The evaluation of Quality of Experience (QoE) [15] which is indicator of the ultimate quality as perceived by end user is also as important as QoS analysis. Various QoE parameters for video quality assessment (VQA) [16, 17] metrics are available to assess multimedia quality. To provide a holistic performance evaluation, computation of both QoE and QoS is crucial.

In this paper, a video transmission over OWC-PON architecture under weakly turbulent regimes is investigated. The QoS metric such as bit error rate (BER) is evaluated. For QoE assessment, metrics of error visibility such as mean square error

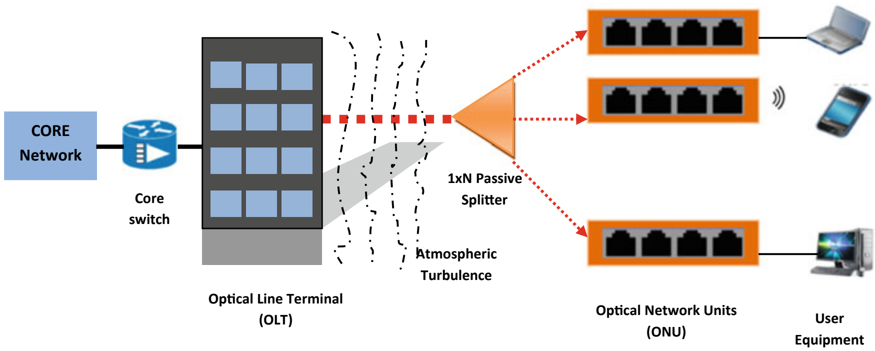


Fig. 1 A typical optical wireless communication passive optical network (OWC-PON)

(MSE) and peak signal-to-noise ratio (PSNR) and structure similarity such as Structure Similarity Index (SSIM) are computed. For measures of association, Pearson's correlation coefficient is computed along channel parameter variation. The channel parameter for weak turbulence is Rytov variance at various channel lengths. The perceivable quality decline is presented by video frames from where degradation can be visually noted. A comparison at limiting cases of QoS and QoE is done to evaluate the difference from both aspects of video transmission.

Methodology

For broadcast transmission, a video with features in Table 1 is taken. Figure 2 shows the screenshot of the proposed video.

The spatial information (SI) [16] and temporal information (TI) [16] are important parameters for defining scene complexity. Spatial information represents the amount of spatial objects carried in a frame. It is calculated as follows [18]:

$$\mathrm{SI} = \max\left\{\mathrm{rms}_{\mathrm{space}}[\mathrm{Sobel}(F(t_n))]\right\} \tag{1}$$

where space is the horizontal and vertical dimensions of image, t_n is the nth time frame at frame rate, rms is the root mean square function, max is the maximum function, Sobel is filter measures edge energy of the spatial content. For this video, SI is listed in Table 1.

Table 1 Features of the broadcasted video

Frame rate	Coding	Total frames	Data rate	SI	TI	O
24 fps	H.264/AVC	228	602 kbps	62.0787	4.9166	2.4841

Fig. 2 Screenshot of the proposed video broadcasted over OWC-PON. *Source* "https://pixabay.com/users/zameenaasman172-20901772/?utm_source=linkattribution&utm_medium=referral&utm_campaign=image&utm_content=72414"

However since video is obviously temporal pooling of video frames at a given frame rate, this temporal change in spatial objects is specified by temporal information. It is calculated as follows [18]:

$$\mathrm{TI} = \max\left\{\mathrm{rms}_{\mathrm{space}}\left[(F(t_n) - F(t_{n-1})\right]\right\} \tag{2}$$

where space is the horizontal and vertical dimensions of image, t_n is the nth time frame at frame rate, t_{n-1} is the previous frame, rms is the root mean square function, max is the maximum function. For this video, TI is listed in Table 1.

Scene complexity (o) is nothing but product of spatial and temporal information [19].

$$o = \log_{10}(\mathrm{SI} \times \mathrm{TI}) \tag{3}$$

The frame-wise SI, TI and scene complexities are shown in Figs. 3, 4 and 5, respectively.

OLT in Fig. 1 houses video pre-processing block where H.264/AVC coded bit stream is converted to suitable NRZ format to be transmitted via optical carrier. This NRZ wave then modulates a continuous wave laser at 650 nm. The output light of power P_T is gathered by a transmit aperture of diameter D_T and launched into atmosphere. While in atmosphere, the light is impaired by scintillation caused due to atmospheric turbulence. Due to uneven heating of earth surface, the air above it gets differentially heated causing random changes in refractive index of air. These are called eddies. With the combined effect of wind and convection of air, these eddies are

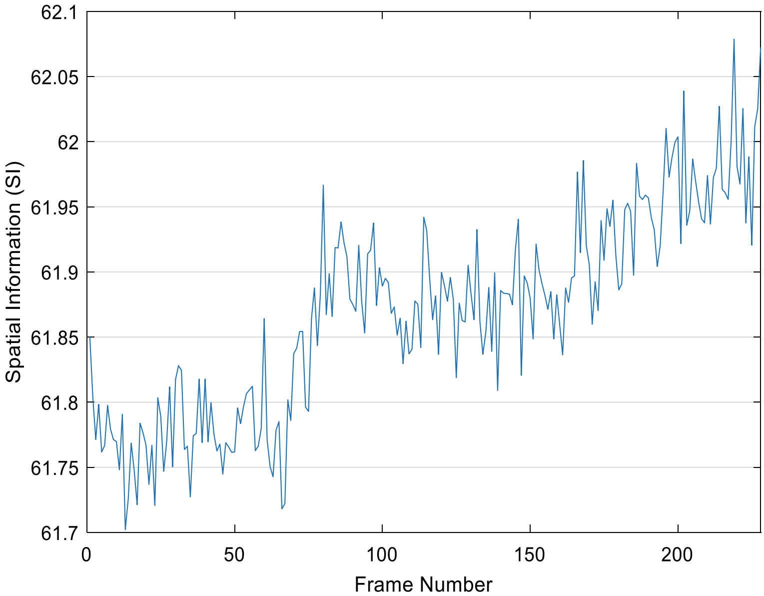


Fig. 3 Spatial information (SI, Eq. 1) versus frame number for transmitted video

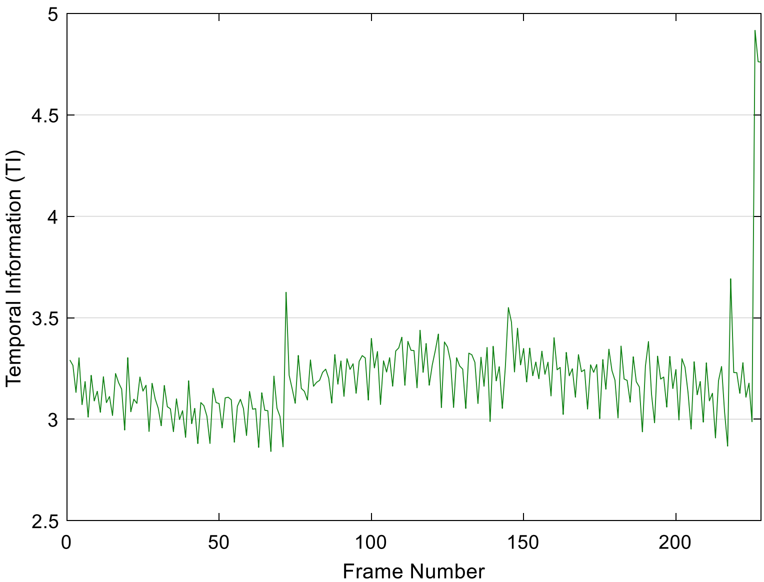


Fig. 4 Temporal information (TI, Eq. 2) versus frame number for transmitted video

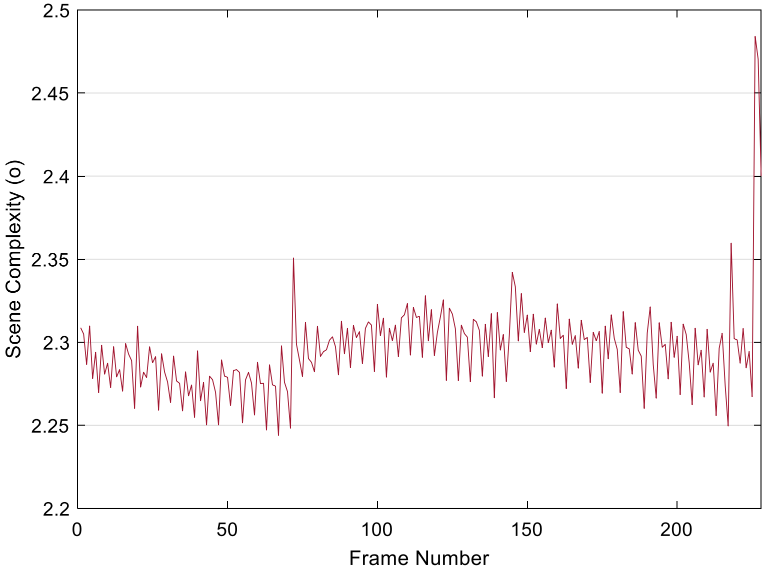


Fig. 5 Scene complexity (o, Eq. 3) versus frame number for transmitted video

further broken down randomly. When light encounters these eddies, it experiences a randomly changing refractive index, due to which it fluctuates at the receiver aperture. This effect is scintillation. A turbulent environment can be characterized by a refractive index structure parameter C_n^2 given as follows [20, 21]:

$$C_n^2 = \left[79 \times 10^{-6} \frac{P}{T^2}\right] C_T^2 \tag{4}$$

where P and T are the pressure (mbar) and temperature (Kelvin) and C_T^2 [21] depends upon the temperature gradient. Another important parameter for a turbulent channel encompassing the channel length, refractive index structure parameter and wavelength of operation is Rytov variance σ_R^2 [22].

$$\sigma_R^2 = 1.23 k^{\frac{7}{6}} C_n^2 L^{\frac{11}{6}} \tag{5}$$

where $k = 2\pi/\lambda$.

For weak turbulence, irradiance function for effect of scintillation is taken as lognormal function [22].

$$f(I) = \frac{1}{\sqrt{2\pi\sigma_R^2 I^2}} \exp\left\{-\frac{[\ln[I] - \mu]^2}{2\sigma_R^2}\right\} \tag{6}$$

where I is received intensity μ is the mean.

This distorted light is then received by a receiver aperture of diameter D_R. The power of received light is P_R. The light falls on photodetector where it is converted into current, and certain types of noises such as shot noise, dark current noise, relative intensity noise, thermal noise are added to the signal. The received signal would have the form:

$$y = \Re P_R I + n \tag{7}$$

where $\Re$ is responsivity of photodetector, I is irradiance, n is sum total of all the noises.

A matched filter and a decision circuit then output the bit stream which is compared against the input bit stream for BER evaluation. Table 2 lists all the parameters used for the system design.

Evaluation of VQA parameters for QoE aspect involves subjective and objective evaluators. The subjective evaluators are based upon classifiers, and mean opinion score is gathered from target viewership in a controlled and sophisticated environment. However, this evaluation is not suitable in real-time environment. The objective evaluation can be classified into no reference versus full reference metrics. In the case of no reference metrics, no ground truth is available for comparison, and the quality is solely based on attributes of received video. In the case of full reference quality metrics, the transmitted video serves as ground truth and comparison is made between ground truth and received video. These metrics have a full reference to back themselves. The full reference VQA metrics used in this paper are as follows:

Table 2 Design parameters

S. No.	Specification	Value
1	Single-channel transmit power	10 mW
2	Transmit aperture diameter	2 mm
3	Receiver aperture diameter	180 mm
4	Wavelength	650 nm
5	Line format	NRZ
6	Modulation	OOK
7	Data rate	2.5 Gbps
8	Sampling rate	50×2.5Gbps
9	Responsivity	0.8 A/W
10	Sensitivity	-40 dBm
11	Absolute temperature	288 K
12	Load resistance	50Ω
13	Dark current	6 nA
14	RIN	-130 dB/Hz
15	C_n^2	8×10^{-16} m$^{-2/3}$

1. **Mean Square Error (MSE)**: For a video sequence having F frames, MSE [23] is defined as follows:

$$e_{\mathrm{MSE}} = \frac{1}{F}\sum_{n=1}^{F}\left\{\frac{1}{MN}\sum_{x=1}^{M}\sum_{y=1}^{N}\left(\widehat{F_n}(x, y) - F_n(x, y)\right)^2\right\} \tag{8}$$

where $\widehat{F_n}(x, y)$ is the distorted nth frame and $F_n(x, y)$ is the reference nth frame. M and N are dimensions of the frame.

2. **Peak signal-to-noise ratio**: PSNR [23] is defined as follows for an 8-bit representation

$$\mathrm{PSNR} = 10\log_{10}\left(\frac{255^2}{e_{\mathrm{MSE}}}\right) \tag{9}$$

These two VQA metrics are good for error visibility but not for the similarity of key attributes of frames such as luminance, contrast and structure which are sharply perceived by human visual system. Hence, a Structure Similarity Index (SSIM) is the metric which encompasses all the above information and gives a VQA metric based on human visual system. It is defined as follows:

3. **Structure Similarity Index (SSIM)**: SSIM is defined as follows [24]:

$$\mathrm{SSIM}\left(I, \hat{I}\right) = \frac{1}{F}\sum_{n=1}^{F}\left\{\left(\frac{2\mu_{xyn}\hat{\mu}_{xyn} + C1}{\mu_{xyn} + \hat{\mu}_{xyn} + C2}\right)\left(\frac{2\gamma_{xyn} + C2}{\sigma_{xyn} + \hat{\sigma}_{xyn} + C2}\right)\right\} \tag{10}$$

where σ_{xyn} is the standard deviation of reference nth frame, $\hat{\sigma}_{xyn}$ is the standard deviation of distorted nth frame, and γ_{xyn} is the covariance between reference and distorted nth frame. $C1 = 0.09 \times 255^2$ and $C2 = 0.81 \times 255^2$ are the weights based on HVS system.

A fourth VQA metric is association-based metric which is Pearson's coefficient of correlation. It represents the possibility of a linear relationship between the transmitted and received images. It is defined as follows:

$$r = \frac{1}{F}\sum_{n=1}^{F}\left\{\frac{\sum_{x=1}^{M}\sum_{y=1}^{N}\left(\widehat{F_n}(x, y) - \hat{\mu}_{xyn}\right)\left(F_n(x, y) - \mu_{xyn}\right)}{\sqrt{\left(\sum_{x=1}^{M}\sum_{y=1}^{N}\left(\widehat{F_n}(x, y) - \hat{\mu}_{xyn}\right)^2\right)\left(\sum_{x=1}^{M}\sum_{y=1}^{N}\left(F_n(x, y) - \mu_{xyn}\right)^2\right)}}\right\} \tag{11}$$

Using above defined parameters and equations, the VQA metrics and BER are evaluated by Monte Carlo approach. A total of 10,000 iterations are used, and the results averaged upon these iterations are presented.

Results

Figure 6 shows the variation in Rytov variance versus channel length. As channel length increases, Rytov variance increases because more and more turbulence gets incorporated into channel. For this Rytov variance, the variation in BER is presented in Fig. 7. As Rytov variance increases so does BER because light has to encounter more severe effect of turbulence at given channel length. The variance of irradiance function increases suggesting more deviation from mean value of intensity meaning more fluctuation. Effect of increasing Rytov variance on MSE is shown in Fig. 8. As Rytov variance increases, MSE increases suggesting more distortion in intensity map of image. Note that at higher values of Rytov variance, the values of MSE shoot up sharply while they do so gradually at lower values. As turbulence gets severe, rate of distortion of video increases.

The variation in PSNR versus Rytov variance is shown in Fig. 9. With increasing Rytov variance, the PSNR falls. However, rate of this decline is sharp at lower values of Rytov variance and gradual at higher ones. This is as expected since MSE and PSNR have reciprocal relationship.

Figure 10 depicts variation in SSIM for Rytov variance. The SSIM decreases in a nonlinear fashion with rate of decline very low at early values of Rytov variance and then a gradual decline at later ones. To assist with the visual distortion represented by decline in SSIM values, Fig. 11 depicts the scenario aptly (Fig. 12).

Figure 10 depicts the variation in Pearson's correlation coefficient versus Rytov variance. The coefficient decreases gradually for lower values of Rytov variance

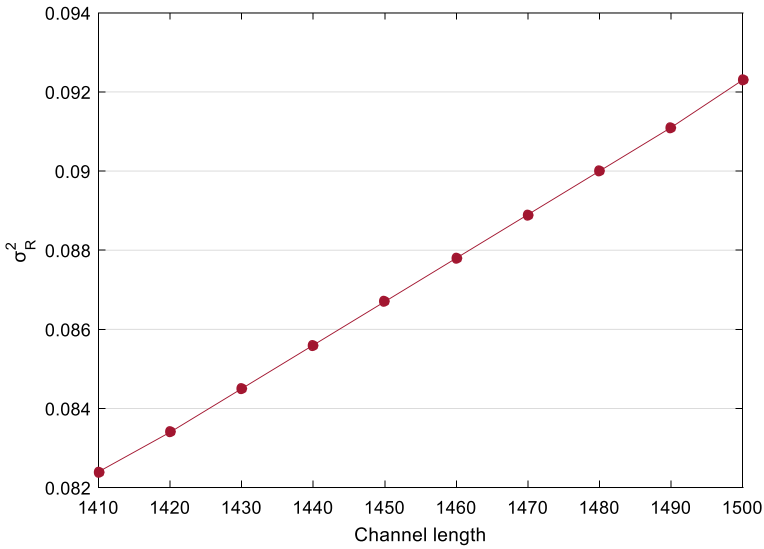


Fig. 6 Rytov variance (σ_R^2) Eqs. 4 and 5 versus channel length for weak turbulence where $C_n^2 = 8 \times 10^{-16}$ m$^{-2/3}$

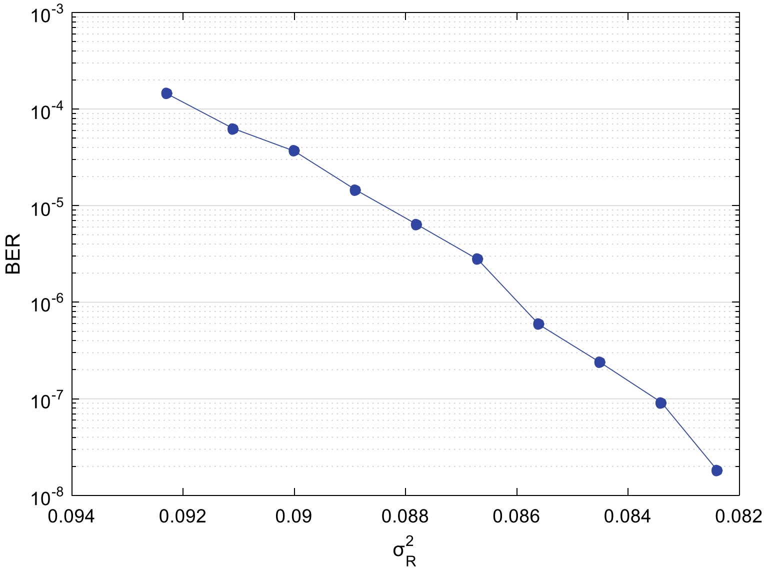


Fig. 7 BER versus Rytov variance for weal turbulence video transmission

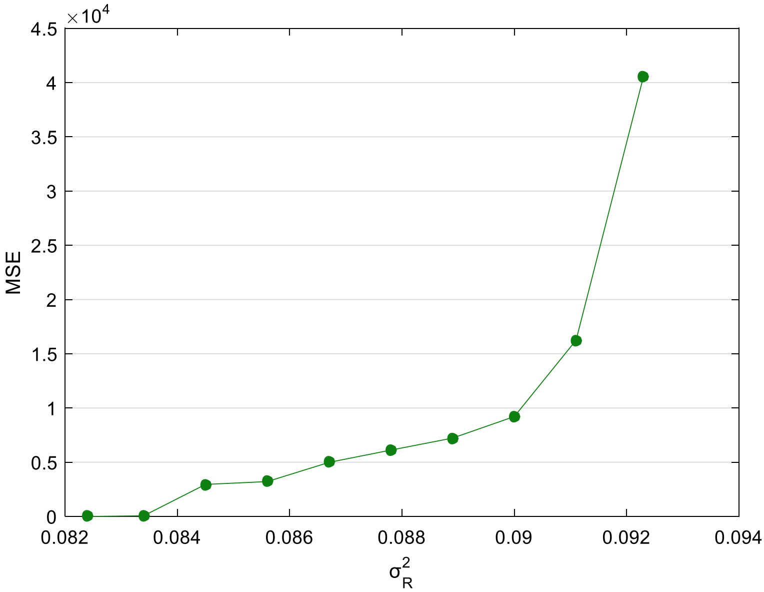


Fig. 8 MSE (e_{MSE} Eq. 8) versus Rytov variance for weal turbulence video transmission

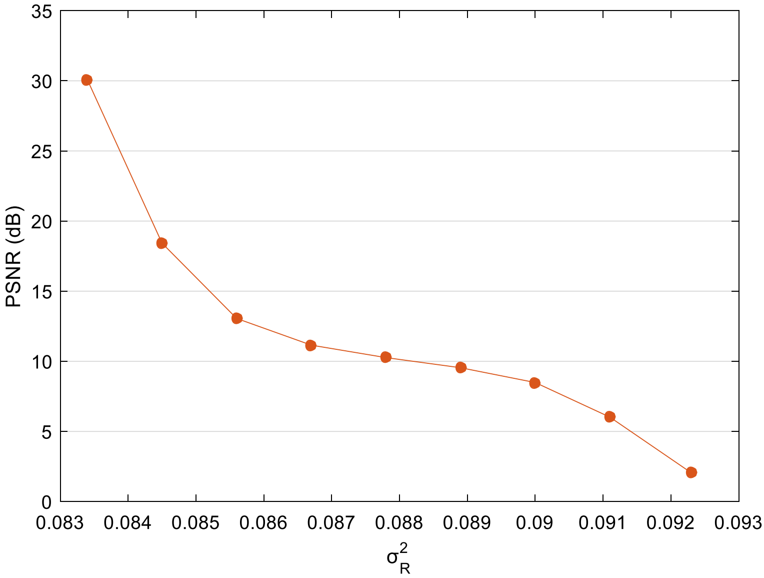


Fig. 9 PSNR (Eq. 9) versus Rytov variance for weal turbulence video transmission

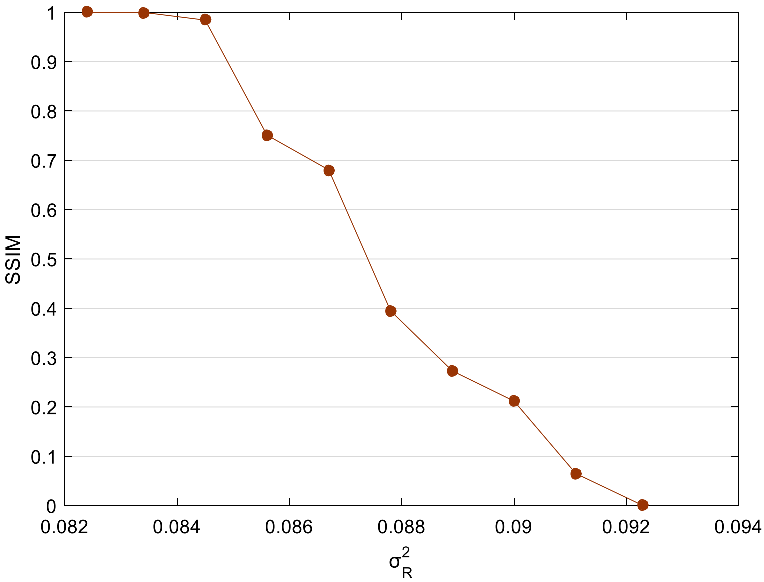


Fig. 10 SSIM (Eq. 10) versus Rytov variance for weal turbulence video transmission

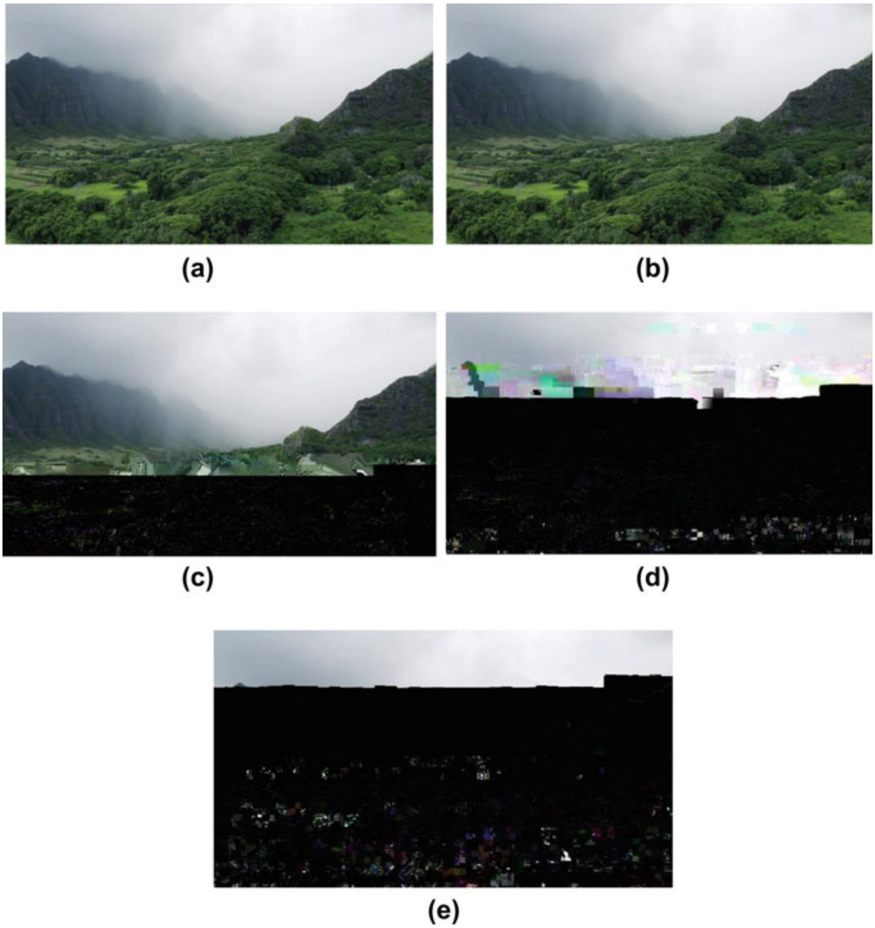


Fig. 11 Received video frames for **a** 1420 m, **b** 1440 m, **c** 1460 m, **d** 1480 m, **e** 1500 m

and sharply for higher values. This indicates that for lower values, the possibility of a linear relationship between transmitted and received video falls gradually and at higher values this possibility falls sharply.

Conclusions

In this paper, we have evaluated various channel performance metrics under both the realms of QoS and QoE. The corresponding variations in all the metrics w.r.t the channel parameter have been presented and discussed in detail. A video transmission over OWC-PON network is successfully transmitted under weak turbulence regime

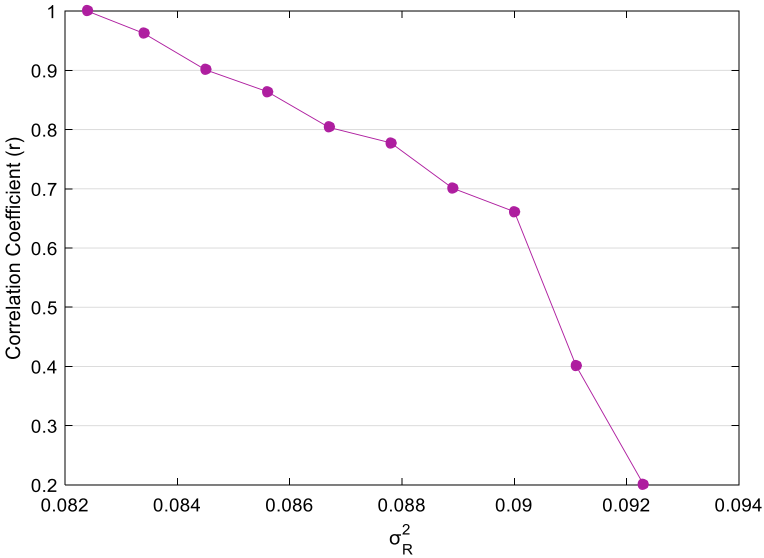


Fig. 12 Correlation coefficient (r, Eq. 11) versus Rytov variance for weal turbulence video transmission

up to a distance 1500 m of limiting case BER. This is the viable length as decided by QoS performance parameters. However in case of QoE parameters such as SSIM, the limiting case for 75% similarity is 1445 m. This is the viable channel length for QoE performance parameter. So in the case of weak turbulence due to QoE visual quality assessment, an additional loss of 55 m is incurred in the system. This loss would be hidden had only been QoS performance evaluated. Hence, both QoS and QoE metric evaluation gives a holistic performance description of multimedia data transmission over OWC-PON networks.

References

1. Gill HS, Singh ML (2021) Performance evaluation of DVB-t image transmission over a MIMO OWC channel at 650 nm under varying turbulence regimes. Wireless Netw 27:1965–1979. https://doi.org/10.1007/s11276-021-02559-5
2. Hasanov MH, Ibrahimov BG, Mardanov NT (2019) Research and analysis performance indicators NGN/IMS networks in the transmission multimedia traffic. In: Wave electronics and its application in information and telecommunication systems (WECONF), Saint-Petersburg, Russia, pp 1–4. https://doi.org/10.1109/WECONF.2019.8840117
3. Giuliano R, Mazzenga F, Vizzarri A (2020) Integration of broadcaster and telco access networks for real time/live events. IEEE Trans Broadcast 66(3):667–675. https://doi.org/10.1109/TBC.2020.2965057
4. Pandian, Pasumpon A (2020) Development of secure cloud based storage using the elgamal hyper elliptic curve cryptography with fuzzy logic based integer selection. J Soft Comput Paradigm 2(1):24–35. https://doi.org/10.36548/jscp.2020.1.003

5. Hong W et al (2021) The Role of millimeter-wave technologies in 5G/6G Wireless communications. IEEE J Microwav 1(1):101–122, winter 2021. https://doi.org/10.1109/JMW.2020.3035541
6. Khiadani N (2020) Vision, requirements and challenges of sixth generation (6G) networks. In: 2020 6th Iranian conference on signal processing and intelligent systems (ICSPIS), pp 1–4, https://doi.org/10.1109/ICSPIS51611.2020.9349580
7. Jiang W, Han B, Habibi MA, Schotten HD (2021) The road towards 6G: a comprehensive survey. IEEE Open J Commun Soc 2:334–366. https://doi.org/10.1109/OJCOMS.2021.3057679
8. Cisco Annual Internet Report (2018–2023) White paper. [online] Available: https://www.cisco.com/c/en/us/solutions/collateral/executive-perspectives/annual-internet-report/white-paper-c11-741490.html. Retrieved on 22 June 2021
9. Kaushal H, Georges K (2017) Optical communication in space: Challenges and mitigation techniques. IEEE Commun Surv Tutor. https://doi.org/10.1109/COMST.2016.2603518
10. Kiycioğlu D, Ünverdi NÖ (2018) Applications of FTTX technology in optical communication systems. In: 2018 26th Signal processing and communications applications conference (SIU), pp 1–4, https://doi.org/10.1109/SIU.2018.8404442
11. Jaffer SS, Hussain A, Qureshi MA et al (2020) Towards the shifting of 5G front haul traffic on passive optical network. Wireless Pers Commun 112:1549–1568. https://doi.org/10.1007/s11277-020-07115-6
12. Alimi I, Shahpari A, Ribeiro V, Kumar N, Monteiro P, Teixeira A (2016) Optical wireless communication for future broadband access networks. In: 2016 21st European conference on networks and optical communications (NOC). https://doi.org/10.1109/NOC.2016.7506998
13. Majumdar AK (2019) Basics of worldwide broadband wireless access independent of terrestrial limitations. In: Optical wireless communications for broadband global internet connectivity. Elsevier Inc., pp 5–38. https://doi.org/10.1016/B978-0-12-813365-1.00002-3
14. Varela M, Skorin-Kapov L, Ebrahimi T (2014) Quality of service versus quality of experience. In: Möller S, Raake A (eds) Quality of experience. T-labs series in telecommunication services. Springer, Cham. https://doi.org/10.1007/978-3-319-02681-7_6
15. Song W (2020) Quality of experience. In: Shen X, Lin X, Zhang K (eds) Encyclopedia of wireless networks. Springer, Cham. https://doi.org/10.1007/978-3-319-78262-1
16. International Telecommunication Union (2008) Subjective video quality assessment methods for multimedia applications. ITU-T Rec. P.910, Apr 2008
17. Smys S, Wang H (2021) Data elimination on repetition using a blockchain based cyber threat intelligence. IRO J Sustain Wirel Syst 2(4):149–154. https://doi.org/10.36548/jsws.2020.4.002
18. Barman N, Khan N, Martini MG (2019) Analysis of spatial and temporal information variation for 10-bit and 8-bit video sequences. In: 2019 IEEE 24th international workshop on computer aided modeling and design of communication links and networks (CAMAD), pp 1–6. https://doi.org/10.1109/CAMAD.2019.8858486
19. Fenimore C, Libert J, Wolf S (2000) Perceptual effects of noise in digital video compression. SMPTE J 109(3):178–187. https://doi.org/10.5594/j08345
20. Li M, Cvijetic M (2015) Coherent free space optics communications over the maritime atmosphere with use of adaptive optics for beam wavefront correction. Appl Opt 54(6):1453–1462. https://doi.org/10.1364/AO.54.001453
21. Priyanka, Singh ML, Gill HS, Singh M, Kaur S (2020) An experimental evaluation of link outage due to beam wander in a turbulent FSO link. Wirel Pers Commun. https://doi.org/10.1007/s11277-020-07333-y
22. Majumdar AK, Luna CE, Idell PS (2007) Reconstruction of probability density function of intensity fluctuations relevant to free space laser communications through atmospheric turbulence. In: Proceedings of SPIE 6709, free-space laser communications VII; 67090M. https://doi.org/10.1117/12.728699
23. Huynh-Thu Q, Ghanbari M (2010) The accuracy of PSNR in predicting video quality for different video scenes and frame rates. Telecommun Syst 49(1):35–48. https://doi.org/10.1007/s11235-010-9351-x

24. Wang Z, Bovik AC, Sheikh HR, Simoncelli EP (2004) Image quality assessment: from error visibility to structural similarity. IEEE Trans Image Process 13(4):600–612. https://doi.org/10.1109/TIP.2003.819861

Analysis of Signal Noise Reduction Techniques

Adwait Kaundanya, Anubhav Anand, Kamal Raisinghani, and Reena Sonkusare

Abstract The basis of modern communication is elimination of unwanted irregular fluctuations that accompany a transmitted signal but are not part of it, also called noise. The noise tends to obscure important or useful information. Thus, the reduction of signal corruption caused due to additive white Gaussian noise (AWGN) becomes very important in order to establish best transmission of signals. Hence, it is essential to reduce this noise and also recognize the best techniques. This paper is a comparative analysis between three popular noise reduction techniques that are used to filter signals for any distortions that may occur. The analysis of the filter is performed using MATLAB which has the ability to easily simulate filters using the in-build functions. The paper lists three different methods to eliminate the effects of noise in signals. The three methods employed make use of the discrete Fourier transformation (DFT) technique, Gaussian filter method, and least mean square (LMS) algorithm. The final results are drawn based on the correlation coefficient as well as the bit error rate of the original and the filtered signal, and final analysis is drawn.

Keywords Noise · Gaussian · Fourier transform · Least mean square · BER

A. Kaundanya (✉) · A. Anand · K. Raisinghani · R. Sonkusare
Sardar Patel Institute of Technology, Munshi Nagar, Andheri West, Mumbai, India
e-mail: adwait.kaundanya@spit.ac.in

A. Anand
e-mail: anubhav.anand@spit.ac.in

K. Raisinghani
e-mail: kamal.raisinghani@spit.ac.in

R. Sonkusare
e-mail: reena_kumbhare@spit.ac.in

V. Bindhu et al. (eds.), *Proceedings of Third International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 844,
https://doi.org/10.1007/978-981-16-8862-1_28

Introduction

One of the issues with digital communication and signal processing is to eliminate the loss of data. As a result, it becomes very important to reduce signal corruption caused due to additive white Gaussian noise (AWGN) in order to transmit the signal effectively. Noise is introduced in a signal due to the transmission channel as well as due to the different filters that the signal passes through. It is difficult to transmit analog signals over large distances, and hence, they are converted into digital signals by sampling it using Nyquist principle. The signal is then quantized and encoded to produce the final signal. Now, when this signal travels through a channel, it faces various disturbances which causes the signal to get distorted due to AWGN. The corrupted signal that we get is basically a combination of the original signal and the noise which degrades the quality of the signal that will be received. The main motive of this paper is to design and analyze some techniques that can be used to filter out or eliminate this noise and get back the original signal in the best quality possible. Each time, the reason of occurrence of noise will not be the same, and hence, one technique of noise reduction may not be suitable for all scenarios. In this paper, we propose three filtering techniques that revolve around different principles to reduce noise and shape the waveform to as to replicate the original noiseless signal. In the literature survey, we have talked about various methods that have been employed to filter signals of different types like ECG, EEG, etc. In our paper, the Gaussian filter method works on the principle of filtering out noise as it is of high frequency. DFT method analyzes the signal in frequency domain and reduces the noise accordingly. The LMS method is a closed loop system which uses concepts like moving averages and autoregression to compare the noise with the current output to achieve the final noiseless signal. The paper in the end also gives a comparative analysis as to which of the three methods worked best in our case using performance parameters like bit error rate and correlation coefficient.

Related Works

Various research papers have been published to design new techniques which help to improve noise reduction in various types of signals. The following literature review is an attempt to study nearly a dozen papers which lists the general algorithm used, their efficiency analysis as well as the final result obtained. Flandrin et al. [1] research about empirical mode decomposition (EMD) as this can be used to represent non-stationary signals adaptively. The paper understands the fact that EMD acts as a filter bank and reports on a number of examples that have additive white Gaussian noise. Boudraa et al. [2] propose a data-driven method to remove noise using EMD. The breaking down of signal into IMFs and then reconstructing them to get noise reduced signal is discussed. Here, the authors experimented on a real signal and two synthetic signals to show that the method used can remove noise effectively. Boudraa et al. [3] have worked on a new signal processing technique to remove noise using the EMD method. The process is simple, fully data-driven, and very effective. The EMD-Soft outperforms other methods of noise removal like median methods, while EMD-SG surpasses wavelet method. The threshold parameters are under focus here because they separate noise from required data signals. A process called sifting is used to break noisy signals into oscillatory components IMFs. These IMFs are later on filtered and thresholded and reconstructed to get the approximate data signal. A new fully data-driven speech noise reduction method is discussed by Khaldi et al. [4]. The method is based on empirical mode decomposition (EMD). Noisy signal is broken down into swaying components called intrinsic mode functions which are then used to reconstruct the signal using shrinkage function. This method turned out to be much better than the wavelet shrinkage method. Kopparapu and Satish [5] state that if the noise statistics are known, then Gaussian filters can be used effectively for signal recovery. More specifically to remove the additive white Gaussian noise (AWGN), the best practice would be to use the Gaussian filter. The characteristics of the filter depends on the relation between the signal noise which determines how accurate the signal recovery will be. The authors have hypothesized certain aspects and concluded that if Gaussian noise has to be removed then Gaussian filter with the assumption that signal-to-noise ratio is known. Singh and Garg [6] describe various types of noises and the patterns that the noises can have with the original signal, the noise can be additive, non-additive. These noises are generally associated with audio signals, these can be canceled or reduced by interference of anti-phase signal which is commonly being used in noise cancelling headphones, it is mainly achieved using digital signal processing techniques. Afroz et al. [7] analyze the NLMS algorithm. Adaptive noise cancellation technique performance parameters, such as number of filter coefficients, number of samples, step size, and input noise level are changed, and its effect on the output signal are compared, and the best fit is determined using graphs and signal-to-noise ratios. Haritha et al. [8] have compared the various methods that have emerged in recent years to remove noise from ECG signals. Removal of noise from ECG is a tough job, but high-quality ECG signal is important in analysis and diagnosis of all heart diseases. Adaptive noise cancellation techniques, wavelet transform denoising methods, empirical mode decomposition (EMD), FIR and IIR filtering, non-local means denoising techniques (NLM), quadrature filtering, variational mode decomposition (VMD), low and high frequency noise removal techniques are the methods discussed in detail. Kumar et al. [9] have used an LMS algorithm for adaptive noise cancellation on audio signals having Gaussian noise. A block was created on GNU radio that operates like an LMS filter. Due to sampling, the noise cancellation signal was applied at different time durations, and the results were compared to the original signal having minimum deviation from the original signal.

Proposed Technique

Many researchers have designed different types of filtering techniques for reduction of noise from a signal. To begin with the comparative analysis, first we generate a sinusoidal signal to represent the information signal. When the signal is received at the endpoint, it contains some amount of noise along with the original signal. Extracting the original signal containing the information is very important and studying which technique is better suited will be our focus. The paper studies three different methods, i.e., Gaussian filter method, DFT technique, and LMS filter. The Gaussian pass filter is a low pass filter and thus eliminates high frequency noise. The DFT makes use of the fact that the amplitude of noise in frequency domain will be low and thus removes the frequencies having less amplitude. LMS filter uses the technique of adaptive filtering. The following are the features that will be varied to perform the comparative analysis-

A. *Standard deviation or sigma for Gaussian filter technique*—The role of sigma is to manage the variation of the impulse response about the mean value and thus determines the frequency that is filtered.
B. *Threshold amplitude for DFT technique*—The threshold amplitude determines the frequencies that will be filtered out, and hence, choosing the optimal threshold is very important.
C. *Order of filter in LMS technique*—The order and step size of an LMS filter are two important features. As step size is automatically determined by the software, we will be varying the order of the filter to get the error to converge.

Methodology

Generating the Noisy Signal

First, we generate a sinusoidal signal to represent the information being transmitted. Due to the noise during transmission, additive white Gaussian noise (AWGN) corrupts the signal to certain extent, and thus, we can obtain the noisy signal by making use of the random number function in order to simulate the effect of AWGN (Fig. 1).

Eliminating Noise Using Gaussian Filter

i. *Generating the Gaussian filter*

The first step involves simulating the impulse response of Gaussian or low pass filter. Gaussian filter is one whose impulse response is a Gaussian curve with the main properties being mean and standard deviation. Here, we will focus

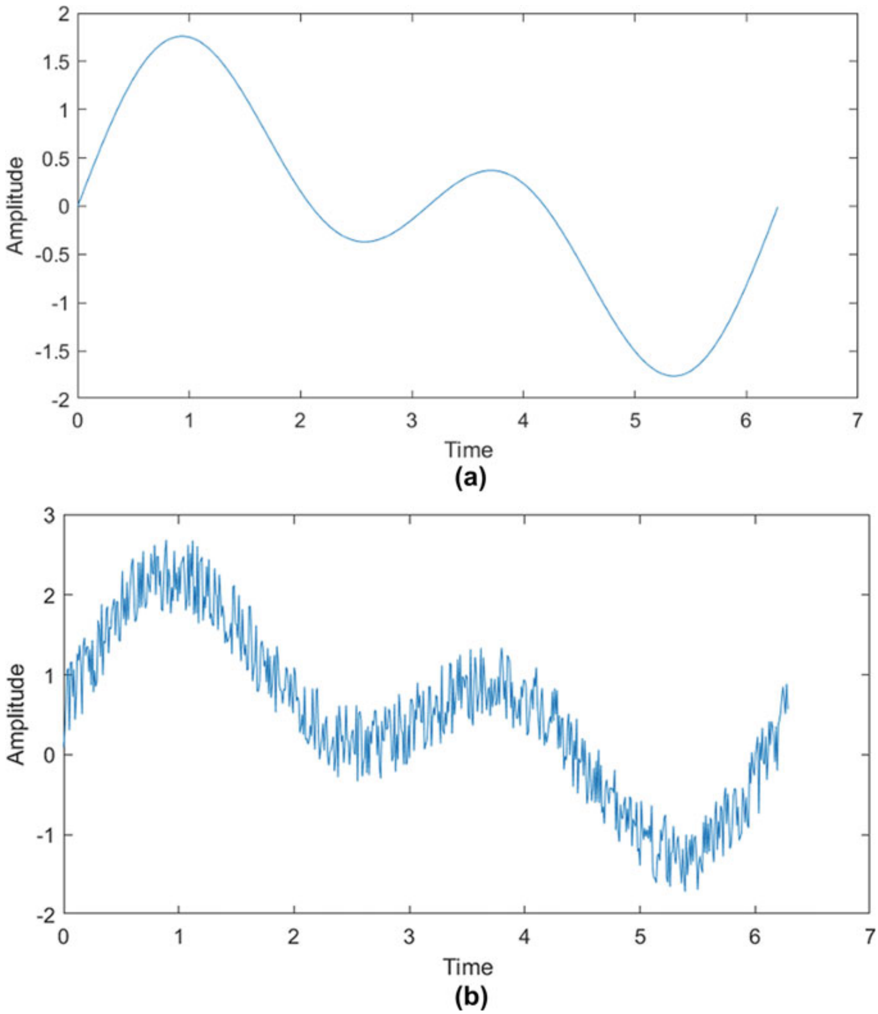


Fig. 1 **a** Original signal, **b** noise corrupted signal

on standard deviation as it decides the variation around the mean value and plays an important role in filtering of the signal. In MATLAB, the filter is generated by using the 'fspecial' function and specifying the value of standard deviation or sigma. The Gaussian filter allows signals with lower frequency to pass and attenuates high frequency signals (Fig. 2).

ii. *Convoluting the noisy signal with the Gaussian filter*

The output that will be obtained after passing the signal through the filter needs to be obtained. Convolution is the technique that gives us the output of the filter if the input signal and impulse response of the filter are both given.

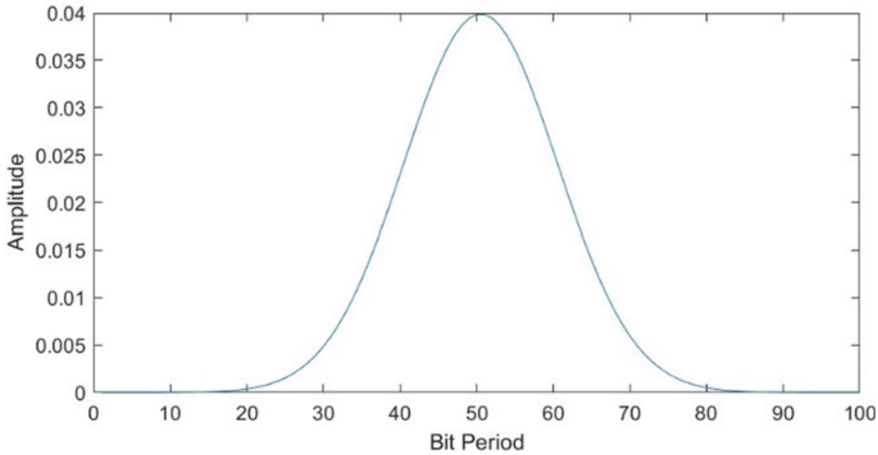


Fig. 2 Response of Gaussian filter

$$y(t) = \int_{-\infty}^{+\infty} x(\tau)h(t-\tau)\mathrm{d}\tau$$

where $y(t)$ = output of filter.

$x(t)$ = input to the filter.

$h(t)$ = impulse response of filter.

The output obtained will be a smooth curve that resembles the original signal. Here, we will be performing convolution by keeping the information signal same and changing the value of standard deviation (sigma) for the Gaussian filter to test for the most optimum performance.

Eliminating Noise Using DFT Technique

i. Obtaining the signal in frequency domain

Fourier transform is used to convert a signal from time domain to frequency domain. Below is the formula for finding the Fourier transform of any signal

$$F[\omega] = \int_{-\infty}^{+\infty} f(t)\mathrm{e}^{-j\omega t}\mathrm{d}t$$

where $F[\omega]$ is signal is frequency domain.

$f(t)$ is signal in time domain.

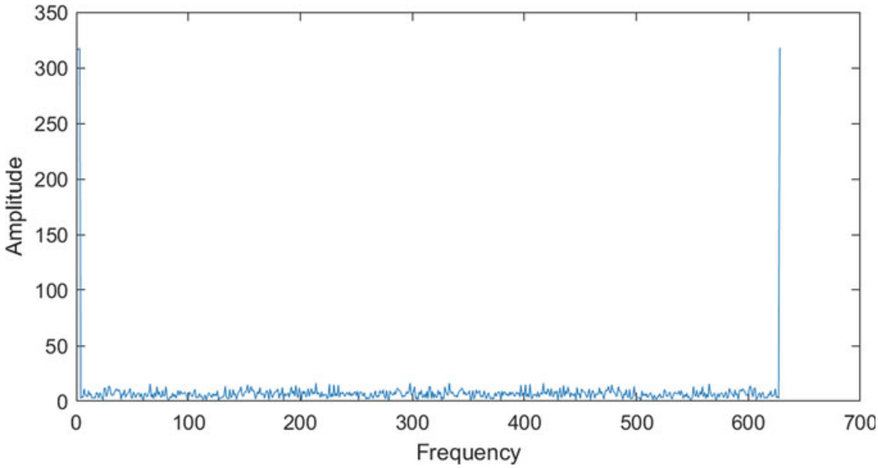


Fig. 3 FFT plot of the noisy signal

We take the discrete Fourier transform of the noisy signal to obtain the frequency spectrum to observe major frequencies that composes the information. In MATLAB, this is done by using the DFT function which automatically converts the noisy signal into its frequency domain counterpart (Fig. 3).

ii. Eliminating frequencies below a certain amplitude threshold

The frequencies that the original signal mainly consisted of, will have high amplitude, while the rest of the frequencies will have lower amplitudes. We decide on a threshold which will only allow the frequencies to exist, and all the others will be attenuated thus removing the noise. The output will again be a smooth curve resembling the original signal. Here, we change the threshold to observe the change in obtained waveform.

Eliminating Noise Using LMS Filter

i. Creating a reference signal

After generating the noise signal, a reference signal is generated that is correlated with the noise signal. This reference signal is used as the base for the removal of noise from the noisy signal using LMS filter.

ii. Constructing the adaptive filter

Least mean square algorithm is used to replicate adaptive filters and reduce the error in the output using the reference signal. The main parameters affecting the LMS filter are its order and step size. MATLAB provides functionality to find the most

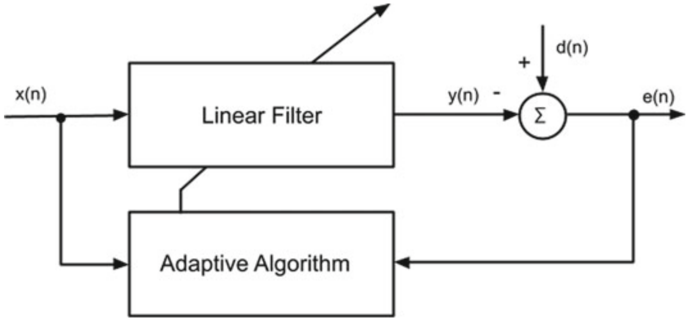


Fig. 4 Block diagram for filtering using LMS

optimum step size, and hence, we focus on the order of the filter. Even though the step size is optimally determined by the software, one can change it as per requirement. Smaller step size usually corresponds to better convergence however takes longer to adapt to the final output (Fig. 4).

Experimentation Result

The performance of the three techniques has been measured using the following performance parameters

i. *Correlation coefficient*—This parameter measures the degree of similarity between the two signals which is the original and the noise reduced signal. This is deduced by using the cross-correlation function in MATLAB.
ii. *Bit error rate*—This parameter gives the percentage of correct bits received after noise reduction as compared to the original signal. This is deduced by first converting the original signal and filtered signal into a digital signal and then comparing each bit. The percentage of correct bits gives us the result.

Gaussian Filter Technique

See Table 1 and Fig. 5.

DFT Technique

See Table 2 and Fig. 6.

Table 1 Observed values of correlation coefficient and BER for Gaussian filter

Sigma	Correlation coefficient	Bit error rate
5	0.7191	0.3084
10	0.7311	0.2899
13	0.7349	0.2893
17	0.7412	0.2862
20	0.7422	0.2909

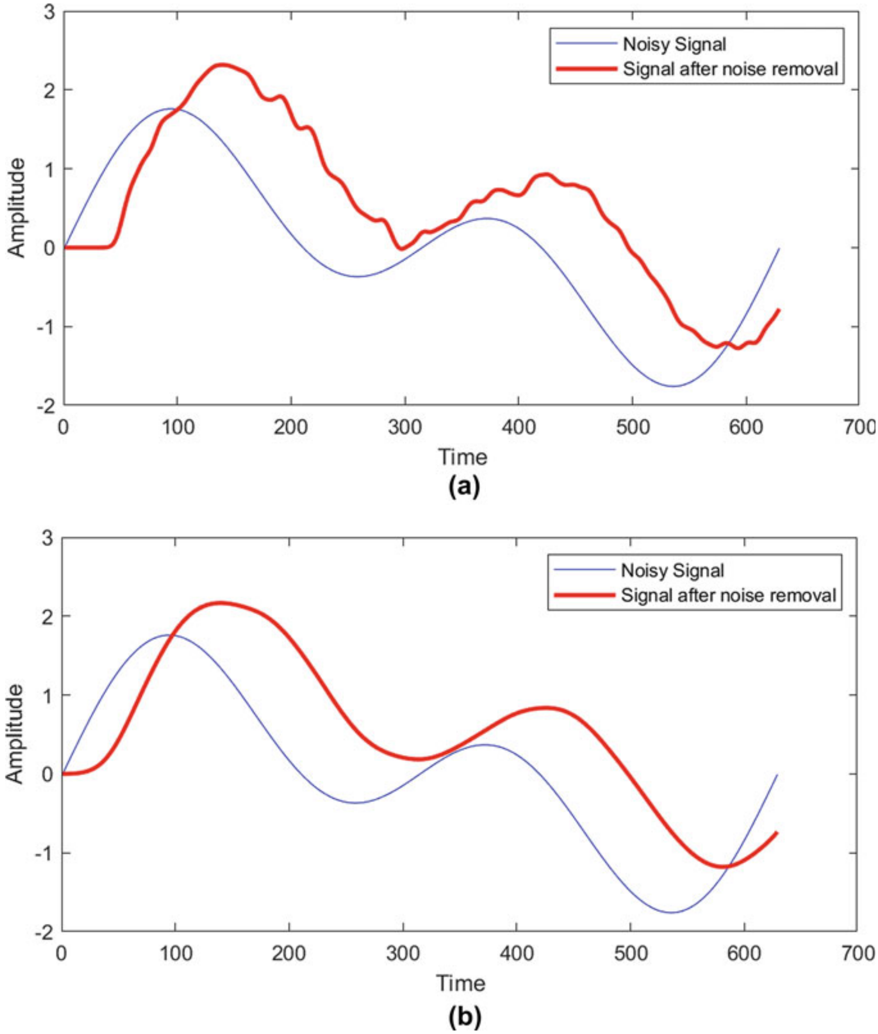


Fig. 5 **a** Original signal (blue) and noise-reduced signal (red) for sigma = 5. **b** Original signal (blue) and noise-reduced signal (red) for sigma = 17

Table 2 Observed values of correlation coefficient and BER for DFT technique

Amplitude threshold	Correlation coefficient	Bit error rate
5	0.9657	0.2067
10	0.9817	0.2114
15	0.9971	0.2401
20	0.9992	0.2369
25	0.9995	0.2417

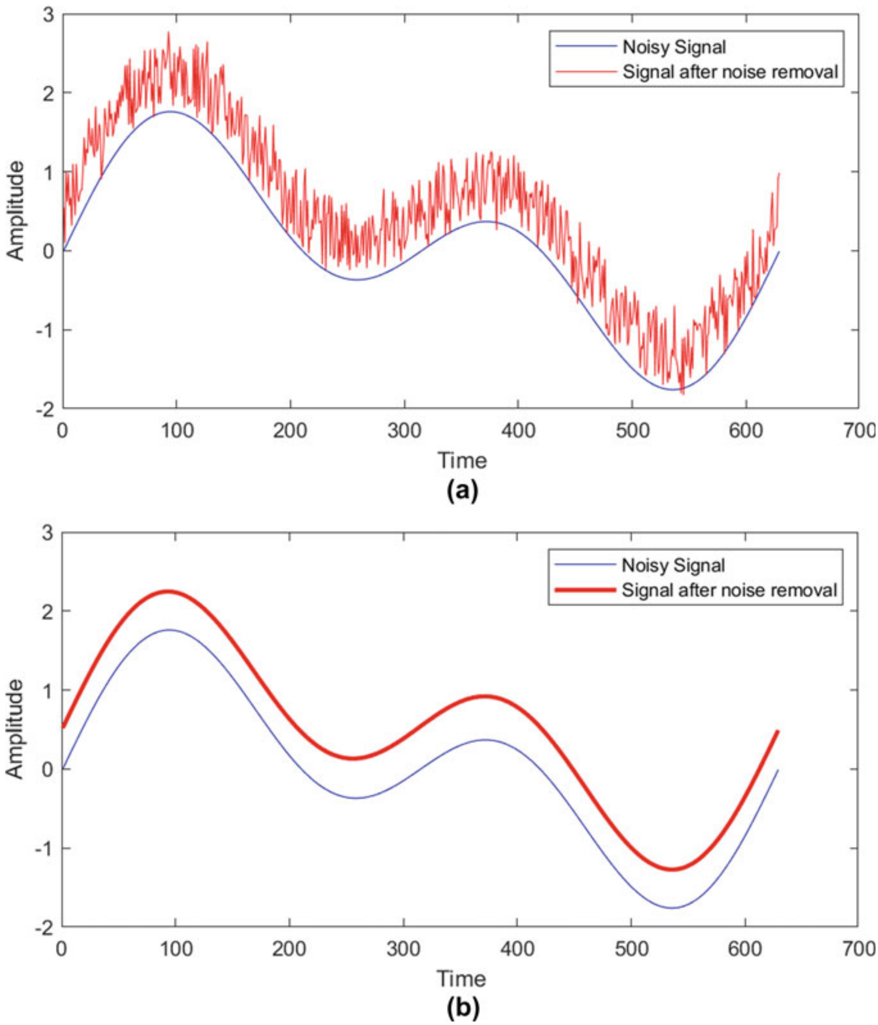


Fig. 6 **a** Original signal (blue) and noise-reduced signal (red) for threshold = 5. **b** Original signal (blue) and noise-reduced signal (red) for threshold = 5

Table 3 Observed values of correlation coefficient and BER for LMS filter

Order of filter	Correlation coefficient	Bit error rate
2	0.9875	0.0986
3	0.9833	0.1145
5	0.9763	0.1288
7	0.9737	0.1463
9	0.9714	0.1304

LMS Filter Technique

See Table 3 and Fig. 7.

Result Analysis

Gaussian Filter Technique

From Table 1, we observe that the correlation coefficient is between 0.70 and 0.75 and increases slightly with sigma while bit error rate is around 0.3 for all values of sigma.

Also Fig. 5 shows that there is a difference in amplitude as well as the phase of the output and hence the low values of correlation coefficient and high values of bit error rate. The filtered output becomes smoother with increase in Sigma.

DFT Technique

From Table 2, we can see that the correlation coefficient is above 0.95 for all values of threshold, while there is no visible pattern for bit error rate with threshold value.

The high value of bit error rate can be attributed the difference in amplitude which decreases the BER especially near the zero crossings as seen in Fig. 6.

LMS Technique

In Table 3, we can observe that the value of correlation coefficient as well as bit error rate is satisfactory for LMS technique and much better than both the other methods.

Figure 7 shows that the output waveform is very close to the input signal and hence gives better results. There is still a small amount of noise that can be observed for LMS technique. The filtered signal resembles the input signal when the order

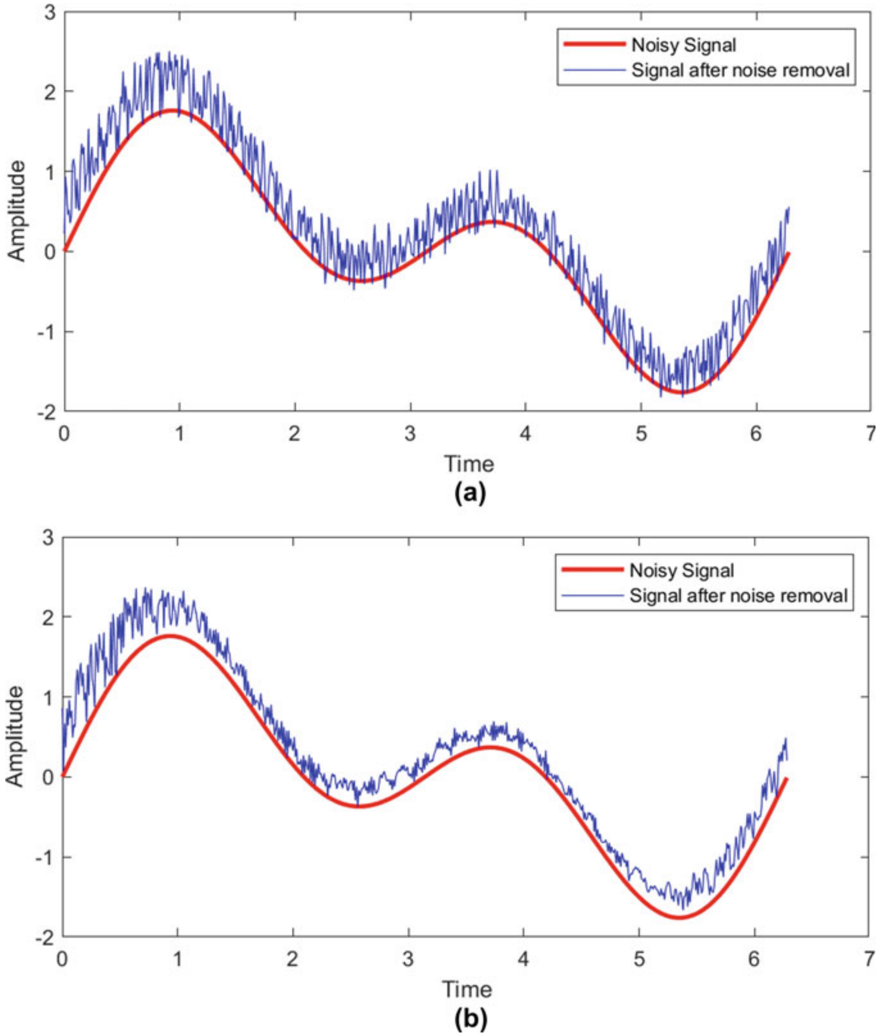


Fig. 7 **a** Original signal (blue) and noise-reduced signal (red) for filter order = 5. **b** Original signal (blue) and noise-reduced signal (red) for filter order = 2

is lower as the error converges and filters the noise more efficiently as compared to higher orders.

Conclusion and Future Work

This paper presents methods to design three types of filters and their comparison using parameters like correlation coefficient and bit error rate. By varying parameters like sigma in case of Gaussian filter method, the amplitude threshold in DFT method, and the order in LMS filter, we observe changes in the overall performance of the filter. As per the result analysis, both Gaussian and DFT technique provide high correlation but appear to not be in phase as well as differ in amplitude which provides poor bit error rate. Comparatively, the LMS filter which is based on adaptive filtering performs much better from the perspective of bit error rate and correlation coefficient proving to be more versatile as compared to the other two techniques.

Future work of this research can include testing various other type of signals using the three algorithms and conducting a similar comparative analysis. More such filtering algorithms can be studied in order to compare and reach a conclusion about the best filtering algorithms.

References

1. Flandrin P, Rilling G, Goncalves P (2004) Empirical mode decomposition as a filter bank. IEEE Signal Process Lett 11(2):112–114. https://doi.org/10.1109/LSP.2003.821662
2. Barsanti RJ, Gilmore J (2006) Denoising via empirical mode decomposition. In: Second international symposium on communications, control and signal processing
3. Boudraa A, Cexus J (2007) EMD-Based signal filtering. IEEE Trans Instrum Meas 56(6):2196–2202. https://doi.org/10.1109/TIM.2007.907967
4. Khaldi K, Boudraa A, Bouchikhi A, Turki-Hadj Alouane M, Diop ES (2008) Speech signal noise reduction by EMD. In: 2008 3rd International symposium on communications, control and signal processing, Saint Julian's, Malta, pp 1155–1158. https://doi.org/10.1109/ISCCSP.2008.4537399
5. Kopparapu SK, Satish M (2011) Identifying optimal Gaussian Filter for Gaussian noise removal. In: 2011 Third national conference on computer vision, pattern recognition, image processing and graphics, Hubli, India, pp 126–129. https://doi.org/10.1109/NCVPRIPG.2011.34
6. Singh M, Garg EN (2014) Audio noise reduction using Butter Worth Filter. Int J Comput Organ Trends 6:20–23. https://doi.org/10.14445/22492593/IJCOT-V6P305
7. Afroz F, Huq A, Sandrasegaran K (2015) Performance analysis of adaptive noise canceller employing NLMS algorithm. Int J Wirel Mob Netw 7:45–58. https://doi.org/10.5121/ijwmn.2015.7204
8. Haritha C, Ganesan M, Sumesh EP (2016) A survey on modern trends in ECG noise removal techniques. In: 2016 International conference on circuit, power and computing technologies (ICCPCT), Nagercoil, India, pp 1–7. https://doi.org/10.1109/ICCPCT.2016.7530192
9. Kumar KR, Anand R, Gandhiraj R (2017) Implementation of noise cancellation using least mean square in GNU radio. In: 2017 International conference on computer communication and informatics (ICCCI), Coimbatore, India, pp 1–5. https://doi.org/10.1109/ICCCI.2017.8117786

Application of Quantum Algorithms for Network Protocols

Vinutna Kolachana, Dolly Upmandewan, Arpit Giri, N. Pavan, Anees Ahmed, M. N. Thippeswamy, and T. R. Vinay

Abstract Cryptographic processes hold immense power in securing data transmissions across the Internet in the modern age. Preparing for future hacking possibilities is deemed essential and crucial to combat data breaches and leaks. With the advent of quantum computing, a field capable of performing complex operations within a short time, breaking into the cryptographic system keys with sheer brute force is visibly possible. This paper aims to create a framework that runs post-quantum cryptographic algorithms shortlisted by NIST for Round 3 on the TLS protocol. The framework is built using Python with the help of LibOQS packages, programmed to work as an API, and invoked with the help of a web application. It allows cross-platform execution as it is contained in a Docker container. The established framework utilizes digital signature cryptography to verify the authenticity of a signed message and key encapsulation mechanism (KEM) for secure communication between client and server. The algorithms can be integrated with various internet-based applications like IoT, blockchain, but the results are demonstrated using a web application for this research.

Keywords Quantum computing · Cryptography · Post-quantum cryptography · TLS · Digital signature · Key encapsulation mechanism · Network protocols

V. Kolachana (✉) · D. Upmandewan · A. Giri · N. Pavan · M. N. Thippeswamy · T. R. Vinay
Department of Computer Science and Engineering, NMIT, Yelahanka, Bangalore 560064, India

M. N. Thippeswamy
e-mail: mntswamy@msrit.edu

T. R. Vinay
e-mail: vinay.tr@nmit.ac.in

A. Ahmed
Unisys India Private Limited, Bangalore, India
e-mail: anees.ahmed@in.unisys.com

V. Bindhu et al. (eds.), *Proceedings of Third International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 844,
https://doi.org/10.1007/978-981-16-8862-1_29

Introduction

Quantum computing is a field of science that deals with quantum mechanics concepts and their application in computing, such that solutions to problems requiring massive computation and operations are made faster and simpler for execution, having lesser energy consumption. The purpose is to develop computer technology based on the behavior of energy and matter at an atomic or subatomic level. Quantum computing could contribute greatly in the fields of finance, military affairs, intelligence, drug design and discovery, aerospace designing, utilities (nuclear fusion), polymer design, AI and big data search, and digital manufacturing [1].

Quantum computing, in short, is the amalgamation of computer science and quantum mechanics, where quantum physical properties of subatomic particles are explored and integrated into computer systems. The most important quantum mechanics principles utilized in this field are superposition, interference, and entanglement. Superposition is the principle wherein a quantum system can exist in multiple quantum states at the same time, until measured at one particular instance, to find its exact state. It is a state where an entity with different measures of a common attribute exhibits all measures simultaneously, making the result a combination of all individual measures [2]. Interference is an extension such that two quantum systems, when superposed, cause the result to either be a sum or difference of their individual measures. Interference can help control the output to lean toward the measure that the user prefers [3]. Quantum entanglement is another property wherein two quantum states are entangled such that their states are not determinable but are correlated. This means that both states are dependent on each other, though they may be in different planes at that instance [4].

In ordinary computers, the simplest functional unit is a “bit” that can have a value of either 1 or 0. In quantum computers, the simplest unit is a “qubit”, with the property of having a value of both 1 and 0 at the same time. This is due to the concept of superposition, which allows the qubit to be in two states simultaneously. The rule of thumb is that the total probability of the qubit having state 0 and state 1 must be 100% at all times. The superposition principle allows qubits to perform computations at exponential rates, thereby contributing to the success of quantum computers as we know them [5].

The advent of this new form of computing has brought its share of pros and cons. The advantage is a wide range of enhanced computations in the fields mentioned above. However, the disadvantage posed is that there would be no control once the computation is made mainstream. With the sheer capacity of the system, it would be very easy for the computer to crack tricky passwords and cryptographic keys even by using a brute force method, which is deemed impossible for classical computers. All privacy measures taken now would be utterly useless once the quantum revolution picks its pace.

Dr. Krysta Svore of Microsoft Research stated—“The RSA-2048 Problem would take 1 Billion Years with a Classical Computer, but a Quantum Computer could do it in 100 s!”. That is the exponential power of the systems. To highlight the

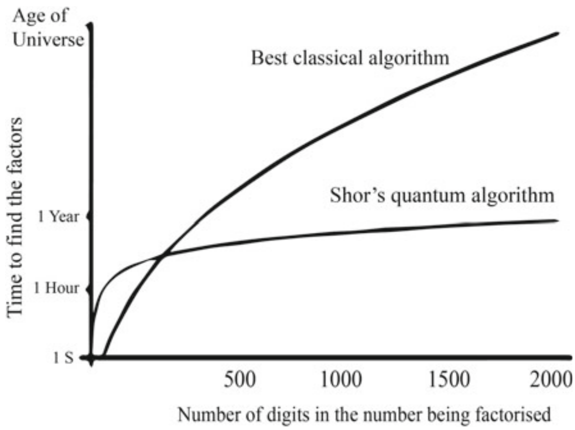


Fig. 1 Comparison of Shor's quantum algorithm with the best classical algorithm

computational efficiency of quantum computers, Fig. 1 shows how much faster Shor's algorithm can decode the prime factor keys of cryptographic algorithms [6].

This brought out a new branch in quantum computing called Post-Quantum Cryptography which focuses on the security of both classical and quantum computers, assuming the wake of a quantum world in the future. It deals with devising encryption or network security algorithms such that both classical and quantum computers are unable to break through the security code. This is required for the current generation when quantum computers are on the rise. It is necessary to build the network's security to avoid breaches of data and information [7]. This field is still relatively unexplored and larger organizations like the National Institute of Standards and Technology (NIST) are researching methods to enhance security. Standardization of such privacy methods would be a breakthrough in the field of computer science and networking.

The novel feature of this research work is that the developed framework utilizes quantum-safe algorithms for cryptography over the Transport Layer Security (TLS) network protocol to encrypt the communication holistically. Not only does the framework differ due to the application of quantum-safe algorithms, but also because it utilizes all of the 7 NIST shortlisted algorithms for encryption, not just one. The framework applies a user-friendly approach as well, by allowing the user to choose what algorithm needs to be applied. An important note is that quantum-safe algorithms do not work the same way as traditional cryptographic algorithms like AES and RSA. They utilize mathematical properties that are different from prime factorization, thereby making key identification by brute force an extremely hard task.

Literature Review

Quantum computing is a new way of computing that is based on quantum mechanics and its unique properties. The studied literature includes the basics of quantum computing, quantum cryptography, network protocols, NIST shortlisted algorithms, identifying the limitations of current cryptographic methods, and many other supportive resources related to the topic.

Firstly, it is necessary to understand the principles, concepts, mathematics, and the working of quantum computers and how they differ from the way classical computers work. To get a comprehensive view of the field, the works of Aaronson of [8] and Rieffel and Polak of [9] prove worthwhile. Considered as the best textbooks by the pioneers of the field, the two sources provide extensive material on the ideologies of quantum computing. Another immaculate source is the book by Sutor [10] which emphasizes the need for quantum computing and how quantum computers work, focusing spectacularly on the computation part of the field rather than just the mathematics.

To understand the need for quantum cryptography, Grau, the author of [11] deems it necessary to know how efficient quantum computers are and how the encryption can be easily broken using algorithms like Shor's algorithm. Mosca's inequality also determines that the encryption algorithm must be ready before quantum computers are commercially available. Even Hoursanov of [12] mentions that quantum-safe algorithms and their implementations must be developed beforehand as development cannot be done instantaneously and getting all protocols to function correctly takes a lot of time.

The authors of [13] give a thorough insight on different cryptographic types like X.509 certificates, IKEv2, SSH, and S/MIME to identify the loopholes of current cryptography. The authors of [14] too describe the current systems in detail and point to the areas that can be infiltrated.

There are a few open-quantum algorithms currently being explored. Lyubashevsky et al. [15] explore lattice-based cryptography to construct security primitives. Quantum-Safe Security Position Paper [16] and Quantum Security Technologies [17] speak of quantum key distribution which is based on quantum physics and allows keys to be exchanged between different locations using the quantum properties of quantum computers. As far as network protocols are concerned, Easterbrook et al. [18] are researching post-quantum TLS, a set of protocols included under transport layer security that limits the vulnerability of key exchange and authentication. Sikeridis et al. [19] speak about the performance of post-quantum algorithms on TLS 1.3 and Sikeridis et al. [20] assess the overhead of post-quantum cryptography in TLS 1.3 and SSH. Crockett et al. [21] attempt to adapt the existing TLS and SSH to incorporate the algorithms.

Lastly, Alagic et al. [22] thoroughly studied to analyze the algorithms that have to be implemented for this work. The entire research runs on the shortlisting done by NIST, its categories, and findings, so this source is of immense importance. NIST has conducted two rounds of shortlisting prior to this research and has arrived at

15 algorithms (7 finalists and 8 alternate selections) among hundreds of entries. The competition is still underway and is expected to conclude in 2023 when a new standard of post-quantum cryptography would be declared for commercial use.

The biggest source of information for this research work is the Open Quantum-Safe (OQS) initiative. Under this collaborative project, there are two segments—LibOQS and protocols/applications. Under the LibOQS scheme, a C library for NIST shortlisted post-quantum algorithms has been developed by quantum computing experts with sponsorship from companies like Microsoft and IBM. The protocols/applications section has certain features that can be incorporated for protocols like TLS, SSH, and X.509 [23]. The Python versions of these algorithms have been utilized for the research.

Brief Overview of the Technologies Used

Quantum Key Distribution (QKD)

Quantum Key Distribution (QKD), also informally called quantum cryptography, uses quantum particles to provide secure symmetric key distribution among the communicating entities. The process of QKD is such that the underlying protocol being used would require to transmit the public key or symmetric key to the receiver of the message for decryption. But this transfer, by itself, can experience eavesdropping and can then threaten the message communication. QKD quantizes this communication such that even if the key transfer is looked upon or is intercepted, the quantum property alters itself and alerts the communicators that the information has been breached. The QKD protocols have been designed to ensure that any eavesdropper in the communication would change the original information being shared, thereby signaling the presence of a third party to the ones communicating with the error being generated. Since qubits are usually made of photons, the quantum properties of light are altered on the interception.

The keys are transmitted among the communicators securely but QKD by itself cannot encrypt the messages as well. Hence, this technology would need to be combined with a conventional symmetric cryptographic encryption like AES to provide end-to-end encryption for messages and make the entire system quantum-safe [24].

Transport Layer Security (TLS)

TLS is the forerunner of Secure Sockets Layer (SSL), and it secures communications over a computer network. This protocol is useful in a variety of applications including email, instant messaging, and voice-over IP.

TLS 1.3, built by Netscape Communications, is the latest version in use to avoid eavesdropping and tampering. It facilitates privacy and data security over the Internet. TLS works mostly on the transport layer which is why it works great at encrypting data from Web sites and applications. To utilize TLS, the system must have a TLS certificate that specifies the domain owner and server's public key and must be issued by a valid certificate authority. The connection between end-users is done through TLS handshaking, a process that incorporates requests and acknowledgements to ensure all data is communicated correctly [25].

OpenSSL

OpenSSL is a robust open-source toolkit for TLS protocol, formerly called the Secure Sockets Layer (SSL). The protocol is built on a full-featured, general-purpose cryptographic library. The SSLeay library, created by Tim J. Hudson and Eric A. Young, is the ancestor of OpenSSL.

The SSL and TLS protocols are implemented in OpenSSL, which is another open-source library. The core C language library implements basic cryptographic functions and provides a variety of utility functions. Wrappers for OpenSSL are available in many programming languages [26].

NIST Shortlisted Algorithms

Anticipating the advent of quantum computers, NIST envisioned a worldwide standard for post-quantum cryptography and opened a global competition for developers to pitch in their quantum-safe algorithms. As per the second round of shortlisting in July 2020, 7 finalists were chosen and 8 were taken as alternative choices. The finale is estimated to be held in 2023. The algorithms all use mathematical domains that differ from traditional prime factorization and can be categorized into five main families, namely code-based, isogeny-based, hash-based, lattice-based, and multivariate system-based [22].

LibOQS, an initiative by the big shots of IT, put together an open-source library containing the codified versions of these algorithms and has paved the way for developers working in the field of post-quantum cryptography to utilize these algorithms effectively by including them in their library. The following two schemes have been mainly used in the research study

Key Encapsulation Scheme—Classic McEliece, NTRU, Kyber, and Saber.

Signature Scheme—Dilithium, Rainbow, and Falcon.

Each algorithm has its own patented technology for creating key pairs and for verifying the corresponding encryption for the key used. Each algorithm uses a certain number of bits for its key pair generation, but the keys that are ultimately utilized for encryption and decryption are 32 bits long.

Table 1 Overview of the NIST algorithm finalists

Scheme	KEM/Signature	Family
Classic McEliece	Key encapsulation	Code-based
Kyber	Key encapsulation	Lattice-based
NTRU	Key encapsulation	Lattice-based
Saber	Key encapsulation	Lattice-based
Dilithium	Digital signature	Lattice-based
Falcon	Digital signature	Lattice-based
Rainbow	Digital signature	Multivariate-based

Table 1 gives an overview of the algorithms selected as the finalists along with their scheme and their family.

LibOQS

LibOQS is a quantum-safe cryptographic library written in C which is open source. LibOQS provides a series of open-source Key Encapsulation Mechanisms (KEM) and Digital Signature algorithms that are quantum-safe.

A common scheme for these algorithms, as well as the test harness and benchmarking routines, are available in the library. Douglas Stebila and Michele Mosca-led Open Quantum-Safe (OQS) initiative aims to build and incorporate quantum-safe cryptography into applications to make implementation easier for real-world problems. Via OpenSSL and OpenSSH, TLS and SSH integrations with LibOQS are established [23].

There is also a Python language wrapper for the library called LibOQS-Python. This allows users to import a specialized package in Python to utilize the classes and functions defined for each of the shortlisted algorithms. It is a derivation from the C library and provides all the essential parameters for the algorithms' usage [27].

Dockerization

Docker is the software for building portable, lightweight containers that help make application development, testing, and deployment easier. Containerization is a form of operating system virtualization that enables users to run multiple applications in isolated containers, all sharing the same operating system. Docker has been used to make sure that no matter what operating system the client is on, data transfer between client and server is always safe from attack by a quantum computer. This has made the system very portable, flexible, and easy to use [28].

Methodology

As specified earlier, the purpose of the research work is to utilize the shortlisted NIST algorithms as they have been defined using LibOQS. This means that the structure and usage of the algorithm are fixed, but the methodology of accessing the algorithm has been defined in this work. Since TLS is being used, the flow has been configured accordingly.

A user's device is considered as the client that is communicating with the server. After applying quantum-safe algorithms for encryption, the client sends the message using the TLS protocol through an HTTPS channel. This channel would ensure security against hackers, quantum/classical computers, and other agencies. The data then goes to the server which performs the decryption and retrieves the message. Figure 2 shows the process in a brief manner. This is a general process that depicts the scope of cryptography.

The architecture of the framework highlighted in Fig. 3 demonstrates the establishment of the environment. The KEM and digital signature algorithms are accessed through the LibOQS-Python package, constituting one segment of the architecture. The other part is that of OpenSSL and TLS wherein the two services are utilized to set up secure communication. A fork of OpenSSL called OQS_OpenSSL_1_1_1 is used to integrate TLS with the quantum-safe algorithms using the certificate and keys generated by them [29]. This entire setup containing the three segments is then enclosed within a container (here, Docker) to run the framework cross-platform. The environment is native to this research work and has been developed solely to facilitate the usage of quantum-safe algorithms.

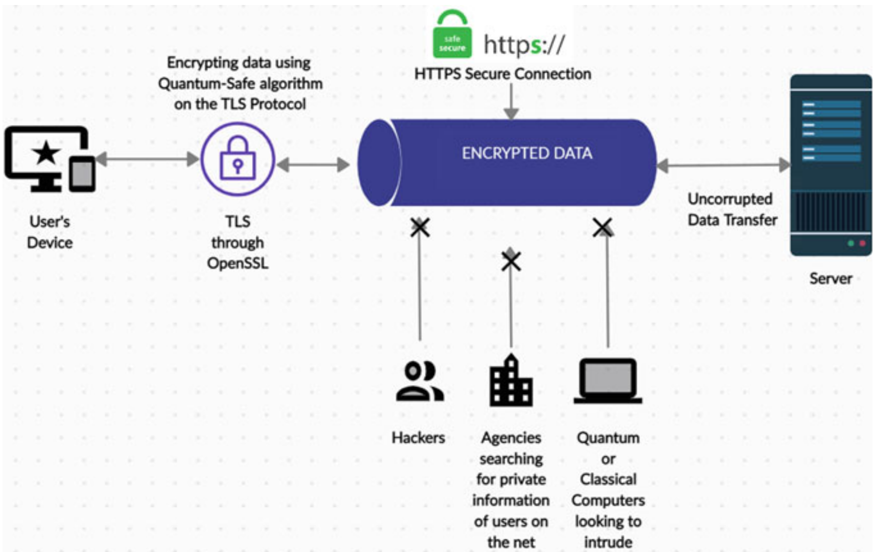


Fig. 2 The design of the system

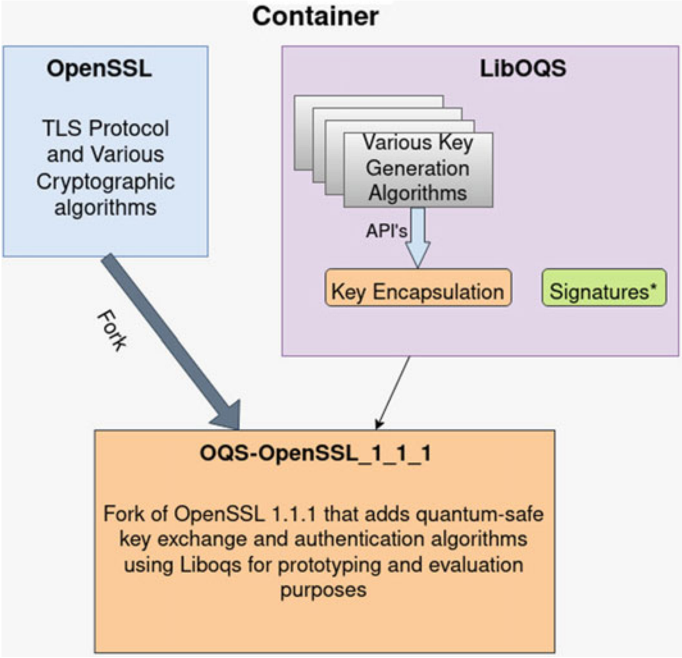


Fig. 3 Architecture of the framework

Additionally, NIST-shortlisted algorithms are divided into two schemes—key encapsulation mechanism and digital signature. It is to be noted that each scheme has its own way of encryption and decryption and that a separate design would be required for each. This means that a separate methodology would be required for the individual schemes as well. Under the following sections, the process of each scheme is elaborated and the approach has been illustrated.

Key Encapsulation Mechanism (KEM)

The Key Encapsulation Mechanism (KEM) scheme is a unique way of communication that deals with two segments—one, the message being sent is to be encrypted/decrypted and two, the keys being used for the message have to be encapsulated/decapsulated. If traditional public-key encryption schemes are to be considered, party A generates the public/private key pair and encrypts its message using the private key. It shares the public key along with the ciphertext to party B, who uses the key to decrypt the ciphertext and retrieve the original message. The drawbacks of this scheme are that the message length may be limited, encryption may not be completely secure, and the keys' mathematics only depend on prime factorization.

To combat the limitations, a scheme called KEM is constructed, containing two layers of communication:

1. A public key layer that establishes a random symmetric key.
2. A symmetric key layer that works to protect the data using the symmetric key generated by the previous layer.

Both these layers work independently to a great extent, thereby assuring the system of greater security. Since there are two forms of keys used to protect the data, even if one set gets compromised, the communicators get notified of the intrusion and appropriate measures are taken [30]. For this research, Advanced Encryption Standard (AES) is being used in the public key layer and the NIST-shortlisted KEM algorithms are used in the symmetric key layer depending on the choice of the user.

For this work, the process mentioned above was tweaked according to the environment established. Certain modifications were made to ensure the end-to-end security of the communication. To illustrate the same, the communication's two endpoints are considered as the client and the server, where the client could be any user's device. The procedure for the communication is as follows:

1. The client would generate the quantum-safe key pair as per the chosen algorithm where the key pair refers to the public key and the private key to be used for communication.
2. The public key is then sent to the server in a secured manner.
3. The server then uses the public key as the input for an encapsulation function that gives an output of a shared secret key and a ciphertext.
4. The server keeps the shared secret key with itself and sends the ciphertext to the client.
5. On receiving the ciphertext, the client uses it as an input for a decapsulation function that produces another secret shared key. The validation is that both the client and the server must match the shared secret keys for the communication to have been intruder-free.
6. Now that it has the shared secret key, the client encrypts the message using the shared secret key for the AES encryption function. It sends the resultant encrypted text to the server.
7. The server finally uses its version of the shared secret key to decrypt the encrypted text and retrieve the original message of the client.

This entire procedure is highlighted in Fig. 4 containing the sequence diagram for the KEM scheme. Figure 5 reflects the corresponding architecture diagram of the scheme. It must be noted that though the communication can be reversed in a real-world situation, this research work contains only the procedure mentioned above.

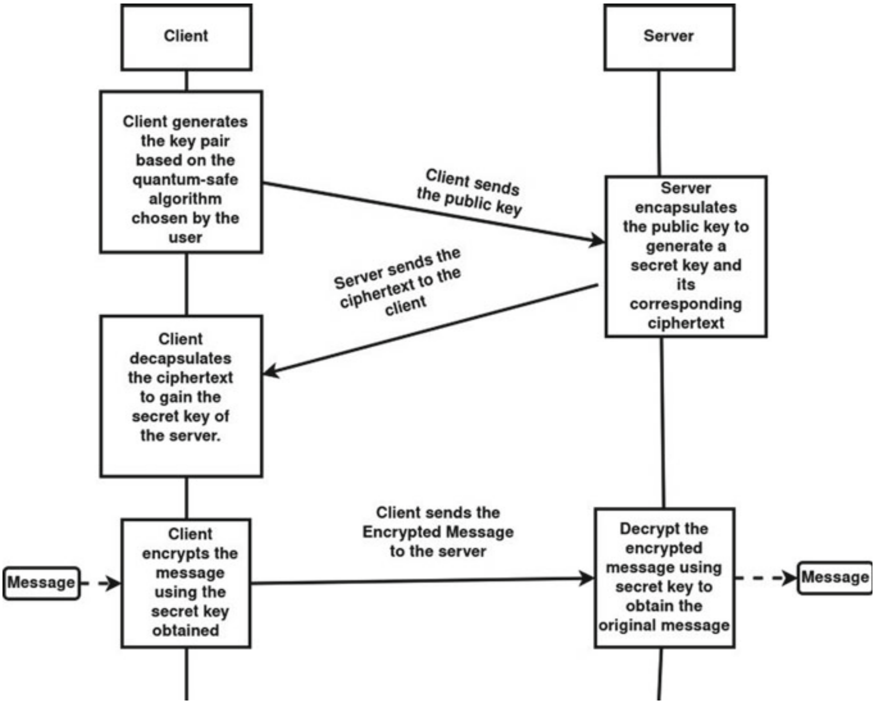


Fig. 4 Sequence diagram of KEM implementation

Digital Signature

Digital signature is a system used to authenticate or verify a message or a document. It works the same way as an e-signature and its purpose is to protect against tampering, deception, and counterfeits. It works on the principle of asymmetric cryptography where two keys, public and private, are used for encryption.

The one who signs the message is referred to as the signer and the other party is referred to as the verifier. The signer signs the intended message using the private key and shares the ciphertext and the public key with the verifier. The verifier then uses it to retrieve the message and then compares it with the original message to see if the signature is authentic. If the verifier finds that the messages are not the same, it means that the message has been meddled with [31].

In this framework, the methodology for digital signature has two endpoints referred to as signer and verifier. The process of digital signature is as follows:

1. The signer generates a key pair based on the chosen quantum-safe algorithm. It generates both a public key and a private key.
2. The signer uses the private key to sign the message and sends the signature to the verifier along with the public key, the selected algorithm, and the original

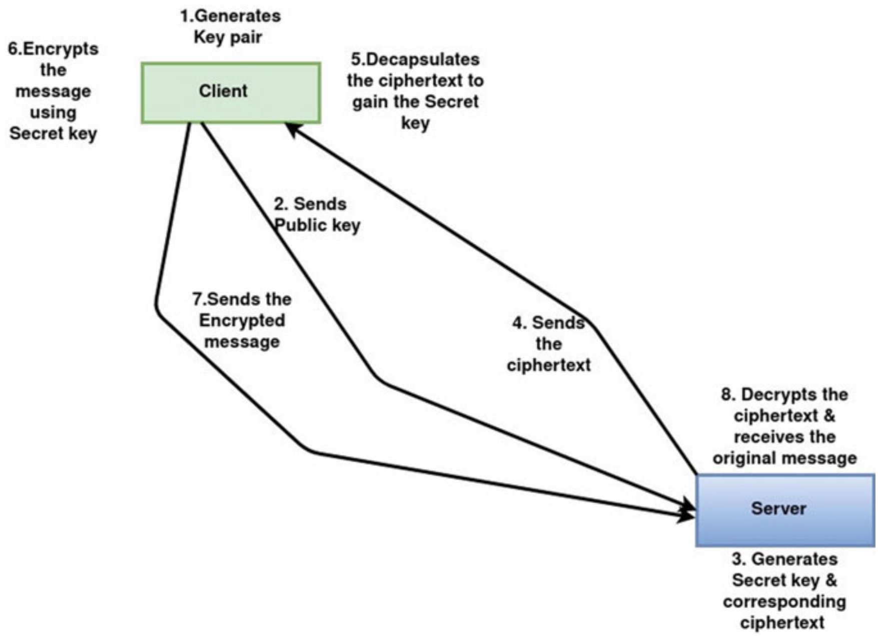


Fig. 5 Architecture diagram of KEM implementation

message for verification. The verifier then retrieves the original message from the signed one using the public key.

3. The verifier compares the original message with the decrypted one to see if they are the same. If they are the same, the signature is accepted as a valid one. Else, the communication has been breached.

The entire sequence has been depicted in Fig. 6 which highlights the process flow. Figure 7 depicts the architecture of the digital signature scheme used in the framework.

Implementation

There are two aspects to establishing the framework—to develop an API with working quantum-safe algorithms and to establish a web application to function alongside the API using TLS 1.3 protocol. The framework is meant to encourage the use of the algorithms in all fields that utilize TLS protocol in any form, like in web applications, cloud platforms, IoT, blockchain, etc. A web application is developed using Flask to demonstrate the usage of the proposed framework.

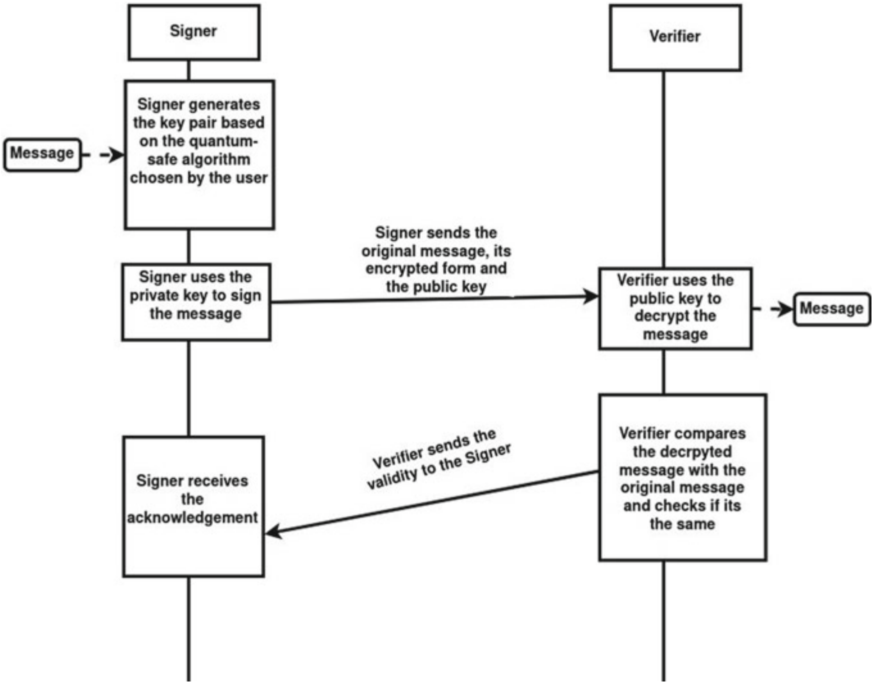


Fig. 6 Sequence diagram of digital signature scheme

API Development

The API includes Python versions of the quantum-safe algorithms and starts with the server–client establishment. In this work, three Open Quantum-Safe libraries have been installed to enable the algorithm functions:

- The OQS_OpenSSL_1_1_1 fork is installed to enable TLS-based certificates and keys that are used to establish an HTTPS connection over TLS 1.3 [29].
- The LibOQS library is installed to get the basic C files of the algorithms along with their mathematical key generation methods [32].
- The LibOQS-Python library is then installed to obtain the Python functions of the algorithms that interact with the LibOQS C library to generate the keys [27].

The LibOQS-Python library provides a Python package called "oqs", which gives the following functions classes and methods on integration:

- Class Signature—To enable the digital signature algorithms and their functions
  - Sigalg variable—A variable containing the algorithm chosen among "DEFAULT" (A default algorithm in case of no choice), "Dilithium2" (Crystals Dilithium algorithm), "Falcon-512" (Falcon algorithm), and "Rainbow-I-Classic" (Rainbow algorithm).

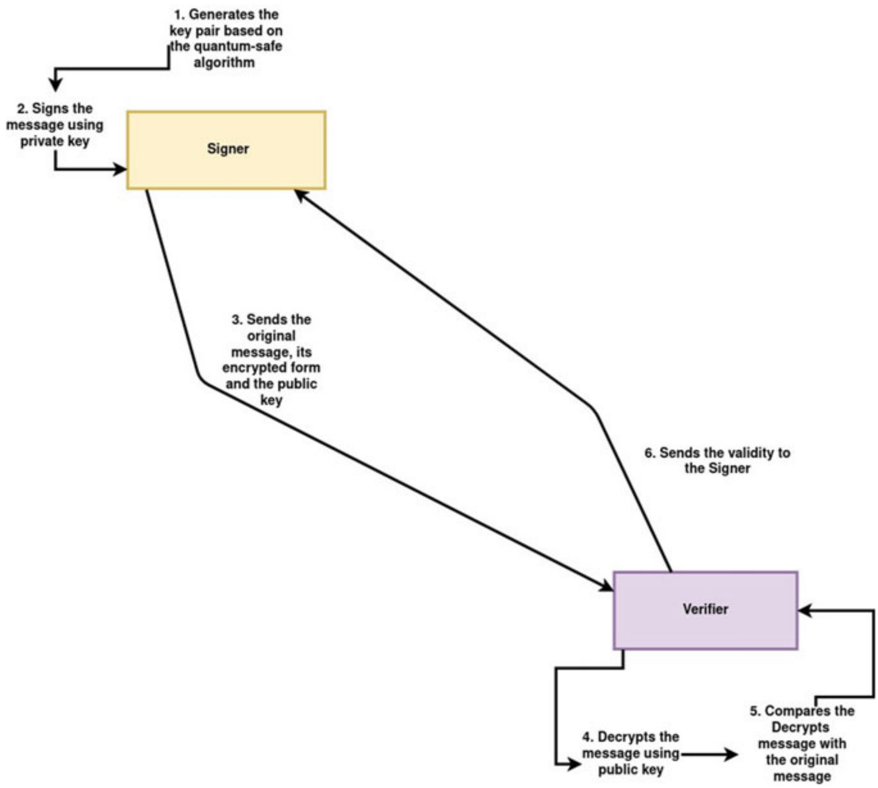


Fig. 7 Architecture diagram of digital signature scheme

- details()—A function that displays the details of the algorithm chosen like name, public key size, signature size, encryption level, etc.
- generate_keypair()—A function that generates the public and private keys used for the signature scheme.
- export_secret_key()—As the previous function generates both public and private keys as outputs, this function extracts only the private key.
- sign()—A function to perform the signature on the message provided by the user. This is used only on the signer's side.
- verify()—A function to verify if the signature is authentic or not. This is used only by the verifier.

- Class KeyEncapsulation—To enable the KEM algorithms and their functions
  - Kemalg variable—A variable containing the algorithm chosen among "DEFAULT" (A default algorithm in case of no choice), "Classic-McEliece-348864" (Classic McEliece algorithm), "Kyber512" (Crystals Kyber algorithm), "NTRU-HPS-2048–509" (NTRU algorithm), and "LightSaber-KEM" (Saber algorithm).

- details()—A function that displays the details of the algorithm chosen like name, public key size, secret key size, encryption level, etc.
- generate_keypair()—A function that generates the public and private keys used for the signature scheme. This is done on the client's side.
- export_secret_key()—As the previous function generates both public and private keys as outputs, this function extracts only the private key.
- encap_secret()—A function on the server's side that uses the public key as an input to generate a shared secret key and a ciphertext. The shared secret key will henceforth be called shared_secret_server.
- decap_secret()—A function on the client's side that take the ciphertext as the input and produces a shared secret key as the output, henceforth called shared_secret_client.

In addition to the KeyEncapsulation class, the AES encryption library also has to be imported and its functions must be used for encryption.

There are two Python files established for the API. One operates as the client for KEM and as the signer for digital signature algorithms (called client.py) and the other operates as the server for KEM and as the verifier for digital signature algorithms (called server.py). The functionalities and processes are established as mentioned in the methodology section between the two files. It must be noted that each file has two separate sections to deal with KEM and digital signature. The entire system is abstracted well to not allow the interference of any other algorithm in the midst of one being executed at that instance.

The methodology is referred to and the corresponding variables/functions provided in the OQS package are utilized. Since the functions and the codified version are defined by the LibOQS library, the only action that must be taken is to ensure the right variables and information is being fed to the functions.

Web Implementation

The web application is deployed using Python Flask. Flask is a web application framework designed to deploy applications easily using Python and make them scalable. Just installing the package helps use its utilities on many protocols, especially TLS [33, 34].

The API is connected to two HTML files, named sig.html and kem.html for digital signature and KEM respectively. A third HTML file called index.html is the landing page of the application and allows the user to choose between the two categories of algorithms, KEM or Digital Signature. The page gets directed to the corresponding web page as per the choice.

- Sig.html
  - Inputs for the web page are:

Message to be signed
Algorithm selection dropdown menu

- Outputs of the web page are:

  The original message
  The message length
  The digital signature algorithm chosen
  The validity of the signature (whether the signature is valid or not)
  The signature appended to the message in encrypted form

- kem.html
  - Inputs for the web page are:

    Message to be communicated
    Algorithm selection dropdown menu

- Outputs of the web page are:

  The original message
  The message length
  The KEM algorithm chosen
  The Validity of the transfer (Whether there is an intrusion or not)
  The secret key used for the encryption
  The encrypted message
  The decrypted message

The entire framework is finally Dockerized to allow it to be deployed on all platforms that support Docker. To run Docker, the Docker package must be installed. It must be built and a container image must be initiated. The container can then be run and launch the application.

In this work, a WSGI server is established to launch the Flask service on HTTPS and TLS [35]. A Web Server Gateway Interface (WSGI) server has a mechanism to configure server communication with its clients and its response to the clients' requests. The configuration has been made to allow multiple client requests at once. Depending on the CPU power and RAM capacity of the system running the server, multiple processes can be stemmed where each process can serve a client request. The systems used here for development could spawn up to 8 processes, and hence, 8 client requests could be handled in a single instance. The entire system is then made available for testing using NGROK which allows the server to be tunneled to HTTP/HTTPS requests made on a local system.

The algorithms use TLS 1.3 by default. Hence, that version is utilized by web browsers to launch the web application.

Results

The result of the work is that communication is established between server and client using OpenSSL and TLS, where data is encrypted using quantum-safe algorithms and sent across nodes using HTTPS.

On launching, the web application shows the index.html page as in Fig. 8. It offers the user the option of going ahead with either the KEM scheme or the signature scheme.

Figure 9 shows that on choosing the signature scheme, the sig.html page is displayed. The first two fields are input fields that take in the message and the algorithm choice from the user. The API is called and it communicates between the signer and the verifier. The remaining fields on the page are the outputs for the scheme. It must be noted that the signature sizes, encryption scheme, and other features change for each algorithm.

Figure 10 shows that on choosing the KEM scheme, the kem.html page is displayed. The first two fields are input fields that take in the message and the algorithm choice from the user. The API is called and it performs end-to-end encryption using both AES cryptography and quantum-safe KEM algorithms between the server and the client. The remaining fields on the page are the outputs for the scheme. It must be noted that the key sizes, encryption scheme, and other features change for each algorithm.

The web browser's security tabs are opened and searched to test if the web application uses TLS for communication. As evident from Fig. 11, the application runs on TLS 1.3 and also uses AES encryption on the HTTPS connection. The page says that it has valid and secure HTTPS because NGROK generates authentic certificates and keys to establish the connection.

The WSGI server also records the processes that are running the different client requests and the web page of the framework is being accessed. The server records are shown in Fig. 12.

Hence, it can be verified that the application is indeed running on TLS and that the research objective has been satisfied. Furthermore, the encryption shows that the quantum-safe algorithms work efficiently and produce validity as true. Thus,

Fig. 8 Web application launch page

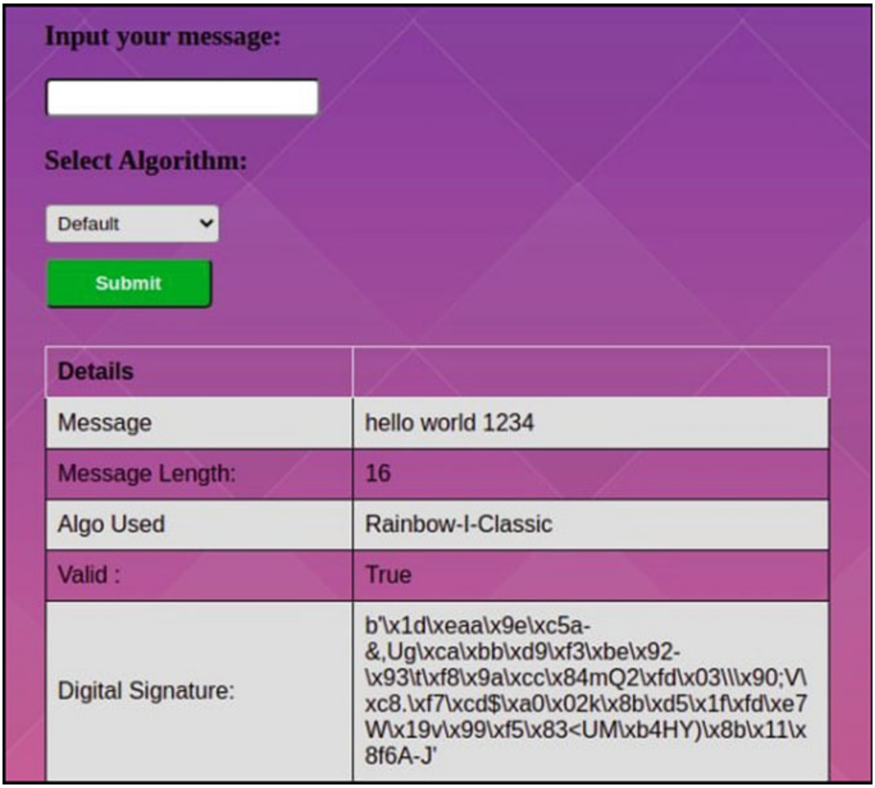


Fig. 9 Web page for digital signature scheme

an efficient framework has been established that satisfies the research objectives completely [36].

Thus, it can be inferred with this research that quantum-safe algorithms can be run effectively using TLS 1.3 protocol and can have authentic communication through web-based applications. The NIST shortlisted algorithms' functionality and working have been defined by LibOQS and this work has utilized their library. The findings are that all shortlisted algorithms can be run efficiently and using the NGROK server, multiple algorithms can be run simultaneously as well, without getting tangled with one another. Also, depending on the size of the message and the chosen algorithm, the key generated changes as well, proving that the input is being taken into account correctly. Hence, if this work is to be used for the information supplied through forms and other mediums, the data would be protected at all times.

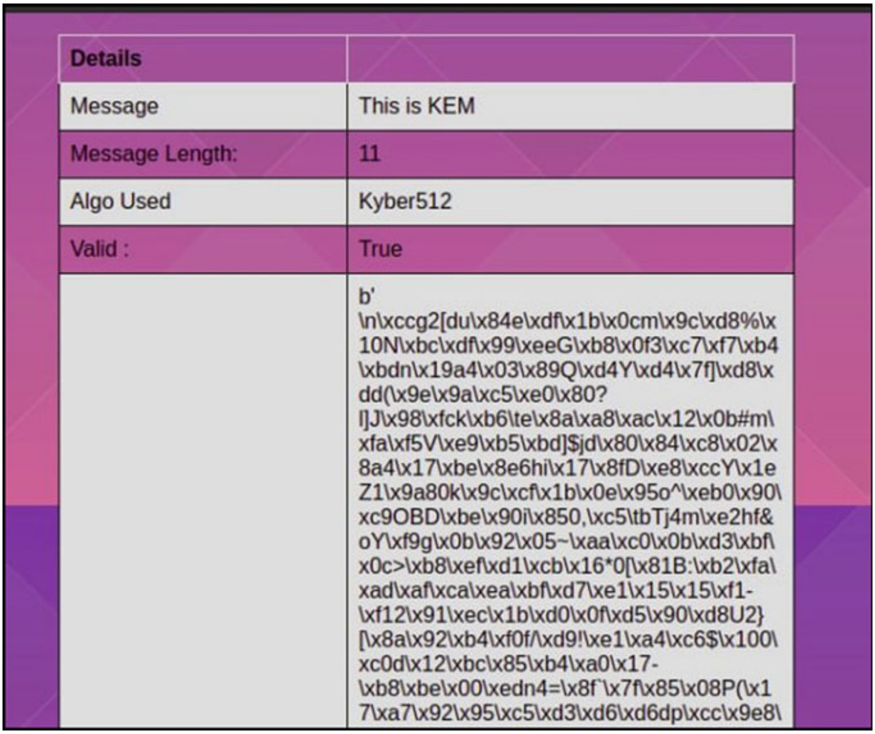


Fig. 10 Web page for KEM scheme

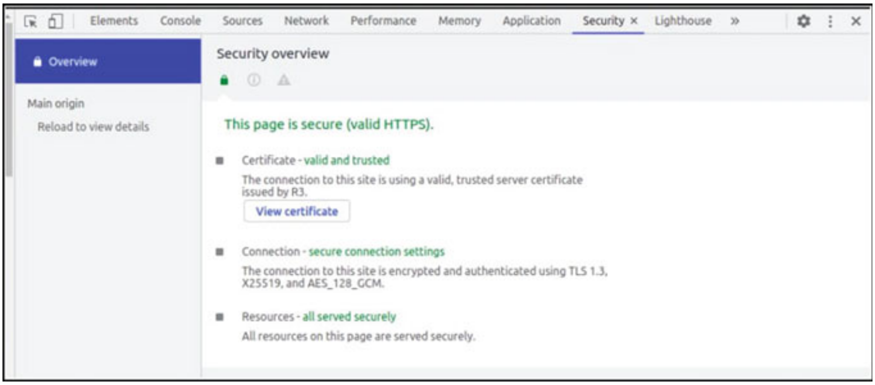


Fig. 11 Security tab of the web browser

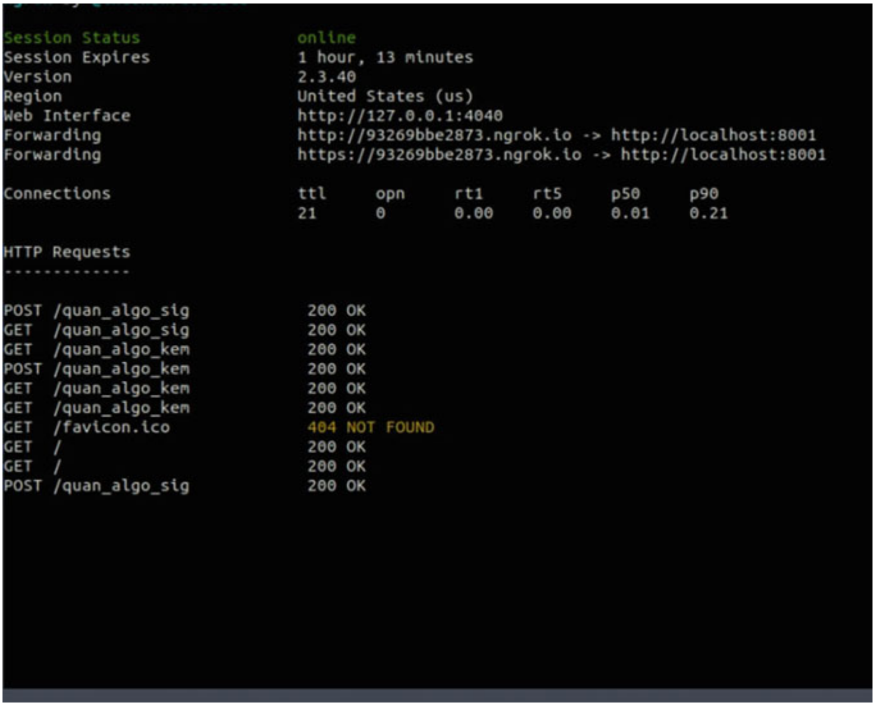


Fig. 12 Server record showing details of multiple requests

Conclusion and Future Work

The research paper focuses on developing a framework that utilizes quantum-safe algorithms on the TLS protocol. The framework takes the help of the LibOQS-Python library to get functions of the algorithms and uses them on a web application to provide end-to-end communication either for digital signature or KEM. To establish TLS handshaking, Flask framework is used with a TLS certificate and key and an HTTPS connection on the web application is established. The user can input the message to be communicated and choose an algorithm depending on the scheme. The output is provided accordingly. The entire framework is Dockerized to allow cross-platform execution.

This implies that a framework has been established using TLS protocol such that all data being communicated via a web application is being encrypted at the client and is decrypted at the server. In the case of digital signature, the data received by the verifier is authentically signed by the signer. Furthermore, unlike the technologies available currently, this framework is capable of executing all the 7 NIST-shortlisted algorithms, perhaps even simultaneously, if multiple clients are deployed.

As a future prospect for this work, it can be extended to include a wider range of applications like IoT, blockchain, or cloud computing. After NIST concludes its competition of quantum-safe cryptography, the selected standard can be established as the only algorithm of the framework to make it an authorized system for industrial usage. More additions can be made to include encryption systems other than AES, like RSA, in the KEM scheme of encryption. The progress in the field of quantum computing would further bring enhanced capabilities in the future to test the framework in a real-world situation.

Acknowledgements The research team sincerely thanks the team of Unisys India Pvt. Ltd, especially Mr. Anees Ahmed, for giving the opportunity to work on “Application of Quantum Algorithms for Network Protocols” and for providing guidance and support during its development. Special thanks to the mentors of the research, Dr. Thippeswamy M.N., and Mr. Vinay T.R. for their valuable insights. The team also extends gratitude to Nitte Meenakshi Institute of Technology for supporting with the required resources and for helping the idea materialize and succeed.

References

1. Frankenfield J (2019) An article on “Quantum Computing. [Online] Available: Investopedia, https://www.investopedia.com/terms/q/quantum-computing.asp. Accessed 3 Dec 2020
2. An article on “Quantum Superposition”. [Online] Available: Joint Quantum Institute, https://jqi.umd.edu/glossary/quantum-superposition. Accessed 10 June 2021
3. Pieper J, Lladser ME (2018) Quantum computation. Scholarpedia. [Online] Available: Scholarpedia, http://www.scholarpedia.org/article/Quantum_Computation#Quantum_Interference_and_Decoherence. Accessed 10 June 2021
4. Orzel C (2017) An article on “How do you create quantum entanglement?. [Online] Available: Forbes, https://www.forbes.com/sites/chadorzel/2017/02/28/how-do-you-create-quantum-entanglement/?sh=6d0ec4ea1732. Accessed 10 June 2021
5. An article on “Quantum Computing 101”. [Online] Available: University of Waterloo, https://uwaterloo.ca/institute-for-quantum-computing/quantum-computing-101. Accessed 3 Dec 2020
6. Sihare SR et al (2017) Analysis of quantum algorithms with classical systems counterpart. MECS, IJIEEB 9:20. https://doi.org/10.5815/ijieeb.2017.02.03
7. An initiative on “Post-Quantum Cryptography” (2020) [Online] Available: NIST Computer Security Resource Centre, https://csrc.nist.gov/projects/post-quantum-cryptography. Accessed 3 Dec 2020
8. Aaronson S (2013) Quantum computing since democritus. Cambridge University Press, ISBN 978-0-521-19956-8. Accessed 25 Nov 2020
9. Rieffel E, Polak W (2011) Quantum computing: a gentle introduction. MIT Press. Accessed 25 Nov 2020
10. Sutor RS (2019) Dancing with qubits. Packt Publications. ISBN 978-1-838-82736-6. Accessed 21 Jan 2020
11. Grau A (2020) An article on “Quantum-safe cryptography”. [Online] Available: Sectigo, https://sectigo.com/resource-library/quantum-safe-cryptography-surviving-the-upcoming-quantum-cryptographic-apocalypse. Accessed 25 Nov 2020
12. Hoursanov A (2020) an article on “Quantum-safe security”. [Online] Available: SAP Innovation Center Network https://icn.sap.com/article/quantum-safe-security-future-proof-data-protection. Accessed 25 Nov 2020
13. Campagna M et al (2015) Quantum-safe cryptography and security. ETSI White Paper No 8. [Online] Available: ETSI https://www.etsi.org/images/files/ETSIWhitePapers/QuantumSafeWhitepaper.pdf. Accessed 25 Nov 2020

14. Lisa et al (2018) Cryptography in a post-quantum world. In: Quantum cryptography whitepaper, Version 5. [Online] Available: Accenture https://www.accenture.com/_acnmedia/PDF-87/Accenture-809668-Quantum-Cryptography-Whitepaper-v05.pdf. Accessed 25 Nov 2020
15. Lyubashevsky V et al (2020) "Security & Privacy", an article in quantum-safe cryptography. [Online] Available: IBM Zurich https://www.zurich.ibm.com/securityprivacy/quantumsafecryptography.html. Accessed 25 Nov 2020
16. "Quantum-Safe Security Position Paper" by Quantum-Safe Security Working Group, 2016. [Online] Available: Cloud Security Alliance https://downloads.cloudsecurityalliance.org/initiatives/qss/What_is_Quantum_Safe_Security_position_paper.pdf. Accessed 25 Nov 2020
17. "Quantum Security Technologies" in White Paper Version 1.0, 2020. [Online] Available: National Cyber Security Center UK https://www.ncsc.gov.uk/whitepaper/quantum-security-technologies. Accessed 25 Nov 2020
18. Easterbrook K et al (2020) Post-quantum TLS. [Online] Available: Microsoft https://www.microsoft.com/en-us/research/project/post-quantum-tls/. Accessed 25 Nov 2020
19. Sikeridis D et al (2020) Post-quantum authentication in TLS 1.3: a performance study. In: NDSS 2020. [Online] Available: Cryptology ePrint Archive https://eprint.iacr.org/2020/071. Accessed 4 Sept 2021
20. Sikeridis D et al (2020) Assessing the overhead of post-quantum cryptography in TLS 1.3 and SSH. In: CoNEXT 2020. [Online] Available: ACM https://doi.org/10.1145/3386367.3431305. Accessed 4 Sept 2021
21. Crockett E et al (2019) Prototyping post-quantum and hybrid key exchange and authentication in TLS and SSH. In: NIST conference. [Online] Available: https://www.douglas.stebila.ca/research/papers/NISTPQC-CroPaqSte19/. Accessed 4 Sept 2021
22. Alagic G et al (2020) Status report on the second round of the NIST post-quantum cryptography standardisation process. NISTIR 8309. [Online] Available: NIST Computer Security Resource Center https://csrc.nist.gov/publications/detail/nistir/8309/final. Accessed 25 Nov 2020
23. Stebila D et al (2021) Open quantum-safe project. [Online] Available: Open Quantum-Safe, https://openquantumsafe.org/. Accessed 20 May 2021
24. An article on "Quantum Key Distribution" (2021) [Online] Available: IDQ Antique, https://www.idquantique.com/quantum-safe-security/overview/quantum-key-distribution/. Accessed 10 June 2021
25. Rescorla E (2018) Standards track—the transport layer security (TLS) protocol version 1.3. RFC 8846, ISSN 2070-1721. [Online] Available: IETF, https://datatracker.ietf.org/doc/html/rfc8446. Accessed 10 June 2021
26. OpenSSL Home (2018) [Online] Available: OpenSSL, https://www.openssl.org/. Accessed 15 May 2021
27. Stebila D et al. (2020) LibOQS Python. [Online] Available: GitHub https://github.com/open-quantum-safe/liboqs-python. Accessed 26 Mar 2021
28. Yegulalp S et al (2021) What is Docker? The spark for the container revolution. [Online] Available: InfoWorld https://www.infoworld.com/article/3204171/what-is-docker-the-spark-for-the-container-revolution.html. Accessed 26 Mar 2021
29. Stebila D et al (2021) OQS-OpenSSL_1_1_1. [Online] Available: https://github.com/open-quantum-safe/openssl. Accessed Mar 26, 2021
30. XML Security Working Group F2F (2009) Key encapsulation: a new scheme for public-key encryption. [online] Available: W3 Lists https://lists.w3.org/Archives/Public/public-xmlsec/2009May/att-0032/Key_Encapsulation.pdf. Accessed Mar 2021
31. Lutkevich B et al (2021) What is a digital signature?. [Online] Available: Tech Target https://searchsecurity.techtarget.com/definition/digital-signature. Accessed 10 Jun 2021
32. Stebila D et al (2021) LibOQS. [Online] Available: https://github.com/open-quantum-safe/liboqs. Accessed 26 Mar 2021
33. Pallets (2021) Flask documentation. [Online] Available: https://flask.palletsprojects.com/en/1.1.x/quickstart/. Accessed May 2021
34. Grinberg M (2021) Running your flask application over HTTPS. [Online] Available: https://blog.miguelgrinberg.com/post/running-your-flask-application-over. Accessed Apr 2021

35. Narendra Harny (2021) Implementing a web-socket using flask. [Online] Available: https://medium.com/swlh/implement-a-websocket-using-flask-and-socket-io-python-76afa5bbeae1. Accessed Apr 2021
36. Beullens W et al (2021) Post-quantum cryptography—current state and quantum mitigation. ENISA. [Online] Available: https://www.enisa.europa.eu/publications/post-quantum-cryptography-current-state-and-quantum-mitigation. Accessed May 2021

A New Method of Reconfigurable ADC with Automatically Optimized Parameters

Jayamala Adsul, J. M. Nair, and P. P. Vaidya

Abstract Reconfigurable analog to digital converter (ADC) has been designed and constructed based on the new method proposed in this paper. The most important part of the reconfigurable ADC is the successive approximation register (SAR) type of ADC whose resolution is controlled based on the maximum signal frequency of the input. This results into the optimum performance with respect to resolution and conversion time of the ADC, since the number of stages of conversion in SAR technique are decided based on the maximum input signal frequency. The rate of change of input is assessed using a circuit designed for this purpose to change the resolution of SAR ADC automatically. An additional flash ADC has been incorporated in the design to extend the range of frequency up to several hundred MHz. The resolution of this hybrid ADC (flash + SAR) is automatically varied from 8-bit to 16-bit according to the maximum input signal frequency. This ADC has been designed and simulated in National Instruments Multisim 14.1, and the simulated results are provided in the paper.

Keywords Flash ADC · SAR ADC · Hybrid ADC · Rate of change of input

Introduction

There are several types of ADCs like flash ADC, delta-sigma ADC, single-slope ADC, SAR ADC, pipeline ADC,. Each type of ADC is applicable in different fields.

J. Adsul (✉)
Department of Electronics Engineering, VESIT, Mumbai, India
e-mail: jayamala.adsul@ves.ac.in

J. M. Nair · P. P. Vaidya
Department of Instrumentation Engineering, VESIT, Mumbai, India
e-mail: principal@ves.ac.in

P. P. Vaidya
e-mail: pp.vaidya@ves.ac.in

V. Bindhu et al. (eds.), *Proceedings of Third International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 844,
https://doi.org/10.1007/978-981-16-8862-1_30

There are many applications of ADC like in the field of Internet of things [1], multi-standard systems [2], detection of neural signal [3], sensors [4] and instrumentation systems [5]. These conventional ADCs with fixed resolution and fixed conversion time are used for limited applications. Flash ADC is used for low resolution (6 to 8-bit) and hundreds of MHz sampling rate applications [6].

Delta-sigma is used for high resolution (up to 32-bit) and in the range of kHz sampling rate applications [7]. Integrating-type single-slope ADC is useful for medium resolution (up to 16-bit) [8] and sampling rate in the range of few Hz. SAR ADC is useful for medium resolution (up to 18-bit) and sampling rate in the range of few kHz to MHz. Pipeline ADC is useful for resolution 10 to 14-bit and continuous sampling rate in the range greater than tens of MHz applications [9]. Also, many researchers have used different methods to reconfigure these ADCs to achieve various application by configuring resolution and conversion time. Flash ADC is reconfigured to vary resolution as 4-bit, 5-bit, and 6-bit to reduce power for wireless applications by just keeping unused comparators in standby mode depending upon the peak of input signal [10]. Also, flash ADC is reconfigured using folding technique for wireless application [11]. Flash ADC is reconfigured using interleaving technique [12]. Pipeline ADC is reconfigured by generating control signal which turns off unused pipeline stages [13, 14]. Also pipeline ADC has been reconfigured based on the amplitude and frequency of the input signal [15]. SAR ADC has been reconfigured by configuring the resolution of digital to analog converter (DAC) and SAR logic [16, 17]. Most of these reconfigurable ADCs are configured for specific applications. In conventional methods, high-frequency limitations are overcome by connecting SAR ADCs in pipeline mode called SAR-assisted pipeline ADC to increase the resolution for high sampling rate application [18, 19].

Figure 1 shows conventional flash and SAR ADC. The flash ADC internally consists resistive string, comparators, and priority encoder. As resolution N increases, the number of resistors and comparators also increases proportional to 2^N-1. So, flash ADC resolution is limited to maximum 8-bit. Figure 1 also shows conventional SAR ADC which internally consists comparator, SAR logic, and DAC. The Vin signal is

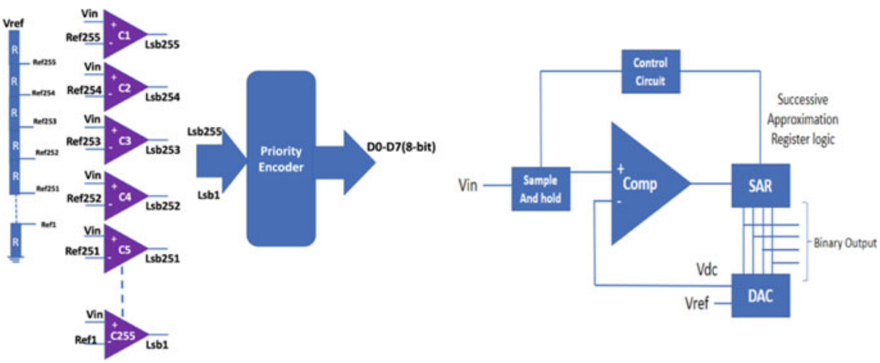


Fig. 1 Conventional flash and SAR ADC [2, 6]

sampled and held constant for conversion. Initially, SAR logic block sets the MSB bit to logic 1 and DAC output to $V_{\mathrm{ref}}/2$. This DAC output is then compared with Vin. If Vin is greater than the DAC output, the comparator output decides MSB bit to 1 and SAR logic set, next LSB bit, and again it is compared with Vin. Then, depending upon the comparator output, the SAR logic decides all the bits to logic 1 or 0. In this SAR ADC, as resolution increases, the conversion time increases linearly, so it is limited to maximum 18-bit. The main reason of speed limitation in SAR ADC is DAC settling time. For example, if the total conversion time of 16-bit SAR ADC is 16us then for 1-bit SAR conversion it is 1us. Conventional SAR ADC gives fixed resolution and fixed conversion time.

Different SAR ADCs for different resolutions and sampling rates are required to be designed for specific applications. Moreover, when the small conversion time is required for high-frequency input signal, the highest possible resolution for this frequency should to be obtained. For this purpose, the maximum rate of change in input signal should be automatically estimated and utilized for deciding maximum possible resolution of ADC. The new method proposed in this paper is based on the principle of deciding the maximum resolution for a given range of maximum input signal frequency which is decided by a circuit designed for this purpose. With this method, ADC can be reconfigured to achieve relatively wide range of applications as shown in Fig. 2. It is a combination of 8-bit flash ADC and reconfigurable 8-bit SAR ADC which is known as hybrid ADC. The unique approach here to reconfigure ADC is that it depends upon the maximum rate of change of input signal itself. This maximum input signal frequency itself decides the resolution of proposed ADC from 8-bit to 16-bit continuously with 1-bit step such as 8-bit, 9-bit, 10-bit, 11-bit, 12-bit, 13-bit, 14-bit, 15-bit, and 16-bit. If the input signal frequency is higher, then it configures the resolution of proposed ADC as 8-bit which is the output of 8-bit flash ADC directly. If the maximum input signal frequency is reduced to half of the maximum higher frequency, then it configures the resolution as 9-bit, and again if the maximum input signal frequency is reduced to one fourth of the maximum higher frequency, then it configures the resolution as 10-bit and similarly up to the resolution of 16-bit. This is how maximum the rate of change of input signal

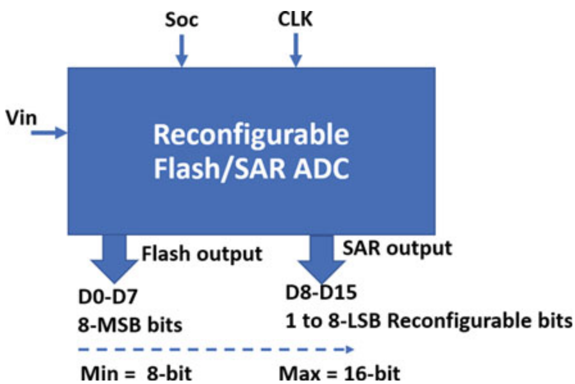


Fig. 2 Hybrid flash and SAR ADC

decides the resolution of proposed ADC which can achieve relatively wide range of applications such as video, data acquisition, and measurement.

New Proposed Design

Figure 3 shows a block diagram of proposed hybrid reconfigurable ADC which consists of a flash ADC with 8-bit fixed resolution and SAR ADC whose resolution is controlled from 1 to 8-bit depending on the maximum input signal frequency as explained above. It consists of sample and hold circuit, 8-bit flash ADC, 16-bit DAC, comparator, and 8-bit SAR logic circuit. For estimation of resolution at highest input signal frequency, the estimation circuit has been designed which consists of differentiator and peak detector. The output of this circuit is digitized using the same flash ADC, the output of which is given to the combinational logic circuit. The output of the combinational circuit is a digital code which is utilized to control the resolution of SAR ADC. This proposed design reconfigures the resolution of 8-bit SAR ADC from 1-bit to 8-bit depending upon the maximum frequency of input signal. As maximum frequency of input signal increases, the resolution of SAR ADC decreases from 8-bit to 1-bit, then in further increase in frequency of input signal, the SAR ADC is made inactive and the 8-bit flash ADC is useful for high-frequency conversions.

The combinational logic circuit output controls the interconnection of 8-bit SAR ADC resolution from 1-bit to 8-bit with 1-bit step increment continuously up to 8-bit. Initially, before the start of conversion, input signal to the differentiator is adjusted to full scale (10 V) and highest of the input signal frequency which depends upon conversion time of SAR ADC.

Then, the peak detector detects the peak of the differentiator output. This peak of differentiator output is then amplified to obtain the full scale(10 V) at the output for

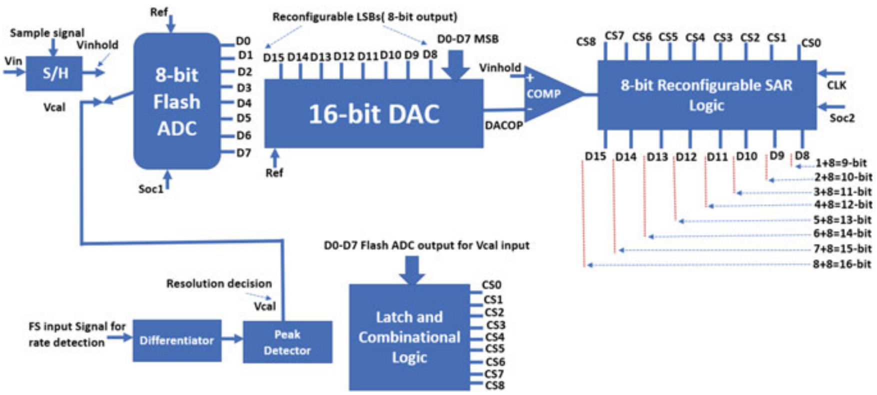


Fig. 3 Block diagram of proposed design

which gain of differentiator is adjusted and the V_{cal} amplified differentiated output is converted by 8-bit flash ADC to generate code D0–D7. This D0–D7 code is latched and given to a combinational logic circuit to generate control signals CS0–CS8 which decides the resolution of SAR logic. This peak amplitude of differentiated output is linearly proportional to the maximum input signal frequency for which the resolution of ADC is configured.

Input Signal Frequency Detector

Figure 4 shows maximum input signal frequency detector which consists of differentiator, peak detector, and amplifier.

The output of differentiator circuit is given as

$$V_o = -R_f C_1 (\mathrm{d}V_{in}/\mathrm{d}t) \quad (1)$$

The gain of differentiator circuit is given as

$$\text{Gain} = -R_f/X_{C1} \quad (2)$$

If the input signal to the differentiator is

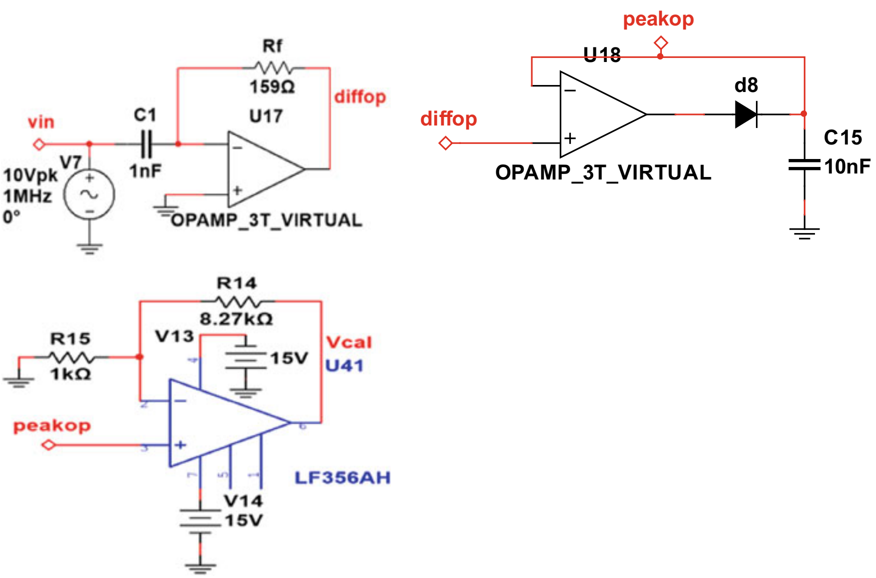


Fig. 4 Differentiator as input signal frequency detector

$$\mathrm{V_{in}} = V_m \sin\omega t \tag{3}$$

Then corresponding output of differentiator is given as

$$V_o = -R_f C_1 V_m \omega \cos\omega t \tag{4}$$

The maximum value given by the peak detector circuit is

$$\text{peakop} = R_f C_1 V_m \omega \tag{5}$$

$$= R_f C_1 V_m \, (2\pi f) \tag{6}$$

The output of amplifier with a gain G

$$V_{\text{cal}} = (2\pi f)\, G R_f C_1 V_m \tag{7}$$

It is clear from Eq. (7) that the values of G, R_f or C_1 can be adjusted for a given maximum frequency such that the value V_{cal} is adjusted to be equal to the value of the full-scale voltage of flash ADC. This maximum frequency is decided by conversion time of SAR ADC and flash ADC depending upon their conversion times.

Equation 4 shows that for the same V_m of the input signal, the output of differentiator is linearly proportional to frequency of the input signal. As frequency increases, the capacitive impedance decreases and gain of the differentiator (Rf/XC1) increases. Consider $V_m = 10$ V for input signal. Figure 5 shows the linear relation between maximum input signal frequency fin and differentiator output diffop. As fin is reduced to half, then the peak of differentiator output also reduces to half.

Conversion Process

After receiving sample and hold command, the Vin signal is held constant (Vinhold), and it is converted by 8-bit flash ADC to produce output D0-D7 and depending upon the maximum frequency of input signal the control signals CS0 to CS8 then decides the resolution of 8-bit SAR ADC by interconnecting the SAR logic blocks. The input signal which is held constant (Vinhold) is converted by the required resolution. Table 1 shows maximum input signal frequency and corresponding resolution for a typical SAR ADC with maximum resolution of 16-bit and conversion time of 16 us. This shows that as the maximum input signal frequency reduced to half, the resolution increases by 1-bit. For 1 MHz maximum input frequency, the SAR ADC resolution is 0-bit and total resolution of the system is 8-bit as decided by flash ADC which produces digital output as D0–D7 and reconfigurable 8-bit SAR ADC will be unused. As maximum input signal frequency is reduced to half, i.e., 500 kHz, then this rate is detected by the differentiator and it sets the resolution of reconfigurable 8-bit SAR

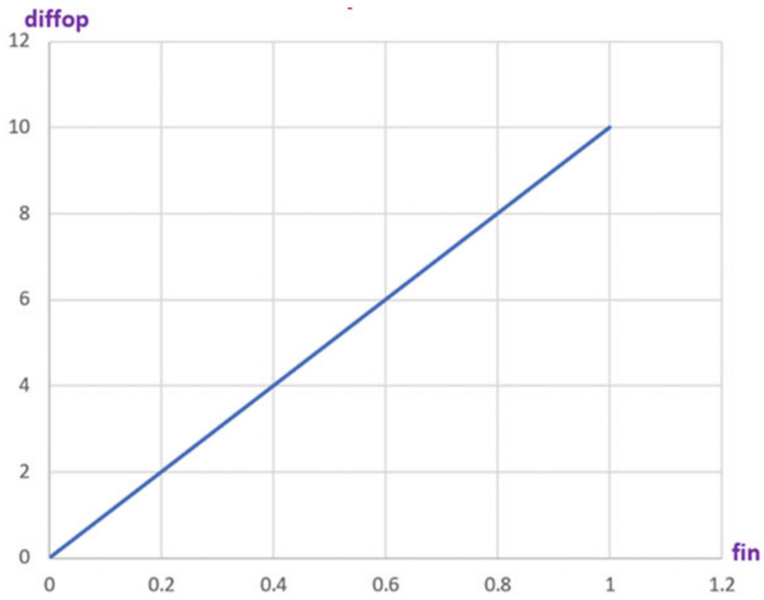


Fig. 5 Differentiator output diffop Vs maximum input signal frequency fin

Table 1 Input signal frequency and resolution of ADC

Maximum input signal frequency	SAR resolution	Total resolution
500 kHz–100 MHz	0	8-bit
500 kHz	1	9-bit
250 kHz	2	10-bit
125 kHz	3	11-bit
62.5 kHz	4	12-bit
31.25 kHz	5	13-bit
15.625 kHz	6	14-bit
7.8125 kHz	7	15-bit
3.90625 kHz	8	16-bit

ADC as 1-bit through control signal CS0 to CS8, then total resolution of the system is 9-bit (8-bit flash ADC D0–D7 MSB bits + 1-bit SAR ADC D8 LSB bit).

Similarly, when the maximum input signal frequency is reduced to 250 kHz, then the resolution of reconfigurable SAR ADC is 2-bit through control signals, then total resolution of the system becomes 10-bit (8-bit flash ADC D0–D7 MSB bits +2-bit SAR ADC D8–D9 LSB bits). As maximum input signal frequency is reduced to 125 kHz, then sets the resolution of reconfigurable 8- bit SAR ADC as 3-bit through control signals, then total resolution of the system is 11-bit (8-bit flash ADC D0–D7

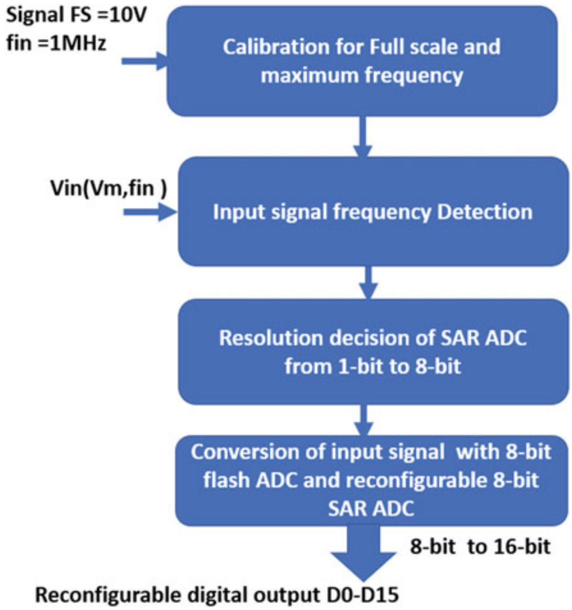


Fig. 6 Process of hybrid reconfigurable ADC

MSB bits +3-bit SAR ADC D8–D10 LSB bits). The process continues up to total resolution of 16-bit as given in Table 1.

By incorporating 8-bit flash ADC in the design, it helps to reduce the total conversion time of hybrid ADC. Figure 6 shows the flowchart for the process of hybrid reconfigurable ADC.

Generation of Control Signal for Reconfigurability

Figure 7 shows combinational logic circuit that has been designed and simulated to generate control signals for reconfigurability CS0, CS1, CS2, CS3, CS4, CS5, CS6, CS7, and CS8. The differentiator output after amplification Vcal is converted using 8-bit flash ADC, and digital code corresponding to this output D0–D7 is given to a combinational logic circuit implemented using logic gates. These control signals CS0 to CS8 activate the stages of SAR logic block to decide the resolution of SAR ADC.

Table 2 shows flash ADC latched output cm1 to cm8 and SAR logic block stages S1 to S8 for different maximum frequencies of input signal. For rate of input signal, i.e., 500 kHz, the cm1 to cm8 code activates only first stage S1 of the SAR logic block which decides the resolution of SAR ADC as 1-bit, so total resolution is 9-bit. The process conversion is implemented up to 16-bit resolution as described earlier.

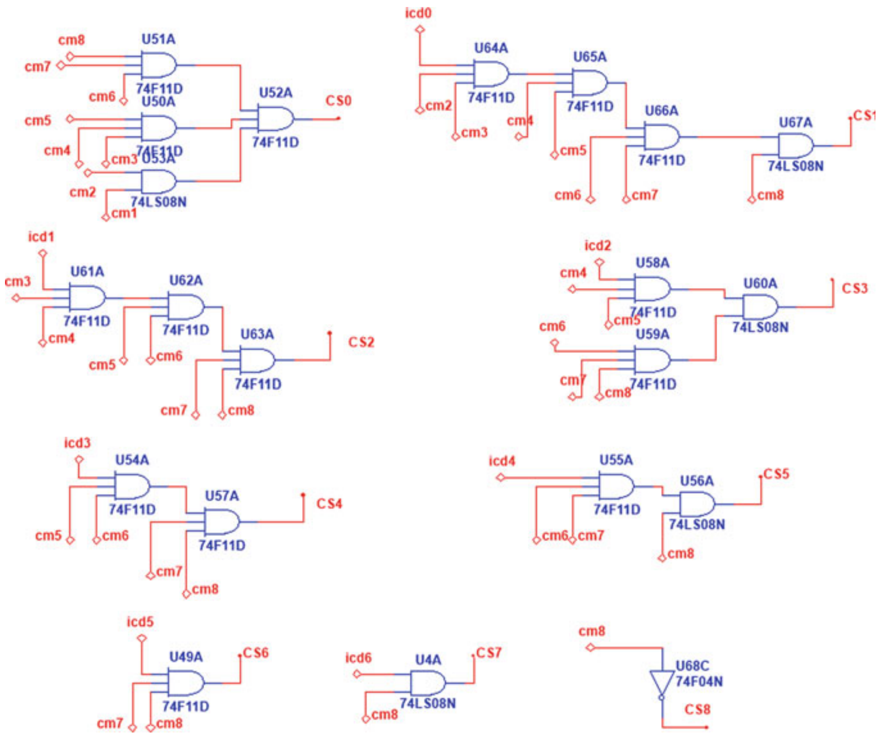


Fig.7 Generation of control signal

Timing Diagram of Reconfigurable ADC

Figure 8 shows conversion timing diagram for various resolutions of ADCs such as from 8-bit to 16-bit, and Table 3 shows the same relationship in tabular form. It shows as resolution increases the conversion time of reconfigurable ADC also increases.

Results

The reconfigurable 8-bit to 16-bit data is presented in Table 4. This Table 4 shows the simulation results for different frequency range, fixed peak value of V_{in}(10 V), and corresponding resolution of SAR ADC. The circuit has been simulated for maximum input voltage range of 0–10 V. For input signal with a frequency of more than 500 kHz up to 100 MHz, only 8-bit flash ADC is active, and reconfigurable SAR ADC is unused, so total resolution of conversion is 8-bit. For input signal (500 kHz, 10 V), 8-bit flash ADC is active and first stage of reconfigurable SAR logic block is active, so total resolution of conversion is 8-bit + 1-bit = 9-bit. Similarly, it has been

Table 2 Maximum input signal frequency and active stages of SAR logic block

Input signal frequency (kHz)	Cm1	Cm2	Cm3	Cm4	Cm5	Cm6	Cm7	Cm8	S1	S2	S3	S4	S5	S6	S7	S8
500	0	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0
250	0	0	1	1	1	1	1	1	1	1	0	0	0	0	0	0
125	0	0	0	1	1	1	1	1	1	1	1	0	0	0	0	0
62.5	0	0	0	0	1	1	1	1	1	1	1	1	0	0	0	0
31.25	0	0	0	0	0	1	1	1	1	1	1	1	1	0	0	0
15.625	0	0	0	0	0	0	1	1	1	1	1	1	1	1	0	0
7.8125	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	0
3.90625	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1

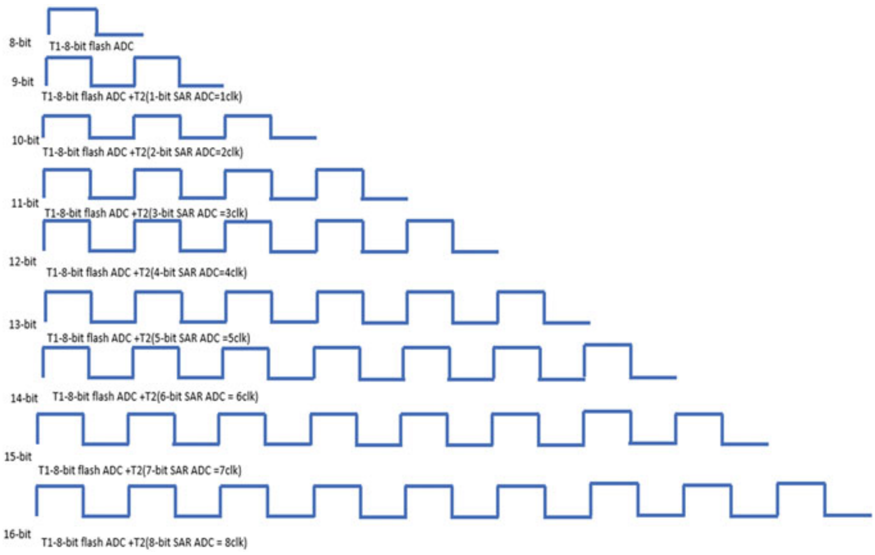


Fig. 8 Conversion time of reconfigurable ADC

Table 3 Resolution and conversion time of ADCs

Resolution	Conversion time
8-bit	Conversion time of 8-bit flash ADC ($T1$ < ns)
9-bit	$T1$ + SAR 1-bit ($T2$)
10-bit	$T1 + 2\,T2$
11-bit	$T1 + 3T2$
12-bit	$T1 + 4\,T2$
13-bit	$T1 + 5\,T2$
14-bit	$T1 + 6\,T2$
15-bit	$T1 + 7\,T2$
16-bit	$T1 + 8\,T2$

verified that as maximum frequency of input signal decreases linearly, the stages of reconfigurable SAR logic blocks become active and the resolution of SAR ADC increases which is shown in bold letter.

Conclusion

The new hybrid reconfigurable flash SAR ADC combines features of the conventional 8-bit flash ADC and the reconfigurable 8-bit SAR ADC useful for high speed and high-resolution applications. This design has been simulated and validated in

Table 4 Results

Maximum frequency fin	Peak value of input signal	Resolution		8-bit flash ADC output								Reconfigurable SAR ADC output							
		Flash	SAR	D0	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	D11	D12	D13	D14	D15
1 MHz	10 V	8-bit	0-bit	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0
500 kHz	10 V	8-bit	1-bit	1	1	1	1	1	1	1	1	**1**	0	0	0	0	0	0	0
250 kHz	10 V	8-bit	2-bit	1	1	1	1	1	1	1	1	**1**	**1**	0	0	0	0	0	0
125 kHz	10 V	8-bit	3-bit	1	1	1	1	1	1	1	1	**1**	**1**	**1**	0	0	0	0	0
62.5 kHz	10 V	8-bit	4-bit	1	1	1	1	1	1	1	1	**1**	**1**	**1**	**1**	0	0	0	0
31.25 kHz	10 V	8-bit	5-bit	1	1	1	1	1	1	1	1	**1**	**1**	**1**	**1**	**1**	0	0	0
15.625 kHz	10 V	8-bit	6-bit	1	1	1	1	1	1	1	1	**1**	**1**	**1**	**1**	**1**	**1**	0	0
7.8125 kHz	10 V	8-bit	7-bit	1	1	1	1	1	1	1	1	**1**	**1**	**1**	**1**	**1**	**1**	**1**	0
3.90625 kHz	10 V	8-bit	8-bit	1	1	1	1	1	1	1	1	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**

This ‘1’ in bold indicates the LSB digital output of 8-bit reconfigurable SAR ADC for different resolutions decided by the input signal frequency while the remaining bits are mentioned as ‘0’. These remaining ‘0’ bits indicate that the stages of the SAR logic block are inactive.

NI Multisim 14.1, which proves the functionality of the proposed concept. The maximum input signal frequency decides the resolution of conversion. This reconfigurable ADC resolution varies from 8-bit to 16-bit in steps of 1-bit increment, and because of this, it is useful for many applications as compared to conventional SAR ADC. As resolution increases, conversion time of ADC increases linearly. This can be extended to further 24-bit by replacing the 8-bit SAR logic by 16-bit and 16-bit DAC by 24-bit, so it can cover wide range of applications.

References

1. Lin H, Wen K (2017) A low power reconfigurable SAR ADC for CMOS MEMS sensor. In: International SoC design conference (ISOCC), pp 7–8
2. Liu J, Liu S, Ding R, Zhu Z (2021) A conversion mode reconfigurable SAR ADC for multistandard systems. IEEE Trans Very Large Scale Integr (VLSI) Syst 29(5):895–903
3. Wang J, Hua Y, Zhu Z (2019) A 10-bit reconfigurable ADC with SAR/SS mode for neural recording. Analog Integr Circ Sig Process 101:297–305
4. Li S, Chen J, Liang B, Guo Y (2020) Low power SAR ADC design with digital background calibration algorithm. Symmetry 12:1757
5. Shen J, Shikata A, Fernando LD, Guthrie N, Chen B, Maddox M, Mascarenhas N, Kapusta R, Coln MC, A 16-bit 16-MS/s SAR ADC with on-chip calibration in 55-nm CMOS
6. Kester W (2021) MT-020: ADC architectures I: the flash converter. Analog devices. https://www.analog.com/media/en/training-seminars/tutorials/MT-020.pdf
7. Kester W (2021) MT-022: ADC architectures III, sigma-delta ADC basics. Analog devices. https://www.analog.com/media/en/training-seminars/tutorials/MT-022.pdf
8. Kester W, Bryant J (2021) MT-027: ADC architectures VIII: integrating ADCs ADC architectures VIII: integrating ADCs. https://www.analog.com/media/en/training-seminars/tutorials/MT-027.pdf
9. Kester W (2021) MT-024: ADC architectures V: pipelined subranging ADCs, analog devices. https://www.analog.com/media/en/training-seminars/tutorials/MT024.pdf
10. Veeramachanen S, Kumar AM, Tummala V, Srinivas MB (2009) Design of a low power, variable-resolution flash ADC. In: 22nd International conference on VLSI design, pp 117–122
11. Wang L, LaCroix M, Carusone AC (2017) A 4-GS/s single channel reconfigurable folding flash ADC for wireline applications in 16-nm FinFET. IEEE Trans Circ Syst II Express Briefs 64(12):1367–1371
12. R. Yousry, H. Park, E. Chen, C. K. Yang: A digitally-calibrated 10GS/s reconfigurable flash ADC in 65-nm CMOS. In: IEEE international symposium on circuits and systems (ISCAS), pp 2443–2447
13. Anderson M, Norling K, Dreyfert A, Yuan J (2005) A reconfigurable pipelined ADC in 0.18 /spl mu/m CMOS. In: Digest of technical papers: symposium on VLSI circuits, pp 326–329
14. Taherzadeh-Sani M, Hamoui AA (2010) A reconfigurable 10–12b 0.4–44MS/s pipelined ADC with 0.35–0.5pJ/step in 1.2V 90 nm digital CMOS. In: Proceedings of ESSCIRC, pp 382–385
15. Ramakrishna P, Kishore KH (2021) A low power reconfigurable ADC for bioimpedance monitroing system. Int J Speech Technol
16. Hu Y, Chen L, Chen H, Wen Y, Zhang H, Liu X (2019) A 100 KS/s 8–10 bit resolution-reconfigurable SAR ADC for bioelectronics application. In: 32nd IEEE International System-On-Chip Conference (SOCC), pp 209–212

17. Yip M, Chandrakasan AP (2011) A resolution-reconfigurable 5-to-10b 0.4-to-1V power scalable SAR ADC. In: IEEE international solid-state circuits conference, pp 190–192
18. Flynn MP, Choo K, Lim Y (2018) Pipeline and SAR ADCs for advanced nodes. In: Harpe P, Makinwa K, Baschirotto A (eds) Hybrid ADCs, smart sensors for the IoT, and sub-1V & advanced node analog circuit design. Springer, Cham
19. The Evolution of SAR ADCs for high sampling rate applications (2019) https://www.dialog-semiconductor.com/sites/default/files/evolution-of-sar-adc.pdf

Use of Speech Analysis in Determining Human Affect: A Proof of Concept

Suja Panicker, Nayan Kakuste, Pushkraj Darpel, and Piyush Dashpute

Abstract Human affect is an important consideration in several cognitive and AI-driven applications. Efficient and timely determination of affect is crucial in detecting various physiological and psychological illnesses affecting humans globally. Cognitive science advocates the use of different modalities such as voice, text, images, and gestures in initiating/promoting human cognition. Thus, multimodality is an inherent nature of cognition. In this work, we have performed experimentation on real-time voice samples collected from four participants and attempted to determine patterns of affect. With suitable use of spectrogram, chroma features, beats, etc., we present various visualizations that depict the inter-individual acoustic differences.

Keywords Affective computing · Speech emotion recognition · Artificial intelligence · Speech analysis · Data visualization · Cognitive science

Introduction

Human affect determination has always intrigued scientists and researchers globally, and the current pandemic has played an ever increasing role in promoting automated/semi-automated systems for affect detection. Predominantly affect has been classified into positive or negative; however, several works are based on Ekman's model [1] of a 5 class problem—angry, sad, joy, disgust, and neutral. An extensive research presented in [2] illustrates the correlation of physiological features (such as blood pressure, heart rate, voice, and pulse.) with human affect. To quote an example, it was noted that increased energy for the speech features accompanied by rapid fluctuations at basic frequency indicates negative affect [3].

Finding patterns in humans through their physiological makeup has always been an interesting research problem [4], and this is the basic motivation for current work wherein we propose to experiment with voice—an important physiological modality in determining human affect.

S. Panicker (✉) · N. Kakuste · P. Darpel · P. Dashpute
School of CET, MIT World Peace University, Pune, Maharashtra, India
e-mail: suja.panickar@mitwpu.edu.in

V. Bindhu et al. (eds.), *Proceedings of Third International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 844,
https://doi.org/10.1007/978-981-16-8862-1_31

This paper is organized as—related work covered in Sect. "Related Work", proposed work covered in Sect. "Proposed Work and Methodology", result analysis is presented in Sect. "Result Analysis", and conclusion covered in Sect. "Conclusion".

Related Work

Owing to the seminal contribution [5], affective computing has been an evergreen area of interest to global researchers, thereby directly impacting society. Deciphering human affect from various media such as tweets [6–8], speech [9], health-related textual posts [10, 11], and images [12] is gaining tremendous attention. In current work, we emphasize on investigating analysis of speech on the human affect. Highlights of latest research are as follows:

Most speech emotion recognition-based systems experiment with following classes of features—prosodic, spectral, voice quality-based, and Teager energy operator-based [13]. Several researchers have worked with wide variety of speech emotions including—happy, rest, sad, surprise, empathetic, fear, disgust, neutral, boredom, etc. [13].

In contrast to most works focussing on analysis of prosodic features [14] more emphasis is required on additional novel features based on intensity and pitch contour along with contextual information. Promising results are obtained on CVRRCar-AVDB and EMO-DB datasets. In [15], an automated emotion detection system that is independent of the lexical or language-based deviations between speech produced during anger, sadness, happiness versus neutral is explored.

In [2], a detailed survey of artificial intelligence-based techniques used by current researchers in the automated detection of human negative affect is presented, along with details of the wide variety of gadgets used in acquisition of physiological data.

Machine learning and deep learning techniques are popularly used in speech research. However, research based on transfer learning and personalization of speech [16], meticulous analysis of speech emotion recognition techniques, datasets used, strategies and trends, and pertinent shortcomings [17] is worth mentioning.

It is noted that various physical, mental ailments could directly affect the human speech, also psychological tension [18] may alter speech for specific time spans. This is an important research gap to be addressed during data acquisition phase. Determining and analyzing the inter-individual and intra-individual acoustic differences under various affect is another consideration.

Speech-based affect detection has various real-time applications such as smart home assistants [19], mobile applications [20], and social media analysis [21–23].

Proposed Work and Methodology

Figure 1 illustrates proposed work.

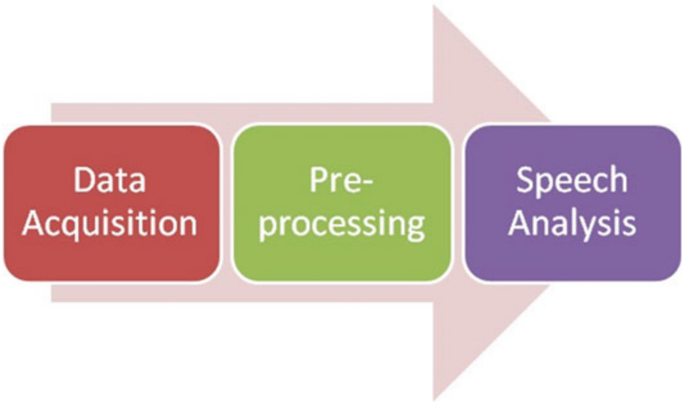


Fig. 1 Proposed work

As illustrated in Fig. 1, the three major modules of current work are real-time data acquisition, preprocessing, and noise removal, followed by speech analysis.

Dataset description: In current work, we collected real-time voice samples of four individuals who consented to participate in this study. The participants are from different demographics and are youngsters in the age group of 20 to 22. To keep proper baselines, we suggested use of the same textual content. These voice samples were collected in real-time experimental setup. To add dynamics to the research, these samples were collected in three different tones—happy, neutral, and sad. Experimental work as described in further section is the outcome of analysis of 12 voice samples. We used Librosa library [24] from Python for basic speech processing and visualizations.

Result Analysis

Experimental results of analyzing the 12 voice samples and the various visualizations obtained thereof are presented herewith.

Figure 2 represents the amplitude and how it varies with respect to time.

Figure 2 depicts that the graph produces high values of amplitude during patches of high audio level and subsequently produces low values where the audio level is low.

All the frequencies present in the audio sample and its magnitude are presented in the power spectrum depicted in Fig. 3

Figure 3 displays coefficients for each frequency.

Figure 4 represents a spectrogram comprising all the frequencies present in the audio sample and its magnitude that changes with respect to time.

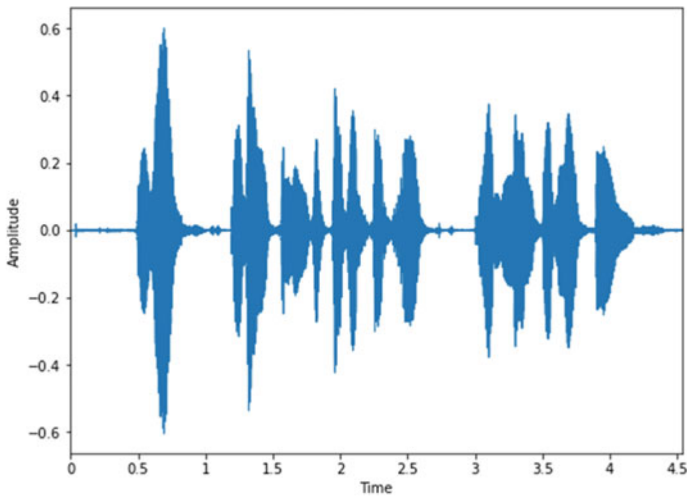


Fig. 2 Voice waveform with respect to amplitude

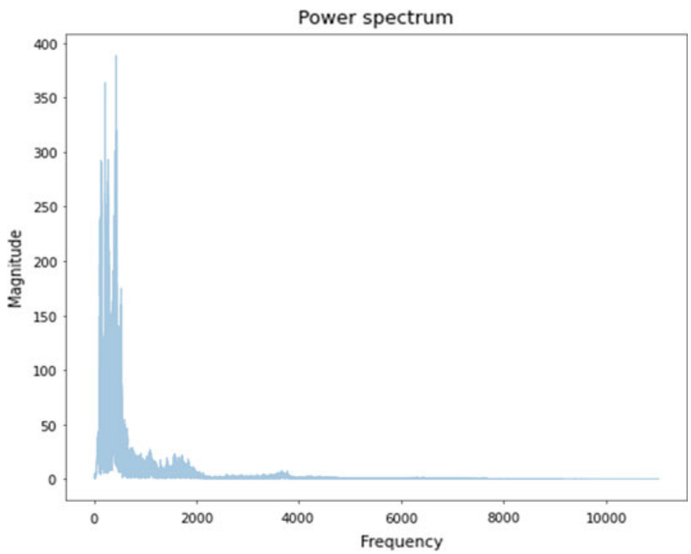


Fig. 3 Power spectrum

The spectrogram of Fig. 4 displays the strength of an emotion over time at various frequencies of a waveform. Mel frequency cepstral coefficient is considered an important marker in acoustic studies [25–27].

$$M(f) = 1125 \ln\left(1 + \frac{f}{700}\right) \tag{1}$$

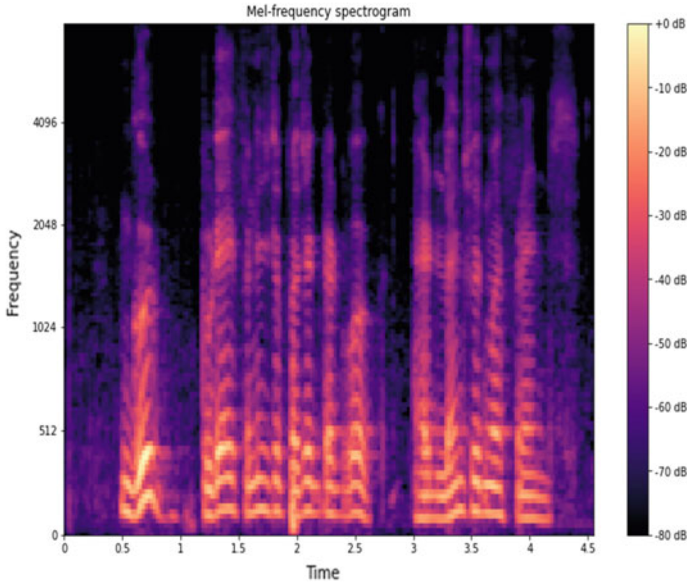


Fig. 4 Spectrogram of four audio samples

In Eq. (1) [25], f = perceived frequency.

We used MFCC feature extraction which includes windowing, DFT, log of magnitude, warping, and inverse DCT. Figure 5 illustrates the role of MFCC in audio extraction.

Figure 5 represents the changing value of MFCC coefficients with respect to time while also considering their respective magnitude.

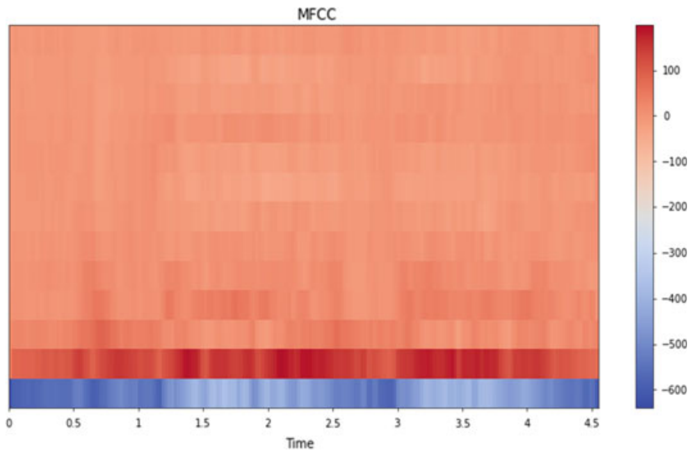


Fig. 5 Role of MFCC

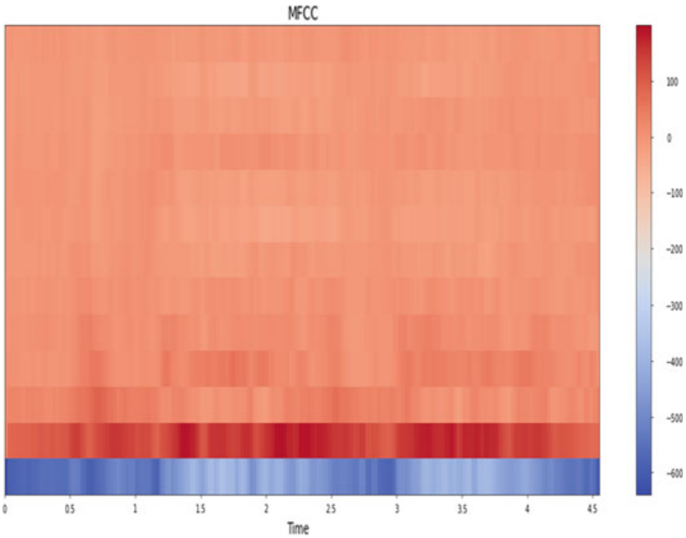


Fig. 6 Harmonic sounds

We also acquired audio samples (songs) from the Music4All dataset [28] and performed feature extraction using the Librosa library to get the following results.

Further we investigated the role of harmonic sounds in Fig. 6. These are the speech signals which one perceives to have certain pitch (example—the one observed during loud singing). The sound from violin is a standard example of harmonic sound. Percussive sounds are the ones which are generated from two crashing objects. To quote another example, note played on piano has a percussive tone (resulting from the sound of hammer that hits strings) preceding the harmonic tone (resulting from the vibrating string).

In Fig. 7, we investigate the beats aspect of the audio samples. Here, the samples are plotted with beat time values—beat time difference and beats per seconds. It is observed from Fig. 7 that audio samples contain high beats at the start and hence we see darkening of the shade from left to right.

From Fig. 7, it is observed that output of beat tracker is estimate of tempo. It is array of the frame numbers and corresponds to the determined beat events.

We further employed chroma energy normalized representation of magnitude of the respective pitch with changing time (Fig. 8).

Figure 8 depicts the 12 element-based representation of spectral energy. Here, bins represent the pitch classes of western music. Further we calculated chroma mean and standard deviation, and these are illustrated in Fig. 9.

It is noted from Fig. 9 that the output of CENS contains the 12 element representation of chroma energy throughout the song. All these elements are summarized using mean and standard deviation across the entire song.

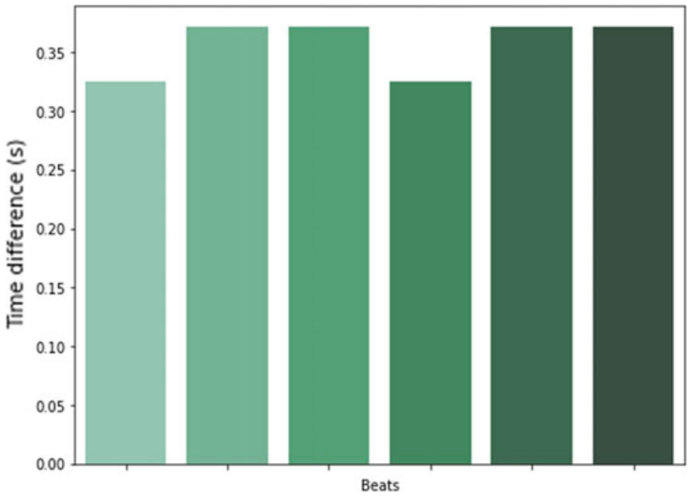


Fig. 7 Investigating the beats

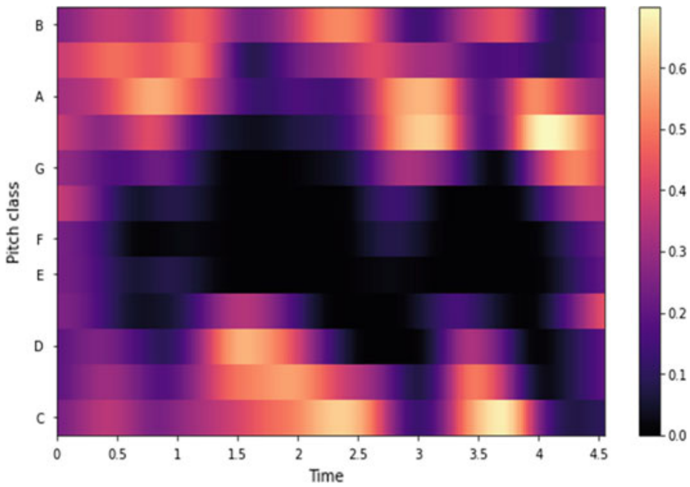


Fig. 8 Visualization of chroma energy

Spectral centroid is the measure used to characterize the spectrum. It indicates center of mass for spectrum [27].

$$\text{Centroid} = \sum_{n=0}^{N-1} f(n)x(n) / \sum_{n=0}^{N-1} x(n) \tag{2}$$

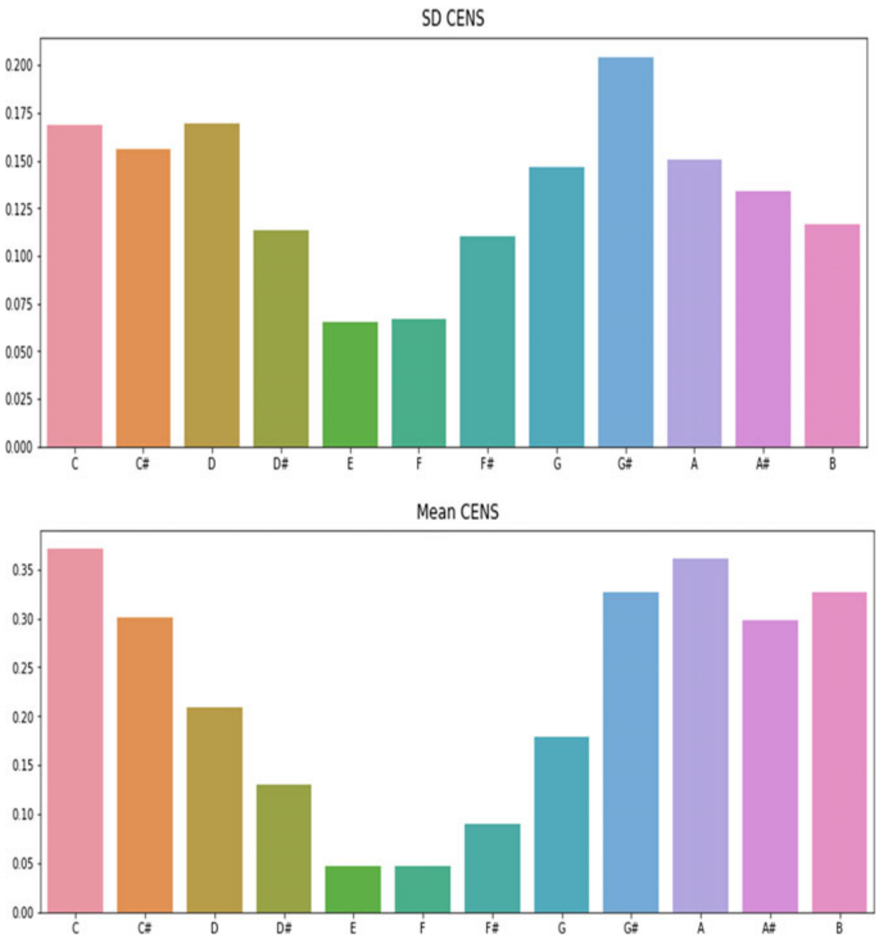


Fig. 9 Enhancing the chroma features

In Eq. (2) [25], $x(n)$ is the weighted frequency of bin n, and $f(n)$ is the center frequency of that bin.

We have depicted spectral centroid in Fig. 10.

As observed from Fig. 10, spectral centroid has strong connection with the "brightness"-related impression of sound.

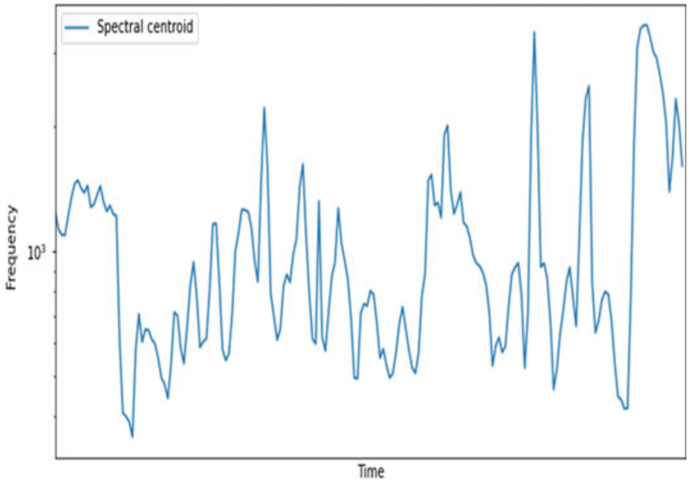


Fig. 10 Spectral centroid of voice samples

Conclusion

Thus, we have performed voice analysis on 12 real-time speech samples. We also experimented with voice samples from Music4All—a newly published and benchmarked dataset in the field of affective computing. Intra-individual acoustic differences hold tremendous knowledge about the individuals and can be further mined for behavioral analysis. Our experimental results suggest that speech can be successfully used as a modality in determining affect, and also, it was observed that the affect happy has higher amplitude than the affect sad. Future work shall include performing user profiling based on the acoustic physiological features.

References

1. Ekman P (1999) Basic emotions. Handbook of cognition and emotion. Wiley
2. Panicker S, Prakasam (2019) A survey of machine learning techniques in physiology based mental stress detection systems. Biocybernetics Biomed Eng 39(2). https://doi.org/10.1016/j.bbe.2019.01.004
3. Sharma N, Gedeon T (2012) Objective measures, sensors and computational techniques for stress recognition and classification. Comput Methods Prog Biomed 108(3):1287–1301
4. Panicker S, Bhujange S, Karne S, Kadam R (2016) Finding patterns in biological parameters. Int J Recent Innovation Trends Comput Commun 4(12)
5. Picard RW, Affective computing. M.I.T media laboratory perceptual computing section technical report no. 321
6. Kunte AV, Panicker SS (2019) Using textual data for personality prediction: a machine learning approach. In: 2019 4th International conference on information systems and computer networks (ISCON)

7. Kunte A, Panicker S (2020) Personality prediction of social network users using ensemble and XGBoost. In: Das H, Pattnaik P, Rautaray S, Li KC (eds) Progress in computing, analytics and networking. Advances in intelligent systems and computing, vol 1119. Springer, Singapore
8. Ahire K, Bagul M, Dhanawate S, Panicker S. A novel proof of concept for twitter analytics using popular hashtags: experimentation and evaluation. In: Goyal V et al. (eds) Proceedings of international conference on communication and artificial intelligence, lecture notes in networks and systems 192. https://doi.org/10.1007/978-981-33-6546-9_31
9. Manoharan S, Ponraj N (2020) Analysis of complex non-linear environment exploration in speech recognition by hybrid learning technique. J Innovative Image Process (JIIP) 2(04):202–209. https://doi.org/10.36548/jiip.2020.4.005
10. Mane VL, Panicker SS, Patil VB (2015) Summarization and sentiment analysis from user health posts. Int Conf Pervasive Comput (ICPC) 2015:1–4. https://doi.org/10.1109/PERVASIVE.2015.7087087
11. Mane VL, Panicker SS, Patil VB (2015) Knowledge discovery from user health posts. 2015 IEEE 9th International conference on intelligent systems and control (ISCO). pp 1–6. https://doi.org/10.1109/ISCO.2015.7282288
12. Salunke V, Panicker SS (2021) Image sentiment analysis using deep learning. In: Ranganathan G, Chen J, Rocha Á (eds) Inventive communication and computational technologies. Lecture notes in networks and systems. Springer, Singapore, p 145. https://doi.org/10.1007/978-981-15-7345-3_12
13. Wani TM, Gunawan TS, Qadri SAA, Kartiwi M, Ambikairajah E (2021) A comprehensive review of speech emotion recognition systems. IEEE Access 9:47795–47814. https://doi.org/10.1109/ACCESS.2021.3068045
14. Tawari A, Trivedi MM (2010) Speech emotion analysis: exploring the role of context. IEEE Trans Multimedia 12(6):502–509. https://doi.org/10.1109/TMM.2010.2058095
15. Kadiri SR, Alku P (2020) Excitation features of speech for speaker-specific emotion detection. IEEE Access 8:60382–60391. https://doi.org/10.1109/ACCESS.2020.2982954
16. Vryzas N, Vrysis L, Kotsakis R, Dimoulas C (2021) A web crowdsourcing framework for transfer learning and personalized speech emotion recognition. Mach Learn Appl 6:100132. ISSN 2666-8270. https://doi.org/10.1016/j.mlwa.2021.100132
17. Prasanth S, Thanka MR, Edwin EB, Nagaraj V (2021) Speech emotion recognition based on machine learning tactics and algorithms. Mater Today: Proc. ISSN 2214-7853. https://doi.org/10.1016/j.matpr.2020.12.207
18. Panicker S, Gayathri P (2020) Investigating the role of machine learning in detecting psychological tension: a brief review. In: Proceeding 2020. International conference on communication and artificial intelligence (ICCAI-2020), Mathura, Uttar Pradesh. https://doi.org/10.1007/978-981-33-6546-9_32
19. Chatterjee R, Mazumdar S, Sherratt RS, Halder R, Maitra T, Giri D (2021) Real-time speech emotion analysis for smart home assistants. IEEE Trans Consum Electron 67(1):68–76. https://doi.org/10.1109/TCE.2021.3056421
20. Hossain MS, Muhammad G (2017) An emotion recognition system for mobile applications. IEEE Access 5:2281–2287. https://doi.org/10.1109/ACCESS.2017.2672829
21. Son LH, Kumar A, Sangwan SR, Arora A, Nayyar A, Abdel-Basset M (2019) Sarcasm detection using soft attention-based bidirectional long short-term memory model with convolution network. IEEE Access 7:23319–23328. https://doi.org/10.1109/ACCESS.2019.2899260
22. Sivaganesan D (2021) Novel influence maximization algorithm for social network behavior management. J ISMAC 3(01):60–68. https://doi.org/10.36548/jismac.2021.1.006
23. Valanarasu MR (2021) Comparative analysis for personality prediction by digital footprints in social media. J Inf Technol 3(02):77–91
24. https://librosa.org › doc
25. Qamhan MA, Alotaibi YA, Seddiq YM, Meftah AH, Selouani SA (2021) Sequence-to-sequence acoustic-to-phonetic conversion using spectrograms and deep learning. IEEE Access 9:80209–80220. https://doi.org/10.1109/ACCESS.2021.3083972

26. Filipe L, Peres RS, Tavares RM (2021) Voice-activated smart home controller using machine learning. IEEE Access 9:66852–66863. https://doi.org/10.1109/ACCESS.2021.3076750
27. Verde L, De Pietro G, Ghoneim A, Alrashoud M, Al-Mutib KN, Sannino G (2021) Exploring the use of artificial intelligence techniques to detect the presence of coronavirus Covid-19 through speech and voice analysis. IEEE Access 9:65750–65757. https://doi.org/10.1109/ACCESS.2021.3075571
28. Pegoraro Santana IA et al. (2020) Music4All: a new music database and ıts applications. 2020 International conference on systems, signals and ımage processing (IWSSIP)

An Overview of Contextual Topic Modeling Using Bidirectional Encoder Representations from Transformers

Pranjali Basmatkar and Mahesh Maurya

Abstract Topic modeling refers to a range of algorithms in natural language processing that gives us an insight into the 'latent' semantic topics or patterns in a collection of documents. These patterns of word co-occurrence are used to determine the hidden 'topics' which are present in the corpus. Topic modeling has been used successfully for information retrieval, classifying documents, summarizing them and for exploratory analysis of large corpora of texts. This survey studies various algorithms that have been used for topic modeling over time including TF-IDF, latent Dirichlet algorithm (LDA), clustering on sentence-level BERT embeddings and a newer hybrid approach of generating contextual topics using a combination of LDA and BERT vectors. This survey will analyze the advantages and limitations of these algorithms.

Keywords Natural language processing · Topic modeling · TF-IDF · Sentence-BERT · Latent Dirichlet algorithm · Contextual topic analysis

Introduction

With a large influx of electronic data in the industry, it is very important to be able to leverage this data and put it to the best of its use. Newer and more accurate automation tools are required to meaningfully extract underlying patterns from this large data and put it to relevant use. These patterns in the data can be analyzed and summarized for better understanding and usage. Topic gathering is one such algorithm in the natural language processing domain that can be used to understand these underlying patterns. Topic gathering can have two sub-categories: topic classification and topic modeling. Topic classification is a type of algorithm where text data is classified into various predefined topics; here, the user will require prior knowledge of the number of unique topics present in the data. For cases where a large amount of unstructured

P. Basmatkar (✉) · M. Maurya
Department of Computer Engineering, SVKMs NMIMS Mukesh Patel School of Technology Management and Engineering, Mumbai, India

V. Bindhu et al. (eds.), *Proceedings of Third International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 844,
https://doi.org/10.1007/978-981-16-8862-1_32

data is present, where there is no prior knowledge of the underlying topic, topic modeling is the algorithm that is used.

Topic modeling is an unsupervised machine learning technique that is capable of scanning a large set of documents, identifying word and phrase patterns within them, and clustering word groups and similar expressions that most accurately describe a set of documents.

Topic models are useful for the purpose of document clustering, organizing large blocks of textual data, information retrieval [1] from unstructured text and feature selection. In the industry, topic modeling is used by online recommendation systems to boost article recommendation engines. Various recruiting companies use this algorithm for extracting latent features of job descriptions and map them to the right candidates. They are being used to organize large datasets of emails, customer reviews and use social media profiles.

Traditionally, various naive algorithms, such as term frequency–inverse document frequency or TF-IDF, have been used to determine frequency counts of words to identify underlying topics. With time, a popular range of topic modeling algorithms such as latent Dirichlet allocation or LDA, latent semantic analysis or LSA [2] and non-negative matrix factorization algorithms became available to users for easy-to-use topic modeling. LDA is a probabilistic model that views the documents as bags of words, determines k topics in the latent space and assigns each word to a topic probabilistically. Despite its popularity, LDA has some shortcomings such as a fixed number of topics that need to be predefined, does not capture correlations and does not take into account the sentence structure or the underlying context.

In recent times, with the evolution of the transformer architecture, bidirectional encoder representations from transformers (BERT) [3] came out as one of the state-of-the-art and strongest models. The BERT embeddings are based on a unique attention-based model that works bidirectionally and also encodes the context of a word with respect to the sentence it is part of while using minimum vocabulary. This technical paper aims to evaluate the advantages and shortcomings of the popular topic modeling approaches and also explores the advantages of a hybrid approach combining the strengths of LDA vectors and BERT embeddings.

In this technical paper, Sect. "Methods of Topic Modeling" provides a literature survey of all the existing topic modeling algorithms; in Sect. "Proposed Methodology", a new hybrid method of topic modeling is proposed called contextual BERT. In Sect. "Experiments", experiments are performed using the defined algorithms using an open-source dataset, which is followed by Sect. "Result Analysis", where the results from the experiments are analyzed and compared, and lastly, in Sect. "Future Works", the future steps are discussed and limitations of the models are addressed for the conclusion.

Methods of Topic Modeling

Term Frequency–Inverse Document Frequency (TF-IDF)

TF-IDF stands for term frequency–inverse document frequency of records. This is a statistical measure algorithm that evaluates the relevancy of a word to a document [4]. This is generally calculated using the frequency of relevant words in a corpus. This kind of frequency can be in turn used to determine the latent meaning of the document. Usually, the relevance of a word increases with the increasing frequency of that particular word in the corpus.

TF-IDF is one of the earliest algorithms in the field of natural language processing for used cases such as text mining [5] and information retrieval. It is also often used by search engines to determine the relevance of a document to the keywords queried. Since time, this algorithm has been used to summarize documents based on the relevant terms generated by it (Fig. 1).

Here, *tf* (*t*, *d*) is the number of occurrences of '*t*' in the document '*d*', '*n*' is the total number of documents, and '*df*' is the number of documents containing the term '*t*'.

Term Frequency—Term frequency stands for the normalized raw term frequency or the weighted value of the number of times a word occurs in a document. The higher the frequency of occurrence of the word, the higher the relevance.

Document Frequency—Document frequency stands for the measure of relevance of the word across all documents in a corpus.

Inverse Document Frequency—Inverse document frequency stands for the weighting method to up weight terms that are less frequent in a corpus.

In total, TF-IDF [6] is a balanced weighting method that assigns weights to each word based on its term frequency (TF) and its balancing document frequency (IDF) and can be used to determine the relevance of a word in a corpus.

While TF-IDF can be used to extract the most descriptive terms from a document, it has some serious limitations, and it is based on the bag-of-word model and hence fails to capture the position of the term in the text, co-occurrences in different documents and also does not capture semantics.

$$tfidf(t,d) = tf(t,d) \times idf(t)$$

$$idf(t) = log\left(\frac{n}{df(d,t)} \right)$$

Fig. 1 Mathematical formula for TF-IDF

Latent Dirichlet Algorithm (LDA)

LDA is one of the most popular topic modeling methods where the probabilistic unsupervised model is used for various use cases such as text mining [7], text summarization and text categorizing tasks [8]. The algorithm assumes that each document is a probability distribution of topics, and each topic is a probability distribution of words from the document. LDA uses the bag-of-words model that treats each document as a vector of word counts. The main feature of LDA is that all the documents in the collection may contain the same set of topics, but each document contains those topics with different amounts. In this iterative process, the documents are observed one after the other while the hidden structure—the available topics, per-document topic distributions and per-word topic assignment—persists.

The basic steps for LDA can be defined as below:

- *k* topics are assumed across the documents
- These *k* topics are distributed across the documents by assigning each word a topic
- Probabilistically, each word w is assigned a topic based on the topics that are prevalent in that document, and the number of the word w has been assigned a particular topic in the past across all documents also known as the beta parameter.
- This process is repeated for each document to get the required results (Fig. 2).

Here, the two main parameters in LDA topic modeling are the alpha and beta values, the alpha value attempts to control the proportion of topics in a particular document '*m*' while the beta parameter is used to control the presence of words in a particular topic '*k*'.

Different variations of LDA models have been prevalent including temporal topic modeling, supervised topic models [9], etc.

Despite its popularity, there are some serious limitations to this approach of topic modeling such that a defined number of topics need to be fixed before modeling which might not be always possible to guess, there is a high chance of uncorrelated

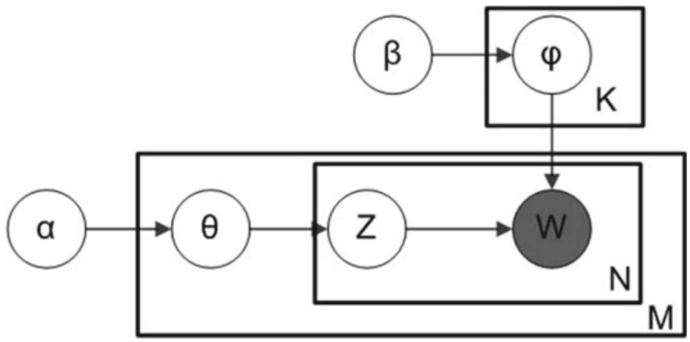


Fig. 2 Plate diagram of LDA topic modeling

topics getting generated, the topics defined are static in nature and do not evolve with time, and since the algorithm uses the bag-of-words model, the semantic relevance is lost and also faces serious performance issues for small-text data.

Bidirectional Encoder Representations from Transformer (BERT)

One of the biggest limitations of the traditional topic modeling approaches is the lack of contextual topics generated. With recent strides of improvement in the language modeling tasks, state-of-the-art approaches such as bidirectional encoder representations from transformers or BERT [3] have been known to generate dynamic contextualized embeddings based on transfer learning where pre-trained language models on large datasets are used and then fine-tuned for specific tasks. One of the biggest advantages of BERT embeddings is that it can use its knowledge of the semantics of the language based on the pre-training. Since BERT uses bidirectional encoding schemes, it is able to capture long-term dependencies and include hierarchical relations.

For the topic modeling use case, one way of leveraging the semantic power of the BERT embeddings is by creating sentence-level embeddings of the dataset such as the sentence-BERT [10]. Here, BERT creates sentence-level embeddings for all the sentences in a given dataset, and these embeddings can be further used to determine sentence similarity [11] using cosine similarity. This method ensures that the order of the words is preserved, semantic data is not lost, and a large pre-trained language model is leveraged.

There are several variations of pre-trained BERT models such as bert-base-uncased, bert-large-uncased, bert-base-cased and many more depending on the number of parameters and the type of data used for training, and some more advanced variations include RoBERTa [12] and StructBERT [13] which can also be employed for better performance.

Proposed Methodology

Contextual Topic Modeling Using BERT

In the previous section, we discussed some leading methods of topic modeling that have been used for generating topics. By analyzing the limitations of these, there is a need for a hybrid approach that attempts to incorporate and leverage the advantages of these methods and aim to create an approach that creates better topics, superior clusters with the least amount of data overlap.

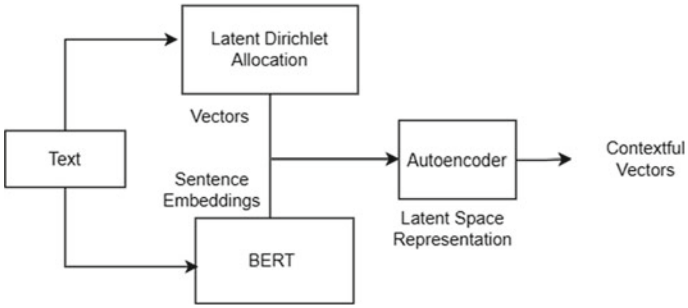


Fig. 3 Contextual BERT workflow

To explore this, in this section a hybrid approach is discussed which utilizes the advantages of probabilistic vectors from the LDA topic modeling and the semantic contextual embeddings obtained using the BERT pre-trained models.

It can be observed that LDA topic modeling is able to use probabilistic methods to efficiently create topics, but could lose semantic meaning owing to the bag-of-words approach followed by it. To bridge this gap, this new method is proposed which aims to leverage the advantages and ease of use of LDA topic modeling and combine it with the semantic power that can be obtained from BERT embeddings. Here, the probabilistic vectors from LDA are concatenated with the embeddings obtained from fine-tuning the dataset with the BERT pre-trained model. Now this concatenated set of vectors can be clustered using clustering techniques to generate topic clusters that can be concise and have high semantic information (Fig. 3).

Experiments

Data

For the purpose of this experiment, open-source data containing Amazon product reviews was used which had about 5000 data points. The dataset contained various features describing the product, the product category and review details. Some of these key features are described below in Table. 1.

Out of these features, the review text feature is a short paragraph of data describing the review. This data will be used to model, understand and create topics based on the experience of the customer with the product.

Table 1 Top features present in the dataset

Sr. No.	Features	Definitions
1	Product ID	Product ID by Amazon
2	Product category	Category the product belongs to
3	Review ID	Review ID as stored by Amazon
4	Review date	Date of review added
5	Review is-recommend	True/false values based on recommendation
6	Review text	Review text as submitted by the customer
7	Review tags	Special tags describing product if any
8	Review username	Username of the customer

Data Preprocessing

The feature 'Review text' consisted of short textual data that described the experience of a customer with the product. This natural language textual data needed to be cleaned and condensed and then transformed into embeddings such that it could be used with all the natural language processing algorithms.

For preprocessing of the data, both word-level and sentence-level processing were conducted.

For **word-level processing**, all basic natural language preprocessing steps were undertaken including:

Tokenization: Here, the text is treated as a string and is split into smaller tokens or words using the NLTK library.

Stop word removal: Stop words or noise words such as the, and, are, etc., were removed using the NLTK stop word dictionary.

Noise Text Removal: Numerical data and punctuation words were removed.

Stemming: Using the NLTK, stemming was performed to bring the words to their canonical root form. This decreases the repetitiveness of text.

For **sentence-level preprocessing**, the text sentences were converted to lowercase. Repetitive words were removed from the sentences to reduce the noise in the phrases. Regular expressions were created for removing content within parentheses, numerical data and special characters to clean the sentences before creating embeddings.

This processed data was used to create a unique word dictionary and corpus that will be used for training models.

TF-IDF with K Means Clustering

In this experiment, TF-IDF vectors were created using sklearn's 'TfidfVectorizer,' and these vectors were used for unsupervised bucketing or clustering using the K means clustering method.

A default number of ten clusters were set for this clustering experiment. Once the ten clusters were created by the model, each of the clusters was analyzed to identify the coherence, the quality of data clustered and the top keywords associated with each of the clusters were used to determine a high-level understanding of the cluster. Using sklearn libraries, the coherence and silhouette score of the clustering model were determined for comparison. The high-level workflow of the experiment is described in Fig. 4.

To further analyze the word assignment to topics, the UMAP library was used to view the cluster distribution described in Fig. 5.

It can be seen that while the words have been evenly distributed to the ten defined clusters, there is no clear boundary or demarcation that separates these data points into individual clusters. The data points from one cluster can be seen overlapping

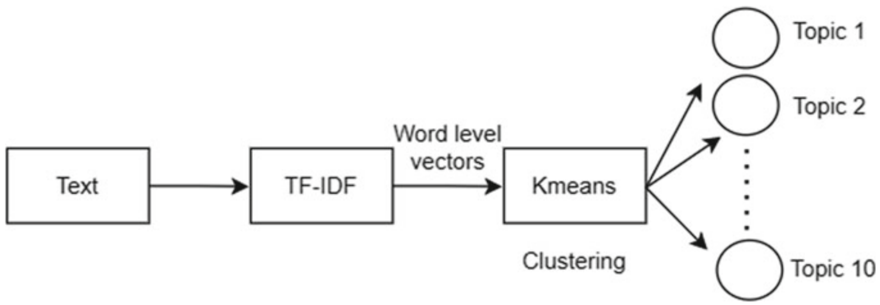
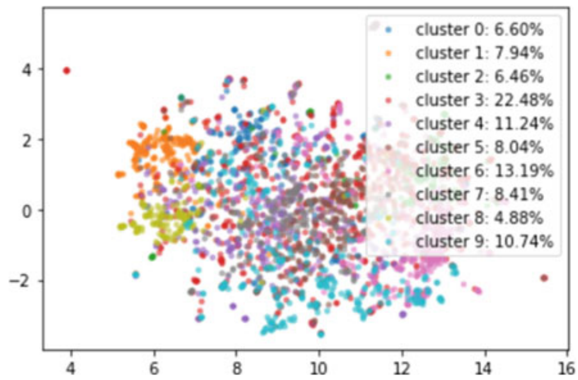


Fig. 4 TF-IDF workflow for topic modeling

Fig. 5 TF-IDF workflow for cluster distribution



with those of the neighboring clusters which indicates that there is no clear separation and the clusters cannot form meaningful topics.

LDA Topic Modeling

In this experiment, the sklearn library was used to build the LDA topic model. The data corpus and the dictionary created during the preprocessing of the data are used as the input to this model. Setting the number of clusters as 10, the clustering model was trained, and the top words of the topics were analyzed to check the quality of the performance of the model. This model was used to determine the coherence score of the topics created for comparison with other methods. The high-level workflow of the experiment is described in Fig. 6.

On observing the topic distribution in Figure 7 of the LDA topic model, it can be seen that the data points are able to be identified within a cluster. The data points within the same cluster are closer to one another and farther away from other clusters.

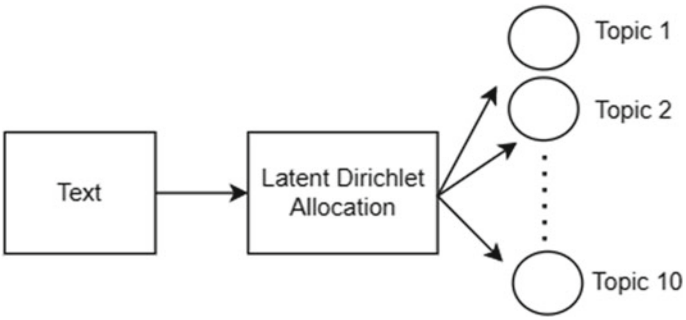
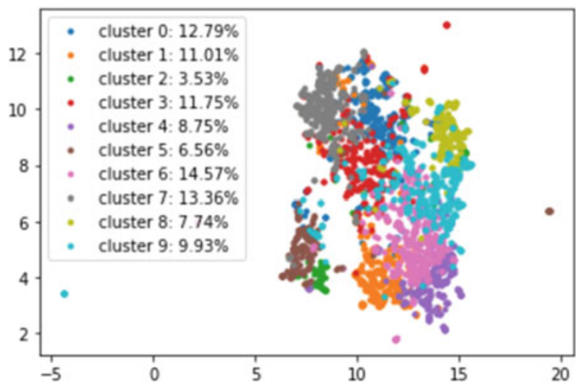


Fig. 6 LDA workflow for topic modeling

Fig. 7 LDA topic model cluster distribution



While clusters can be identified, there is a large amount of overlap seen between the clusters. In conclusion, although the quality of topics [14] has increased, there is a large overlap between these clusters which could dilute the quality of the clusters.

BERT Sentence Embeddings with K Means Clustering

For this experiment, to explore the capabilities of the BERT sentence transformer library, sentence-level embeddings were created using the sentence_transformer library to employ the 'bert-base-nli-max-tokens' pre-trained model. Further, these sentence-level embeddings were clustered together in an unsupervised way using K means clustering. The top words associated with each of the clusters were analyzed, and the coherence scores were generated. Again, a default number of ten clusters were set for uniformity and fair comparison between the methods. The high-level workflow for the experiment is described in Fig. 8.

On analyzing the topic distribution in Fig. 9, we can see that while the topics created do form visible and distinct clusters, the intra-cluster distance of the data points has increased, and the inter-cluster distance remains relatively the same. This indicates that there is a lack of clear distinction between the topics generated.

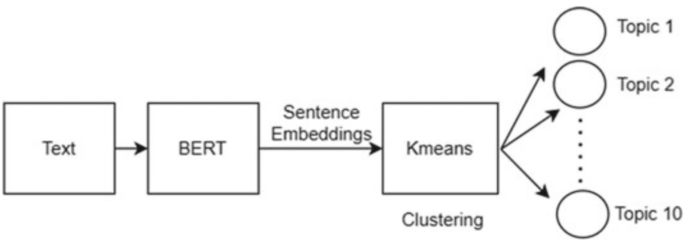
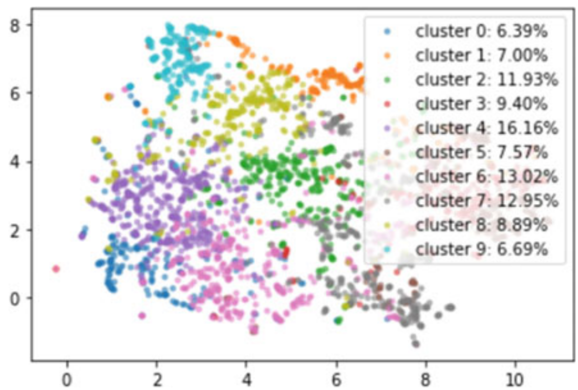


Fig. 8 BERT sentence embeddings and clustering workflow

Fig. 9 BERT clustering cluster distribution



Contextual BERT

For this experiment, to explore the capabilities of the LDA and BERT vectors together, an LDA model was trained using sklearn, and the topics and their probabilities [15] were stored. In parallel, sentence-level embeddings were created for the sentences in the corpus using a pre-trained BERT [16] model. These vectors were concatenated together. An auto-encoder was used to reduce the dimension space, and these vectors were clustered together using K means clustering [17]. A default of ten clusters was created using this method.

This cluster model was used to determine the coherence score obtained for topics. The high-level algorithm for this approach is shown in Fig. 10.

Upon observation of the topic distribution in Fig. 11, it can be seen that the data points are clubbed together in distinct clusters such that the intra-cluster distance between the clusters is low and the inter-cluster distance is high. This allows the creation of distinct clusters where the least number of data points can be seen overlapping on the cluster boundaries.

The defined boundary can be seen as an indication of improved cluster distribution that reduces the chance of ambiguity among the top representative words of those clusters.

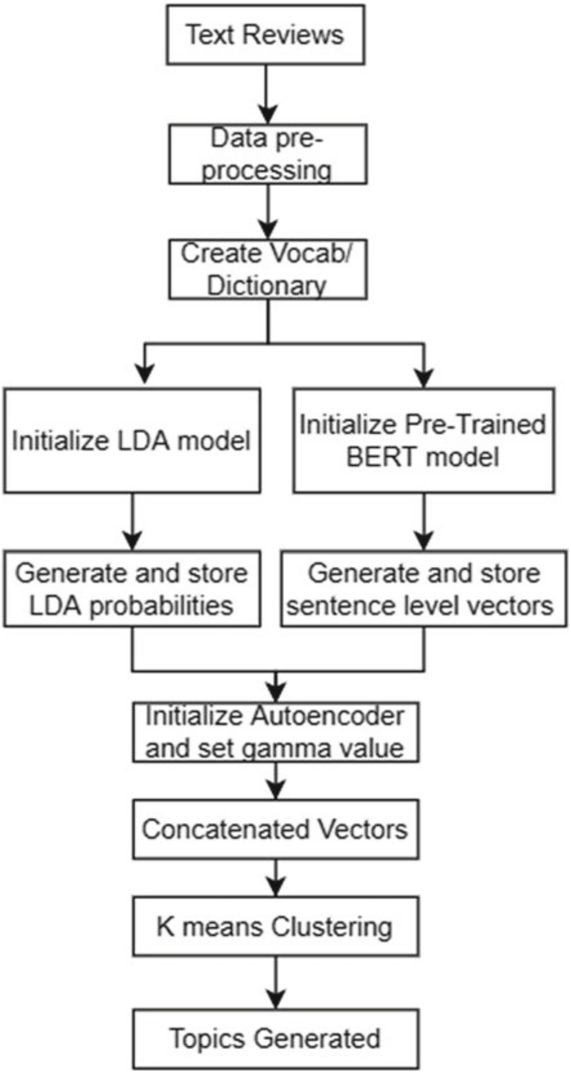
Result Analysis

After experimenting with the four algorithms to obtain meaningful topics, the performance of each of the algorithms was judged based on three factors: the coherence score, the silhouette score and the cluster distribution. To confirm the best results of the topic distribution, the top words from a few of the topics generated were analyzed to get a high-level understanding of a few of the clusters.

The coherence score [18] measures the semantic similarity between the high importance words within a topic. It can be used to access the quality of the topics generated. A higher coherence score indicates better and coherent topics. There are various coherence measures present like C_umass, C_v and C_p. In this experiment, we have used the C_v coherence measure that is based on a sliding window protocol of the top words of a topic to determine the co-occurrence between these words which is in turn used to determine the similarity between these top words. It ranges between 0 and 1. The silhouette score is used to infer the performance of the clustering algorithms, and it studies the separation distance present between the defined clusters. The score is in the range of -1 to $+1$, where $+1$ indicates that samples in a cluster are far away from samples of the neighboring cluster. The topic distribution is another visualizing method to understand the sample distribution of each topic and also interpret the topic boundaries.

Based on the experiments, the coherence scores for the TF-IDF K means, LDA topic modeling, BERT K means and contextual BERT are documented. For cluster

Fig. 10 High-level workflow for the contextual BERT topic modeling algorithm



analysis, the silhouette scores for TF-IDF K means, BERT K means and contextual BERT are documented in the table below (Table 2).

From the scores generated, we can see that while the LDA model and the BERT K means model can generate good coherence scores for the topics, the contextual BERT model which is a combination of LDA vectors and sentence-level vectors from BERT has the best performance for coherence and silhouette scores.

To further investigate the performance of the contextual BERT model, the top words for each of the topics were checked to determine the high-level topic defined by the cluster. Some of the topics and their top words are listed in Table 3.

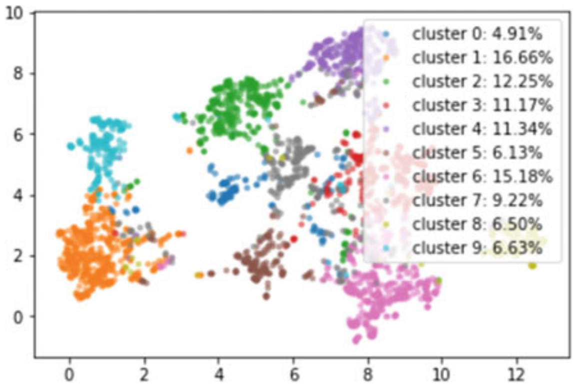


Fig. 11 Contextual BERT cluster distribution

Table 2 Performance comparison based on coherence and silhouette scores

Model type	Coherence score	Silhouette score	Topic observations
TF-IDF K means	0.396	0.01	Undefined cluster topics generated with no clear boundaries
LDA	0.409	–	Defined cluster topics with overlapping data points
BERT K means	0.374	0.06	Defined clusters with overlapping data points
Contextual BERT (BERT and LDA)	**0.451**	**0.195**	Defined and distinct clusters with the least overlapping data points

Table 3 Top words for topic clusters

Topic	Topic-1	Topic-2	Topic-3	Topic-4
Top keywords	Kindle	Product	Sound	Son
	Book	Echo	Speaker	Tablet
	Screen	Music	Quality	Control
	Battery	Amazon	Echo	Parent
	Size	Home	Device	Lock
	Product	Video	Item	App
	Read	Screen	Music	Kid
	Life	Camera	Hub	Device
	Reader	Control	Product	Feature
High-level topic	Reader/tablet Info	Amazon home control features	Sound/speaker quality	Device parental control

Considering the topic distribution as shown in Fig. 11 based on the UMAP topic distribution, the top words of each cluster and the performance metrics, we can see that the quality of clusters is improved by the contextual BERT approach.

Future Works

While comparable results for topic modeling can be achieved from the proposed method contextual BERT, there are certain limitations to the proposed method, such that the BERT pre-trained language models are trained on a large set of general data, while topic modeling might give good results on a general dataset, and the performance may fall for domain-specific data. The proposed model is not tuned on the parameters. To further improve the quality of topic modeling, the following further steps can be taken.

Explore Larger Pre-Trained BERT Models

There are various types of BERT pre-trained models [19] available that have been trained on larger and more specific data sources. A larger language model can be used to identify a larger vocabulary of words and be able to correctly identify the semantic meanings of each term and effectively group them together.

Fine-Tune the Contextual BERT Model for the Number of Cluster Parameter

The model is currently trained on a default set of ten clusters. Optimizing this parameter on the basis of silhouette score on the given data will improve the quality of clusters and also improve the coherence score of the model.

Adding Additional Data for Training

Currently, for this experiment, due to limited resources and computational power, a set of 5000 data points were used to observe the results. Adding more data for this experiment will likely confirm these results and also increase the performance of the model.

Exploring Ways to Use These Topic Words to Create a Small Descriptive Summary

In an attempt to use the topics generated by the topic models, for further research, techniques can be explored to use these top relevant words within each topic to summarize and create descriptive tags for each topic for better understanding and usage.

Conclusion

After observing the result from the four experiments, it can be noticed that the contextual BERT model based on the LDA and sentence-level vectors from BERT has outperformed the other algorithms. The coherence score is much higher, and the silhouette score is also higher indicating that the clusters formed are distinct and hence distinct topics can be generated from the data. The topic distribution of contextual BERT also indicates a higher quality of defined clusters, and the topics generated are contextual with the least overlapping relevance. Hence, this approach is able to capture the semantic meaning of the phrases and leverage the probabilistic vectors for creating improved clusters for topic modeling and can be used improved topic modeling.

References

1. Yi X, Allan J (2009) A comparative study of utilizing topic models for information retrieval. In: Boughanem M, Berrut C, Mothe J, Soule-Dupuy C (eds) Advances in information retrieval. ECIR 2009. Lecture notes in computer science, vol 5478. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-00958-7_6
2. Mohammed S, Al-augby S (2020) LSA & LDA topic modeling classification: comparison study on E-books. 2502–4752. https://doi.org/10.11591/ijeecs.v19.i1.pp%25p
3. Devlin J, Chang M-W, Lee K, Toutanova K (2019) BERT: Pre-training of deep bidirectional transformers for language understanding. In NAACL-HLT
4. Qaiser S, Ali R (2018) Text mining: use of TF-IDF to examine the relevance of words to documents. Int J Comput Appl 181. https://doi.org/10.5120/ijca2018917395
5. Jing L-P, Huang H-K, Shi H-B (2002) Improved feature selection approach TFIDF in text mining. Proceedings. International conference on machine learning and cybernetics, vol 2. pp. 944–946. https://doi.org/10.1109/ICMLC.2002.1174522.
6. Oren N (2002) Reexamining tf. idf based information retrieval with genetic programming. In: Proceedings of SAICSIT 2002. pp 1–10
7. Tong Z, Zhang H (2016) A text mining research based on LDA topic modelling. Comput Sci Inf Technol 6:201–210. https://doi.org/10.5121/csit.2016.60616
8. Kolla B (2019) Categorizing research papers by topics using latent dirichlet allocation model. Int J Sci Technol Res 8:1442–1446
9. Alghamdi R, Alfalqi K (2015) A survey of topic modeling in text mining. Int J Adv Comput Sci Appl 6(1):7. https://doi.org/10.14569/IJACSA.2015.060121

10. Reimers N, Gurevych I (2019) Sentence-bert: sentence embeddings using siamese bert-networks. In: Proceedings of the 2019 conference on empirical methods in natural language processing, EMNLP
11. Agirre E, Banea C, Cardie C, Cer D, Diab M, Gonzalez-Agirre A, Guo W, Mihalcea R, Rigau G, Wiebe J (2014) SemEval-2014 task 10: multilingual semantic textual similarity. In: Proceedings of the 8th international workshop on semantic evaluation (SemEval 2014). Dublin, Ireland. Association for Computational Linguistics, pp 81–91
12. Liu Y, Ott M, Goyal N, Du J, Joshi M, Chen D, Levy O, Lewis M, Zettlemoyer L, Stoyanov V (2019) RoBERTa: a robustly optimized BERT pre-training approach (cite arxiv:1907.11692)
13. Wang W, Bi B, Yan M, Wu C, Bao Z, Peng L, Si L (2020) StructBERT: incorporating language structures into pre-training for deep language understanding. ArXiv, abs/1908.04577
14. Krasnashchok K, Jouili S (2018) Improving topic quality by promoting named entities in topic modeling. In: Proceedings of the 56th annual meeting of the association for computational Linguistics, vol 2: short papers. Melbourne, Australia. Association for Computational Linguistics, pp 247–253
15. Blei DM (2012) Probabilistic topic models. Commun ACM 55(4):77–84. https://doi.org/10.1145/2133806.2133826
16. Liu Y, Ott M, Goyal N, Du J, Joshi M, Chen D, Levy O, Lewis M, Zettlemoyer L, Stoyanov V (2019) RoBERTa: a robustly optimized BERT pre-training approach. arXiv preprint arXiv:1907.11692
17. Reimers N, Schiller B, Beck T, Daxenberger J, Stab C, Gurevych I (2019) Classification and clustering of arguments with contextualized word embeddings. In Proceedings of the 57th annual meeting of the association for computational Linguistics. Florence, Italy. Association for Computational Linguistics, pp 567–578
18. Korenčić D, Ristov S, Šnajder J (2018) Document-based topic coherence measures for news media text. Expert Syst Appl 114. https://doi.org/10.1016/j.eswa.2018.07.063
19. Liu Y, Ott M, Goyal N, Du J, Joshi M, Chen D, Levy O, Lewis M, Zettlemoyer L, Stoyanov V (2019) Roberta: a robustly optimized bert pre-training approach. ArXiv, abs/1907.11692

Smart Traffic Controller Based on Traffic Density and Prioritized Emergency Vehicle Clearance

N. Satheeskanth, M. Mathushan, J. Joy Mathavan, A. Kunaraj, and G. Daisan

Abstract Traffic lights play a key role in avoiding traffic congestion especially in cities in peak time. The traffic congestion occured during peak hours is the result of mismanagement of traffic in the off-peak hours is the bitter truth. In this paper, the application of the ultrasonic sensors and microcontroller prioritizing the flow of traffic in the road based on vehicle density is widely discussed. The emergency vehicles are identified using RFID in the proposed prototype and given the top priority. Portable smart traffic management system is built as a prototype using inductive sensing elements, RFID and microcontroller. The prototype of the smart traffic control system was built and tested in the controlled scenarios, comprehensive study was made, and the results were thus obtained. The obtained results clearly show that smart traffic controller is highly effective in clearing the traffic compared to the traditional traffic control system.

Keywords Smart traffic controller · Traffic density · Microcontroller · RFID

Introduction

The first traffic light system was designed by Lester Farnsworth and installed in Salt Lake City in 1912 [1]. Over these years, the traffic light system has undergone tremendous changes with the invention of microprocessor in late 1960s and with the invention of the light-emitting diode in 1980s [1]. The idea of introducing smart traffic controllers to regulate the traffic flow within the city is not new as the engineers around the world are trying to simulate the best possible mechanism to effectively eliminate the traffic congestion. To deal with the dynamic changes in the logistics and consider the cost of implementation, smart traffic controller using microcontrollers

N. Satheeskanth · M. Mathushan · J. J. Mathavan (✉) · A. Kunaraj · G. Daisan
Faculty of Technology, University of Jaffna, Jaffna, Sri Lanka
e-mail: joy@tech.jfn.ac.lk

A. Kunaraj
e-mail: kunaraj@tech.jfn.ac.lk

V. Bindhu et al. (eds.), *Proceedings of Third International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 844,
https://doi.org/10.1007/978-981-16-8862-1_33

and ultrasonic sensors are suggested. Traffic situation was observed around the two major junctions in Jaffna peninsula, Sri Lanka, for about a month; hence peak and off-peak times were determined. Then the data collection, experimental evaluation and analytics were performed to arrive at the conclusion. It was estimated in 2011 that 32 billion of LKR was lost due to traffic congestions in Sri Lanka which is 1.5% of GDP of the country [2]. Critical amount of resources can be conserved by concentrating on regulating traffic flow in major cities not in Sri Lanka but also around the world.

Literature Review

It is very much obvious that conventional traffic control system in the modern world also contributes to the traffic congestion in cities and hence needs to be replaced with smart and intelligent traffic control system that can take decisions depending on the on-the-spot situations. Advantages of smart traffic system for the developing countries which are struggling with traffic congestion are clearly described by Ranjith. A et al. He points out that emergency vehicle clearance in traffic congestion is a major struggle even in developed countries and suggests that these issues can be overcome by smart traffic control system [3]. But, the cost of such smart traffic control system and requirement of skilled people for the operation and maintenance becomes the downside of the smart traffic control system. In a view to overcome the downsides of smart traffic control system, Mohit Dev Srivastava et al. presented the preliminary steps in the application of a smart traffic light control system based on programmable logic controller (PLC) technology. The density of vehicles is measured using total weight of the vehicles. From the total weight, a number of vehicles on the road were obtained [4]. This method is far from being efficient, and it is also problematic for a traffic police to observe the whole situation within a stipulated time. Later, Xu Li, et al. in their paper used mobile sensing networks to identify traffic situation of the road and take decision based on the signals transmitted back from those sensing networks [5]. But this method can be slow depending on the process time and mobile network. This system was also affected by the weather changes. Poor signal level also causes further delay in signal transmission. The complication of this system requires high installation and maintenance cost and a lot of hardware implementations. In an attempt to solve this, C. Barz et al. present the smart traffic control system that can function with PLC using various sensors [6]. The entire traffic control system coordinates the traffic flow, providing a path which depends on coordination of combination of sensors and functions of radar to inform the traffic situation. But using radar for smart traffic control in every junction is practically impossible in developing countries. To overcome this challenge, Ashit S Chitta et al. proposed a design of smart traffic system using ARM LPC 2148 microcontroller. CC2500 RF tags are used to identify the emergency vehicles. GSM module is used to communicate with controller through mobile networks. In the conventional traffic control system, congestion can be caused by large delay in red light. The ARM microcontroller-based traffic control system proposes a traffic light control and monitoring system that can reduce the traffic congestions, caused by traffic lights. This system works based on IoT and is controlled by ARM7 microcontroller [7]. At the same time, Roxanne Hawi et al. describe about insufficient space and resources to build new roads with the increase in number of vehicles. But in developing countries, it is difficult to accommodate growing demands of vehicles within limited resources, so some solutions are required to avoid traffic congestion [8]. One such upcoming solution is to involve smart traffic control systems (STCS) to make traffic congestions pretty less. These systems use real-time data to eliminate traffic congestions and improve traffic management. Nikhil R. Chitragar et al., in their paper, mentioned that traffic signals are the most suitable method of regulating traffic in a junction. Current traffic system fails to control the traffic congestions effectively. When a particular road has got more traffic congestions than the other roads, the smart traffic control using PLC calculates the vehicle density in a particular road, and then the system decides to control the signal depending on the priority. The above proposed method has number of flaws including sensors positioning and lack of planning for emergency vehicle clearance system [9]. Rajeswari S et al. in their research work proposed installing RFID to every vehicles thereby counting the number of vehicles that passes the junction and thereby determining which lane to be given priority. RFID was installed even in ambulance and was detected by the controller yet the biggest drawback is installing RFID to billion of vehicles is ideal for a prototype but is not feasible and also complicate the controlling process [10]. Priyanka sharma et al. in their paper proposed to reduce the traffic congestion on the road by traffic diversion system based on weighted data [11]. This work is done based on weight sensing sensors whose output will be fed to a PLC, which in turn will control the traffic light switching decision. According to this method, measuring the weight along the road is not possible. Only a short distance of the road can be measured. At the same time, weight measurement technique of vehicle is very complicated because of the measuring mechanism and high initial cost. Hence, Muhammad Arshad Khattak in his paper describes about the use of photoelectric sensors to measure the density of vehicles [12]. If the number of vehicles is high, the density of light is low. But this idea of measuring the light density is inaccurate, since the sensor works properly only if it is close to the vehicle, and this system will not work in the night time. This system is also affected by the climatic changes. Sultane Shubham et al. describe about the manually operated emergency vehicle clearance system which may result in number of human errors [13]. So, an automatic system that can identify the emergency vehicles and clear the path for emergency vehicles needs to be identified. In the meantime, Wei-Hsun lee and Chi-yi Chiu described about building a cloud-based advanced traffic controller for a smart city using emergency vehicle signal preemption (EVSP), public transport signal priority (TSP), adaptive traffic signal control (ATSC), eco-driving supporting and message broadcasting [14]. However, this system is very complex. Integrating this proposed system with IoT and existing systems in developing countries like Sri Lanka is almost impossible. Similarly, Hitiyaremye et al. in their research work accept the failure of the traditional traffic light system and agree the need of smart traffic system to eliminate the delays and traffic congestions. They propose a system in which vehicles communicate with each other using vehicle to infrastructure V2I, vehicle to vehicle V2V and vehicle to device V2D communication protocols [15]. The proposed system as described is complicated and does not comply with the existing infrastructure available in Sri Lanka. In an attempt to improve the efficiency of smart traffic controller, Roxanne hawi et al., describe about involving fuzzy expert system FES and artificial neural network ANN [16]. The proposed system feeds the vehicle counts obtained using RFID to the system. The decision to operate the traffic lights is obtained based on the output of the processed data using the FES and ANN algorithms. At the same time, Sakuna prontri et al. in an attempt to solve the existing traffic congestion in the smart cities proposed fuzzy learning vector quantization FLVQ and fuzzy learning quantization particle swarm optimization FLVQ-PSO-based smart traffic controller [17]. The proposed algorithms learn and then forecast the traffic flows in the streets and make decisions based on the predicted outcomes. The proposed algorithm is flexible to the changing environment and hence proved to be effective. Yet, the fuzzy and neural algorithms are complex to design and implement. Khaled Kamel et al. in their research work described the importance of the effective parking management system, lack of which also contributes significantly for chaos and congestions in the city traffic. This paper describes about parking slot identification using MobileNet classifier which is a lightweight deep neural network [18]. Subarna Shakya in his research work analyses the novelty about integrating the transportation using edge computing to establish collaborative services and preserving individual privacy. The concept of integrating the services in a smart city is also a good way to reduce traffic in a city [19]. The existing methods discussed in the literature are constructed using artificial algorithms (fuzzy logic, FLVQ and ANN) and PLC, which are complex to design and implement in developing countries with limited resources and technical expertise.

The proposed method in this current research work though utilizes autonomous algorithm is implemented using simple and easily available microcontroller AT mega 328p-pu. It detects the density of the vehicle using ultrasonic sensors placed at regular intervals, and the emergency vehicles are identified using RFID implemented in particular vehicles. It aims at providing a smart, simple solution by developing a system that can be integrated with the existing infrastructure.

This research paper is written in a way that Sect "Introduction" introduces the current traffic scenario, and Sect. "Literature Review" has the detailed literature review of the steps taken by different authors to solve the traffic congestion issue. Section "System Description" describes the current system designed and developed in this research work to solve the above-mentioned issue. Section "Research Methodology" explains the working methodology of the developed system, Sect. "Results and Discussion" analyses the results obtained, and the last section mentions about the conclusion of this research work.

System Description

At Thirunelveli Junction and Ariya Kulam Junction, Jaffna, Sri Lanka, usually in peak hours, following scenario was observed for a long time. These are the two crucial junctions which experience maximum traffic in Jaffna Peninsula always. The usual fixed time intervals of traffic lights cause long queues and finally to traffic jam. Based on the data shown in Tables 1 and 2, there are some irrelevant patterns that were identified. These data proved the need of smart traffic control system in high traffic roads.

The following data is collected for a week in the city of Jaffna; on high rush weekdays, average is taken, and projection is obtained for the month and verified. Jaffna town, being a regional hub, comprises number of schools, banks, government and private offices and draws huge amount of people on weekdays. Weekends usually record less traffic and hence ignored in general. When one observes the data given in Tables 1 and 2 closely, the problem with the existing system can be captured. For example, in Table 1, for day 5, when 27 vehicles are waiting in road 1 for 72 s, the green light is on for 15 s just for 11 vehicles in road 4.

Same pattern can be observed in Table 2, day 2. When 42 vehicles are waiting in road 1 for 72 s, the green light is on for 15 s just for three vehicles in road 4. This setup needs to be revised considering the rush created by number of vehicles. It is discussed in detail in this paper.

Table 1 Average of vehicle density per rotation of signal light at Thirunelveli junction, Jaffna, Sri Lanka

Road 1	Duration for which lights are on	Number of vehicles				
	72 s (Red)	Day 1 13	Day 2 16	Day 3 29	Day 4 18	Day 5 27
Road 2	52 s (Red)	4	11	18	5	7
Road 3	32 s (Red)	12	6	7	4	3
Road 4	15 s (Green)	11	25	30	9	11

Table 2 Average of vehicle density per rotation of signal light at Ariya Kulam Junction, Jaffna, Sri Lanka

Road 1	Duration for which lights are on	Number of vehicles				
	72 s (Red)	Day 1 27	Day 2 42	Day 3 19	Day 4 22	Day 5 27
Road 2	52 s (Red)	10	17	39	16	9
Road 3	32 s (Red)	2	6	21	7	11
Road 4	15 s (Green)	6	3	6	11	7

Research Methodology

Figure 1 shows the selected four-way junctions taken for study (Thirunelveli Junction and Ariya Kulam Junction). These two were the junctions where traffic lights were installed at the initial stage when traffic lights were introduced to Jaffna population. This shows the importance and rush on these particular junctions. Figure 2 shows the flowchart of the operation of proposed smart traffic controller. The switching decisions, namely normal mode (time-based control), auto mode (density based control) and emergency mode, are made based on the output of the sensors as depicted in the flow diagram. Figure 3 illustrates the block diagram of proposed density-based smart traffic control system using ultrasonic sensors and RFID as input elements and microcontrollers as controlling element. The block diagram of traffic controlling system is designed for a four-way junction. This system has the following add-on features comprising four parts, namely compilers and ultrasonic sensors (for vehicle density measurement), microcontrollers and RFID modules (for emergency vehicle identification), control unit and traffic signal lights. Vehicle density measurement system has controlling microcontrollers one for each road. The working principle of vehicle density measurement system is as follows. Each microcontroller circuit has two ultrasonic sensors whose arrangement is in one side of the junction, and these two sensors are decided to be fit at 20 m distance. If the queue of vehicles is in the range of 0–20 m, then one ultrasonic sensor will be activated and will be fed to the microcontrollers.

But if the queue of vehicles exceeds 20 m, then both ultrasonic sensors will be activated and will be fed to the microcontrollers. In this case, microcontroller works as AND gate. If both signals are logical 1, then compiler will generate logical 1 as a signal and send the signal to control system. Through the relay, the microcontroller output signal is stepped up from 5 V DC to 24 V DC since the output signal of the compiler is 5 V, input signal of selection microcontroller is 24 V, and also the relay works as a switch. If both signals are not logical 1, then compiler will give logical 0 as signal to the selection microcontroller through the relay, and at this situation, the relay will not step up the output signal from 0 to 24 V. In this situation, microcontroller

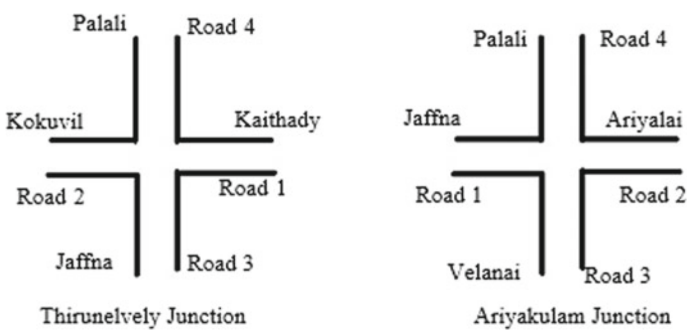


Fig. 1 Selected four-way junction for study

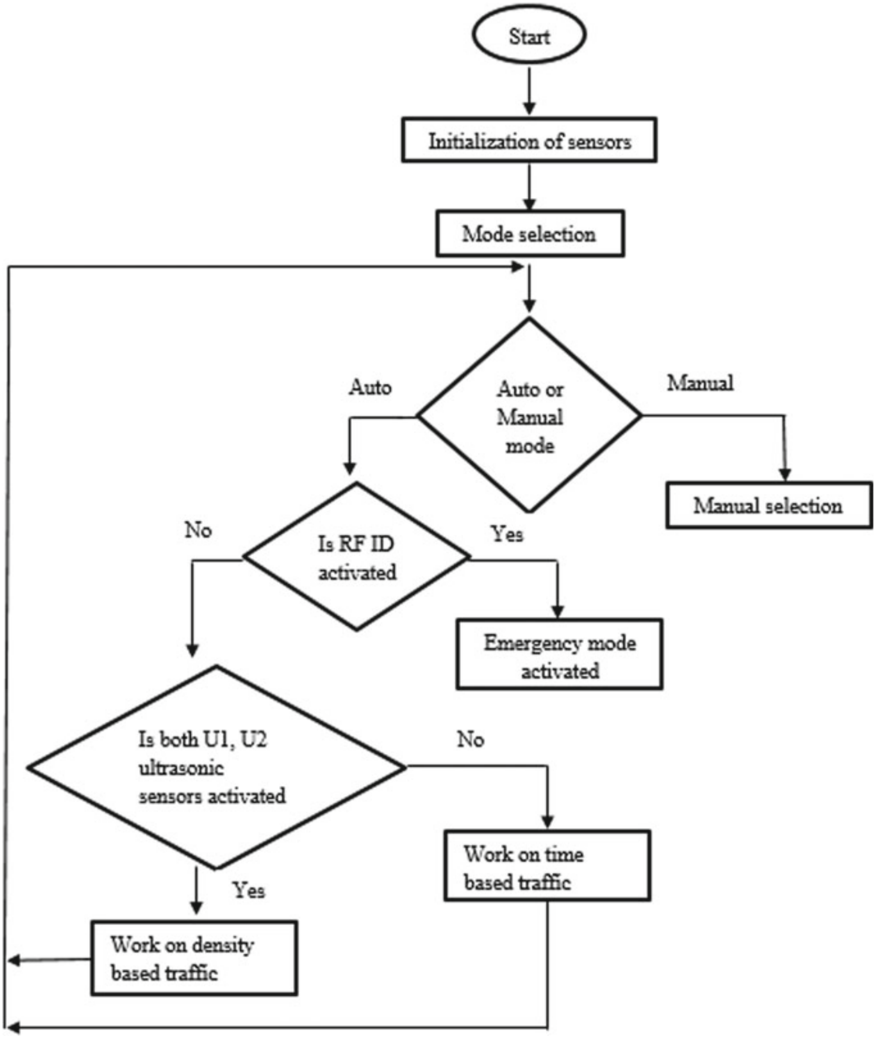


Fig. 2 Flowchart showing the operation of proposed smart traffic controller

will not receive logical 1 as signal so the microcontroller will not trigger to auto mode. If it triggers only, it will change the signals based on 64 combinations. In this manner, if the queue of vehicles exceeds 20 m, that road would be given priority to green signal. Once the range of queue of vehicles decreases below 20 m, the next road which have the queue of vehicles above 20 m would be given priority. If no road has length of queue of vehicles above 20 m, then the signal system works normally without consulting this newly introduced system. Higher priority is expected to be given to emergency vehicles, if there is one entering into the range. Briefly the operation of traffic controller under auto mode can be explained as, when the vehicle

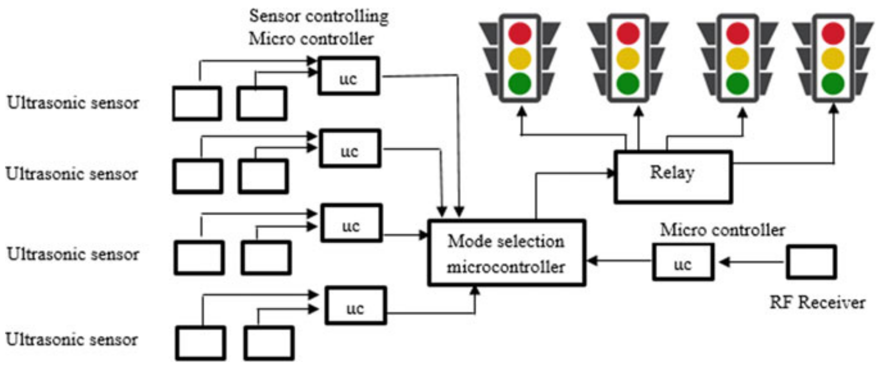


Fig. 3 Block diagram of overall function

density is identified using the ultrasonic sensors, it would be fed into the system. The controller decides which road to be given priority and duration of green signal for that particular road is more, until the traffic is cleared up to the threshold level. Once the density of that road falls below the threshold level, the road with the second-highest density would get the next priority and get the chance to open up. Figure 6 shows the circuit diagram of ultrasonic sensor in Sect. "Ultrasonic Sensor Circuit Diagram", the algorithm for the above-mentioned scenario is shown in Sect. "Algorithm for the Working of Selection Microcontroller Based on the Output from the Ultrasonic Sensors", and the hardware components utilized in the prototype building are listed in Table 3.

If the operation of traffic controller is under manual mode, traffic police officer can change the traffic signal according to the requirements. When the smart traffic

Table 3 Specifications of hardware components utilized in the proposed system

Hardware components used	Specification
AT mega 328p-pu microcontroller	28-pin AVR microcontroller Flash program memory: 32 Kbytes EEPROM data memory: 1 Kbytes SRAM data memory: 2 Kbytes I/O pins: 23 Timer: Two 8-bit/one 16-bit
Ultrasonic sensor HC-SR04	Working voltage: DC 5 V Working current: 15 mA Working frequency: 40 Hz Max range: 4 m Min Range:2 cm
RFID sensor and tag MFRC-522	Operating voltage: 2.5 V to 3.3 V Communication protocols: I2C, SPI, UART Data rate: 10 Mbps Range: 5 cm

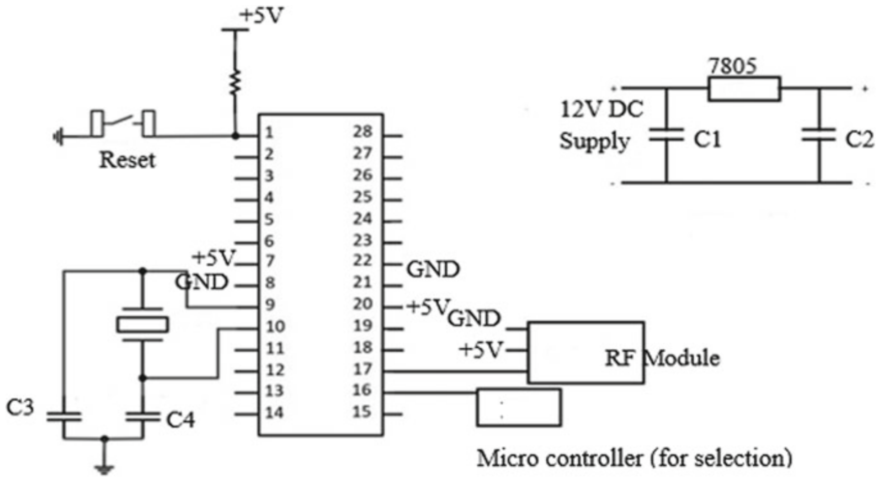


Fig. 4 RFID sensor circuit diagram

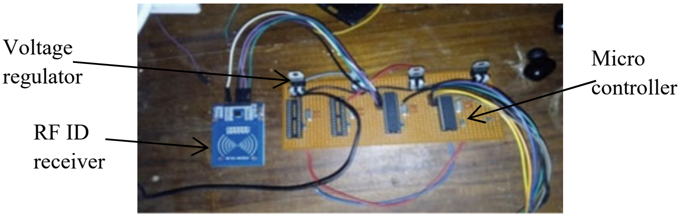


Fig. 5 RFID hardware implementation

controller works in the emergency mode, the operations take place as follows. RFID, according to this research work, fixed to the emergency vehicles is activated when the emergency vehicles enter the junction. The RF receiver shown in Figs. 4 and 5 picks up the signal, and the following algorithm gets activated to give priority to the emergency vehicles. The emergency vehicles include ambulance, fire brigade and VIP vehicles.

RFID Circuit Diagram

RFID circuit used for prioritizing the emergency vehicle clearance along with the algorithm is shown in Fig. 4. Figure 5 shows the hardware of the RFID system.

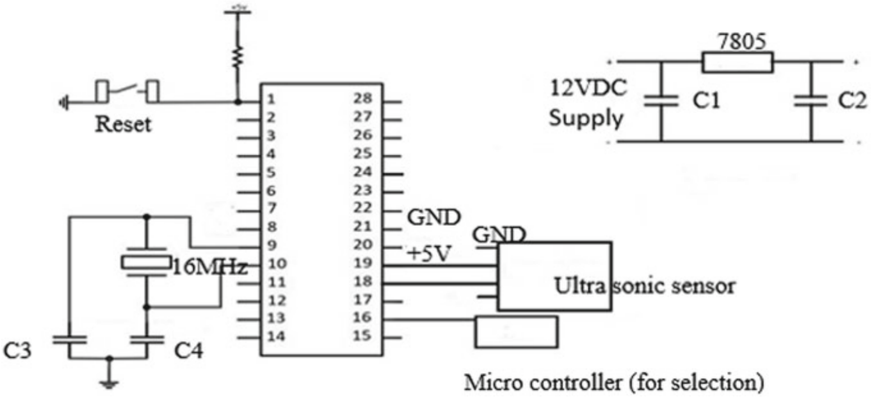


Fig. 6 Ultrasonic sensor circuit diagram

Algorithm When RFID is Activated by an Emergency Vehicle on Road 1

When RFID receiver is activated by the emergency vehicle approaching the junction on road 1, and if green light is ON on road 1, it will remain green till the emergency vehicle leaves the range of RFID. Otherwise, in order to give priority to road 1, green light on all other roads than road 1 turns to yellow in order to switch to red, while red light on road 1 turns to yellow in order to switch to green.

Output: Green light on the road 1
    If (green light already on in road 1)
    Continue (green light until emergency input is on)
    Else
    Change (red in road 1 to yellow)
    Change (green in any other road to yellow)
    Turn on (green light on road 1 and red light in all the other roads)
End

Ultrasonic Sensor Circuit Diagram

Ultrasonic sensor circuit diagram used for prioritizing the road depending on the traffic density is shown in Fig. 6, and the algorithm is shown in Sect. "Algorithm for the Working of Selection Microcontroller Based on the Output from the Ultrasonic Sensors".

Table 4 Functioning of the proposed system

Range of vehicle	Emergency vehicle	Sensors activated	Sensor controlling microcontroller	Mode selection microcontroller	Road priority
0–20 m	No	0	0	0	No
0–20 m	No	1	0	0	No
0–20 m	No	2	1	1	Road get high priority
0–20 m	Yes	0	0	1	Road with emergency vehicle get high priority

Algorithm for the Working of Selection Microcontroller Based on the Output from the Ultrasonic Sensors

```
duration = pulseIn(echoPin, HIGH);
distance M = (duration * 0.0340 / 2)/1000;
if (distance M&lt; 3)
{digital Write (pin, LOW);
}
else
{Digital Write (pin, HIGH);
}
Input: ultrasonic distance is less than 3m
Output: Green light on the road 1
If (green light already on road 1)
Continue (green light until emergency input is on)
Else
Change (yellow to other roads than road 1)
Turn on (green light on road 1 & red light on other roads)
End
```

Thus, the signal lights are switched on based on to the selected mode-by-mode selection microcontroller and is shown in Fig. 8. These data are shown in Table 4 briefly.

Results and Discussion

1. In case if there is no emergency vehicles in the proximity of the junction and none of the ultrasonic sensors are activated, then no outputs can be obtained from

the sensors to be fed to the controller, none of the roads will get the priority, and the system continues to work as the time-based traffic control system.

2. In case no emergency vehicles enter the junction and only one ultrasonic sensor is activated, still no roads will be given any priority. The traffic control system will continue to function as time-based traffic control.
3. In case, both the ultrasonic sensors are activated simultaneously and no emergency vehicles are in the proximity, then time-based traffic control system is interrupted, and priority of clearing the traffic is given to the particular road which has high traffic density.
4. In case emergency vehicles are near the junctions, time-based traffic control system and priority-based traffic clearance are interrupted, and priority is given to the road with the emergency vehicles.

Conclusion and Future Scope

The prototype of smart traffic light system was designed as shown in Fig. 7, in which if both sensors are activated by the line of vehicles, then that particular lane gets the highest priority than other lane, as in other lanes, the traffic density is comparatively less. In practical application, the two sensors are kept 20 m apart, and depending upon the requirement, more number of sensors can be incorporated in between. When an emergency vehicle having RFID installed in it approaches within the detectable range of the sensor as shown in Fig. 8, that particular lane gets the top priority irrespective of the traffic density in the other lanes. Once the

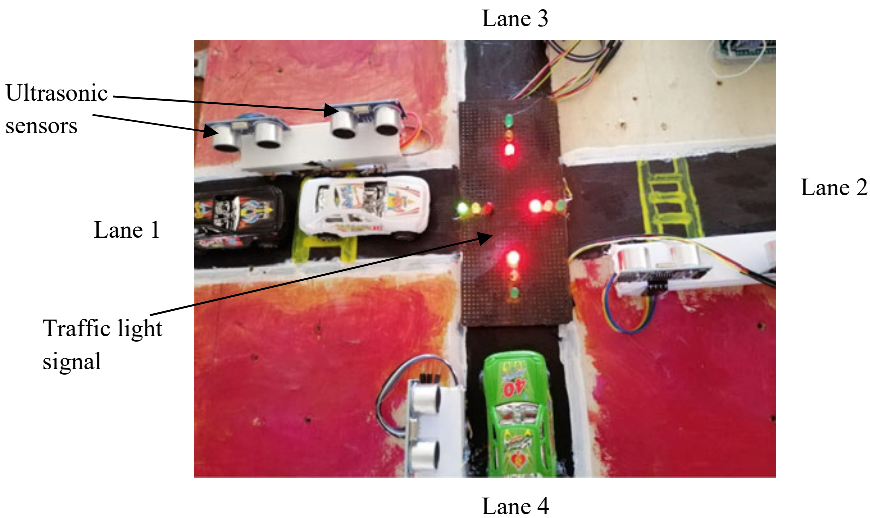


Fig. 7 Working example

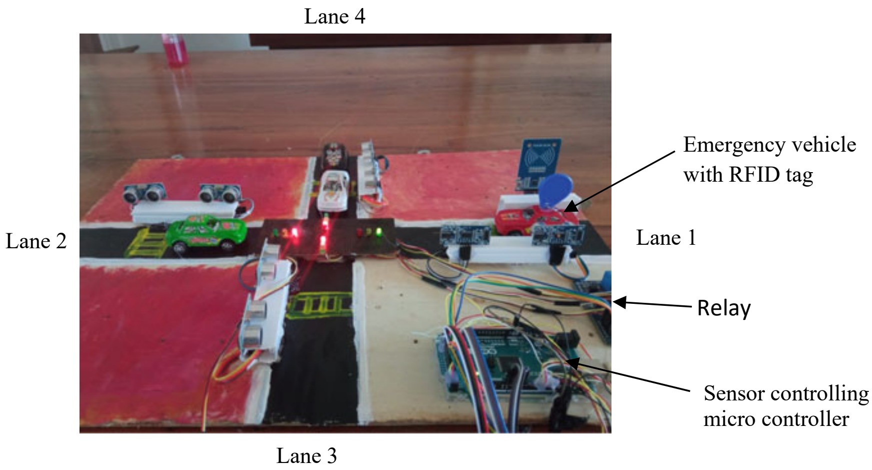


Fig. 8 Working example when emegency vehicle is detected

emergency vehicle passes the junction, again the controller would be switched to the density-based operation. Thus, the above prototype tested internally in the electrical laboratory and redundancy that occurs when using time-based algorithm is effectively eliminated and thereby increasing the efficiency to a greater extent. Yet the real efficiency can be derived only from real-time application by incorporating the proposed prototype in the existing traffic light controller. The sensor unit is synchronized with microcontroller to enhance its endurance and performance. This prototype of smart traffic controller can easily be attached with any real-time four-way traffic signal system simply disengaging the existing controller and connecting the proposed smart controller with ultrasonic sensors placed on proper intervals on the road. The controller is designed as a prototype to synchronize with the existing models. Ultrasonic sensor was used to provide better output while identifying the density of vehicles. Increasing the number of sensors may help to improve the identification of accurate density of vehicles. Another added enhancement of this system is emergency vehicle identification. The Emergency mode of operation is designed using RFID for robust and unique identification of emergency vehicles. Using the proposed method of controlling traffic, based on the traffic density, proves to be smart and effective way to avoid traffic congestion.

However, permission to test the prototype on the real-time system is yet to be approved, and hence, it is included in the future scope. Smart traffic control system will include the breakdown mode in the future to ensure regular flow of traffic even if the traffic control signal failed to operate properly due to various reasons including power failure. Synchronizing the smart traffic controller over the IoT might allow different smart traffic controllers interact with each other for communication and further effective control.

References

1. Castro C, Horberry T (2004) The human factors of Transport signs, 1st edn. CRC Press. https://doi.org/10.1201/9780203457412
2. https://www.sundaytimes.lk/110313/BusinessTimes/bt03.html
3. Ranjith AV, Rajendra S, Shivashankara BS, Sharath HK, Sundar R, Hebbar S, Golla V (2018) E-Intellectual 4way traffic control system using PLC and SCADA
4. Srivastava MD, Prerna SS, Sharma S, Tyagi U (Dec 2012) Smart traffic control system using PLC and SCADA. Int J Innovative Res Sci, Eng Technol 1(2)
5. Li X, Shu W, Li M, Huang H-Y, Luo P-E, Wu M-Y (2009) Performance evaluation of vehicle-based mobile sensor networks for traffic monitoring. IEEE Trans Veh Technol
6. Barz C, Oprea C, Erdei Z, Pop-Vadean A, Petrovan F (2014) The control of an industrial process with PLC. International conference on applied and theoretical electricity (ICATE). Craiova, Romania, pp 1–4
7. Chitta AS, Dinesha P (2016) Priority management of emergency vehicles using IOT approach. Int J Adv Res Comput Commun Eng 5(9)
8. Hawi R (2015) Techniques for smart traffic control: an ın-depth review. Int J Comput Appl Technol Res 4(7):566–573
9. Chitragar NR, Ramesh GB (Dec 2016) Smart traffic control system based on vehicle density using PLC. Int J Innovative Res Electr, Electron, Instrum Control Eng 4(12)
10. Rajeshwari S, Hebbar S, Golla V, Implementing intelligent traffic control system for congestion control, ambulance clearance and stolen vehicle detection. IEEE Sensor J https://doi.org/10.1109/JSEN.2014.2360288
11. Sharma P, Goen A (2018) Smart traffic control system using weighted data. Int J Adv Res Sci Eng 7(6)
12. Khattak MA (2013) PLC based intelligent traffic control system. Int J Electr Comput Sci IJECS-IJENS 11(6)
13. Sultane S, Chawda H, Karani K, Gajjar A, Bhagirathi D (March 2017) PLC based traffic density control using sensors. National conference on emerging trends, challenges & opportunities in power sector
14. Lee W-H, Chiu C (2020) Design and implementation of a smart traffic signal control system for smart city applications. Sensors mdpi
15. Eric H, Baptiste MJ, Gilbert U (2018) Impact of intelligent road traffic control in a technically connected vehicle environment. Int J Comput Sci Trends Technol (IJCST) 6(3)
16. Hawi R, Okeyo G, Kimwele M (2015) Techniques for smart traffic control: an in-depth review. Int J Comput Appl Technol Res 4(7):566–573
17. Prontri S, Wuttidittachotti P, Thajchayapong S (2015) Traffic signal control using fuzzy logic. 12th İnternational conference, electrical engineering/electronics, computer, telecommunications and ınformation technology (ECTI-CON)
18. Kamel K, Smys S, Bashar A (2020) Tenancy status ıdentification of parking slots using mobile net binary classifier. J Artif Intell 2(03):146–154
19. Shakya S (2021) Collaboration of smart city services with appropriate resource management and privacy protection. J Ubiquit Comput Commun Technol (UCCT) 3(01):43–51

Analysis and Design Approach of Footstep Power Generation Using Pressure Sensors

SundeepSiddula, K. VenkataRamarao, and S. K. Mohammad Hasheer

Abstract As global energy demand is increasing drastically, this increase, in the demand causes concern pertaining to the global energy crisis and allied environmental threats. Renewable energy sources have the potential to be an excellent medium for resolving the current energy shortage. This paper is going to explain the generation of electrical energy from non-conventional energy sources simply footsteps. Non-conventional energy sources are very essential at this time to our nation. Non-conventional energy using footstep these are input source to generate the output of the electrical power. A lot of mechanical energy is available in our environment in different forms which are not utilizing. Instead of wasting, utilize those vibrations for generating energy that can electrify the untouched and remote areas and small power consumption devices.

Keywords Piezoelectric sensors · Piezoelectric effect · Footsteps · Non-conventional energy · Microcontroller · Rechargeable battery · LCD display

Introduction

At present-days, electricity has become a most important for the human population. Its demand is increasing day by day. Modern technology needs a huge amount of electrical power for its various operations [1].

SundeepSiddula (✉)
Department of Electrical and Electronics Engineering, Vignana Bharathi Institute of Technology, Hyderabad, India

K. VenkataRamarao
Department of Electronics and Communication Engineering, Chalapathi Institute of Engineering and Technology, Guntur, India

S. K. Mohammad Hasheer
Department of Mechanical Engineering, R.V.R. & J.C. College of Engineering, Guntur, India

V. Bindhu et al. (eds.), *Proceedings of Third International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 844,
https://doi.org/10.1007/978-981-16-8862-1_34

With the advancement of trends in the technology, the electrical utilization has been increasing rapidly, this advancement or trend in technology is directly proportional to the utilization of electrical energy [2]. As the usage of fossil fuels and other material which are conventional sources of power generation has very high adverse effects on environment with increase in pollution and generation of harmful gases, etc., using the conventional sources leads to large requirement of land and equipment for power generation also increases the cost of implementation. In order to generate the maximum amount of energy with low cost of implementation the renewable energy sources have been implemented such as solar, thermal, hydro, etc., even though they are economically friendly they can't be used for small appliances because of their ratings and limits of usage [3, 4]. To use the generated electrical energy for small loads and appliances, piezo materials have been implemented which are small disk-like structures. These sensors when exerted by the pressure the voltage is generated, and the energy is being utilized by the loads such as street lights, lamps, small drives, and low-rated house hold appliances [5–8].

The word piezo was obtained from Greek word piezein, which means to press or to squeeze. The first demonstration was in the year 1880 by the curie brothers. They obtained the piezoelectricity by the combined knowledge of pyro electricity along with the crystal structures to study the crystal behavior and demonstrated the same using different crystals of quartz, topaz, cane sugar, Rochelle salt and most of the piezo sensors are of quartz type [9, 10]. When the sensor has subjected to pressure or stress the quartz has changed its dimensions and produce the electric charge. The most widely used piezo material was perovskite structured Pb which is extensively used in various types of electromechanical sensors and actuators, etc. [11, 12].

Constructional Features

The components that are used for the foot step power generation are piezo sensors, microcontroller, rechargeable battery, LED, and LCD display. The connections are made according to the below Fig. 1.

Piezoelectric Sensors

Figure 2 shows the piezo sensors which work on the principle of piezoelectric effect, i.e., when the pressure is exerted on the sensors, they produce an electric charge on the crystal surface like quartz. These sensors produce different rates of electric charges based on the position of pressure applied on the sensor. In order to obtain high amount of voltage, these sensors are connected in series connection. With the increase in amount of no. of series connection of sensors the voltage increases [13]. Figure 2 shows the design of a piezoelectric sensor.

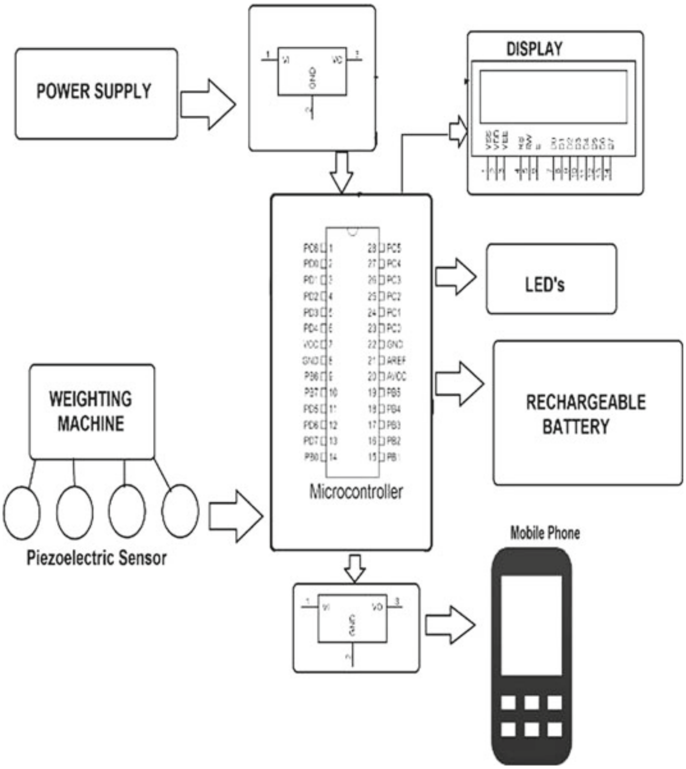


Fig. 1 Hardware circuit of implementation

Fig. 2 Piezoelectric sensor

When the sensor is excited to pressure, it produces an electrical output by the rearrangement of dipoles in them. The rate of voltage generated is directly proportional to the pressure or stress that is being applied on the sensor.

The output voltage that is generated from the sensor depends upon the position of pressure applied on the sensor if the pressure applied on the midpoint of sensor then maximum voltage is obtained and is utilized [14, 15]. Figure 3 shows the voltages at different pressure points on the sensor.

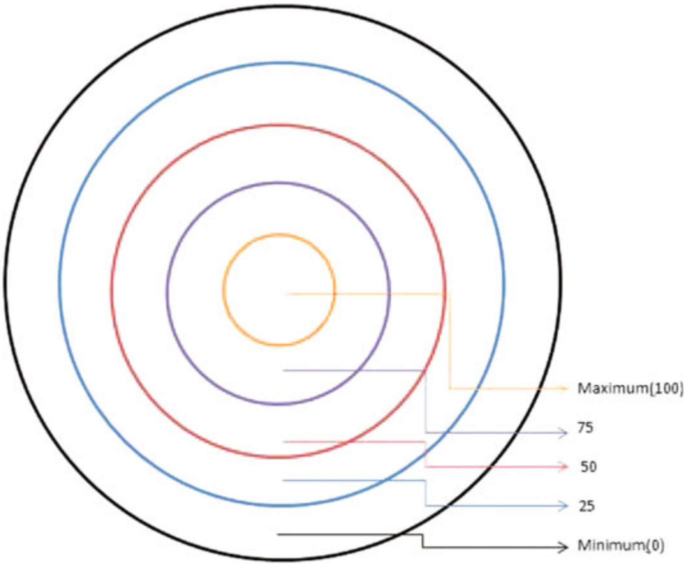


Fig. 3 Voltage pressure points on the piezo sensor

Microcontroller

The below Fig. 4 shows the microcontroller circuit device which acts as the main CPU of the system, i.e., central processing unit which has I/O ports and analog

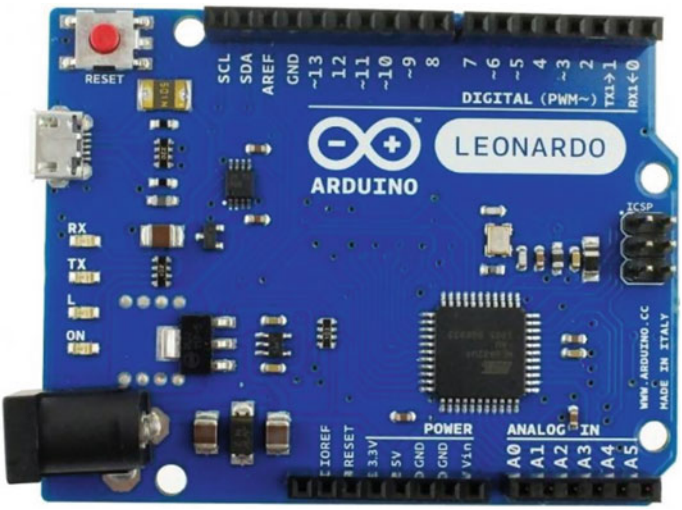


Fig. 4 Microcontroller

to digital converter pins, etc. This microcontroller is used to measure and display the no. of footsteps and the voltage generated when piezo sensor is squeezed. This also measures the battery voltage level using analog to digital converter present and based on the SOC status of battery the voltage generated by the piezo sensor is used to charge the battery.

Rechargeable Battery

Rechargeable battery is used as storage device or a backup device where the generated output power is stored and is utilized when there is an interruption in power supply or failure of the system.

LCD Display

Liquid crystal display is a flat display which uses light modulating properties of liquid crystals. This LCD is used to display the count of no. of footsteps and the voltage that is obtained overall with piezoelectric sensors.

Led

LED belongs to the family of diodes, where the light is being emitted. These LED's are placed on the piezo sensors which are used to indicate that the voltage is generated by glowing the LED when the pressure is being applied on the sensor, if the LED doesn't glow that indicates the voltage is not generated.

Working

The footstep power generation works on the principle of piezoelectric effect principle. This can be explained as the ability of materials for generating electric charges in response to the applied force or stress on the sensor. As the sensors are made of quartz material which contains crystals in them which are neutral in nature. Atoms present inside them are not symmetrically arranged but are equally charged. When the pressure is being applied on the sensor it gets compressed and when the pressure is removed it gets decompressed. During the process of compression and decompression the mechanical vibration is being sensed by the crystals present in them thus, the voltage is being produced by both net positive and negative charges that appear on the opposite face of crystals. The voltage is obtained by indicating the

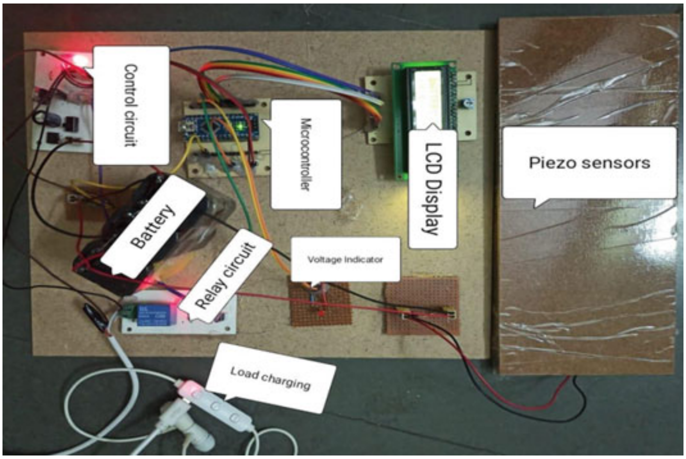


Fig. 5 Hardware implementation of footstep power generation

LED glow mentioning that voltage is generated. The voltage that is obtained is fed to the rechargeable battery and loads through the microcontroller which acts as a central processing unit where the battery voltage level is measured by analog and digital converters present in the controller.

Based on the SOC (state of charge) status level, it indicates that whether the battery is to be charged or not. This SOC status is known as threshold voltage, if the voltage obtained is more than the threshold voltage the load lamp glows and if the obtained voltage is less the load lamps doesn't glow. And thus, the electrical power is generated and utilized. The obtained voltage from the footsteps from the piezo sensor is displayed on the LED indicating no of footsteps and overall voltage. The below Fig. 5 shows the hardware implementation of footstep power generation.

Observation Results

The below Table 1 describes the detailed values of the complete hardware implementation.

Table 1 Steps required to charge the battery up to 1V

Applied weight (kg)	Steps require to charge 1V
1.2	115
1.6	102
1.8	93
2.0	80

Table 2 Hardware requirements

Equipment	Type/Rating	Quantity
Microcontroller	ATMega328	1 No
Piezoelectric sensors	PMN-PT	7 No
LCD display	16×2	1 No
Rechargeable battery	4 V, 1.5Ah	3 No
LED's	-	As required

Conclusion

This paper concludes that as the usage of electrical energy is increasing with the demand, in order to utilize the energy for small application like streetlights, led lamps, etc., which in-turn saves the energy produced by other sources which can be used for large appliances. Thus, the piezo sensors are used for generating the energy and to utilize it in a proper perspective manner. Unlike other sources, the mechanical energy is of abundant and can be reusable as it is produced by the humans daily by their working environment. Thus, it can be concluded that efficient energy can be produced and utilized from the piezo sensors.

Appendix

Table 2 shows that the hardware requirements of this research work.

References

1. Minami Y, Nakamachi E (2012) Development of enhanced piezoelectric energy harvester induced by human motion. || Engineering in medicine and biology society (EMBC), 2012 annual international conference of the IEEE
2. Karuppusamy P (2020) Synchronization of reactive power in solar based DG and voltage regulated elements using stochastic optimization technique. J Electr Eng Autom 2(1):50–59
3. Suma V (2021) Power efficient time-division random-access model based in wireless communication networks. IRO J Sustain Wireless Syst 2(4):155–159
4. Karrupusamy P. Modified symbiotic organisms search for plug-in hybrid electric vehicles through renewable micro-grids. J Electr Eng Autom 2(2):76–83
5. Triono AD et al. (2018) Utilization of pedestrian movement on the sidewalk as a source of electric power for lighting using piezoelectric censors. In: 2018 3rd IEEE international conference on intelligent transportation engineering (ICITE). Singapore
6. Patil A, Jadhav M, Joshi S, Britto E, Vasaikar A (2015) Energy harvesting using piezoelectricity. 2015 International conference on energy systems and applications. Pune, India
7. Jin Y, Sarker S, Lee K, Seo HW, Kim DM (2016) Piezoelectric materials for high performance energy harvesting devices. 2016 Pan Pacific microelectronics symposium (Pan Pacific). Big Island, HI, USA

8. Dharmambal V, Nisha KCR, Bhavana C (2016) Piezo film based renewable energy system. 2016 international conference on circuit, power and computing technologies (ICCPCT). Nagercoil, India
9. Songsukthawan P, Jettanasen C (2017) Generation and storage of electrical energy from piezoelectric materials. In: 2017 IEEE 3rd international future energy electronics conference and ECCE Asia (IFEEC 2017—ECCE Asia). https://doi.org/10.1109/ifeec.2017.7992403
10. Li T, Xia Z, Chuan J, Luan H, Analysis of the characteristics of piezoelectric sensor and research of its application‖. 18th IEEE international symposium on the applications of ferroelectrics on 23–27 August 2009. pp 1–4
11. Ottman GK, Hofmann HF, Bhatt AC, Lesieutre GA (2002) Adaptive piezoelectric energy harvesting circuit for wireless remote power supply ‖. IEEE Trans Power Electron 17(5):669–676
12. Kumar D, Chaturvedi P, Jejurikar N, Piezoelectric energy harvester design and power conditioning ‖, IEEE students' conference on electrical, electronics and computer science (SCEECS) on 1–2 Feb 2014. Circuits. ‖ Applications of ferroelectrics
13. Design study of piezoelectric energy (Feb 2010) Harvesting devices for generation of higher electrical power using a coupled piezoelectric-circuit finite element method IEEE transactions on ultrasonic's. Ferroelectr, Freq Control 57(2)
14. Panghate S, Barhate P, Chavan H (Feb 2020) Advanced footstep power generation system using RFID for charging. Int Res J Eng Technol 7(02)
15. Kamboj A, Haque A, Kumar A, Sharma VK, Kumar A (2017) Design of footstep power generator using piezoelectric sensors. In: 2017 International conference on innovations in information, embedded and communication systems (ICIIECS). Coimbatore, India

Development of Automated Protection and Monitoring System for Poor Railway Infrastructure

Tareq Anwar Shikdar, Md Moontasir Rashid, Fahad Bin Ayub, and Sekh Faisal

Abstract This research is based on developing different sub-models for maximizing railway protection. The main motive of developing the railway protection and monitoring system is to help the railway coordinators and locomotive drivers to ensure proper safety to the railway passengers. Three individual sub-models are accumulated to develop the entire model. Firstly, in order to minimize level crossing accidents, the gateman will be notified about the exact arrival time of the train. In addition, by using sensors, the crossing gate will be automatically closed if any train crosses the level crossing. Moreover, the proposed model can scan the train routes in order to avoid collision accidents with other trains or any obstacles across the route. This entire model used many modern technologies, and it is almost errorless. The countries which possess poor rail infrastructure can implement this cheap and reliable protection and monitoring system in order to avoid major rail accidents.

Keywords Vibration sensor · Arduino · Proximity sensor · SIM 800A GSM module · Solar panel · Servo motor

Introduction

The railway system is one of the major transportation systems for moving passengers and freights. Bangladesh railway carried 65 million passengers and 2.5 million tons of freights [1]. Besides, those railway-related accidents are more dangerous than any other transportation [2]. One of the causes of rail accidents is the presence of a railway system manually; this system can be automated for reducing the accidents at a high ratio. Also, the management of train in Bangladesh is so weak that scheduling can be a cause of accidents. Finally, as a result, there are so many reasons for railway accidents, especially four factors that have major impacts. The factors of train accidents are namely human and external, facilities, infrastructure, but human factors are the highest factors among all of these [3]. By human errors

T. A. Shikdar · M. M. Rashid (✉) · F. B. Ayub · S. Faisal
Department of Electrical and Electronic Engineering, Leading University, Sylhet, Bangladesh

V. Bindhu et al. (eds.), *Proceedings of Third International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 844,
https://doi.org/10.1007/978-981-16-8862-1_35

railway level crossing, train collision-related accidents occurred frequently. On 8th December, 2010, ten people were killed by collision; on 23rd June, 2019, several bogies of Dhaka-bound Upaban Express train from Sylhet veered off the tracks at around 11:40 pm after a culvert over the Barochhara Canal broke down, 200 yards off Baramchal Railway Station. At least, five people died, and hundreds were injured. Locals who worked overnight with the first responders claimed at least 10–15 people have been killed. On 12 November, 2019, at least 16 passengers died and score other injured as Dhaka-bound Turna Nishita Express. Rammed Chattogram-bound Udayan Express from Sylhet at Mandobagh Railway Station in Kasba. Train driver's fault took sixteen lives [4]. In Bangladesh, frequent accidents are caused by human errors in level crossing, collisions due to train driver's faults and mechanical problems in the construction of trains. Railway level crossing accidents and collisions between two trains are the major incidents in Bangladesh railway. Focusing on this section, this paper is proposing three individual models for resolving this scenario. For concluding the accidents in railway level crossing, this paper proposed two systems called (i) notifying gateman regarding the arrival of the train (train alerting system) and (ii) automatic railway level crossing gate. Also, for diminishing collision between trains, a system is proposed called train collision prevention system. For reducing the death ratio by rail accidents, Bangladesh Railway System should be automated. The aim of this paper is to provide various methods and devices to reduce accidents and implement automation in Bangladesh Railway System.

Literature Review

With the advancement of technology and better transportation management, train accidents are decreasing day-by-day in everywhere the world. In an underdeveloped country, there is no automated systems and advanced technology for reducing the accident ratio. But, there are so many works in railway which has been done by railway workers. Opening and closing railway level crossing is one of them. This process can be automated. For reducing human errors in this field, the railway system needs some algorithm for detecting the trains [5, 6]. In 2017, Amin Saruhjono proposed a system that provides advance notification to the coordinator for pulling down the railway level crossing gate by sensing the vibration of the train [7]. In 2019, a research team proposed a system for automatic railway crossing gates by automata [8]. The unmanned system was also designed in research that no human participation is required in pulling down the gate of railway level crossing [9]. Authors also discuss some processes in automatic railway control using IR sensor and USB UART board [10]. Before this, in 2014, Mingjian Sun and his team were researching the detection of high-speed train defects by rail vibration [11]. Additionally, some researchers had worked on accelerometer sensors for detecting the train [12, 13]. Detecting trains for automatic gate crossing were top ahead researches for railway protection. There are so many surveys which helps to create a different, unique, and affected system such that machine language and image processing methods [14–16].

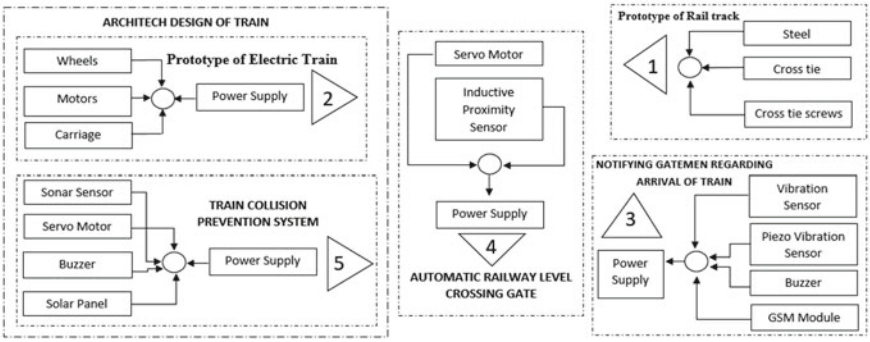


Fig. 1 System design of the proposed automated railway protection and monitoring system

Like level crossing subsystem with image processing [17] and OpenCV with buzzer and led bulb were done by the researchers [18]. In 2019, Aniqa and her team placed a proposal for the automatic gate that presents UML and format model of RGCS model using DFA [19]. In the sector of reducing anti-collision, RFID was used [20]. Authors have proposed several kinds of research for anti-collision between trains by GPS mechanism [21–24]. Before that research, Frederic proposed a system-based vision for anti-collision [25]. But, our work is different from others such that the technologies and sensors are used in this paper which are unique, low cost, and more accurate. However, all this research works are contributed to improve the railways and minimize accidents.

Methodology

Figure 1 shows the entire model of the proposed system. The components, motors, and sensors were implemented as the system design described. First of all, a prototype of rail track and electric train was built. The device for reducing the accidents in railway level crossing depends on vibration of train which is detected using two sensors. An automatic level crossing gate is also implemented. Lastly, in order to prevent train collision, a device is built using an ultrasonic sensors so that it can scan the pathway. İn addition, GPS can be used for viewing real-time monitoring of trains. Figure 2 shows the system process of the proposed system.

Train and Railroad Infrastructure Construction

Rail tracks have some basic configurations like rail geometry, track measurement, corrugation, and vehicle dynamics [26]. In this paper, a rail track and train were built

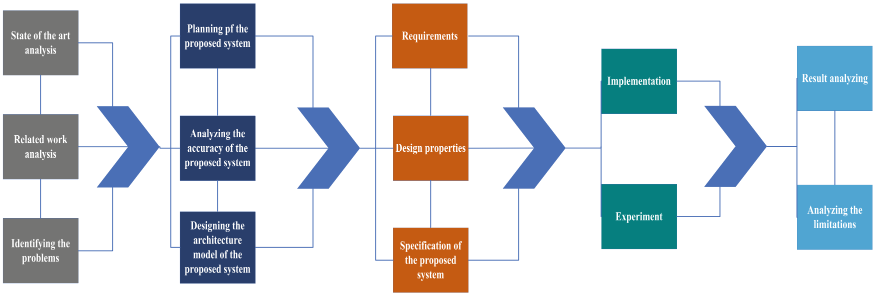


Fig. 2 System process of the proposed automated railway protection and monitoring system

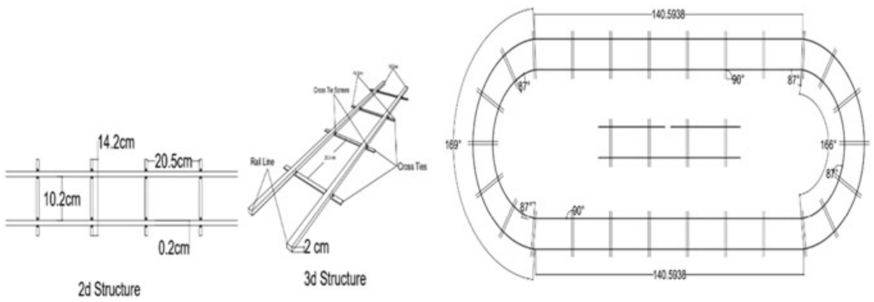


Fig. 3 Structure and measurement of the railroad

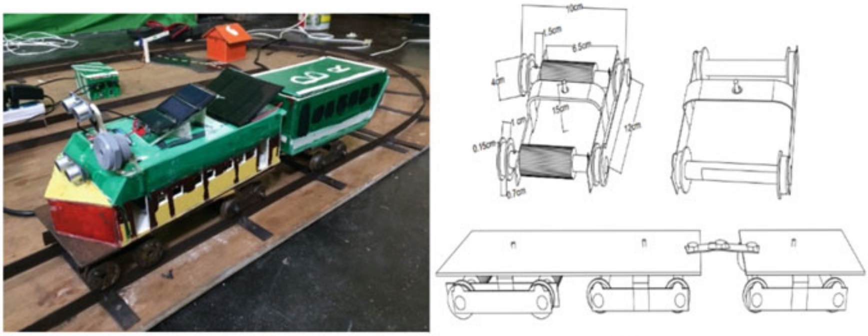
as a prototype to experiment the system and how it works on the train over the rail track.

Prototype of Rail Track

A 4.5 m long rail track made of pure steel on a wooden board was built. The distances between cross-ties and rail widths are 20.5 cm and 10.2 cm, respectively, shown in Fig. 3. Rail track was fit over cross-ties by using cross-tie screws. On this board, the prototype electric train and three individual models were experimented, and further, analyses were taken for improvement of this proposed system.

Prototype of Electric Train

A 42 × 13 cm electric locomotive train is built with 12 wheels, and a speed control system is illustrated in Fig. 4a. A mini-solar panel is attached to the roof of the train with sensors and a buzzer. Two DC motors are connected to the wheels. Each of the



(a) (b)

Fig. 4 **a** Prototype of electric train **b** Structure of the electric train

gear motors is rated 12 V and high torque of 300 rpm. Each carriage contains four wheels which are connected by a connecting rod that is fixed with the backside of the wheels shown in Fig. 4b. One lithium polymer battery is connected with the motors to supply power in the first carriage. To connect the carriages, a chassis over the wheels is added shown in Fig. 5. This prototype train is designed for taking the result when the proposed system is running. For noticing the accuracy of the proposed system, this train is used for observing real-life scenarios. Prototype train helps to collect the experimental calculation of each proposed system.

Fig. 5 Connections of servo motor and wheels over the chassis

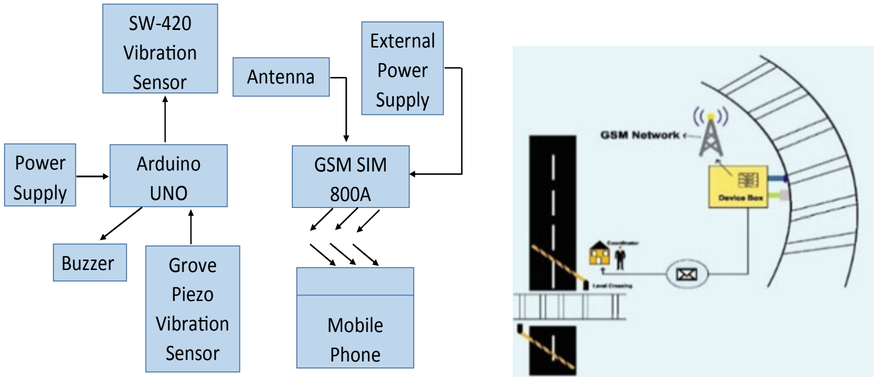


Fig. 6 Block diagram of connections of train alerting system (left) and general scheme (right)

Notifying Gateman Regarding the Arrival of Train (Train Alerting System)

The proposed system consists of an alerting system that uses the vibration of the train and sends messages to the crossing level coordinator/gateman in order to close the crossing level gate.

Block Diagram of the Train Alerting System

In Fig. 6, the block diagram of train alerting system is depicted. One Arduino is used in this device for the operation of following functions:

1. To sense the vibration of the train by using vibration sensor as input
2. To detect the train's arrival as input
3. Send the status of the train's arrival to the crossing level coordinator as output.

Flowchart

Figure 7 describes the flowchart that was followed while designing and experimental testing.

Circuit Diagram of Alerting System

In Fig. 8, Piezo vibration sensor has two digital pins which is connected to Arduino's digital pin, and the SW-420 vibration sensor is connected to the Arduino as well as

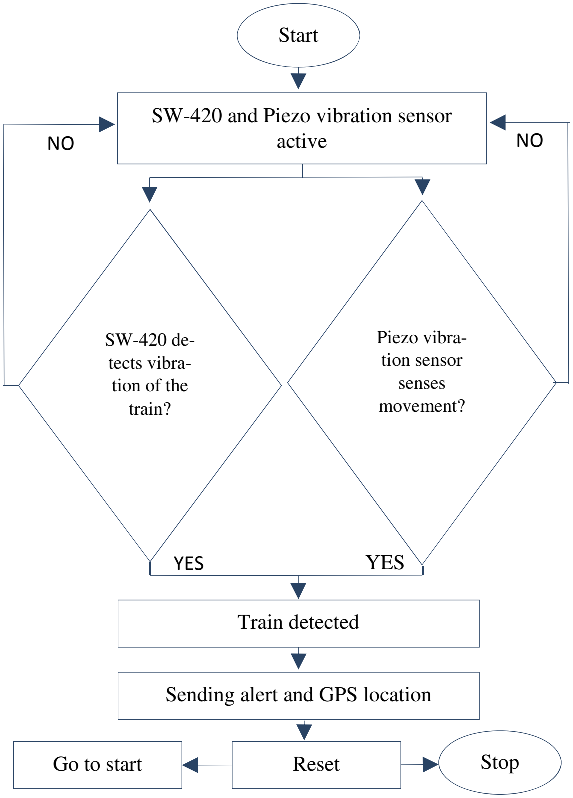


Fig. 7 Flowchart of the train alerting system

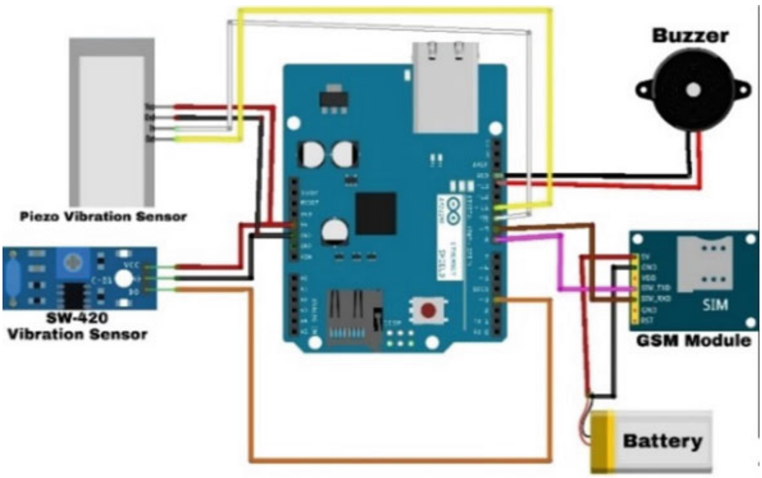


Fig. 8 Circuit diagram of the train alerting system

the GSM module with an external power source, and also, a buzzer is connected to the digital pin of the Arduino including GND pin. All the sensors get power from the battery through Arduino using Vin port and GND port. Including, for network establishment, local SIM was inserted.

Hardware Development for Detecting the Vibration and Movement Produced by Train

The alerting device consists of following items that is described below:

- Vibration Sensor Module: In the device, SW-420 motion vibration sensor is used. It switches alarm electronic module for Arduino at 3.3 V–5 V with digital switching output. When this sensor measures the vibration value produced by the train, it transfers the signal to the Arduino for further operation. This sensor is placed in the nearest track line of the level crossing.
- Piezo Vibration Sensor: For the verification of the train's arrival toward the level crossing, Grove-Piezo vibration sensor is used. This sensor is based on PZT film sensor LDT-028, and it is specially designed for detecting or measuring touch and vibration. In the track, when the train touches this sensor, the Arduino takes input from it for completing the further operation. This sensor is placed in the nearest place where the SW-420 vibration sensor was placed.
- Buzzer: One buzzer was used for producing a high sound alarm. This buzzer was connected to the Arduino. When SW-420 and Grove-Piezo vibration give a signal to the Arduino, the buzzer beeps with a high sound alarm as output. The buzzer always beeps, if two sensors give signals at a time.

Hardware Development of GSM Module

In this device, SIM 800A was used in a GSM module. When the SW-420 vibration sensor detects the vibration of the train and train touches the Grove-Piezo vibration sensor at a time, it gives a signal to the Arduino and commands the GSM module to send a message to the coordinator for closing the crossing level gate. Figure 9 shows the alerting system and message sent by the GSM module.

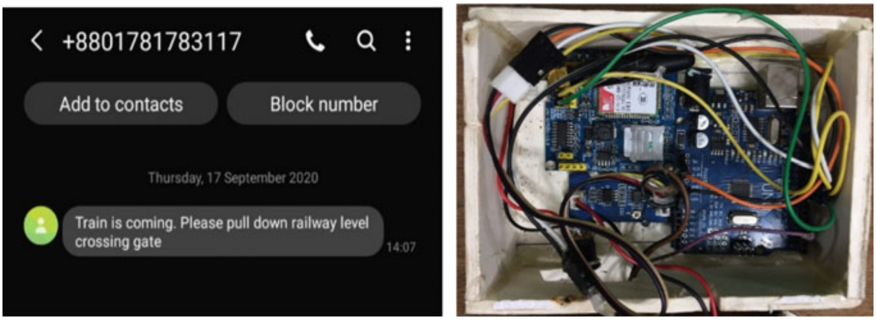


Fig. 9 Message sent by GSM module and train alerting system device box

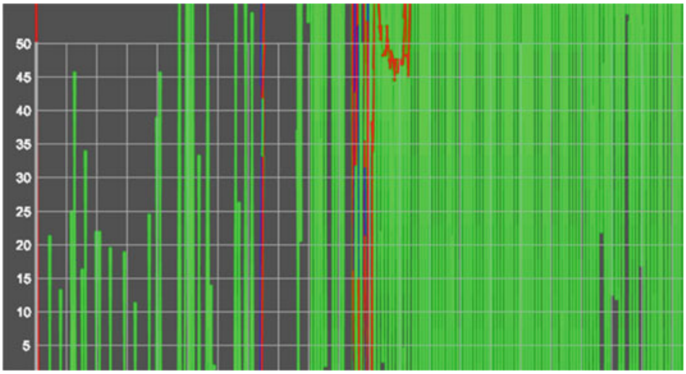


Fig. 10 Seismic graph of the vibration of the train

Experimental Result

The vibration produced by the train is analyzed and programmed into Arduino with the help of a certain logarithm, where the X-axis is the time, and the Y-axis is the vibration level. Figure 10 shows the seismic graph.

Performance Analysis

Overall, the performance of the alerting system is excellent. From Table 1, it can be seen that on the 3rd time, the SW-420 vibration sensor couldn't detect the train as it was not getting proper power supply from the Arduino. This error has been minimized by rechecking the circuit connections.

Table 1 Performance analysis of the train alerting system

Exp. No.	Vibration of the train	Piezo vibration sensor	Buzzer	Alert message
1	Detected	Touched the train	Beep	Sent
2	Detected	Touched the train	Beep	Sent
3	Not detected	Touched the train	No beep	Not sent
4	Detected	Touched the train	Beep	Sent
5	Detected	Touched the train	No beep	Sent
6	Detected	Touched the train	Beep	Sent
7	Detected	Touched the train	Beep	Sent
8	Not detected	Touched the train	No beep	Not sent
9	Detected	Touched the train	Beep	Sent
10	Detected	Touched the train	No beep	Sent

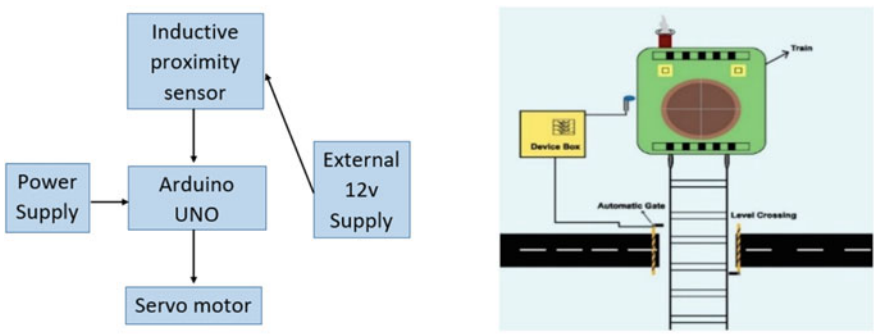
Automatic Railway Level Crossing Gate

The proposed system consists of a device which is called the automatic gate. When a train crosses the level crossing, the gate will be pulled down without any human interface. A surveillance camera will be set up to monitor the gate continuously.

Block Diagram of Automatic Gate

Figure 11 shows the block diagram of the device where Arduino functionalized these as input and output:

1. To sense the arrival of train



(a) (b)

Fig. 11 **a** Block diagram of automatic railway level crossing gate **b** General scheme

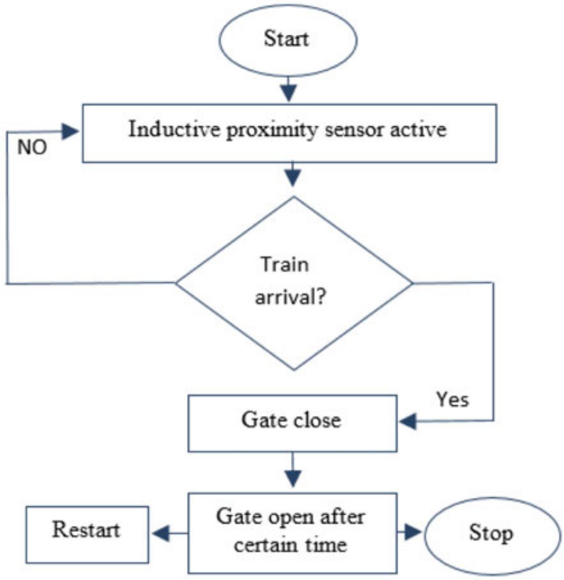


Fig. 12 Flowchart of automatic level crossing gate

2. To pull up and pull down the crossing level gate automatically by using two servo motors.

Flowchart of Automatic Gate

Figure 12 describes the flowchart that was followed while designing and testing.

Circuit Diagram of Automatic Gate

In Fig. 13, two switching modes were implemented for the buck module and servo motor. Servo motor and triggered circuit of inductive proximity sensor were connected to Arduino.

Power Supply for All Components of This System

- Power Supply: Two adapters were used in this system to covert AC to DC at 12–2A
- Buck Module: One buck module was used to step down the voltage. This buck module was used to supply 5 V to the system.

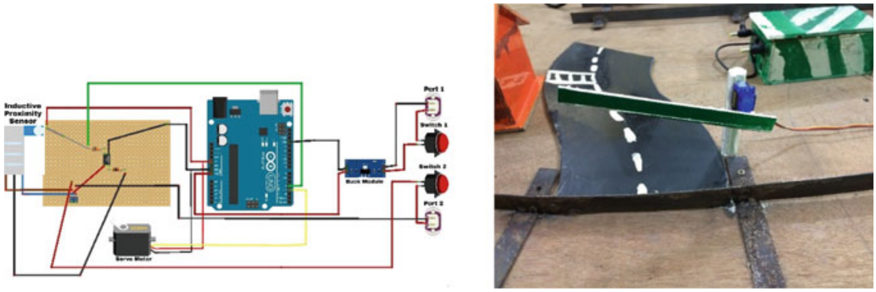


Fig. 13 Circuit diagram of automatic railway level crossing gate

Hardware Development for Detecting Train

One NPN inductive proximity sensor was used to detect the train and to pull down the level crossing gate automatically. Figure 14 shows the placement of the sensor beside the rail track. This sensor can detect any metallic objects from a 4 mm distance. As the train's body is metallic, so the sensor could detect the train and send a signal to the Arduino. The operating speed and voltage of one servo motor are 0.12 s/60 degrees and 4.2–6 V. When Arduino takes input from NPN inductive proximity sensor, the servo motor rotates at a fixed degree as Arduino.

Train Collision Prevention System

This device is an anti-collision device that notifies the locomotive driver when any obstacle comes to collide with the train. When this device detects the obstacle, the high sound alarm will be turned on in the engine room so that the driver could take the control of the train. There are two ways to detect obstacles by the anti-collision device. One is by using an ultrasonic sensor, and another is by using GPS.

Fig. 14 Sensor for detecting train beside the rail track

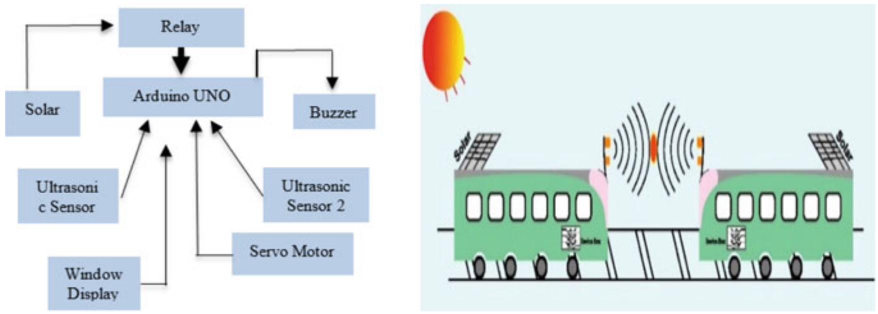


Fig. 15 Block diagram (left) and general scheme (right) of anti-collision device

Block Diagram

Figure 15 shows one Arduino is used for performing below functions:

(1) Scanning the railway path to find out any obstacles or train ahead
(2) Alerting the locomotive driver.

The ultrasonic sensor is used as input. For a visual representation, a radar is included by using an ultrasonic sensor and servo motor.

Flowchart and Circuit Diagram of Anti-collision Device

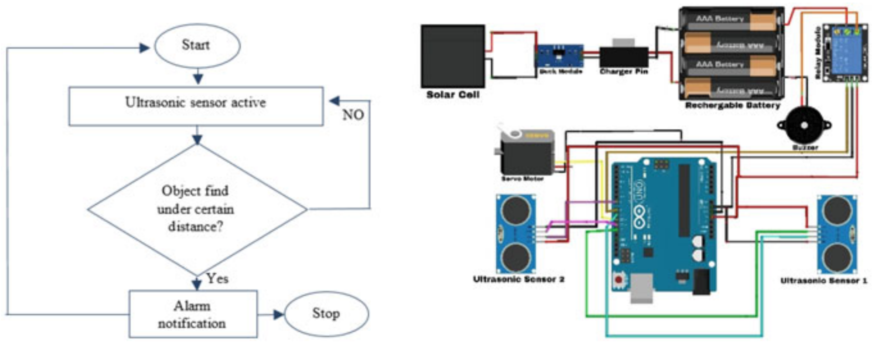
One Arduino was used to functionalize these:

1. To obtain a signal from the accelerometer and emergency switch
2. Send output signal as GPS location by message to particular numbers.

Figure 16 shows the flowchart and the circuit diagram of this device.

Power Supply for All Components of Anti-collision Device

- Power Supply: One 4 s Li-ion battery was used in this system. Each cell batteries were 3 V.
- Buck Module: One buck module was used in this system for converting electricity coming from the solar panel and stores into the battery.
- Solar Panel: Four solar panels were used in this system. One solar panel is rated 9 V, and rest of the solar panels were of 3 V.



(a) (b)

Fig. 16 **a** Flowchart and **b** Circuit diagram of anti-collision device

Hardware Development for the Detecting Obstacles

- Ultrasonic sensors: Two ultrasonic sensors were used to find out any obstacles. The sensor was fitted in front of the train. The sensors can sense about 1 foot. If sensors sense any obstacle at this certain distance, it sends a signal to the Arduino for further operation. Components used in this device are described below:
- Buzzer: One high sound alarm buzzer was used for alerting the driver. When ultrasonic sensor 1 senses the obstacles, then Arduino takes input signal from it and processes the output signal as a high sound alarm.
- Relay: One channel 5 V relay module was used in this system. This relay is connected with Arduino. Ultrasonic sensor 1 senses obstacles in every 10 s with the help this relay module. The relay was used only for the ultrasonic sensor.
- Servo motor: One servo motor was used for rotating the ultrasonic sensor 2 in order to scan the path
- Window display: A window display is used for ultrasonic sensor 2. As this sensor obtains any input, window display shows the visual scanning mode of obstacles under a certain distance.

Figure 17a shows the window display of ultrasonic sensor 2 where situation 1, 2, 3, and 4 describes the distances between train and obstacle in different angles. Also, red and green colors define the obstacles and range of scanning. Figure 17b shows the placement of sensors over the roof of the train.

Anti-collision Device by Using GPS

For reducing train collision, GPS methods can be used [26, 27]. All local and intercity trains are provided with a GPS device which will provide real-time monitoring

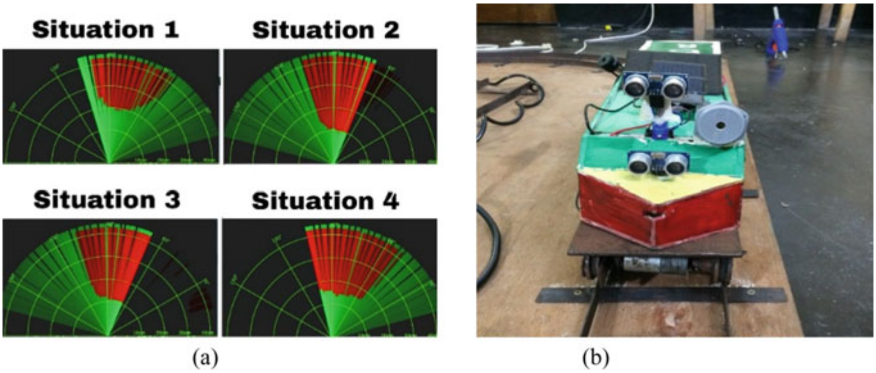


Fig. 17 **a** Radar screen shows results provided by ultrasonic sensors 2 **b** Sensors and solar panels installed over the roof of the train

system. All the locomotive drivers are provided with an application in cellphones where they can find the location of each trains. Figure 18 shows the general scheme with GPS system. Figure 19 shows the location of a train on the map of Bangladesh.

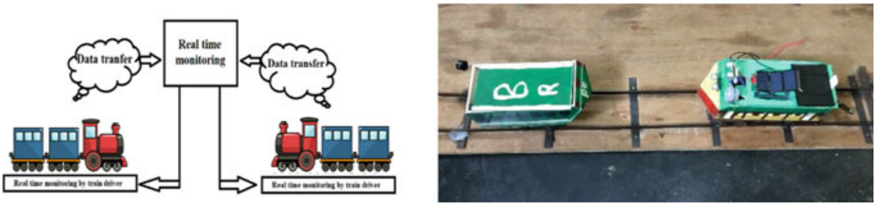


Fig. 18 General scheme of anti-collision device with GPS module

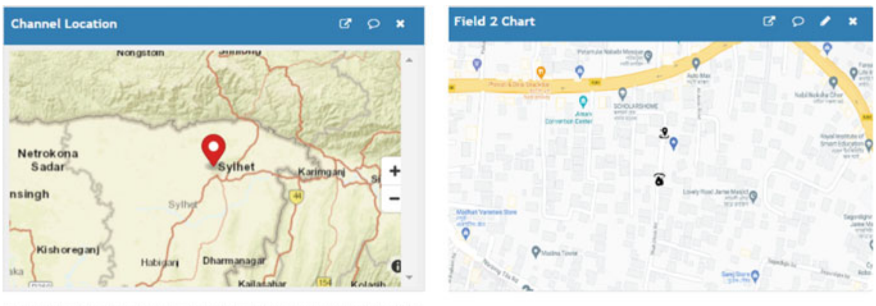


Fig. 19 Location tracking with GPS module

Table 2 Performance analysis of anti-collision device

Exp. No.	Distance between train and obstacles (cm)	Alarm notification status
1	2	Worked
2	4	Worked
3	5	Worked
4	7	Worked
5	1	Worked
6	15	Worked
7	20	Worked
8	25	Worked
9	26	Not worked
10	30	Not worked

Performance Analysis of Anti-collision Device

Table 2 is the observations taken while experimenting with this device. The sensors worked properly under 25 cm. At the range of 30 cm, the sensor didn't work properly because of its limited capacity.

Overview of the Automated Railway Protection and Monitoring System

Table 3 shows the apparatus and equipment required for building this system. Figure 20 shows the entire prototype built to implement the idea and increase safety where poor railway infrastructure is used (Table 4).

Future Work

In this proposed system, the proposal of each system has some limitations; such monitoring system can't be controlled anywhere so that IoT system can be implemented in future work. As well as for reducing the cost efficiency, the overall system can be established in a PLC system and connected the overall system from each railway station and each monitoring station in the software-based monitoring system.

Table 3 Required components with rating and quantity

Serial No.	Components name	Rating	Quantity
1	DC gear motor	12v, 300RPM	2
2	Micro-controller	Arduino Uno	2
3	GSM module	SIM900A controls with AT command and SIM800A	2
4	Antenna	Qaud-banned antenna, operates four frequencies	3
5	SIM network	Local mobile network	2
6	Wheels	Metal	12
7	Chassis	Metal	4
8	Rail track	Metal	As required
9	Cross-tie and screws	Metal	As required
10	Inductive proximity sensor	NPN type, range 0.4 cm	1
11	Ultrasonic sensor	Range 1 feet	2
12	Vibration sensor module	SW-420	1
13	Grove–Piezo vibration sensor	PZT film sensor LDT-028	1
14	Servo motor	Operating speed and voltage of the servo motor are 0.12 s/60 degrees and 4.2–6v	2
15	Buck module	DC-DC,12v to 5v	3
16	Relay module	1 channel, 12v	1
17	Charger pin	B-type	1
18	Battery (1)	4 cell at 3v	1
19	Power source	Lipo battery, 12v and 1800 mAh and 2200 mAh	3
20	Wire	Silicon wire (max current up to 70A)	As required
21	Solar panel	6v–9v	4
22	PVC board	10 square feet	1
23	Cardboard	12 square feet	1

Conclusion

The system had been tested on a realistic scenario involving a rail mechanism to obtain accurate results. Difficulties had been faced cause of the unavailability of some elements in our country. The main problem is the scarcity of open-source research resources. The main intention of this paper is to increase and raise awareness about the railway's hazards, and thus, even the smallest contribution can make a big impact. Bangladesh Railway Ministry appraised the idea of a railway protection and monitoring system and advised us to improve this project and test it several times so

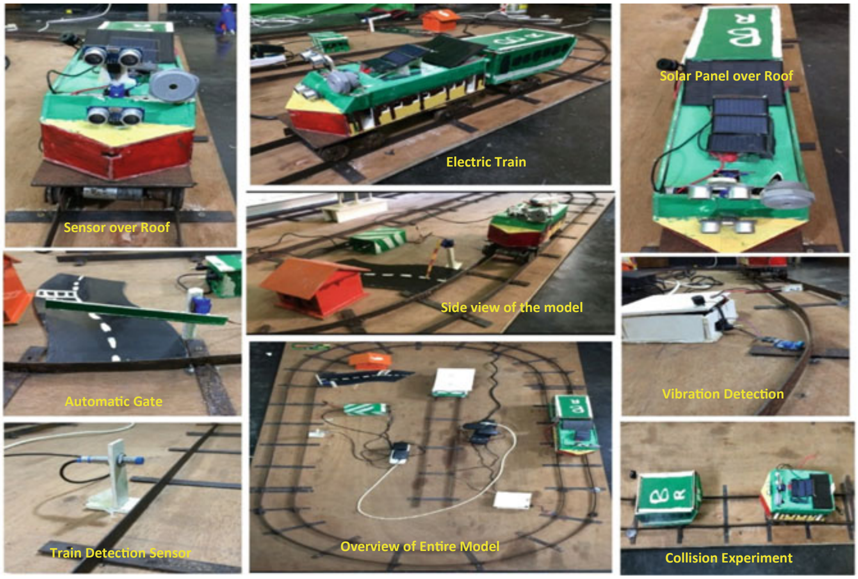


Fig. 20 Prototype of the automated railway protection and monitoring system

Table 4 Overall success rate of the proposed system

Proposed system	Experimental times	Success rate
Alerting system	**10**	80%
Automatic level crossing gate	**10**	100%
Anti-collision device	**10**	87.5%

that it can be accredited by renowned institutions for practical use. The aim of this project is to provide 100% safety to railway passengers. This research can be used to find solutions and can be implemented and modified to suit the needs of the system it is incorporated into.

References

1. (20 April 2017) "Analysis of problems" Bangladesh railway-government of the People of Republic Bangladesh
2. Pwint HNY, Tun ZM, Tun HM (May 2014) Automatic railway gate control system using microcontroller. Int J Sci, Eng Technol Res (IJSETR) 3(5)
3. Fatahurrohman R (2012) Human error dan Kaitannya pada kecelakaan Kereta Api di Indonesia_APLIGO
4. Wikipedia (19 Sep 2020) "Bangladesh railway". Retrieved from https://en.wikipedia.org/wiki/Bangladesh_Railway

5. Palmer J (10 March 2017) The need for train detection. 11th IET professional development course on railway signaling and control systems (2006). https://doi.org/10.1049/ic.2010.0088 Web
6. Gasparini F, Schettini R, Gallina P (2001) An innovative algorithm for train detection. Internet Imaging III. 4672:280–286
7. Saruhjono A, Wardihani ED, Hardani YFK, Analysis of the vibration signal detection for rail train arrival. 2017 IEEE ınternational conference on communication, networks and satellite (Comnetsat)
8. Rehman A, Latif S, Zafar NA (2019) Automata based railway gate control system at level crossing. Int Conf Commun Technol (ComTech) 2019:30–35. https://doi.org/10.1109/COMTECH.2019.8737833
9. Banuchandar J, Kaliraj V, Balasubramanian P, Deepa S, Thamilarasi N (Jan–Feb 2012) Automated unmanned railway level crossing system. Int J Mod Eng Res (IJMER) 2(1):458–463
10. Chowdhury KBQ, Khan MR, Razzak MA (2020) Automation of rail gate control with obstacle detection and real time tracking in the development of Bangladesh railway. 2020 IEEE 8th R10 humanitarian technology conference (R10-HTC). pp 1–6. https://doi.org/10.1109/R10-HTC49770.2020.9356986
11. Sun M, Wang Y, Zhang X, Liu Y, Wei Q, Shen Y, Feng N, Feature selection and classification algorithm for non-destructive detecting of high speed rail defects based on vibration signal
12. Amjath MIM, Kartheeswaran T (2020) An automated railway level crossing system. Int Conf Image Process Robot (ICIP) 2020:1–7. https://doi.org/10.1109/ICIP48927.2020.9367346
13. Singh G, Kumar P, Mishra RK, Sharma S, Singh K (2020) Security system for railway crossings using machine learning. 2020 2nd ınternational conference on advances in computing, communication control and networking (ICACCCN). pp. 135–139. https://doi.org/10.1109/ICACCCN51052.2020.9362976
14. Smitha TV (2021) A study on various mesh generation techniques used for engineering applications. J Innovative Image Process 3(2):75–84
15. Pandian AP (2019) Revıew of machıne learnıng technıques for volumınous ınformatıon management. J Soft Comput Paradigm (JSCP) 1(02):103–112
16. Sungheetha A, Sharma R (2020) A comparative machine learning study on IT sector edge nearer to working from home (WFH) contract category for improving productivity. J Artif Intell 2(04):217–225
17. Angrisani L, Grillo D, Moriello RSL, Filo G (2010) Automatic detection of train arrival through an accelerometer. 2010 IEEE Int Instrum Meas Technol Conf I2MTC 2010—Proc (i):898–902
18. Jiang B, Wang Z, Railway vibration monitoring system based on ARM and acceleration. Ind Electron …, (213019):998–1004
19. Rehman A, Latif S, Zafar NA, Automata based railway gate control system at level crossing. 2019 ınternational conference on communication technology (ComTech)
20. Zhao J, Wu L, The ımprovement of RFID anti-collision algorithm. 2012 24th Chinese control and decision conference (CCDC)
21. Ruhai G, Huiyan W (2010) Research on the anti-collision mechanism based on GPS. Second WRI Glob Congr Intel Syst 2010:55–57. https://doi.org/10.1109/GCIS.2010.40
22. Wang W, Zhou R (2009) Research on the anti-collision system of surface coal mine based on the highly accurate GPS location technology. 2009 9th International conference on electronic measurement & ınstruments. pp. 3-196–3-199. https://doi.org/10.1109/ICEMI.2009.5274323
23. Chen Y-J, Chen C-C, Wang S-N, Lin H-E, Hsu RC (Oct 2006) GPSenseCar—a collision avoidance support system using real-time GPS data in a mobile vehicular network. ICSNC '06, ınternational conference on systems and networks communications. pp. 71–7
24. Morioka Y, Sota T, Nakagawa M (2000) An anti-car-collision system using DGPS and 5.8 GHz ınter-vehicle communication at an off-sight ıntersection. Technical report of IEICE ITS2000-4. pp. 19–24
25. Maire F, Vision based anti-collision system for rail track maintenance vehicle. 2007 IEEE conference on advanced video and signal based surveillance

26. Cui Y, Ge SS (Feb 2003) Autonomous vehicle positioning with GPS in urban canyon environments. IEEE Trans Robot Autom 19,110(1):15–2s
27. Ammoun S, Nashashibi F, Laurgeau C (Sep 2006) Real-time crash avoidance system on crossroads based on 802.11 devices and GPS receivers. In: Proceedings of the IEEE ITSC 2006, IEEE ıntelligent transportation systems conference Toronto, Canada

Psychoinformatics: The Behavioral Analytics

Sparsh Nimje, Jayesh Katade, Nachiket Dunbray, Shreyas Mavale, Siddhivinayak Kulkarni, and Sally Firmin

Abstract Human behavior is very complex and cannot be explained using traditional mathematical models. Intermediate forms, such as those obtained from personality data, can be used to predict behavioral aspects of a person, creating the hypothesis that arbitrating psychological models can be drawn directly from recordings of behavior. In recent years, smartphone addiction has increased to a great extent. Since the excessive use of smartphones has negatively affected our daily life, many applications to reduce dependence on smartphones have been developed around the world. Personal attributes or personality types can be extracted from data obtained directly from smart phones without the interaction of participants who may have social or health interventions. Many people who excessively use their smartphones have an uncontrollable urge to use the Internet. Internet addiction refers to uncontrolled use of the Internet which causes hindrance in our daily life. Due to its negative impact on the education and lives of people, it is necessary to detect tendencies of people toward addictive behavior and provide them with preventative support and treatment. Similarly, the development of social media has seen rapid growth in its usage. People often find themselves overusing utilities such as virtual communication, texting, and sharing information which have also caused various behavioral problems. This study provides a summary of the various methods and studies done on these behavioral problems and to analyze different techniques, and machine learning models are used to predict addictive personality types.

Keywords Human behavior · Smartphone addiction · Internet addiction social media addiction · Behavioral analytics · Informatics · Big five personality traits

S. Nimje (✉) · J. Katade · N. Dunbray · S. Mavale · S. Kulkarni
School of Computer Engineering and Technology, MIT World Peace University, Pune, India

S. Kulkarni
e-mail: siddhivinayak.kulkarni@mitwpu.edu.in

S. Firmin
School of Information Technology and Physical Sciences, Federation University, Ballarat, Australia
e-mail: S.Firmin@federation.edu.au

V. Bindhu et al. (eds.), *Proceedings of Third International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 844,
https://doi.org/10.1007/978-981-16-8862-1_36

Introduction

Psychology is a field that studies the human brain's activity and analyzes the causes of various changes happening to a human brain. Psychologists and therapists mainly rely upon various traditional methods for analyzing human brains such as research, questionnaires, and surveys; however, these haven't been able to gather a lot of information. In such scenarios, psychoinformatics comes into play. Psychoinformatics is the field which gauges various parameters in human psychology with the help of computer science [1]. With recent advancements in computer science and information technology, psychologists and researchers have been able to track down these changes happening to people in a more efficient way. It allows psychologists and therapists to analyze and correlate different parameters such as cognitive ability, skills, time spent on the smartphone, how many hours they spend on social media, and how it has affected their brains. Since the origin of the Internet, there have been several innovations in the field of computer applications of technologies. Invention of smartphones, social media platforms are some of the prominent innovations. These technologies have a huge impact on the human psyche.

Mobile phones have revolutionized the world in a different way. The world needed long-range communication which led to the invention of mobile phones. It changed the way people used to communicate with each other. Now, the concept has totally been directed to smartphones. A smartphone can do everything possible in its reach. It has emerged as a multitasking application where people can use it as a calculator, listen to music, play videos, click pictures, and many other tasks. But, recent analytical studies have shown a problem with the people (especially teenagers) using smartphones. Out of 2.6 billion smartphone users, over 87% of millennials keep their phones always by their side [2]. Around 66% of the population of the world showed a sign of smartphone addiction. Nomophobia is a psychological condition where a person has a fear of detachment from their smartphones [1]. Various psychological disorders have been observed in nomophobic people such as anxiety issues, social isolation, and depression.

As more teenagers are addicted to their smartphones, surveys show that 1 in every 25 teenagers have an "unrestricted urge" to use the Internet. Since its arrival, the popularity of the Internet has grown exponentially. People use it for several purposes like gathering information, communication, and social interactions. The term Internet addiction arrived from overuse of the Internet and online activities. People found themselves staring at their computer or mobile screens for hours surfing the Internet. The repetitive and long usage of the Internet can cause various behavioral disorders such as depression, anxiety, and problems of impulsivity. Various psychological studies have been carried out which are concentrated on finding out the internal reasons for Internet addictions.

In the past few years, social media has become a buzzword, and its usage has seen serious growth as it has integrated in people's life all over the world. It has provided a new medium of communication which is social networking sites. Social networking sites have created a way for people to communicate with each other, share information with each other, and get in touch with a person in a more efficient and easy manner. Recently, people have started using it more often and for a long period of time. The long usage of social media has impacted people in a negative manner causing various behavioral disorders such as disrupted sleep patterns, low esteem, anxiety, and impaired decision-making skills. Although social media addiction is new to the world, it might become one of the topics of research in the near future.

All these problems can become complex for various behaviorists and psychologists, and thus, providing a data-centric solution can be beneficial. These problems can be resolved by having an analytical approach. Collection of data from the smart devices, analyzing the data, and implementing various data mining models can help in finding out many hidden patterns and correlations between these behavioral disorders. The aim of this work is gathering information about various data mining and analytical models and bringing out quantitative results regarding these modern day behavioral problems.

Motivation

Evolution of technologies brought many changes to the existing world. Long-range communication with the help of smartphones, access of knowledge to anyone through the Internet, various technological innovations are some of the influential technological innovations. But, in the past years, these technologies have resulted in many behavioral problems faced by us. Excessive use of smartphones caused smartphone addiction, people using excessive Internet showed Internet addiction, and the recent uprising of social media is making people addicted to it. These addictions are causing serious problems in people's behavioral personalities such as anxiety, withdrawal, impulsive nature, anger issues, and isolative behavior. The aim of this paper is to draw out conclusive techniques and methodologies which can be used by specialists to counter these problems. Section "The Big Five Personality Traits" contains the description of the big five personality traits. The frequency of behavioral addictions is measured with the help of these traits. Further, Sect. "Literature Review on Behavioral Addiction" contains the literature review of the behavioral addictions discussed in the paper. Descriptions of various techniques which were used to minimize the effects of these addictions are discussed in Sect. "Techniques to Minimize the Addiction" of the paper.

The Big Five Personality Traits

There has been much research in present times where researchers have recognized five important personality traits which determine a persons' behavior, i.e., openness, conscientiousness, extraversion, agreeableness, and neuroticism [3, 4]. These personality traits form the basic structure of the human mind. It was developed to

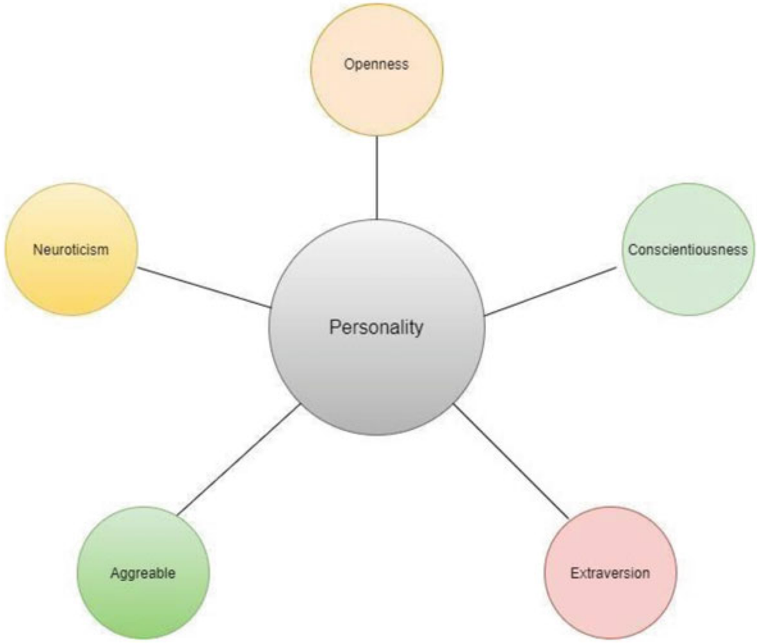


Fig. 1 Big five personality traits

measure the relation between personality and intellect of a person, but today, the five-factor model is primarily in psychological aspects. The influencing factors of smartphone addiction, Internet addiction, and social media addiction on the human psyche are discussed in this paper (Fig. 1).

Openness

People with this trait show signs of high imagination and have great insights about a problem. These types of people have good decision-making skills in comparison with others. Openness describes the depth of a person's mental life. They try to learn new things and experience new adventures. A person who always tries to have newer experiences like ordering a different dish every time they go to a restaurant shows openness behavior.

Conscientiousness

Mostly, goal-oriented people inhibit this personality trait. They have a well-structured and organized way of doing things. They show the ability to delay gratification and try to maintain socially acceptable behavior. An example of conscientious behavior is when a person asks about your whereabouts, takes care of your wellbeing. Most of the conscientious people work in science or finance areas as these require structured and organized behavior as a skill set.

Extraversion

This trait mainly deals with the interaction of a person, whether they are introverted or extroverted. An extroverted person usually gets energized when they are around people, whereas an introvert would try to avoid social gatherings. An introvert rather than being a center of attention likes to maintain solitude in social meetings. An extrovert would perhaps like to lead rather than standing in the crowd and doing nothing.

Agreeable

Agreeable behavior of a person indicates how well they are getting along with others. Agreeable people show signs of empathy and have a helping nature. They exhibit more prosocial behavior, trying to help others, and thus, they usually find their career where they get to help other people. Charity workers and mental health volunteers are the people who are high in agreeableness.

Neuroticism

Neuroticism is a trait which indicates one's emotional stability and temper. Neuroticism is basically the physical and emotional response to one's anxiety and stress. People who tend to see the change in their behavior due to stress are considered neurotics. Neurotics often find themselves overthinking about any situation which leads to more stress and these people may find themselves in depression.

Literature Review on Behavioral Addiction

Various studies and research are working on the problem of smartphone addiction, Internet addiction, and social media addiction. The methods used to solve these problems involved the use of machine learning algorithms. The insights from the results obtained can help psychologists and behaviorists to help people overcome these addictions.

Smartphone Addiction

Smartphones showed very promising signs when they were invented. Long-range communication was just a click away. Apart from the communicative properties, smartphones were diverse in many features such as using it as a calculator, playing music, watching videos, and surfing the Internet. Recent statistics show the negative impact of long usage of smartphones on the human brain. Out of 2.6 billion smartphone users, over 87% of millennials keep their phones always by their side. Around 66% of the population of the world showed signs of smartphone addiction. Research done by Kambham et al. [5] used the social media data to analyze participants' regular app usage and SMS engagements. The use of Pearson and Spearman correlation, in 2018, between relevant features and five personality traits: openness, conscientiousness, extroversion, agreeableness, and neuroticism gave root mean squared errors of 12.7%, 15.4%, 20.4%, 15%, 22.2%, respectively. In research done in a US university [6], data collection was done using an open-sensing framework which was designed to collect phone call data and text message data. The support vector machine (SVM) classifier was used to find the relation between the big five personality traits and features extracted. They achieved an accuracy of 54%, 61%, 51%, 51%, and 49% on five personality models, i.e., neuroticism, extraversion, conscientiousness, agreeableness, and openness, respectively. In a study by Yasudomi et al. [7], the collected data were from various questionnaires which showed the subject's psychological characteristics and their phone's data which demonstrates their behavioral characteristics. The questionnaire data and log data were constructed to form two models. These models predicted whether the subject uses the app restriction function or not. The first model only used questionnaire data and the second model used the user's log data. Accuracy achieved by the first model was 61.2%, and the second model was 67.1%. In 2018, Lee et al. [8] collected data from 125 students on their campus. 125 student's data collected were based on a questionnaire survey, and 64 student's personal device's data were collected. The information gathered from the user's smartphone extracted the following features: usage time, usage time per day, data usage, screen turns, usage time per app, number of executions per app, and frequently used apps. The decision tree model was implemented to determine whether the users have an addiction or not. The accuracy achieved was 89.57% which confirmed a strong relation between smartphone usage and addiction symptoms.

Internet Addiction

The Internet in itself is a big name which was built to help everyone have access to any information around the globe. In the past years, however, some people have shown signs of Internet addiction which is causing many psychological problems. Many researchers and psychologists are working on efficient ways to reduce the effect of Internet addiction. One of the studies done by Peng et al. [9], they collected university student's data and implemented three Internet addiction models. These models were the linear model (LIA), the neural network model (NIA), and the clustering-based model (CIA). The accuracy obtained by them was 71%, 63%, and 71%, respectively. One of the studies done on school students aged between 14 to 19 years old was based on various Internet addiction tests [10]. The positive and negative affect scales were used as data collection tools. The study showed that 19% of people showed Internet addiction. Iran University of medical science [11] conducted a study on university students. They used a stratified random sampling method to collect data from various disciplines. Measuring tools included the study questionnaire, the drug addiction scale, and the general health questions. The data which were collected was analyzed by multiple linear regression using SPSS version 22. The study showed 31% of students had Internet addiction [11].

Social Media Addiction

Surfing and scrolling on social media platforms have become a very popular activity of a persons' daily life. Many people use social media in a non-problematic way; it becomes an issue for people who use it more excessively and compulsively. It is important to find an efficient solution to the problem of social media addiction. In recent years, researchers have used behavioral analytics to overcome this problem. In a study done by Langde and Namrata [12], the collected data were distributed into three groups of different ages. The age groups were, namely: 12–18 age group, 18–24 age group, and 24–30 age group. The overall results showed more behavioral impact on 18–24 age groups. Symptoms analyzed were expectancies, withdrawal, denial, impulsivity, and annuity. We showed the impact of cognitive and emotional problems. In a study [13] done by Savci et al., problematic social media use (PSU) was modeled using artificial neural network (ANN) and SVM. Fifteen different features were evaluated on 309 university students. The estimation using five-fold cross-validation predicted the problems using ANN and SVM, each with an accuracy of 61%. Again, with the help of a forward selection technique, five important features were extracted, and the prediction accuracy involving these five features was 62% using ANN and 63% using SVM. With the advancements in machine learning and analytical models, various approaches have been developed to predict behavioral impacts of social media. Tadesse et al. show the predictions based on personality traits of various Facebook users [14]. They analyzed four machine learning models

Table 1 Literature analysis of various algorithms

Authors and papers	Algorithm used	Characteristics
Kambhan et al. [5]	Pearson and spearman correlation	Regular app usage and SMS engagement
Yves-Alexandre et al. [6]	Support vector machine	Phone call and text message
Yasudomi et al. [7]	Classification algorithm	App restriction function
Peng et al. [9]	Linear, neural, and clustering algorithm	University students' Internet engagements
Mohammad et al. [11]	Questionnaire, drug addiction scale, and general health questions	Stratified random sampling
Savci et al. [13]	Artificial neural network and support vector machine	Forward selection technique
Tadesse et al. [14]	XGBoost	Facebook users' personality traits

and found correlation between various feature sets and various personality traits. The personality prediction model built on XGBoost gave the highest prediction of 74.2%. The social network analysis feature set achieved an accuracy of 78.6% [14]. Long hour usage of social media can cause depression and anxiety. These problems can be prevented if they are detected early [15]. A standard analysis on the dataset obtained from social networking sites is done in [15]. Various algorithms and their combination are used in the prediction of depression such as decision tree method, random forest, support vector machine, Naive Bayes algorithm. An interest-based maximization algorithm is developed to identify the influence of social media on particular users [16]. One of the important discussions in this social media-driven world is to keep track of behavioral characteristics (Table 1).

Techniques to Minimize the Addiction

Analyzing Smartphone Data

The methodology mainly focuses on sensing methods for smartphone data collection such as app usage, SMS engagement. Use of the Wi-Fi system to track the location and represent movement to study only those times when the participants are in a specific area. Rather than a classification model (which is generally used in prediction of behavioral traits), a regression model is being deployed as it gives more accurate results [5, 17]. The data which are being collected is converted into feature variables for further analysis. The features mostly include the device's several sensors including Wi-Fi, Bluetooth, battery percentage data, screen turning (on/off), accelerometer sensor readings, location data, and which apps were used by participants and for how long. Each participant's personality traits [3] were determined by using the extracted features and applying regression analysis.

Usage of App Restriction Function

In the study [7], Yasudomi et al. collected data in two categories. One in the form of a questionnaire which represented the psychological features and the other in the form of log's data which represented behavioral features. A total of 160 student's data were collected through their smartphone. 85 students were asked to install the home app with the app restriction function. The remaining students install the home app without any app restriction function. The former was termed as an intervention group, and the latter was termed as a non-intervention group. App restriction function is commonly available in smartphones nowadays. The home app had provided the functionality for the user to switch between two modes: normal and study mode as shown in Fig. 2. The figure illustrates the transition from normal mode to study mode and vice-versa. The study mode has restrictions while using some apps. This study implicates the effectiveness of the app restriction function in the home app.

Home app transition function. The home app transition is a restriction function which can only be enabled by the user as mentioned [7]. The functionality of this home app transition let the user switch between the normal mode and study mode. By enabling the study mode, users can focus on their work more efficiently, eradicating the unnecessary pop-up notifications. The study mode gives access to limited applications; thus, the home app transition function is one of the efficient and positive ways to reduce smartphone addiction.

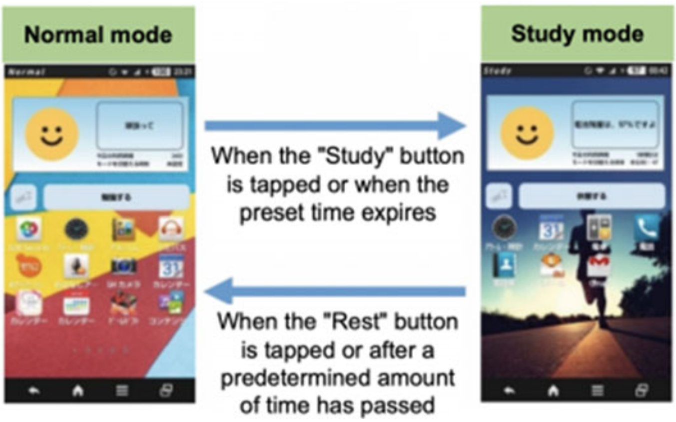


Fig. 2 Home app transition

Machine Learning Algorithms

To study Internet addiction in university students, three machine learning models were used: linear model, neural model, and clustering-based model [9]. This research was done on a small scale among university students in China. The personal attributes of the students were also included such as their college department, age, and gender. The consumption records consisted of a student's profile, time, place, and the amount of one consumption. Students' Internet authentication records were collected from the campus Wi-Fi which indicates the amount of time spent by students in the campus. Student's online time and their average gap between Internet accesses could be extracted. The daily consumption behavior at college restaurants and online behavior can be represented as a vector. The linear model determined the linear relation between students' behavior and features extracted. The neural model is used to find non-linear relations and also to handle the high order relative terms. For the CIA model, the students who spent more time online were thought to be addicted to the Internet. So, a clustering method was devised to find out the normal online time, and the difference between student's online time was proposed. A study by an Iran university was to determine Internet addiction among the students [11]. A stratified random sampling method was used to select the appropriate students for the study. The data collected from the students included questionnaires, the drug addiction scale, and questions related to general health. Pearson product-moment correlation model and multiple linear regression models were analyzed using the SPSS version.

Analysis of overuse of social media is mostly done by many machine learning models and techniques. Many researchers are working on predicting social media addiction among people, more importantly among teenagers and students. The design of the linear regression model using least square method [18] is one of the different and efficient methods implemented by Valakunde and Ravikumar [18]. Figure 3 illustrates the visualized version of the least square method in which the random error is calculated for each data point. They trained the linear regression model with the .csv file and plotted the graph. For the testing of the model, data were collected from the participants, and a graph was plotted using the data points. The methodology of this technique was finding out the distance between these data points and the lines plotted earlier to determine the degree of addiction and the machine to comment on the behavior of the participant.

LIWC, SPLICE, and SNA Technique

User behavior on social networking sites such as Facebook is influenced by the presence and behavior of other users. The interaction between the users can have an impact on the transition of new information or new behaviors through groups. There are many applications for understanding the occurrence and spread of such behavior. In the study [14], all information in the dataset was classified into two groups. The

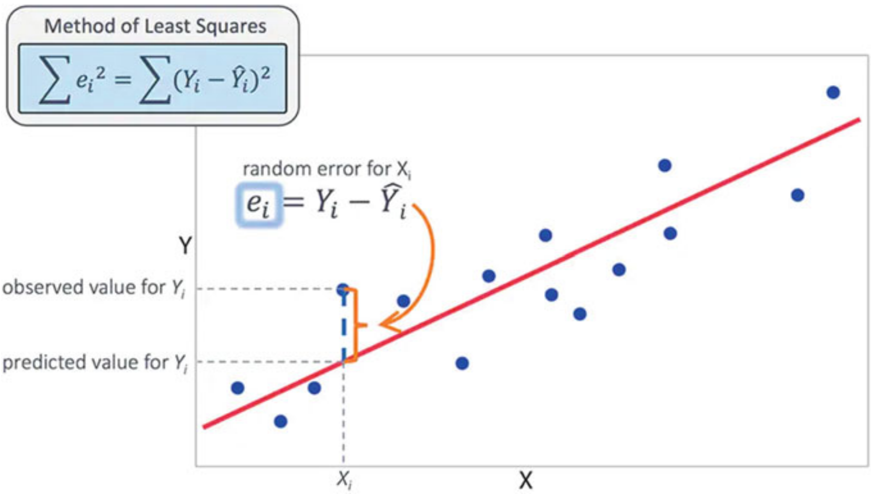


Fig. 3 Linear regression using least square method

first group is to extract text features that reflect the language habits of Facebook users and contain the number of phrases and number of topics. To analyze the status texts on Facebook, they used two dictionaries, namely the linguistic inquiry and word count (LIWC) and structured programming for linguistic cue extraction (SPLICE) [14]. The second group is the analysis of social interaction behavior, including network size, density, broker, and sensitivity. These pieces of information reflect the basic social media behavior of users on Facebook. LIWC is used to extract features which are subcategorized into standard count, psychological process, relativity, personal concerns, and other linguistic dimensions. LIWC2015 is used to analyze multiple language files efficiently. Another technique which is social network analysis (SNA) analyzes the social form which is the combination of user relations of a given population and their interaction with nodes and ties. It is used to quantify and examine the patterns in relationships among interacting individuals. Features related to social network users in association with their personality traits are considered such as their network size, betweenness, density, brokerage, and transitivity. Network size describes the number of connections. Betweenness suggests the wide variety of shortest related paths among people who aren't related directly. Density suggests the extent of interpersonal relations between two individuals. Brokerage is the quantity of ties that an individual gets from the other individual. Transitivity is based on the fact "friends of my friends are also my friends" scenario.

Figure 4 gives a pictorial representation of the steps involved in behavioral analysis. Majorly, there are four important steps involved in the analysis.

Data collection. The right and important data must be collected which will help in increasing the accuracy of the techniques. This data can be collected by conducting surveys, collecting personal questionnaires, mobile's log data, social media data, and Internet usage data.

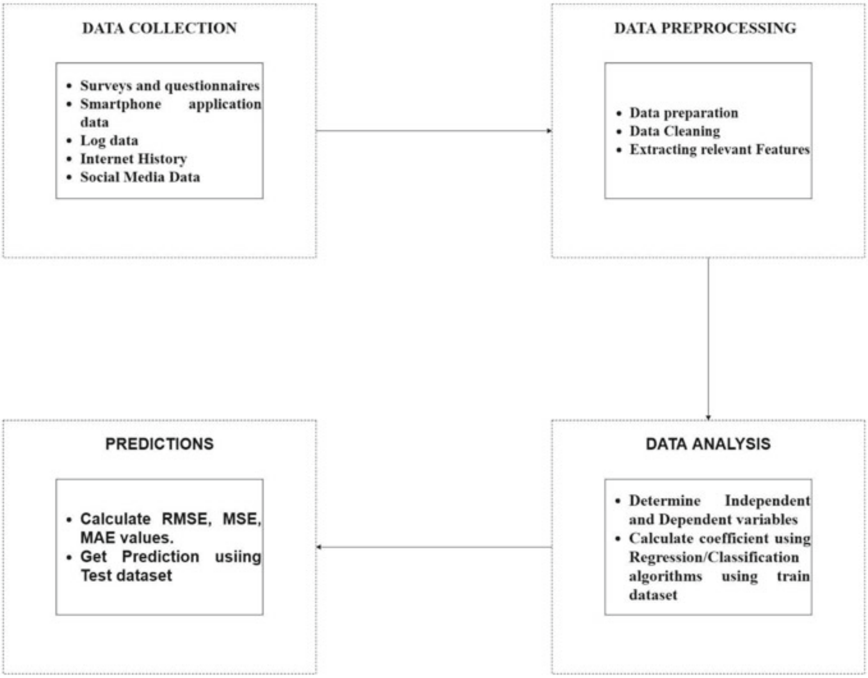


Fig. 4 Architectural flow for behavior analysis

Data preprocessing. The collected data need to be processed beforehand in order to provide clean and valuable data to the machine learning algorithm. There are several methods and techniques which are involved in data preprocessing such as data cleaning, label encoding, and feature scaling.

These steps are followed to provide well labeled and preprocessed data to machine learning algorithms.

Data Analysis. This step involves the use of machine learning algorithms to analzye the collected data. Firstly, the dataset is splitted into two sets: training data and testing data. Then, with the help of training data, classification and regression algorithms are trained.

Predictions. Lastly, with the testing dataset, the values of important variables such as root mean square error (RMSE), mean square error (MSE), mean absolute error (MAE) are calculated, and predictions are made using these values.

Discussion

Many observations can be drawn out while analyzing these addictions related to personality traits. The five-factor model [4] explains the range between two extremes. The effect of behavioral addiction from smartphone, Internet, and social media can be analyzed using the five-factor model.

Users who showed a high range in openness are seen to use high frequency function words. They tend to use longer expressions than usual. Tentatively, they are more open to talk about potentially sensitive objects and highly correlate imagery as they are usually seen as imaginative people. People who are high in openness use the activation words more often. They often correlate with complex composite words, whereas people who are high in agreeable and extraverts use them less likely. Users high in openness show positive correlation with brokerage and betweenness as they have diverse interests. People with higher openness appreciate adventure, art, and creative ideas, whereas people with low openness are conventional and less creative.

Conscientious users correlate negatively with negative emotions and words. They generally avoid the discussion about sad subjects. People who showed high frequency in consciousness tend to talk with other people and often tend to discuss what they see or hear. Conscientious people do not use informal words. Conscientious users highly indicate that they don't get manipulated so easily by others and often maintain their usual plans. Conscientious people believe in hard work and stay organized. Low conscientious people are creative, more tolerant, and less likely to be bound by rules.

Extraverted users often update their status with high frequency words, socially accepted words, and common adjectives. Regarding text length, extroverts prefer to write text messages because they are negatively correlated with word length (-0.08) and sentence length indicating that people who have high scores on extraversion tend to use fewer words and shorter sentences in their status posts. On the other hand, they also use words of more positive emotion, indicating that the more outgoing they are, the more likely they are to talk about their achievements. Extraversion measures a persons' ability to use words of more positive emotion. Extroverted people indicate a more outgoing and energetic nature. They are more likely to talk about personal acquaintances and achievements. Whereas, introverted people tend to be more reserved and show less energy in social life.

Agreeable people often use interrogative sentences and question marks. They don't tend to share information about social activities. They tend to use more action words and have a positive correlation between complex composite words. Agreeableness is a personality trait which is commonly used to measure the focus of a person on having a positive relationship. People who are highly agreeable tend to be more compassionate. They most likely trust other people, while those who are low on agreeableness are more focused on themselves, and they show less compassion.

Neurotic users updated their statuses with more negative words such as anger, anxiety, and affective processes. People who show neuroticism use lengthy sentences, and they show a positive relation with lengthy words stating that people who are neurotics will be linked to more negative emotions such as anxiety and anger. On the scale, people with higher neuroticism are more prone to stress and nervousness, whereas people with lower neuroticism show patience and self-confidence. In comparison to other personality types, neurotics are anxious, insecure, and hostile people. They exhibit negative emotions which are likely to be avoided by other individuals. Table 2 explains the range of five personalities and lists the different properties a person might have if they have a corresponding personality.

Table 2 Big five personality ranges

Personality trait	High frequency	Low frequency
Openness	Imaginative curious experimental	Practical conventional resist change
Conscientiousness	Disciplined organized reliable	Spontaneous flexible negligent
Extraverted	Social enthusiastic adventurous	Introverted reserved self-sufficient
Agreeable	Co-operative trustworthy empathetic	Dominant independent antagonist
Neuroticism	Anxious self-conscious irritable	Composed resilient confident

Conclusion and Future Works

The comparative study of three behavioral addictions correlates with each other. Internet addiction is correlated with people using their smartphones continuously. The use of social media applications is mostly done on smartphones. The behavioral problems such as withdrawal, anxiety, anger issues all are common characteristics of people having these addictions. People who are introverts feel the use of social media as they find it difficult to share their thoughts openly which in turn causes psychological problems such as withdrawal and intolerance. In contrast to this, the people who are extroverts have high correlation with openness traits which makes them cyber-space oriented, and they have an urge to share each and every moment on social media. People started emphasizing virtual relationships as they find it easier to connect to anyone. The problem starts when they become addicted to it. People have started relying upon these commodities and thus have started experiencing symptoms like withdrawal and intolerance. The Internet provides a lot of information where people could surf the Web, watch videos, play games, but the study showed that it lowers the productivity of a person. People start isolating themselves from the world.

The methods discussed in this study can be used in healthcare departments, working environments as in the organization and industries to better understand the psychological problems people are going through. The insights gathered by using behavioral analysis can provide many out of the box solutions. The app restriction function discussed in the paper can be very beneficial as it restricts participants at work using some apps which create distraction. It could be implemented as a daily purpose application which could be used at work to reduce the distraction of the people. People can help themselves by increasing their interactive capability in-person. They can understand the trigger warning which makes us use these applications. The results obtained by using these studies and methods can be utilized by various psychologists and therapists. It can also be used to predict someone's personality on the basis of their social media use by analyzing their personality along with their behavior on social media. Many ill-minded social media groups provoke chaos in the world, and thus, by using these methods, we can predict their behavioral personality. These methods can also come in handy with a person who knows the problem but is not able to solve it. This study can help them understand their problem and find solutions to it.

References

1. Montag C, et al. (2016) Toward psychoinformatics: computer science meets psychology. Computational and mathematical methods in medicine 2016:2983685. https://doi.org/10.1155/2016/2983685
2. Shoukat S (4 Feb 2019) Cell phone addiction and psychological and physiological health in adolescents. EXCLI J 18:47–50. PMID: 30956638; PMCID: PMC6449671
3. Bai S, Hao B, Li A, Yuan S, Gao R, Zhu T (2013) Predicting big five personality traits of microblog users. 2013 IEEE/WIC/ACM international joint conferences on web intelligence (WI) and intelligent agent technologies (IAT). pp 501–508. https://doi.org/10.1109/WI-IAT.2013.70
4. McCrae RR, Costa PT (1987) Validation of the five-factor model of personality across instruments and observers. J Pers Soc Psychol 52(1):81–90. https://doi.org/10.1037/0022-3514.52.1.81
5. Kambham NK, Stanley KG, Bell S (2018) Predicting personality traits using smartphone sensor data and app usage data. 2018 IEEE 9th annual information technology, electronics and mobile communication conference (IEMCON). pp 125–132. https://doi.org/10.1109/IEMCON.2018.8614854
6. de Montjoye Y-A, Quoidbach J, Robic F, Pentland A (2013) Predicting personality using novel mobile phone-based metrics. In: SBP. Springer, pp 48–55
7. Yasudomi K, Hamamura T, Honjo M, Yoneyama A, Uchida M (2021) Usage prediction and effectiveness verification of app restriction function for smartphone addiction. 2020 IEEE international conference on E-health networking, application & services (HEALTHCOM). pp 1–8. https://doi.org/10.1109/HEALTHCOM49281.2021.9398974
8. Lee MS, Han MMC, Pak JG (2018) Analysis of behavioral characteristics of smartphone addiction using data mining. Appl Sci 8:1191. https://doi.org/10.3390/app8071191
9. Peng W, Zhang X, Li X (2019) Intelligent behavior data analysis for internet addiction. Sci Program 2019:1–12. https://doi.org/10.1155/2019/2753152
10. Savci M, Aysan F (2016) The role of attachment styles, peer relations, and affections in predicting internet addiction. Addicta: Turkish J Addictions 3:416–432. https://doi.org/10.15805/addicta.2016.3.0028
11. Langarizadeh M, Naghipour M, Tabatabaei M, Mirzaei A, Vaghar M (2018) Prediction of internet addiction based on information literacy among students of Iran university of medical sciences. Electr Phys 10:6333–6340. https://doi.org/10.19082/6333
12. Landge MB, Mahender CN (2018) Understanding the impact of social media on cognition and the emotional aspects using IA test. 2018 3rd international conference on communication and electronics systems (ICCES). pp 831–834. https://doi.org/10.1109/CESYS.2018.8724019
13. Savci M, Tekin A, Elhai JD (2020) Prediction of problematic social media use (PSU) using machine learning approaches. Curr Psychol https://doi.org/10.1007/s12144-020-00794-1
14. Tadesse M, Lin H, Xu B, Yang L (2018) Personality predictions based on user behavior on the facebook social media platform. IEEE Access. pp 1–1 https://doi.org/10.1109/ACCESS.2018.2876502

15. Smys S, Raj JS (2021) Analysis of deep learning techniques for early detection of depression on social media network-a comparative study. J Trends Comput Sci Smart Technol (TCSST) 3(01):24–39
16. Sivaganesan D (2021) Novel influence maximization algorithm for social network behavior management. J ISMAC 3(01):60–68
17. Antal M, Szabó LZ, Nemes G (2016) Predicting user identity and personality traits from mobile sensor data. International conference on information and software technologies. pp 163–173
18. Valakunde N, Ravikumar S (2019) Prediction of addiction to social media. 2019 IEEE international conference on electrical, computer and communication technologies (ICECCT). pp 1–6. https://doi.org/10.1109/ICECCT.2019.8869399
19. Bachrach Y, Kosinski M, Graepel T, Kohli P, Stillwell D (2012) Personality and patterns of facebook usage. In: Proceedings of the 4th annual ACM web science conference (WebSci '12). Association for computing machinery, New York, NY, USA. pp 24–32. https://doi.org/10.1145/2380718.2380722
20. Stachl C et al. (2020) Predicting personality from patterns of behavior collected with smartphones. Proc Nat Acad Sci U.S.A. 117(30):17680–17687. https://doi.org/10.1073/pnas.1920484117
21. Montag C, Reuter M, Markowetz A (2017) The impact of psychoinformatics on internet addiction including new evidence. https://doi.org/10.1007/978-3-319-46276-9_13

Machine Learning: A Study and Modelling of Different Techniques

M. Rosemary Binoy, S. Spandana, and J. Sangeetha

Abstract In our world today, machine learning has found applications in major domains such as business, entertainment, health and so on. Adequate understanding and knowledge are inevitable in order to bring out the most of these machine learning techniques. In our work, we have studied and applied popular and commonly used machine learning algorithms categorized into regression, classification and clustering techniques to real world datasets namely the air pollution dataset, mushroom dataset and abalone dataset, respectively. Our work tackles existing issues, the techniques used to solve them, the parameters of the chosen techniques to be analysed and the expected outcomes of the same. The performance of the algorithms under each of the aforementioned techniques are compared and evaluated using metrics and visualizations. We then conclude on the best performing models for the chosen datasets.

Keywords Machine learning · Regression techniques · Classification techniques · Clustering techniques

Introduction

The impact of enabling a computer to learn systematically and automatically and enhance from experience with limited or no human arbitration is finding growth in various domains. Major industries like medicine, space, marketing, etc., are being revolutionized by machine learning which falls as a subset of Artificial Intelligence. The primary aim of which is to allow the computer to learn on its own with no human intervention or assistance and enhance from experience accordingly. The learning process begins with input data, followed by an analysis and training of the

M. Rosemary Binoy (✉) · S. Spandana · J. Sangeetha
Department of Computer Science and Engineering, M S Ramaiah Institute of Technology, Bangalore, India

J. Sangeetha
e-mail: sangeethakirank@msrit.edu

V. Bindhu et al. (eds.), *Proceedings of Third International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 844,
https://doi.org/10.1007/978-981-16-8862-1_37

model using the input data. After which the model seeks to uncover some underlying pattern or make a prediction and learn from feedback. The model is then evaluated using suitable metrics such as accuracy, precision, recall, f1 score, etc. Machine learning can only deliver the best performance when the appropriate model is applied to the appropriate data. Supervised and unsupervised learning are the two popular categories of machine learning. Identifying if an email is spam or not is a classification problem that falls under supervised learning because we first teach the model using labelled input and output data. On the other hand, grouping similar news articles is an unsupervised problem that does not learn from pre-labelled data. There are again different algorithms available under supervised and unsupervised learning that can be used to infer knowledge. However proper understanding of the algorithm that one takes to use and the understanding of the data is vital to bring out the best and realistic results. Many factors such as computation power required, accuracy, precision, execution time should be considered depending on the data being handled for the selection of the most suited algorithm.

In the past, researchers have used regression techniques [1] with real world datasets as in Ghazali et al. [2] have used multiple linear regression techniques to predict the Ozone concentration in the atmosphere. Kingsy et al. [3] have performed analysis of the quality of the air using an enhanced version of the K-means [4] clustering algorithm. Sivakumar and Selvaraj [5] have proposed predictive modelling of using supervised classifiers such as (Naive-Bayes) NB, (Decision tree) DT, (Support Vector Machine) SVM, (K-nearest neighbour) KNN and Improved DT on academic performance of the student data with improved decision tree doing the best in accuracy. Naeem et al. [6] have proposed hierarchical clustering algorithms. Verma and Dutta [7] have used different algorithms namely (adaptive network-based fuzzy inference system) ANFIS, (Artificial neural network) ANN and Naive Bayes to work on the mushroom classification problem. Husaini [8] proposes to apply Naive Bayes, Bayes Net and (Ripple-Down Rule learner) RIDOR algorithms on the mushroom dataset along with using ensemble methods to understand the performance of the considered algorithms. Wibowo et al. [9] propose three algorithms C4.5, Naive Bayes and SVM to predict the class of mushroom species [10]. C4.5 and SVM produced better accuracy, however, C4.5 showed faster execution time. Khan et al. [11] proposed clustering techniques namely Expectation Maximization, Farthest Fast and K-means to be applied on the mushroom dataset. It was observed that the farthest fast algorithm performed better in terms of speed and K-means performed well in terms of accuracy.

In this paper, we have reviewed the commonly applied algorithms namely regression, classification and clustering techniques on the real world datasets and compared the results. Every algorithm has advantages as well as drawbacks. Certain algorithms like linear regressions are subject to assumptions and the outcome deviates when the assumptions are not met, in other cases such as artificial neural networks, interpretability is a major issue. In these cases between input and output is a black box that makes it difficult to verify and validate the computations made. Also data is a major ingredient, the model can only be as good as the data. Biases in the data or incomplete data or lack of sufficient data can lead to wrong outcomes even if the model is trained as effectively with the data as possible. Hence, one main observation is that the right selection of the algorithm to be applied to the use case largely depends on your objectives and more importantly on the data. Therefore, adequate considerations of trade-offs of different algorithms available to be applied to a problem and essential understanding and transformation steps of the data have to be followed to obtain the most out of machine learning.

The rest of the paper is organized as follows: In Sect. "Methodology", we have discussed methodology. Results and discussions are explained in Sect. "Results and Discussions" and finally conclusion in Sect. "Conclusion".

Methodology

In this research work, we are comparing three major machine learning techniques namely regression, classification and clustering techniques which are elaborated in the sections below.

Regression Techniques

Regression analysis is a process where the relationship between dependent variables and independent variables (target variable) is estimated. It is a type of supervised learning and mainly used for prediction. Some of the types of regression techniques include univariate linear regression, multivariate linear regression, polynomial regression, logistic regression, ridge regression and lasso regression. The three regression techniques used in our work are univariate linear regression, multivariate linear regression and regularized linear regression, which have been elaborated below and applied on the air quality dataset to understand the relationship between concentrations of the various pollutants. Table 1 shows a simple comparison of the studied regression techniques. The dataset [12] consists of 9358 instances with 14 attributes such as date, time, (Carbon monoxide) CO, PT08.S1 (tin oxide), Non Metallic Hydrocarbons, benzene and NO_2, titania, NOx in ppb, tungsten oxide, indium oxide, temperature, RH relative humidity and AH absolute humidity. The attributes in the dataset contain real values for the concentration of pollutants as well as humidity and temperature values.

Univariate Linear Regression. Univariate linear regression models the linear relationship between the independent and dependent variables. The linear relation is given by Eq. (1):

$$y = mx + c \tag{1}$$

In the above equation, x is the independent variable, y is the dependent variable, m is the weight, and c is the bias. The obtained result is the line that most closely fits

Table 1 Regression techniques—comparison of considered algorithms

	Univariate	Multivariate	Regularized
What is it?	One dependent variable prediction using one independent variable	One dependent variable prediction using two or more independent variables	One dependent variable prediction using constrained coefficients for two or more (generally many) independent variables
Example Use Case	In business to understand the relationship between advertisement spending (x) and revenue (y)	Scientists identify how different amounts of fertilizer and water affect growth of crops	Estimation of economic asset returns from many financial parameters known for having strong correlation
When to use it?	Linear relationship between x and y attributes [visually inspect using scatter plot]. Residuals for all x must be same, independent and normally distributed [visually inspect using residual plots]	Linear relationship between dependent and independent variables [Use scatter plots to check]. No high correlation between independent variables. Residuals must have constant variance, independence and normal distribution for every observation	Large number of independent variables. Number of observations is lower than the number of independent variables. High correlation between independent variables. When one requires a sparse solution, i.e. solution that uses only a few attributes

the data given. Error is defined by the distance between the point and the regression line and the line with the least total prediction errors is the line that best fits.

Multivariate Linear Regression. Multivariate linear regression involves multiple independent variables that help in predicting the values of the dependent variable. Hence, multiple coefficients have to be calculated and result in high computation. Relation is given by Eq. (2):

$$Y_i = b_0 + b_1x_1 + b_2x_2 + \ldots + b_kx_k + \varepsilon_i \tag{2}$$

In the above equation Y_i is the predicted value of ith component of dependent variable y with k dependent variables, and x is the independent variable and b is the coefficient.

Regularized Linear Regression. Regularization methods are extensions of the linear regression method. The main outcome of these methods is the minimum sum that is obtained by adding the squared error of the model. lasso and ridge linear regression are few examples of regularization methods. Minimization of the sum of coefficients (Absolute value) is the target in lasso regression, whilst ridge regression targets the minimization of the squared value of the same.

Classification Techniques

Classification is a type of supervised learning where the objects are classified into one or more classes. The idea here is to predict the target class label. The classification model learns from the input values given for training and predicts the class labels for new data. It is mainly used to predict categorical class labels. Viral pattern detection (Yes/No Pneumonia) from chest *x*-rays for Covid-19 is a recent real use case of classification [13]. Another use case is melanoma diagnosis with help of SVM using dermoscopic images [14]. Out of the many classification algorithms, Support Vector Machines (SVM), Logistic Regression (LR), Decision Tree (DT), Genetic Algorithm (GA), Adaboost algorithm, Artificial Neural Networks (ANN), Naive Bayesian classifier and K-Nearest Neighbour (KNN) have been mainly used in our work, which are elaborated below and applied on the mushroom dataset. Mushroom classification is given the data of a mushroom, identifying if it is poisonous or edible. Various classification techniques have been applied to solve this problem using the mushroom dataset [15]. This dataset consists of details of mushrooms belonging to the Lepiota and Agaricus Family (23 species) and contains 22 attributes. Each species is classified into either edible, or poisonous. In this study, we have not considered the stalk-root attribute, as this attribute had a lot of missing values. The attributes in the dataset are categorical in nature. The class attribute contains the predicted value whether the mushroom is poisonous or edible.

Logistic Regression. This algorithm models the data using the sigmoid/logistic function. The function is given by Eq. (3):

$$g(z) = \frac{1}{(1 + e^{-z})} \tag{3}$$

Usage of this classification technique is when we want to do binary classification like yes/no or 0/1. It usually requires a fairly large number of samples for better results because maximum likelihood estimates do not give the best results on a rather low size dataset.

Decision Tree. In this study, we have used the ID3 algorithm. The metrics considered in ID3 are entropy and information gain. In the ID3 algorithm, entropy of the entire dataset is calculated first. Entropy of every categorical value is calculated for each attribute and averaged, followed by the calculation of gain for the current attribute. Then, selection of the attribute with highest information gain is done and this iteration repeats till the desired tree is obtained. Calculation of entropy is done for every remaining attribute and the algorithm uses the attribute with smallest entropy to split the set. As the entropy gets higher, it helps in constructing a better classification tree [16]. Decision tree deals with nonlinear data effectively. Deciding whether a borrower will defaulter not based on historical information is an example of a decision tree problem.

Artificial Neural Networks. A simple neural network has input, hidden and output layers with each layer having one or more nodes. In our work, we have considered the backpropagation algorithm. In the backpropagation algorithm, we assign random weights, followed by activation rate of hidden nodes and output nodes. This is followed by the calculation of error rate, altering the linkages between the nodes and the error is back propagated to the hidden nodes and finally the weights between the hidden nodes and the input nodes are also modified. This entire process repeats until it meets the criteria [17]. Handwriting recognition, image compression, speech recognition are a few use cases that can be solved using artificial neural networks.

Support Vector Machine. This is an algorithm that classifies two-group problems and works relatively well when there is a clear margin of separation between the classes. The data points are represented in n-dimensional (n-no of features) space. Then, the hyperplane is found that differentiates the two classes. The hyperplane is a n-dimensional boundary that segregates the classes in the n-dimensional space. This algorithm aims at maximizing the margin between the data points and the hyperplane. The optimal hyperplane is the hyperplane with maximum margin. Margin maximization and loss can be balanced by adding the regularized parameter and the parameters of SVM can be tuned to improve the model performance. One use case of SVM is anomaly detection by using normal and anomalous labelled samples to generate a predictive model.

Genetic Algorithm. This algorithm finds the individuals through natural selection that can reproduce and produce next generation offspring. The initial population consists of a set of genes which is given by a string. The fitness score is then assigned by the fitness function. The probability of selecting the population for reproduction is determined by this fitness score. The selection phase has two pairs of populations with good fitness scores selected and their genes are passed to the next generation. In the crossover phase, a crossover point is chosen at random for each pair of parents, and the genes are then exchanged until the crossover point is reached which results in the creation of the offspring and added to the population. The algorithm terminates when the offspring produced is the same as offspring of the previous generation [18].

Adaboost Algorithm. Adaboost or Adaptive Boosting is an ensemble technique. The focus of Adaboost is to obtain a strong classifier from a set of weak ones. The algorithm starts with assigning equal weight to every instance of the data. A weak classifier is now trained twice on this data and also simultaneously calculating the errors. The weights are adjusted during the second time of the training by decreasing the weight value of the rightly categorized instances and increasing the weight value of the wrongly categorized instances. This results in a strong classifier and is used to predict the final results.

Naive Bayesian Classifier. The Naive Bayesian model is essentially a classification algorithm that functions on the principles of Bayes theorem. A very important point to note is the assumption taken by this algorithm is that the features are independent and equal. Mathematically, Bayes theorem is given by Eq. (4) where we aim to calculate the probability of y given X. For our mushroom dataset, we find the class probability in order to decide the classification result. There are two classes for which we calculate using Eq. (4), i.e. y is poisonous or edible.

$$P(y|X) = \frac{P(X|y)}{P(y)P(X)} \tag{4}$$

The classification result is the class that shows the highest class probability for the input data tuple. Text classification is a popular use case of Naive Bayes.

K-nearest Neighbour. The K-nearest neighbours algorithm is also called a lazy learner meaning that it does not do computational work until we ask it to generate a prediction. Basically, it has no training period. Only two parameters have to be worried about in KNN namely the value of K and the distance function. The distance functions include Euclidean distance, Manhattan distance and Minkowski distance. A validation versus error curve can be used to find the most suited K value. Initially, the input data tuple is assigned a class owned by the majority of its k-nearest neighbours that are found by measuring the distance between them using various distance functions. If $K = 1$, then the input data tuple is given the same class as its nearest neighbour. The KNN algorithm then runs multiple times with different values of K and chooses the best value of K that minimizes the number of errors. KNN finds application in recommendation systems for identifying similar products to show the user, especially in the case of e-commerce sites.

Clustering Techniques

Clustering is a method of unsupervised learning that involves grouping of data points. Each data point is assigned into a specific class. In our work, we have used two clustering techniques namely k-means clustering and hierarchical agglomerative clustering that have been elaborated below and applied on the abalone dataset. Abalone is a sea snail whose commercial value increases with age. Grouping the abalone into clusters will help us isolate the abalone of higher age, this will simplify and reduce the time taken by the pre-existing process and be of less labour to man. Hence, we have used the abalone dataset in our work in order to predict the age of the abalone by using the given physical measurements. The dataset [19] has 4177 instances. It contains 9 attributes such as gender, length, diameter, height, whole, shucked, viscera and shell weight and rings (integer). The age of an abalone can be estimated from the number of rings as they are directly proportional.

K-means Clustering. It is an unsupervised learning that clusters the data into K clusters using euclidean distance as a measurement. To discern the best K values techniques such as the elbow method can be used. The algorithm works as follows: Initially, K-means are assigned with random values. Iterating through the items, the item is assigned to the closest mean and the mean is updated. Given the means, each item is then classified to its closest cluster. The means hold the mean values of the items categorized in it. One technique that can be used to find the most favourable K value is the elbow method, which is a plot of within cluster sum of squares vs the number of clusters (K). The X-axis value for which the plot takes a turn can be inferred as the ideal K value. We have used the elbow method in our modelling.

K-means++ method can be used to select centroids with specific probabilities. The initial set of centroids obtained using K-means++ will give better clustering results on account of the probability based specific selection of the initial centroids. Finally, evaluation of the quality of the cluster is done using silhouette scores.

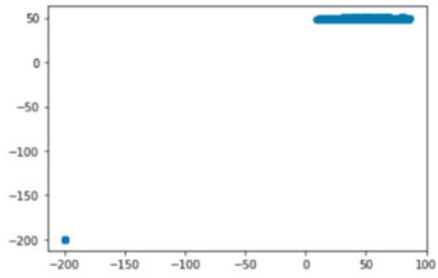
Hierarchical Agglomerative Clustering. This method constructs a hierarchy of clusters. In this technique, the proximity matrix is computed first and each data point is first represented by an individual cluster. Similar clusters are then combined with other clusters and the proximity matrix is updated at each iteration. This repeats until one or K clusters are formed. The main operation is the calculation of the proximity of two clusters. Merging or dividing the clusters requires calculation of similarity between the clusters. The similarity between two clusters can be calculated using certain approaches such as MIN and MAX.

Results and Discussions

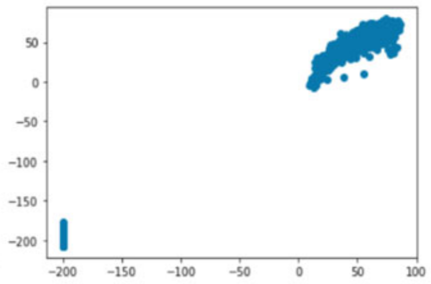
For our study, we have broadly classified major machine learning algorithms into three categories namely regression techniques, classification techniques and clustering techniques. This work aims to bring out the comparison between the algorithms under the aforementioned categories in order to give a comprehensive understanding of the performance and working on the chosen datasets. Our work is organized into three phases that follow, the first one handling regression techniques, second classification techniques and third clustering techniques. In our work, we have considered three datasets that fall into the everyday domain. They have been applied to different machine learning algorithms namely regression, classification and clustering techniques, respectively. Each dataset was first subject to suited basic data preprocessing steps such as handling missing, handling null values, etc., after which every algorithm under every technique was applied to the respective preprocessed data using standard libraries in the Python 3.7 language in jupyter notebook.

Regression Techniques

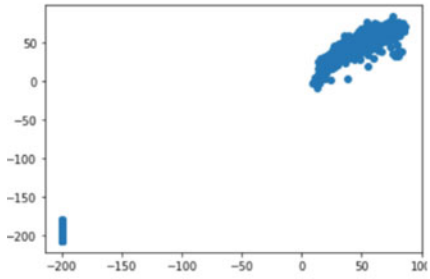
Regression techniques are our first area of study. In our research, we have applied various (machine learning) ML models to air quality dataset. The reason being that the output required to better understand the condition of the atmosphere is numerical and continuous in nature. The regression techniques were applied on this dataset split into 70% train and 30% test. The attributes taken to perform univariate linear regression are the Relative Humidity [RH] and Absolute Humidity [AH]. Figure 1 shows the plot for actual vs predicted values of univariate, multivariate, lasso and ridge regression algorithms, respectively. In regression techniques ideally, we want the predicted value to be as close to the actual value as possible. And hence the ideal plot would show the scatter plot close to the 45 degree angled straight line.



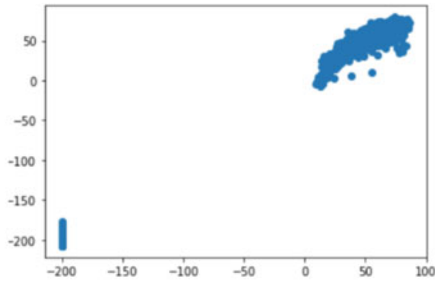
a) Univariate Linear Regression



b) Multivariate Linear Regression



c) Lasso Linear Regression



d) Ridge Linear Regression

Fig. 1 Plot (X-axis: Actual value, Y-axis: Predicted value by algorithm) of actual versus predicted values of test data in regression techniques where a good plot would mean the scatter dots are along the 45-degree line and that the actual and predicted values match correctly. As visible, multivariate, lasso and ridge regression show better plots in comparison with univariate

In this case the plots show that our prediction is more or less close to this line. However, the plots of multivariate, lasso and ridge are better than linear, as the scatter points are closer to the 45 degree line. As given in Table 2, the R2 obtained on applying univariate linear regression is 89.40%, and multivariate linear regression is 97.54%. Higher R2 value shows that the model fits the data more precisely. The table also shows the comparison of the regression algorithms based on max error and mean squared error. Multivariate and ridge regression showed the minimum mean squared error of 66.1643. Regularized linear regression techniques such as lasso and ridge regression have been applied and each produced an R2 of 97.47%

Table 2 Regression techniques—comparison of performance on the air quality dataset

Regression Techniques	R2	Max error	Mean squared error
Univariate linear regression	0.8940	39.1801	285.5892
Multivariate linear regression	0.9754	44.7339	66.1643
Lasso linear regression (Alpha = 0.5)	0.9747	47.6030	67.9459
Ridge linear regression (Alpha = 0.01)	0.9754	44.7339	66.1643

and 97.54%, respectively. However, multivariate and ridge regression give the best outcome in comparison with the others and give higher accuracy with 97.54. The reason for multivariate regression performing well on the dataset is because the linear relationship of the dependent variable is best captured by not a single independent variable but a combination of all independent attributes in the dataset. Lasso does feature selection by setting certain coefficients to zero. However, we notice that when this happens the max error increases, which shows the importance of having all the attributes as in multiple regression. Since ridge regression does not get rid of any attribute but only minimizes their effect, the max error does not increase in case of ridge regression. Hence, we can say that selection of a suitable regression model depends on the relationship of the prediction and predictor variables, especially their correlation, number of dependent variables considered and number of observations in the dataset.

Classification Techniques

In this section, we undertake classification techniques as our area of study. We have used the mushroom dataset in our research in order to classify the mushrooms into edible and poisonous. The dataset was split into 70% train and 30% test and the classification algorithms were applied. Table. 2 shows the accuracies and the entities of the confusion matrices obtained for the respective algorithms taken. One can understand the performance of a model from the confusion matrix by looking at the entities. A good model will have more true positives and negatives in comparison with false positives and negatives. Figure 2 shows the decision tree obtained by

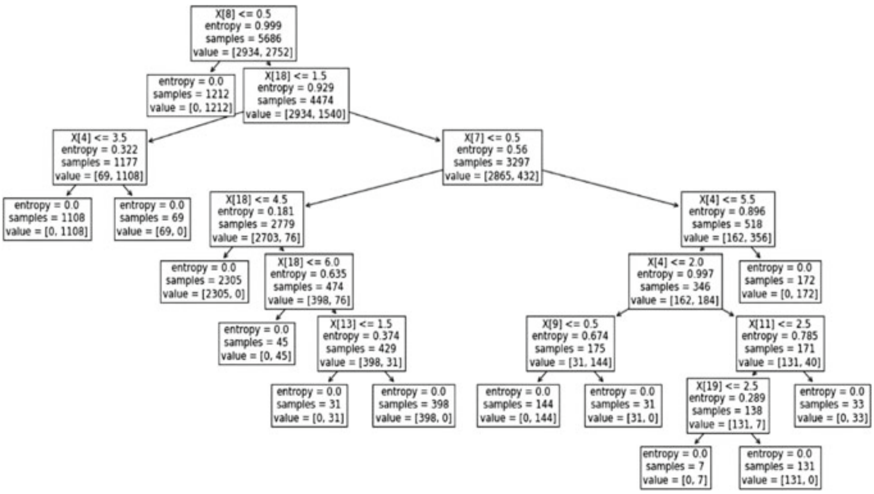


Fig. 2 Visualization of decision tree obtained on training the model with entropy criterion

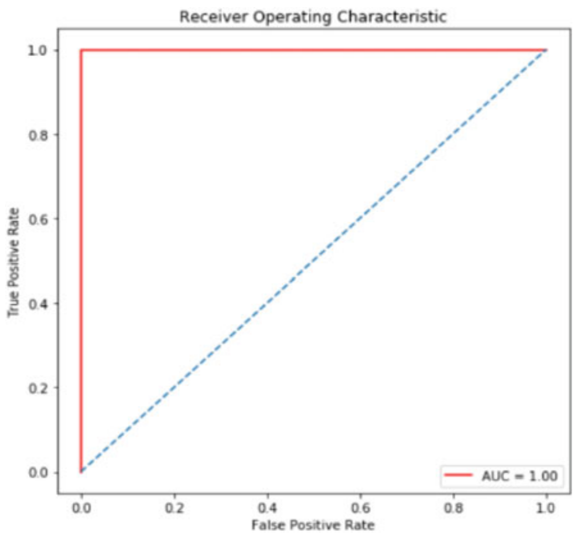


Fig. 3 ROC curve for SVM obtained after genetic algorithm

taking entropy as the criterion. Every node shows the attribute that was selected as the split variable. For example X [9] which is the eighth column gill-colour is our root node or the attribute with the highest entropy. Both gini and entropy criterion were considered, entropy gave the better accuracy with gini giving 99.98%.

The parameters of the SVM algorithm were tuned to improve the performance. It is important to note that the performance of the SVM algorithm depends on the tuning of its parameters as brought out in [20]. The ROC curve of the SVM model is graphed in Fig. 3. It shows that the SVM classifier has captured the information in the data to perfectly distinguish between the two classes poisonous and not poisonous. Fitness in genetic algorithms is a function that produces how good our candidate solution is. The fitness of the genetic algorithm per iteration is shown in Fig. 4. Notice that fitness improves with every iteration. The adaboost model took 100 estimators and the base estimator to be a decision tree classifier. Gaussian, Bernoulli and multinomial distributions were considered for our Naive bayesian classifier. Gaussian gave the better result in comparison with the other two. Our KNN algorithm was implemented with $K = 3, 4, 5, 6$ amongst which 5 nearest neighbours gave 100% accuracy.

For clarity, the reader can read Table 3 to understand the comparative performance of different classification algorithms on the same dataset. It also shows how much of a difference in performance is seen depending on the algorithm for the same problem. Hence, it once again reiterates the fact that choosing the correct algorithm based on the dataset is vital for the best outcome. Table 3 shows that out of the eight algorithms considered decision tree, SVM, adaboost and KNN algorithms are considered to be best suited for this dataset and give very high accuracy with 100. The reason for such high accuracy could be due to high correlation between the attributes and also since the majority of the attributes are categorical and not continuous. This dataset having many attributes and being nonlinear in nature is also a reason to which models like

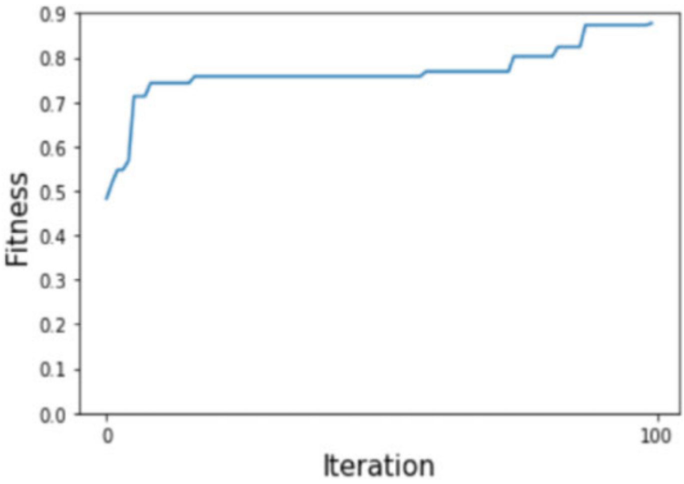


Fig. 4 Fitness versus iteration in training of parameter tuning

Table 3 Classification techniques—comparison of performance on the mushroom dataset

Classification technique	Accuracy	True positive	True negative	False positive	False negative
Logistic regression	94.55	1332	1223	66	60
Decision tree	100	1297	1163	0	0
Artificial neural network	91.7	–	–	–	–
Genetic algorithm	87.69	–	–	–	–
Adaboost algorithm	100	1274	1163	0	0
Support Vector Machine	100	1394	1287	0	0
Naive Bayesian classifier	93	1192	1075	88	82
K-nearest neighbour	100	1274	1163	0	0

SVM and decision tree and adaboost which is a combination of weak learners in this case decision trees have done well as they are known for effectively dealing with nonlinear data. It is advised to the reader that the aforementioned four algorithms are a good fit model for this binary classification problem of mushroom into edible or poisonous and hence should be preferred over the others when handling the same or very similar use case.

Clustering Techniques

We look at clustering techniques that are a form of unsupervised learning under this section. In our research the clustering algorithms mainly, K-means and hierarchical agglomerative clustering have been applied on the abalone dataset. For the K-means algorithm, our entire dataset has been used. A problem with K-means is to find the appropriate K value for the best clustering results. One of the techniques used to handle this issue is the elbow method as used in [21] to find the optimal number of clusters, which we used and found the best K value to be 4. Figure 5 shows the result of the elbow method. It shows the curve becoming parallel to the X-axis as it approaches the value 4. We can consider the elbow to be at this point. Two types of centroid initializations have been used on the dataset namely random initialization and k-means++ initialization.

With $K = 4$ and k-means++ initialization, it took only 0.13 s to converge and had a silhouette score of 0.57 which was better than random initialization that took 0.14 s and gave a 0.54 silhouette score. We can understand from this that by choosing the better initial centroids and not falling into the random initialization trap, (that random choice can result in selection of two centroids that belong to the same cluster) we can reduce the time for convergence especially as the dataset gets larger in size and get better clusters as visible in the silhouette scores. Figure 6 shows the data points distribution. X-axis shows length and Y-axis show visceral weight of the shellfish. Figure 7 shows the centroids and the 4 clusters obtained for the K-means clustering. The exact values can be inferred from the graph's X-axis and Y-axis which represent length and visceral weight, respectively. We can see that the visceral weight of abalone can be put into four groups, the mean of each group being the centroid

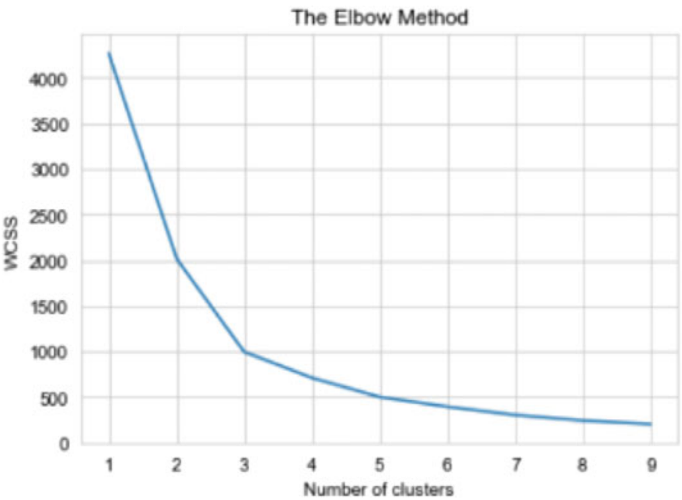


Fig. 5 Results of the elbow method (X-axis: Length, Y-axis: Viscera weight)

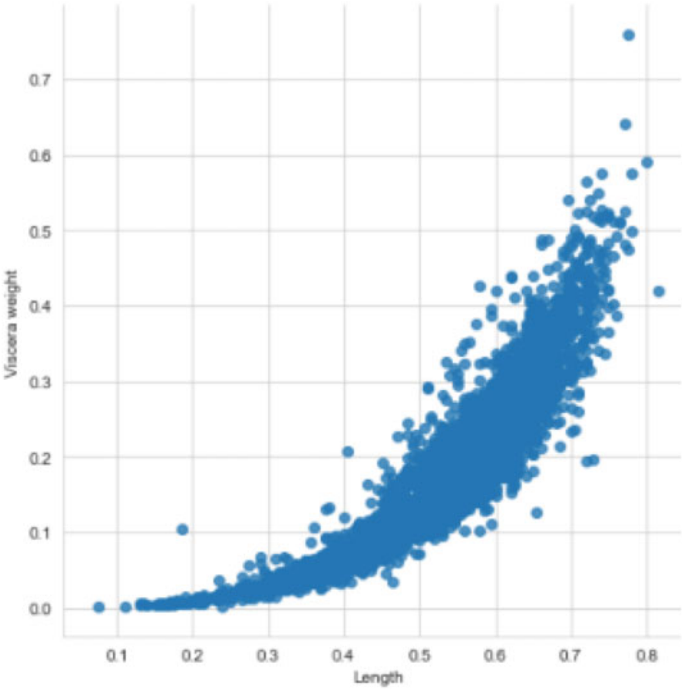


Fig. 6 Distribution of data points

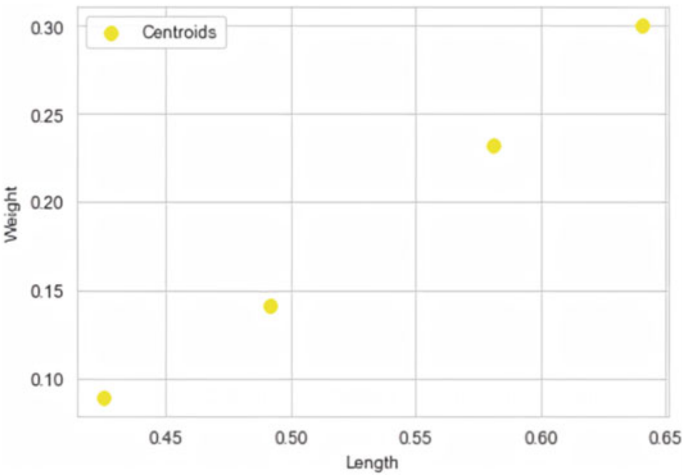


Fig. 7 Selected centroids after clustering

that are 0.0893876, 0.14160037, 0.23218466 and 0.30003503. The other clustering technique used is hierarchical agglomerative clustering on the same dataset taking ward linkage, the time taken to execute was 1.95 s which was greater than *k*-means, and the resultant silhouette score was 0.56. Higher the silhouette score is to 1, better the clustering. Figure 8 shows the generated dendrogram and the appropriate number of clusters was seen to be 3. This figure brings out the hierarchy of the clustering. We can see that we have an even split or equal dissimilarity when we have 3 clusters. Table 4 shows the comparison of both the algorithms with respect to execution time and silhouette score. This table brings out the differences of *K*-means, *K*-means++ and Agglomerative clustering on the factors of execution time and silhouette score. Both the clustering techniques performed more or less with the same results on the basis of quality of clusters formed. However, as the size of the dataset increases, *k*-means might be a better option considering the execution times.

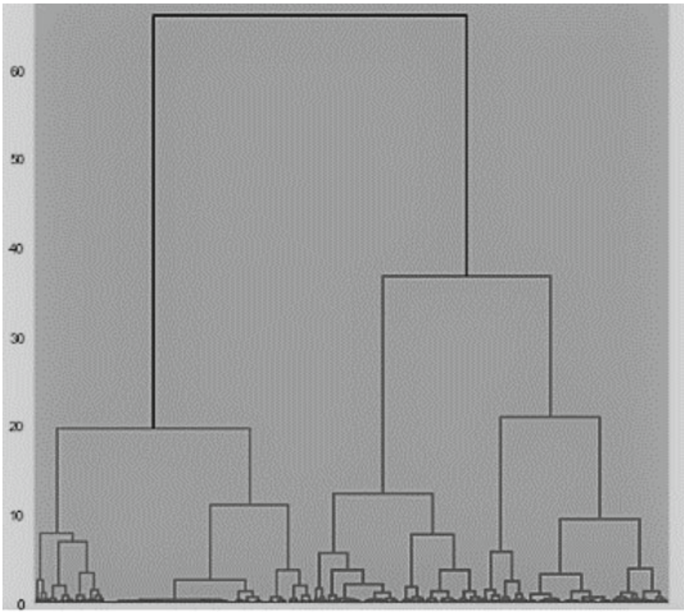


Fig. 8 Dendrogram obtained for agglomerative clustering

Table 4 Comparison between *K*-means and hierarchical clustering

Clustering technique	Silhouette score	Execution time
K-means [random initialization]	0.54	0.14 s
K-means [*K*-means++ initialization]	0.57	0.13 s
Agglomerative clustering	0.56	1.95 s

Conclusion

We have used the air pollution dataset to study regression techniques. Techniques applied included univariate, multivariate, lasso and ridge linear regression techniques. Multivariate and ridge regression performed best for the air pollution dataset with an R2 value of 0.9754 (both) whereas the linear and lasso regression performed moderately well but not as well as the aforementioned.

We used a popular mushroom dataset for our study of classification techniques to classify as edible or poisonous. Eight different classification techniques namely Support Vector Machine, Adaboost, Naive Bayesian Classifier Logistic Regression, Decision Tree, Genetic algorithm, Artificial Neural Network and KNN were used. Out of which four algorithms namely decision tree, Support Vector Machine, adaboost and k-nearest neighbour had the best performance giving 100% accuracy, whereas the other classification techniques did not perform well on the dataset. However, if we had to recommend a classification model for this use case we would suggest SVM as it is known for performing well generally on binary classification.

Another food dataset is Abalone which we have used under clustering techniques. Clustering techniques namely K-Means and Hierarchical clustering technique were applied on the abalone dataset, and the results have shown that K-means gave the best silhouette score (0.57) with a relatively smaller execution time of 0.13 s, whereas the Hierarchical Agglomerative clustering technique took 1.95 s for execution. The most favourable number of clusters for the K-means was found to be 4 for the abalone dataset, whereas it was found to be 3 for hierarchical clustering. As the hierarchical clustering took more time than K-means, K-means is recommended if the size of the dataset is large. Overall multivariate linear regression, nonlinear classifier models like decision tree and k-means are the better optimally performing model under each of the three techniques considered for the chosen datasets. In this work, we have only seen the major classic algorithms under each technique, in the future work, one can focus on the concluded best performing algorithms and their many variants, hybrids and optimized versions, to identify the most ideal match for the chosen use cases.

References

1. Larkin A, Geddes JA, Martin RV, Xiao Q, Liu Y, Marshall JD, Brauer M, Hystad P (2017) Global land use regression model for nitrogen dioxide air pollution. Environ Sci Technol 51:6957–6964
2. Ghazali NA, Ramli NA, Yahaya AS, Yusof NFFM, Sansuddin N, Al Madhoun WA (2010) Transformation of nitrogen dioxide into ozone and prediction of ozone concentrations using multiple linear regression techniques. Environ Monit Assess 165:475–489
3. Kingsy GR, Manimegalai R, Geetha DM, Rajathi S, Usha K, Raabiathul BN (2016) Air pollution analysis using enhanced k-means clustering algorithm for real time sensor data. In: IEEE region 10 conference (TENCON). IEEE, Singapore, pp 1945–1949
4. Bansal A, Sharma M, Goel S (2017) Improved k-mean clustering algorithm for prediction analysis using classification technique in data mining. Int J Comput Appl 157:0975–8887

5. Sivakumar S, Selvaraj R (2018) Predictive modeling of students performance through the enhanced decision tree. In: Kalam A, Das S, Sharma K (eds) Advances in electronics, communication and computing, vol 443. Lecture notes in electrical engineering. Springer, Singapore, pp 21–36
6. Naeem A, Rehman M, Anjum M, Asif M (2019) Development of an efficient hierarchical clustering analysis using an agglomerative clustering algorithm. Curr Sci 117:1045
7. Verma SK, Dutta M (2018) Mushroom classification using ANN and ANFIS algorithm. IOSR J Eng (IOSRJEN) 8:94–100
8. Husaini M (2018) A data mining based ensemble classifier classification approach for edible mushroom identification. Int Res J Eng Technol (IRJET) 5
9. Wibowo A, Rahayu Y, Riyanto A, Hidayatulloh T (2018) Classification algorithm for edible mushroom identification. In: International conference on information and communications technology (ICOIACT). IEEE, Yogyakarta, pp 250–253
10. Sameh A, Moghayer M, Mohanad G, Mohammad A (2019) Classification of mushroom using artificial neural network. Int J Acad Appl Res (IJAAR) 3
11. Khan AAR, Nisha SS, Sathik MM (2018) Clustering techniques for mushroom dataset. Int Res J Eng Technol (IRJET) 5:1121–1125
12. UCI machine learning repository. https://archive.ics.uci.edu/ml/datasets/Air+Quality
13. Chen J-Z (2021) Design of accurate classification of COVID-19 disease in X-ray images using deep learning approach. J ISMAC 3(02):132–148
14. Balasubramaniam V (2021) Artificial intelligence algorithm with SVM classification using dermoscopic images for melanoma diagnosis. J Artif Intell Capsule Netw 3(1):34–42
15. UCI machine learning repository. https://archive.ics.uci.edu/ml/datasets/Mushroom
16. Phyu TN (2009) Survey of classification techniques in data mining. In: Proceedings of the international multiconference of engineers and computer scientists, vol 1. Hong Kong
17. AnalyticsVidhya.https://www.analyticsvidhya.com/blog/2014/10/ann-worksimplified/
18. Towards data science. https://towardsdatascience.com/introduction-to-genetic-algorithms-including-example-code-e396e98d8bf3
19. UCI machine learning repository. https://archive.ics.uci.edu/ml/datasets/abalone
20. Tuba E, Mrkela L, Tuba M (2016) Support vector machine parameter tuning using firefly algorithm. In: 26th International conference radioelektronika (Radioelektronika). IEEE, Kosice, pp 413–418
21. Marutho D, Handaka SH, Wijaya E (2018) The determination of cluster number at k-mean using elbow method and purity evaluation on headline news. In: International seminar on application for technology of information and communication. pp 533–538. IEEE, Semarang

Internet of Things: Security Mechanisms

Syeda Sabah Sultana and J. Sangeetha

Abstract The Internet of Things (IoT) is changing cities industries and homes. Each of these benefits, however, comes with an enormous risk of security issues and privacy loss. One of the problems is to operate the IoT devices safely and securely, because the complexity increases with the increased number of IoT devices. This expanded complexity makes new security challenges such as safety, privacy, and usability far beyond the difficulties that individual face to provide security to an individual device. This paper highlights security aspects, challenges, importance of security at different layers in IoT, and the existing solutions. Also, different approaches proposed to address security and privacy issues concerning security requirements: authentication, confidentiality, and access control are discussed and analyzed explicitly in the realm of the Internet of Things.

Keywords Security · Privacy · Confidentiality · Authentication · Encryption

Introduction

The concept of Internet of Things (IoT) was initially introduced in 1999 [1] and the definition of IoT remains subjective to various viewpoints. According to generation, the IoT is to be the future Internet which unites diverse ranges of technologies. However, a significant number of challenges exists. Among them, security plays an elementary role to ensure to secure IoT against attacks and breakdowns. However, security includes more extensive scope of tasks, services, availability, integrity, data confidentiality, access control, etc. The various IoT devices brings entirely exceptional attacks against various security, privacy elements. At the perception layer of IoT, sensing devices/sensing technologies have extremely restricted calculative

S. S. Sultana (✉) · J. Sangeetha
Department of Computer Science and Engineering, M.S. Ramaiah Institute of Technology, Bengaluru 560054, India

J. Sangeetha
e-mail: sangeethakirank@msrit.edu

V. Bindhu et al. (eds.), *Proceedings of Third International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 844,
https://doi.org/10.1007/978-981-16-8862-1_38

capacity and energy supply, thus making it difficult to provide security. The middle layer, where the network layer exists, IoT depends on networking and communications which provides different types of attacks. At the transport layer, to avoid leakage of personal information and risky actuating tasks, transmission of data and peer authentication are the important aspects. At the application layer, aggregation of data and encryption to moderate the scalability and vulnerability issues are achieved. A framework which consists of system-level security analytics and compatible security is required to build a trustworthy IoT [1]. Today, every infrastructure major concern is security. When talking about home infrastructure, we are mainly concerned with access control, identification, and authentication. Internet of things interconnects every device to the Internet, allowing interaction between the physical and digital world. Due to lack of security, the data leakage and data manipulation are leading to misuse, robbery of assets. Smart home technology is evolved beyond basic comfort functionality such as lights automatically operated and door openers providing numerous benefits, for example: water flow sensors and energy-smart meters effectiveness. IP-enabled cameras and related motion sensors door locks provide greater home safety protection. Nonetheless, smart devices can be exploited by attackers to inflict physical harm, political, and mental damage. Burglars, for example, will aim the connected door lock to create secret access codes, and arsonists can target a smart oven to trigger a victim's fire home [2]. Early intelligent home systems had complicated configuration procedures for computers. Several enterprises have recently implemented new systems which are simpler for users to install, are cloud-based, which provide a platform for programming third party developers to create applications that understand the benefits of smart home.

In requisites of security, IoT is held with indefinite challenges because of following reasons:

- Due of various loopholes which are open to attack, IoT is ordinarily viewed as extension of current Internet to several technologies. The various technologies include mobile broadband, wireless sensor network (WSN).
- Devices in IoT are connected to Internet which makes access to hackers for breach and remote code extension because Internet is always in an unsecured mode.
- Enterprises are equally threatened if IoT leads to trouble/expose to weakness as devices are connected to IoT [3].

The objective of this paper is to emphasize on the issues concerning security and privacy encompassing Internet of Things. This paper aims to shed light on the study of issues and challenges concerning security requirements such as authentication, confidentiality, and access control; importance of security at different layers in IoT and on the constrained devices; and the existing solutions in response to the emerging research of IoT security.

Further the paper is organized as follows. Section 2 discusses security issues and their solutions at several layers of IoT architecture. Section 3 describes the key requirements of security in IoT. Section 4 explains the challenges faced by the constrained devices in IoT, and Sect. 5 provides the review of the articles.

Security at Various Layers of IoT

In this section, security issues and their current solutions are addressed in various layers of IoT systems. Figure 1 show the four layers architecture of IoT. IoT architecture consists of the following:

1. IoT perception layer security
2. IoT Network layer security
3. IoT transport layer security
4. IoT application layer security.

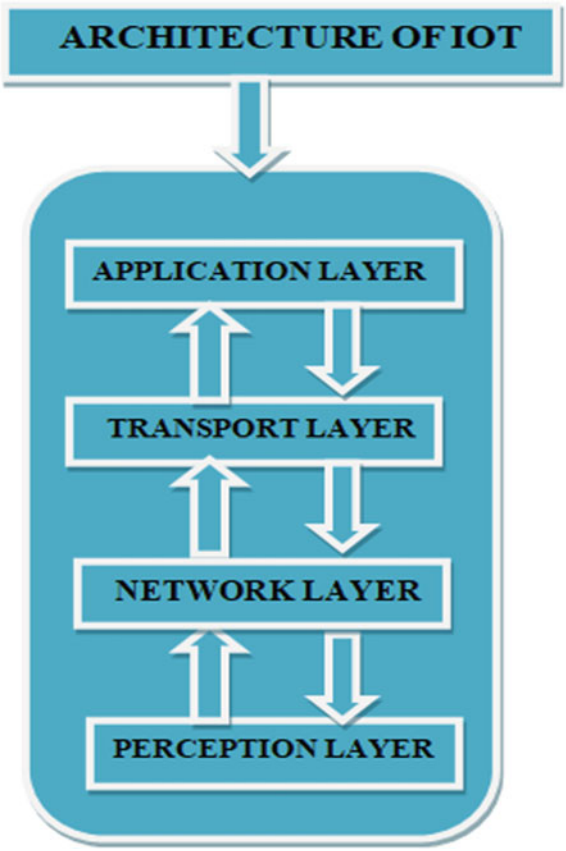


Fig. 1 Four-layer architecture of IoT

IoT Perception Layer Security

Perception layer is implemented as the base layer in IoT architecture. IoT devices and components in this layer interact through devices such as radio frequency identification (RFID), sensors, and actuators [4]. It contains various sorts of gathering and controlling modules. This layer's main responsibility is to collect useful information/data from things or the environment (such as WSN, heterogeneous devices, sensor-type real-world objects, humidity, and temperature) and transform them in a digital setup. The main purpose of objects is unique address identification and communication between short-range technologies such as RFID, Bluetooth, near-field communication (NFC), and 6LoWPAN (low-power personal area).

The perception layer is classified into two parts: the first consists of perception node utilized for data possession and control, perception network that communicates with transportation network in such a way that it sends collected data or sends control instructions to gateway and controller. Detecting of abnormal sensor node is one of the security issues in the perception layer. This occurs when they are attacked by cyber-attacks. Identification of defective nodes is required and takes measures to evade deprivation of service and to guarantee quality of service [5].

In [5], the author proposed a localized algorithm known as fault detection algorithm to detect defective nodes in WSN. Another security problem is the cryptographic algorithms. For node authentication, public key algorithm can be considered conveniently, because of its extensive scalability and the overall network can be secured without complicated key management protocol. There are some privacy concerns to IoT users when providing data to the collector at the server side. It is necessary to remove the data some time before, so that the collector does not follow to the other end.

IoT Network Layer Security

Network layer is implemented in IoT architecture as the middle layer. Also it is called as transmission layer. It processes the information received from the sensor layer and finds routes for data transmission to the devices, IoT hub, and application through integrated networks. Communication of data occurs through interfaces or gateways, applications, other protocols, and communication technologies [4].

This layer's main responsibility is to help and secure data transmission between the application and perception layer of IoT architecture [6]. This layer mainly collects information and delivers to the perception layer toward several applications and servers. This layer also ensures unique addressing and routing abilities to the unified integration of uncountable devices in a single cooperative network. Various types of technologies are contributed for this phenomenon such as wired, wireless, and satellite.

In network layer, for IoT devices to facilitate secure communication with WSN and to provide end-to-end security, it is necessary to enable Internet protocol version 6 (IPv6) over low-power wireless personal area networks (6LoWPAN) to allow Internet protocol security (IPSec). This method is beneficial as the endpoints need not be modified on the Internet. Also, without the need for a trustworthy gateway, the security is implemented [5].

In [7], for the communication of the endpoints, the author proposed an end-to-end communication between sensor network allowed by IP and the conventional network. To facilitate secure integration [8] of IP-enabled WSN's and allocate end-to-end security, a model is proposed by the author for secure connection and mechanism. The model enables sensor nodes and hosts for security by introducing 6LoWPAN security headers. Also, to selectively control energy expansion with security operations on the wireless network, the author [5] provides a survey for IoT on the prerequisites and features used for mobility support. The researcher [9] provided an efficient solution of lightweight mobile IPv6 with IPSec. This solution is suitable for constrained environments and considered for constrained devices. Also, for dynamic ecosystems, it is conscious of the requirements of IoT, thus providing a finest solution for efficiency and security.

IoT Transport Layer Security

The transport layer security ensures mechanisms for secure transmission of data over the Internet. In this layer, end-to-end latency, message integrity, authentication, and confidentiality are the key challenges. The transport layer focuses on end-to-end communication and provides features such as reliability, congestion avoidance, and guaranteeing that packets will be delivered in the same order that they were sent. In [10], a standard two-way authentication architecture based on security for the IoT is introduced. This proposed architecture for IoT provides end-to-end latency, message integrity, authentication, and confidentiality with reasonable energy. The architecture is based on existing Internet standards like data transport layer security (DTLS), and this security scheme is based on RSA (Rivest–Shamir–Adleman) exchanging X.509 certificates. It is performed at the time of fully authenticated DTLS handshake and is designed for 6LoWPAN.

In order to provide security in IoT, constrained application protocol (CoAP) proposes DTLS. The maximum transmission unit is 127 bytes for IEEE 802.15.4 and also the transmission and reception of bit is more costly in low-power wireless sensor networks. However, in [11], 6LoWPAN protocol in IoT devices can be used for compression of long IP layer header of DTLS protocol, reducing significant DTLS header sizes by 64 bits (62% of the record header).

In [12], end-to-end security in homogenous networks between two devices was discussed. It also considers this security such as securing of the constrained network which is vulnerable to attacks (flooding and replay) in the transport layer against the low power and lossy network (LLN) and also focuses on approach to ease such attacks by mapping the protocol transport layer security (TLS) to DTLS protocol between nodes in a LLN to ensure end-to-end security. To avoid leakage of personal information and risky actuating tasks, transmission of data and peer authentication is the important aspects. Although the security solutions exist, but there is demand for lightweight security mechanisms for smart objects, and hence, for peer authentication, the author introduces use of certificates. In [13], to reduce the overheads of the DTLS handshake, the author proposed three designs based on prevalidation, session resumption, and handshake delegation.

IoT Application Layer Security

Application layer is implemented as top layer also known as business layer in IoT architecture. Smart-grid, smart-transportation, and smart-cities each having different requirements exist in this layer. In application layer, communication of received data occurs from network layer which uses this data for granting the necessary operations and services. The services provided are the storage service and analysis service to backup data received in database or for evaluation for predicting the condition of physical devices [4]. This layer's main responsibility is to link the major gap between the users and applications. This IoT layer combines the industry to attain the high-level intelligent applications type solutions such as the disaster monitoring, health monitoring, transposition, fortune, medical and ecological environment, and handled global management relevant to all intelligent type applications.

Most modern IoT devices [5] are running complex software and hence facing security risks. They will be infected by computer virus (Trojan), if IoT devices are connected to Internet. IoT created a new environment, where powerful botnets are created by malware. Mirai, a malware used for creating botnet can gain shell access to secure shell (SSH) accounts and default passwords of Telnet. As the access is obtained: delayed processes, delete files, other malware installation on the system can be created. Another malware family Internet relay Chat (IRCTelnet) was also designed to carry out massive DDoS attacks. These are used to infect insecure IoT devices which are based on Linux and turn these devices into a botnet.

IoT Security Requirements: Authentication, Confidentiality and Access Control

The three key requirements of security in IoT [2] includes authentication, confidentiality, and access control. Invariable transmission and sharing of data is enabled by IoT between things and users. In such a sharing environment, to ensure secure communication, these requirements are important and required.

A. **Authentication and confidentiality**

Authentication ensures that in a network data delivered is valid. Also devices or applications requesting the data are also valid. In the security of IoT devices, authentication is a vital requirement. However, to execute the cryptographic operations, devices have inadequate memory and less CPU power essential for an authentication protocol. To outsource an expensive computations and storage of a fog device, these IoT constrained devices will help in execution of the authentication protocol [14].

To improve IoT application development capabilities, the approach used is of a custom encapsulation mechanism. It establishes secure communication by combining cross platform communication (signature, encryption, and authentication) among different things mentioned [2]. Two-way authentication schemes for IoT is implemented based on RSA and are outlined for IPv6 over 6LoWPANs. The evaluation of this scheme shows that the architecture provides confidentiality, message integrity, authenticity, memory overhead, and end-to-end latency [2]. In [10], DTLS unreliable user datagram protocol (UDP) acquires the security properties of TLS, as DTLS is a variation of TLS. There are some disadvantages of using security protocol DTLS of application layer as compared to network layer security protocol: end-to-end communication security is not provided by the lower layer security protocols. Data on reception is decrypted, and on forwarding, data is re-encrypted in a multi-hop network. Access is obtained through a node in a network to all data text by an attacker. To overcome these issues, a secured connection on each hop to form a mesh network should be established as scalability is an issue for these protocols. In an end-to-end security protocol, some cryptographic overhead on the sender and receiver side is also occurred. Without a secured connection, the nodes are able to forward packets to its destination because of the payload protection in routing algorithms.

Confidentiality ensures that data is available throughout the process to only authorized users, and that unauthorized users cannot eavesdrop or violate. In IoT security, it is a vital concept, because measuring devices (RFID, sensors, etc.) that are large in number can be incorporated into IoT. Thus, ensuring that the data obtained by a measuring tool does not disclose sensitive information to its neighboring devices is important. Methods that are enhanced, such as secure key management systems, should be initialized and utilized to attain confidentiality.

B. **Access control**

Access control means granting permissions to the utilization of resources in an IoT network. Data holders and data collectors are identified as two types where the users and "things" act as data holders. These data holders should provide data regarding a specific target to data collectors.

Identification or authentication of users and "things" as legitimate holders should be done by data collectors. In [2], a large quantity of streaming data needs to be processed and not as the usual database systems, where discrete data is processed. In [14], access control is also used as a security practice to guarantee that access to IoT device and collected data is accessed by an authorized entity. It is required that a given action is performed by trusted parties. The IoT introduces new challenges limited power and bandwidth. Real-time requirement cannot be fulfilled with conventional cloud computing. To overcome these challenges, fog computing is used as an ideal solution and also provides services with elastic resources at a minimal cost. It is in similar domain as of trust. It is usually cryptographically implemented for outsourced data [15], due to the distinct features of cloud computing.

IoT-Constrained Devices' Security Challenges

For existing solutions [16], many challenges in IoT security require new fundamental challenges. Most of them are usually designed from resource-constrained devices. Thus, security solutions provided are unfeasible. Implementation of solutions on a huge number of devices results in difficulties although many devices have sufficient resources. These inconveniences can be fixed with the used of cloud-based services, but unfortunately this result in major delays and require significant bandwidth for many systems and applications. For example: to prolonged battery life, it is a significant challenge for IoT manufacturers and software developers to design a footprint of some memory 64 KB to 640 KB. These developers need to keep the design simple while leaving the space for security software to defend the security threats. Like any other security mechanisms, encryption and decryption are resource-intensive tasks. They require significant processing and storage capacities that IoT devices lack.

Updating of IoT Devices Are Difficult and Constrained

Updating of IoT devices are difficult and demanding as they are circulated massively. Brute force arrangements are the methods which are utilized as occurrence reaction to unravel security issues. In order to restore files and have the system to be rebooting, the systems have to be offline sometimes which will result in significant business loss [16].

IoT devices are distributed geographically; hence, updating them would be demanding and difficult to control which will result in business loss and interference due to the maximal number of actions to be taken such as shutting down, re-installing or rebooting, and replacing the potential system. Such a change is not suitable for IoT systems. Unlike some of the Internet devices that does not always manage to upgrade essentially. For the IoT environments, system that uses well built-in security mechanisms are not suggested. To provide higher security level, fog computing [16] is used on IoT devices.

Firewalls is Infeasible and Impractical for Usage of Security in IoT Devices

IoT devices (sensors, wearable devices, vehicles, drones, etc.) are exposed to physical environment which is not safe. Consequently, it is a feasible task to access IoT devices through wired/wireless local network. Moreover, devices that primarily rely on firewalled castles are difficult to implement security solution as they are focused on perimeter-based protection. Thus, it is necessary to place the system next to firewalls to provide intrusion detection and intrusion prevention. Each of the devices must execute these solutions, and each individual host should be provided by threat protection. However, casting a firewall [16] to every IoT device can be unmanageable, costly, and complex. The result shows that the security approach used is not suitable.

Public Key Infrastructures Are not Suitable for IoT Environments

Public key infrastructures (PKI) [16] come with a number of drawbacks. It involves encryption decryption functionality which contains complex and bandwidth intensive algorithms. For some IoT environments, these functionalities require large capacity of power and bandwidth to perform encryption and decryption of data. This can slow down IoT devices that work on large amount of data. In order to create claims regarding properties of its software, hardware, or runtime environment, a device utilizes its public key certificates. These claims are validated cryptographically by the verifier. However, given that the certification authority is a third party, the organizations using PKI for IoT environments cannot recognize the possibility of cybercriminals issuing false certificates to the environments.

Related Work in IoT Security

In addition to increasing the number of interconnected devices and the diversity of IoT applications, IoT is rising rapidly across diverse industry verticals. Nevertheless, IoT solutions are not yet mature, and several problems remain to be addressed. The most significant among them is security. There are billions of connected devices and sensors and are increasing day by day which needs them to be secure and achieve reliable connectivity. Thus, security architectures that are well-designed are needed by IoT technologies adopted by companies and organizations. The scope of the IoT threat is emergent: any IoT system may be a potential target of attack as the surface of the attack is very large. Some IoT devices situated at untrusted locations, can be accessed physically and even get control of the system by the attackers. Most IoT systems follow fewer amounts of criteria of safety which includes less privileged access or role-based access. For example, many IoT devices communicate over the network without encryption including smart home devices such as webcams, home thermostats, TV's, remote power outlets, door locks, door openers, home alarms, and does not provide an option to user for well-built passwords. IoT devices are resource-constrained and built to consume less power while providing all the features needed at a reasonable cost. As a consequence, protection is an afterthought, sometimes put in the development lifecycle at the base of the priority list. IoT attack vectors will target computers, gateways, SIM/cell, transceivers, wearable, and exploit poor passwords, backdoors, a poor encryption, etc. The large array of IoT specific such as operating system, custom configurations, and firmware versions make it difficult to build general IoT security solutions. Monitoring different IoT operating system is a challenging task. The deployment of security solutions to an exponentially growing number of different IoT devices should also be highly scalable. Increasing numbers of IoT devices pose new security problems.

In [5], the author provides the survey in four aspects on IoT security and privacy issues. The primary aspect explains about the related limitations and their existing solutions. Secondly various different types of attacks existing in IoT are summarized. Several IoT authentication schemes proposed are explained. Access control scheme and architectures available in IoT are well discussed. Finally, analysis of the security issues and mechanisms in four layers of IoT system have been studied.

The increased potential risk in IoT devices comes when injecting sensing and intelligence into every device. It is necessary to function them securely. Increase in the number and connectivity of such devices, managing this collection of devices becomes difficult, which results in new challenges such as safety, privacy, security, and usability. As IoT systems have been used in homes and other areas, it is necessary to provide safety to them. Also, it is necessary to simplify how people interact with these devices. In order to better understand, this paper includes some study of use-case scenarios and safety, security, and privacy measures [17].

There are many key challenges in IoT devices [18], and IoT devices need simple security solutions that mostly function at minimum energy levels and with lesser number of capabilities, thus resulting in hindrance of difficult security solutions. This hindrance is caused due their memory and computational requirements such as cryptographic protocols. One way of securing devices against emulation attacks is the use of environmental based fingerprinting. To authenticate devices, device fingerprinting is a technique that uses unique features extracted from the objects transmitted signals and environment. However, this technique faces three main limitations: First, across all the objects in the system, finger printing mechanism is same; secondly, connection between devices and central sensing node; finally, as objects have various finger printing features, it is assumed that fingerprints remain constant and does not vary over time. Hence, to validate these various objects, transfer learning mechanism is used. This approach is used in the framework that combines knowledge from different fingerprints.

Infinite Gaussian mixture mode (IGMM) is used to realize first phase, that is fingerprints of every object follows multivariate Gaussian distribution. The final stage is preceded with the comparative analysis between the results of IGMM and the expected device cluster shape computed by Bhattacharyya distance. Hence, to avoid false alarms during attack detection, environmental factors have to be considered. The simulation results show authentication performance of up to 8% increase with respect to conventional authentication techniques.

The author [19] proposed an authentication scheme on IoT devices. To support the authentication, all set of devices communicate with a gateway. There is connection between controller and the gateway which can access to the central data. The access can be provided by passing the authentication scheme through gateway and controller. There are three levels where the message flows between: things, gateway, and the controller. The first phase is required to obtain a public key certificate by a gateway through the controller. The second phase starts by things sending an authentication request to the gateway. The last stage is requisition of authentication from IoT device to gateway. Testing is completed by the tool AVISPA. Evaluation of result shows the identity-based authentication scheme is opposed to various attacks.

The researcher [6] explains the controlled switch-based appliances through human speech, and there are human perception problems that come from human speech. Also, the missing of necessary parameters (ambiguity in human speech) is a problem to understand for identifying an object by a computer. To overcome the inherent issues, context information is employed. Through context information, it can provide indication to understand human speech to be used as a control command for home appliances. Thus, the paper concludes with the monitoring of proposed control system by sensors at home.

Three steps are required for composition of this system: the first is the speech detection that requires monitoring of sound to detect the occurrence of speech from all microphones and to record the sound when it detects human speech. Analysis of features such as time and frequency and ratio of silence can detect human speech. The second step is speech extraction and meaning from the first step that is from the recorded sound. Index number is also received in the first step to detect user location for the command executor. The final step is the identification of the requested command known to be command execution from the speech and control home appliances. This paper concludes that the accuracy of speech recognition from human speech is still low. This should be improvised as the future work [6].

In [20], the author addresses challenges involving improvising the availability of smart power sockets and switches for identifying smart sensors and signals for turning the devices. It also includes the ways of reducing the energy consumption and raise the human safety in homes. The system proposed allows access control, energy-harvesting, storage for electronic components, and wireless communication. The objective is to design and implementation of security in such a way that authorized person is only able to change the state of power and to add a locking attribute for the building power supply. Smart power, switches, and Zig-bee communication protocol is being used for effortless control of energy utilization.

Although IoT [21] is one of the promising technologies, but there is also a fact that IoT is not secure from threats and attacks. However, the security mechanisms should not resist the operation of the device. The radio frequency identification (RFID) authentication protocol is designed with a well-built security mechanism.

RFID has many advantages as it can execute many computations over barcode, and also at the same time, it can read more tags. The author discusses about the protocol proposed using error correcting code (ECC), hash functions, and pseudo random number generator (PRNG). A temporary key is assigned at the beginning of authentication session by the server. The tag reader initiates the authentication session by authenticating tag and reader and on the other end the server authenticating tag and reader. PRNG and ECC generate public and private keys at each session. ECC provides more scalability and security as compared to other techniques, thus providing confidentiality and authentication.

In [22], the author explains about the IoT security based on the novel cognitive and systematic approach. The approach explains the responsibility of every component and interactions with the other components of the proposed scheme, and their impact on the entire system is described. Security-related queries have been discussed. The paper also highlights on some standardization activities that are occurring currently, important solutions, and open issues to guarantee privacy of IoT components and its applications.

Table 1 presents the survey on security of IoT studied in the previous articles, which describes the various security issues with respect to the different perspectives on IoT.

The author [23] explains that RFID is vulnerable to attacks on security and privacy. This is because, any request via wireless communication from a reader, the RFID responds to its unique ID. Due to the non-selective response of RFID tags to all reader queries, the items recognized with tags may reveal information that is insightful. Through this, the adversary can attain trace goals rely. Physical attacks, DoS attacks, cheat tags, eaves dropping, and communication flow analysis and other security issues are all faced by RFID systems. Tag reading by attackers can be executed without suitable control solutions. Thus, all various attacks must be opposed by these systems. Security mechanism that is existing has been discussed.

In [24], security protocol plays an important role in securing the network and privacy. For communication of business, protocol Information Systems Security Architecture Professional (ISSAP) is used for building client and service platform. To grant users a transparent manner of IoT smart business, an open platform United Stated Patent and Trademark Office (USPIOT) for business is established. The author explains the security issues that the IoT communication demands and requires ISSAP agreement in a secure protocol.

The author [25] explains about the RFID systems and its operation. It describes the security and privacy to the minimal cost IoT devices. Research work and security mechanisms have been proposed. It includes various generation of efficient algorithms, which includes symmetric encryption, hash functions, etc., which will assist security and privacy. For resources to be held with security, manufacturing of RFID is required along with circuit fabrication. Resistive protocol is required to be developed. Development of secure systems can be done by following research work on RFID security in future.

Numerous elements are contained in an IoT solution that effects the implementation features of security and privacy. This results in functionality issue. Some elements such as open-source and proprietary are also among those that users are unable to control them. However, a smart home device can be controlled by user via application of smart home remotely. This consists of embedded devices linked to the cloud. In [26], to grant digital entity, a lightweight identity stack is proposed for IoT with the devices and users that interact with them. An authentication scheme is used for FIDO (Fast Identity Online), every time a FIDO authentication receives request from the user, a keep-alive protocol is used.

Centralized structures can provision services where information is obtained, executed by central entities, whereas distributed architecture provides information at the edge of the network in a dynamic way. The author mentions various challenges faced by distributed approach to understand the future of Internet of Things. The challenges include management of entities of authentication and authorization. However, there are numerous advantages as well. Distributed entities manage data thus also providing implementation of privacy policies. Thus, the author [27] presents a research study of distributed approach based on features and security challenges of IoT.

Discussion

Security at both the physical devices and service applications is critical to the operation of IoT, which is indispensable for the success of IoT. Open problems remain in a number of areas, such as security and privacy protection, network protocols, standardization, identity management, and trusted architecture. In this paper, challenges of IoT devices and their way of resolving them have been reviewed. Also, the security mechanisms which have been proposed already for IoT are discussed. The security necessities are reviewed including authentication, confidentiality, and access control schemes. Analysis of the security issues in the four layers of IoT architecture is discussed. The security challenges in enabling technologies of IoT also are reviewed. In future research, the security strategies for IoT should be carefully designed by managing the tradeoffs among security, privacy, and utility to provide security in multi-layer architecture of IoT.

Conclusion

Internet of Things has become a major research topic. It focuses on combining various devices and integration of different sensors to communicate with each other. IoT has been increasing rapidly with security concerns and requirements for the large-scale development such as the Smart-e-health, Smart-home automation, middleware, and software-defined network. Security is key issue and also a major challenge on the Internet of Things. In this paper, challenges of IoT devices and their way of resolving them have been reviewed. Also, the security mechanisms which have been proposed already for IoT are discussed. The security necessities are reviewed including authentication, confidentiality, and access control schemes. Analysis of the security issues in the four layers of IoT architecture is discussed. The approach toward this study was to understand the features and challenges of security on the Internet of Things.

Acknowledgement We would like to thank M.S. Ramaiah Institute of Technology, Computer Science and Engineering, for supporting this research work.

References

1. Li S, Tryfonas T, Li H (2016) The internet of things: a security point of view. Internet Res 26(2):337–359
2. Sicari S, Rizzardi A, Grieco LA, Coen-Porisini A (2015) Security, privacy and trust in internet of things: the road ahead. Computer networks, vol 76, pp 146–164, ISSN 1389-1286
3. Balamurugan S, Ayyasamy A, Suresh Joseph K (June 2018) A review on privacy and security challenges in the internet of things (IoT) to protect the device and communication networks. Int J Comput Sci A Inf Sec (IJCSIS) 16(6)
4. Lin J, Yu W, Zhang N, Yang X, Zhang H, Zhao W (2017) A survey on internet of things: architecture, enabling technologies, security and privacy, and applications. IEEE Internet Things J 4(5):1125–1142
5. Yang Y, Wu L, Yin G, Li L, Zhao H (2017) A survey on security and privacy issues in internet of things. IEEE Internet Things J 4(5):1250–1258
6. Han Y, Hyun J, Jeong T, YOO J-H, Hong JW-K (2016) A smarthome control system based on context and human speech. 18th international conference on advanced communication technology (ICACT)
7. Raza S, Duquennoy S, Chung T, Yazar D, Voigt T, Roedig U (2011) Securing communication in 6LowPAN with compressed IPsec. International conference on distributed computing in sensor systems and workshops (DCOSS). Barcelona, pp 1–8
8. Granjal J, Monteiro E, Sa Silva JS (2010) A secure interconnection model for IPv6 enabled wireless sensor networks. IFIP wireless Dyas. Venice, pp 1–6
9. Jara AJ, Fernandez D, Lopez P, Zamora-Izquierdo M, Skarmeta A (2014) Lightweight MIPv6 with IPSec support. Mobile Inf Syst 37–77
10. Kothmavr T, Schmitt C, Hu W, Bruenig M, Carle G (2013) DTLS based security and two-way authentication for the internet of things. Ad Hoc Networks
11. Raza S, Trabalza D, Voigt T (2012) 6LoWPAN compressed DTLS for CoAP. IEEE 8th international conference on distributed computing in sensor systems. Hangzhou, pp 287–289
12. Brachmann M, Keoh SL, Morchon O, Kumar S (2012) End-to-end transport security in the IP-based internet of things. pp 1–5
13. Hummen R, Ziegeldorf J, Shafagh H, Raza S, Wehrle K (2013) Towards viable certificate-based authentication for the internet of things. Proceedings HotWiSec workshop. pp 37–42
14. Alrawais A, Alhothaily A, Hu C, Cheng X (Mar–Apr 2017) Fog computing for the internet of things: security and privacy issues. IEEE Internet Comput 21(2):34–42
15. Yi S, Qin Z, Li Q (2015) Security and privacy issues of fog computing: a survey. pp 685–695
16. Wissam R, Daniele S, Kouichi S (Oct 2017) A new security middle architectture based on fog computing and cloud to support IoT constrained devices. pp 1–8
17. Fu K, et al. (2016) Safety, security, and privacy threats posed by accelerating trends in the internet of things. Technical report computing community consortium
18. Sharaf-Dabbagh Y, Saad W (June 2016) On the authentication of devices in the internet of things. IEEE 17th internattional symposium on a world of wireless, mobile and multimedia networks (WoWMoM). pp 1–3

19. Salman O, Abdullah S, Elhajj IH, Chehab A, Kayssi A (June 2016) Identity-based authentication scheme for the internet of things. In: IEEE symposium on computers and communication (ISCC). pp 1109–1111
20. Zungeru AM, Gaboitaolelwel J, Diarra B, Chuma JM, Ang L-M, Kolobel L (2019) A secured smart home switching system based on wireless communications and self-energy harvesting
21. Tewari, Gupta BB (2018) A robust anonymity preserving authentication protocol for IoT devices. IEEE international conference on consumer electronics (ICCE). Las Vegas, NV, pp 1–5
22. Sfar AR, Natalizio E, Challal Y, Chtourou Z (2017) A roadmap for security challenges in internet of things
23. Feng H, Fu W (2010) Study of recent development about privacy and security of the internet of things. International conference on web information systems and mining. Sanya, pp 91–95
24. Zhao YL (Oct 2013) Research on data security technology in internet of things. Applied mechanics and materials, vol 433–435. Trans Tech Publications, Ltd pp 1752–1755
25. Weis S, Sarma S, Rivest R, Engels D (2004) Security and privacy aspects of low-cost radio frequency identifications systems, vol 2802
26. Chifor B-C, Bica I, Patriciu V, Pop F (2018) A security authorization scheme for smart home internet of things devices. Future Gener Comput Syst 740–749
27. Roman R, Zhou J, Lopez J (2013) On the features and challenges of security and privacy in distributed internet of things. Comput Netwroks 57:2266–2279
28. Balo F (2016) Internet of things: a survey. Int J Appl Math, Electr Comput
29. Gubbi J, Buyya R, Marusic S, Palaniswami M (2012) Internet of things (IoT): a vision, architectural elements, and future directions
30. Shahid N, Aneja S (2017) Internet of things: vision, application areas and research challenges. International conference on I-SSMAC (IoT in social, mobile, analytics and cloud) (I-SMAC). Palladam, pp 583–587
31. Aggarwal C, Ashish N, Sheth A (2013) The internet of things: a survey from the data-centric perspective
32. Said O, Masud M (2013) Towards internet of things: survey anad future vision. Int J Comput Networks 5:1–17
33. Perera C, Zaslavsky A, Christen P, Georgakopoulos D (2014) Context aware computing for the internet of things: a survey. IEEE Commun Surv Tutorials 16(1):414–454
34. Granjal J, Monteiro E, Sa Silva J (2015) Security for the internet of things: a survey of existing protocols and open research issues. IEEE Commun Surv Tutorials 17(3):1294–1312
35. Gil D, Ferrandez A, Mora H, Peeral J (2016) Internet of things: a review of surveys based on context aware intelligent services

Hand Gesture Mapping Using MediaPipe Algorithm

Ravi Kishore Veluri, S. Rama Sree, A. Vanathi, G. Aparna, and S. Prasanth Vaidya

Abstract Hand gestures are a type of nonverbal communication that may be deployed in a variety of situations, including communication between deaf and deaf-mute individuals, robot control, and human–computer interface home automation, and medical applications, among others. A wide range of approaches have been used in hand gesture-based research publications, as well as those based on designed to operate sensing technology and Internet vision. As a result, the hand sign may be divided into several categories, such as position and gesture, but also dynamic and stable, or a combination of the two. When it comes to enhancing human–computer interaction, gesture recognition has emerged as a critical component to consider. In this paper, we are concentrating on developing the system that simplifies the process of writing code in order to use hand gestures which is now performed manually. Angular and Django were used to create the hand gesture mapper, which is a Web application that provides an graphical user interface via which the user can easily assign functionality to a gesture with a few of mouse clicks. Once the user has completed the setup of his or her gestures, he or she can request the Python code that was created for that configuration. This paper provides the ability to pick any one of the workstation GUI tasks and assign motions to it, as well as the reverse of this functionality. In order to link hand gestures and graphical user interface functionality, the hand gesture mapper reduces the need to write code. It does this by

R. K. Veluri (✉) · S. R. Sree · A. Vanathi · G. Aparna · S. P. Vaidya
Aditya Engineering College, Surampalem, India
e-mail: ravikishore1985@aec.edu.in

S. R. Sree
e-mail: ramasree_s@aec.edu.in

A. Vanathi
e-mail: vanathi.andiran@aec.edu.in

G. Aparna
e-mail: aparna.ganni@aec.edu.in

S. P. Vaidya
e-mail: prasanthvaidya.s@aec.edu.in

V. Bindhu et al. (eds.), *Proceedings of Third International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 844,
https://doi.org/10.1007/978-981-16-8862-1_39

automatically creating code for the user movement configurations that are supplied in the application.

Keywords Gestures · Communication · Recognization · Interfacing

Introduction

Hand gestures mapper is designed to reduce the requirement to write code for the purpose of connecting hand gestures with hotkey functions, which is now required. Recent years have seen significant advancements in the technology of hand detection and tracking. It is now possible to detect and follow the hand in a real-time video feed; thanks to advancements in hand gesture recognition technique, which makes use of computer vision. However, in order to make advantage of this technology, some level of coding skill is required, which means that it is still out of reach for some segments of the user population. Hand gesture mapper was developed to fill this gap and help people to get more familiar with his technological innovations. It is possible for users to establish and reconfigure their actions using the HGM Web-interface, without attempting to create any code. This allows users to quickly and simply create and use hand gestures. The major objective of studying hand gestures is to develop a system capable of detecting and utilizing specific hand gestures for the purpose of conveying information or command and control. As such, it encompasses not just the detection of hand movement, as well as its interpretation as meaningful orders. Generally, these methods are used to interpret movements in HCI applications. This paper provides a console (GUI), which allows the user to map hand movements to specific GUI functions that he or she is interested in.

Objective

The new concept entails developing a Web application with Angular and Django and making it available to users. This paper provides a GUI, through which the user may map his or her movements to certain GUI features that he or she finds useful. By only a few mouse clicks, the user may quickly and easily design his or her own gesture setup in minutes. The code will be created automatically by the hand gesture mapper, Web-interface once the user has saved his gesture preferences and clicked on the generate code button.

The following are some of the benefits of using this application:

- The creation and re-configuration of actions has been made simpler and faster.
- The user can design his or her own gestures.
- If the hotkey combinations of the user's choosing are not mentioned in the hotkey functionalities that are given, the user can define it.

- If the transaction name is not accessible among the given application names, the user can specify the name of the application process.

Literature Review

Hasan [1] To recognize hand motions using non-geometric features, a multivariate Gaussian distribution was used. The input hand image is segmented using two methods [2]: skin color-based segmentation using the HSV color model and clustering-based threshold algorithms. Some operations are performed to capture the shape of the hand in order to obtain hand feature; the modified direction analysis algorithm is used to find a connection among statistical parameters [1] from the data, and it is used to compute object (hand) slope and trend [1] by shaping the future of the hand gesture [1], as shown in Fig. 1.

Kulkarni [3] Using a neural network method, distinguish static postures in American sign language. The input image is transformed to an HSV color model, downsized to 80 × 64, and certain image preprocessing operations are performed to separate the hand [3] from a uniform environment [3], and features are retrieved using the histogram technique and the Hough algorithm. Feeding back for gesture classification, three-layer neural networks are used. The system achieved 92.78% recognition rate using MATLAB language with eight samples for each 26 characters in sign language, five samples for training and three samples for testing [3].

Hasan [4] Scaled normalization relying on brightness factor comparison was used for gesture identification. The input image is segmented using a threshold technique with a black background. Any segmented picture is normalized (trimmed), and the image's center mass [4] is computed, so that the coordinates are adjusted to match as centric of the hand object at the *X*- and *Y*-axis origins [4]. Because this method is dependent on the object's center mass, the resulting images are of varying sizes. Reference [4] see Fig. 2, as a result, a scaled normalization operation is used to address

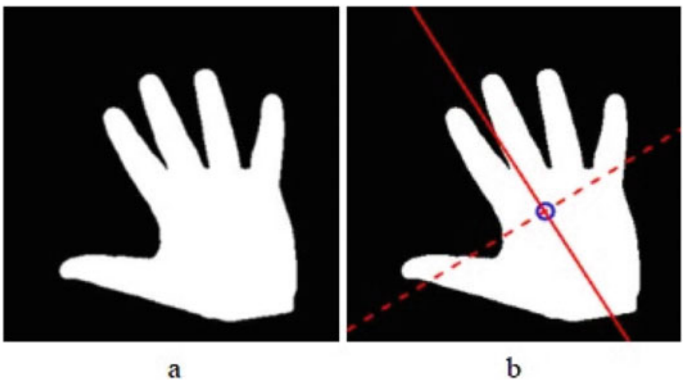


Fig. 1 Computing hand direction

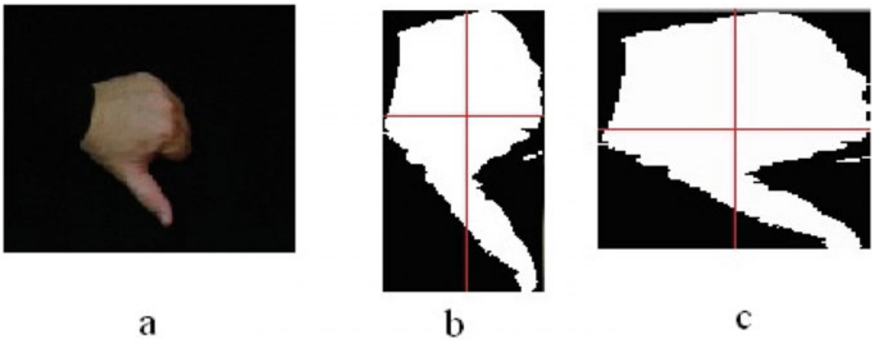


Fig. 2 Application of the input image cutting procedure, followed by the normalization process scaling

this difficulty, which maintains picture dimensions as well as time [4], where each block of the four blocks is scaling with a factor that is distinct from the factors of the other blocks. For feature extraction, two methods are used: first, edge mages, and second, normalized features, in which only the luminance values of pixels are evaluated, and other black pixels are ignored to minimize the size of the feature vector [4]. The database contains six different motions, each with ten examples, five for training and five for testing. The recognition rate for the normalized feature problem outperformed the normal feature technique, with 95% recognition rate for the former and 84%.

Wysoski et al. [5] The boundary histogram was used to illustrate rotation invariant postures. The input image was captured using a camera, a filter for skin color detection was applied, and a clustering procedure was utilized to locate the border for each group in the clustered image using an ordinary contour-tracking algorithm. The image was grid-divided and the boundaries were normalized. The boundary was represented as a chord size chain, which was then employed as a histogram by dividing the image into *N* radial areas based on a specified angle. Neural networks MLP and dynamic programming DP matching were utilized in the classification procedure. Many experiments have been carried out on various feature formats, as well as using different chord size histograms and chord size FFTs. A total of 26 static postures from American sign language were employed in the experiments. The piece used a homogeneous background. For hand gesture recognition, Stergiopoulou [6] proposed a new self-growing and self-organized neural gas (SGONG) network. A color segmentation technique based on skin color filter in the YCbCr color space was used for hand region detection, and an approximation of hand shape morphology was detected using (SGONG) network; three features were extracted using finger identification process, which determines the number of raised fingers and hand shape characteristics, and Gaussian distribution model was used for recognition.

Lv et al. [7] Prosthetic control requires predicting human hand movement intent. Surface electromyography (sEMG) is now widely used in portable exoskeleton cognitive state and human–computer interaction. It is still difficult to extract data from sEMG signals as well as evaluate the willingness of human hand movement. An algorithm due to self-mapping networks (SOM) and radial basis neural networks (RBF) was proposed for selecting features and classification recognition, followed by structural equation modeling (PCA) to reduce feature vector size and finally pattern categorization from image sequences to hand motion. The Euclidean distance between data centers has been used to calculate the variability of the node and find the optimal center and diameter of the radial basis function, in terms of improving the academic achievement of RBF network. The MYO armband sensor sampled real sEMG signal data from six volunteers under eight gestures. A maximum probabilistic model of 100% is achieved by the proposed algorithm, with an ordinary recognition accuracy of 96.8752.7296% and a response time of 0.43 s. We also compared the proposed method's effects on hand motion acknowledgement with RBF, K-nearest neighbor, and multilayer perception with compressed conjugate gradient classifiers. 95.833 3.3244% (RBF to k-means), 94.583 2.243% (KNN), and 88.89 1.1324% (KNN) (MLP with SCG). Compared to other gestures, the technique proposed in this study has the following benefits: (1) This study selects the PCA technique and threshold value technique on the STAE used to identify the active section of sEMG signal, and while maintaining actual improvement, it can correctly detect gestures that are easily distracted, implying that this classification has a good prospect in dynamic stretches.

Gao et al. [8] Wi-Fi signals have been used to detect human activities for decades. One common assumption is that human activities as well as Wi-Fi earned signal patterns are identical. This assumption is not valid when the user performs actions in various locations and orientations. When the user's relative location and orientation to the transceivers changes, the transmitted power patterns of the same action become inconsistent, leading to unpredictable sensing performance. The position-dependent problem prevents the implementation of Wi-Fi-based application areas. To address this fundamental issue, we suggest a new stance sensing strategic plan and then use gesture recognition as an application example. The key idea is to observe from the hand instead of the transceiver and extract features independent of position. So, we build a position-independent feature called Motion Navigation Primitive (MNP). MNP captures the hand's changing direction patterns whenever the person uses the very same gesture to different position-specific factors. The MNP pattern converts gestures into stroke sequences that are easily recognized. We created gesture, a Wi-Fi gesture recognition scheme, to test the proposed strategy. Our system outperforms the state-of-arts in various settings. We believe that the proposed method represents a major step toward the gesture recognition and will inspire future alternatives to position-independent activity recognition.

Thakur et al. [9] Sign language is used by deaf and dumb individuals all over the globe. Communication between a verbally impaired person and a regular person has always been difficult. Sign language recognition is advancement for deaf-mute people. The commercial exploitation of affordable as well as accurate gesture recognition is currently a worldwide concern. For this reason, image processing as well as neural network-based sign language recognition methods is useful over gadget systems. This paper aims to develop a user-friendly and precise sign language system using neural networks to generate text and speech. Also presented is a text to sign new training model that allows two-way communication without a translator.

Vivekanandam [10] In the COVID-19 situation, task monitoring in online team meeting is required. During the lockdown, most classes were taught via online Web applications. The class size is too large for a single tutor to manage. The implementations are also created to display only a few faces in a window. The listener's activity must be verified to enhance the quality of online classes. The paper evaluates artificial intelligence-based deep learning techniques to find a suitable approach for real-time listener monitoring.

Algorithm/Technique

Multivariate Fuzzy Decision Tree (MFDT)

Bien et al. [11] learn and categorize hand gestures by employing the recursive partitioning method (see Resources). In order to separate the data using fuzzy inclusion, the nodes must be increased in the fuzzy decision tree of categorizing hand motions is relatively enormous. Low performance rates are caused by a large number of nodes. Efficient control of a fuzzified decision boundary in the MFDT leads to reduction of intra-person variation. The decision tree learning approach approximates discrete valued signaling pathways and represents learnt functions as a decision tree. Other well decision tree algorithms are ID3 and C4.5. In contrast to fuzzy decision trees, decision trees in MFDT are generated using a multivariate concept, whereas fuzzy decision trees are formed using a multivariate split concept.

Hidden Markov Models

Richarz and Fink [12] state that hidden Markov models is one of the most extensively used tools for time-series analysis since it can show temporal correlations between multiple models and samples, as well as segmentation and classification. A variant variation of this capacity is widely utilized for training and interference hidden Markov models. Yang et al. define that hidden Markov models are considered to be among the most prominent methods for recognition of dynamic hand motion. Hidden factors ensure that multiple states and variants are evident in HMM. HMM output also consists of occult state sequence information, and although each phase is not obvious, potential output can be distributed.

Dynamic Time Warping (DTW) Framework

The [13] DTW framework is a time classification method which tackles the problem of mobility velocity variability by using matching techniques. Using distance matrices, this method looks for the optimal warping among sequences of gestures movement. Different dimensions or referential misalignment can occur in the sequences. DTW uses robust target functions for signal locations. DTW can also be used to increase overall performance with the other algorithms, like HMM.

Latent Regression Forest (LRF)

Chang et al. [14] present LRF from a single image space an algorithm for calculating 3D hand posture is developed. This technique is primarily designed for depth picture searches that require an ordered coarse-to-fine search. Binary decision trees are learned in this framework by continually splitting the input data into subregions, each of which corresponds to a certain skeletal joint. LRF can be thought of as a collection of binary datasets that have been trained on a bootstrap sample. Split nodes, split nodes, and leaf nodes are the three types of nodes that make up the tree. The route is defined by various nodes, the search object is divided into different objects, and the data points are the final finish points which function as skeleton joints.

Most Discriminating Segment Longest Common Subsequence (MDSLCS)

Stern et al. [15] address the LCS approach, DTW's precursor. Because of its constant dimensions data strings, LCS benefits from Euclidean and Manhattan LCS similitude measure which is also immune to noise. Noisy path movement components will not be compared in LCS. The authors propose using the MDSs as a classification algorithm, MDSLCS, based on the LCS approach. The primary idea of the MDSLCS is to identify and infer MDSs automatically, making it a better classification than removing whole gestures. Each gesture is described by MDSs analogous to speaker phonemes or calligraphy strokes. In comparison to 89.5% recognition using HMM, a recognition rate of 92.6% was attained.

Support Vector Machine (SVM)

Glette et al. [16] use technique of SVM gesture recognition SVM is based on the notion of structural risk minimization, which is a benefit that leads to good generalization performance. SVM provides a solution by determining an ideal hyper plane via a nonlinear transformation known as feature space. Optimal hyper planes have the greatest margins; margins are the shortest distance between separate hyper planes and closest data points, which are termed Support vectors.

Implementation

The analytical phase is followed by the system design step. Design is responsible for keeping track of proof design sections and laying out a plan for the development process. The link between systems development and system implementation is design. The paper has fundamental functionality is to make the audience to create and change his gesture configurations, as well as to produce Python code based on the gestures specified by the user to make the gestures easier to use. With help of the options given in the HGM online interface, the user may rapidly set up any gesture settings. When the user has finished configuring the gesture, he can acquire produced code by pressing the "generated code" button on the home screen. The "view created code" page displays the generated code. To use the specified gestures, the user may copy and execute the code shown in Fig. 3.

The basic functionality of the paper lies in helping the viewer to create and modify his gesture configurations and to produce Python code based on the actions of the user to make use of gestures easier shown in Fig. 4. With the help of the options

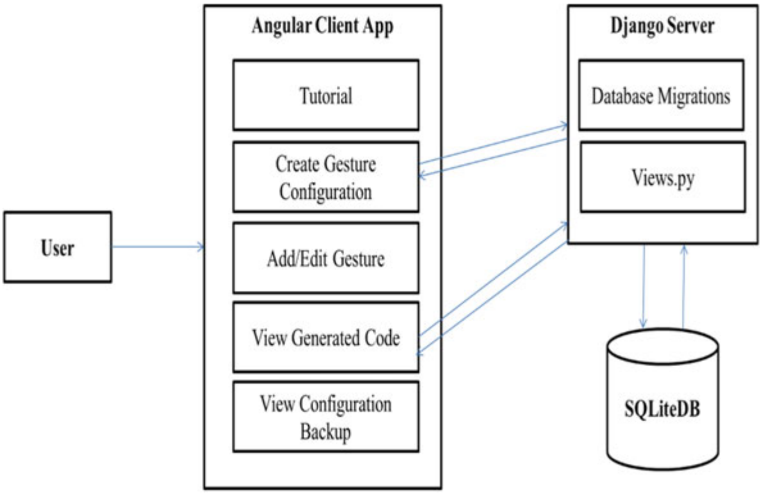


Fig. 3 Architectural diagram

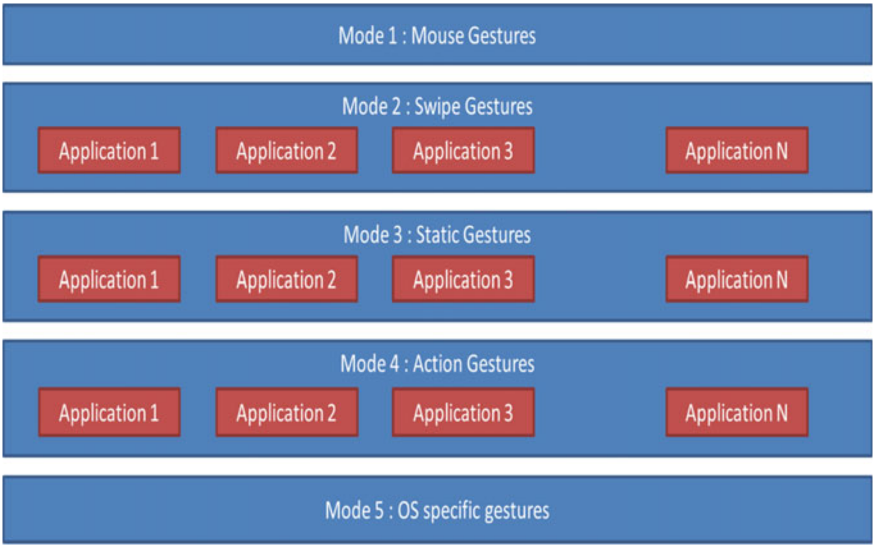


Fig. 4 Architecture of the algorithm

given in the HGM online interface, the user may rapidly arrange the gesture settings. When the user has finished configuring the gestures, he can acquire produced code by clicking the “generated code” button on the home screen. The “view generated code” page displays the generated code. To use the specified gestures, the user may copy and execute the code.

The client-side interface allows the user to quickly configure a set of hand motions by choosing from a menu of possibilities. It also allows users to create a new action, i.e., a hot key sequence, if the action they want is not included among the available alternatives. If the program of the user’s choice is not included in the list of possibilities, the user has the opportunity to define it. In a tabular style, the online interface displays gestures and associated information, during the configuration of gestures.

When the page reloads, the server is able to transmit the list of built-in hand gestures to a client. The server creates and delivers the automatic code for both the specified hand gestures response to the customer after the user completes the gesture setting. Automation code is done via Python, which generates automatic code for manual gestures received on the server, in Python. The automation includes the pre code for MediaPipe connectors to make manual detection and real-time video streaming easier. It also includes OpenCV code for capturing webcam data in real time. The code is already prewritten in the module for mouse tracking.

The user can view this same tutorials already supplied and use the hand motion mapping design relies on the tutorials. The user can also insert a different gesture or modify an already configured gesture. And the user can set gestures. After the user has added his/her requirements, the generated code can be viewed and the generated code can be copied and executed. The user also provides the flexibility to upload and back up the existing configuration. By default, it is in the interest of the user to edit or delete certain gesture configurations.

Users can introduce a new static gesture or even a spin gesture or action gesture when adding a new gesture. It is regarded as a static gesture whenever the swipe direction is established to none. The action to be carried out should be specified by the user as well as a gesture name should be selected for application. The user can make a new scroll lock combination with flexibility. Moreover, the implementation name as well as application process name can be specified.

Five different modes are available: mouse gestures, gestures of swipe, action, static, and operating system specific gestures. Whenever the user is moving his/her hand in a *x*-direction across the screen, there is a swipe gesture. The direction may be a swipe to the right, or swipe to the right. The swipe should never be any when adding a new gesture. If it is not, it was seen as a static gesture; mouse gesture uses gesture recognition to control the mouse cursor, and the actions are based on a hand’s initial and final setup. When adding a new action gesture, the user should enter the first gestural status and the final gesture status once they are specified, and the client can specify charges to be brought and application to be applied and other details. The specific actions of the OS are gestures associated with the operating system. The attribute specifies the hand gesture status to contribute a static gesture, then not swipe direction as well as give both these details such as application name.

Summary Results of the Research

The following tables summarize several systems for recognizing hand gestures. In Table 1, a comparison of hand gesture recognition techniques is made. In Table 2, the potential applications and invariant vectors of several hand gesture recognition processes are summarized.

Comparision of gestures recognization accuracy score with database.

As sign language is utilized during conversation for the interpretation and explanation of a particular topic, particular emphasis has been paid to it. Many systems to recognize gestures using various kinds of sign languages have been suggested. For instance, American Sign Language ASL is detected with border statistics, MLP neural network with dynamic compatibility. Japanese JSL sign language is known to include 42 alphabet and 10 words utilizing the recurrent neural network. Arabic sign language detected utilizing the neural network of two different types, partial and fully-recurrent [4].

Because sign language was used during the discussion for the presentation and analysis of a certain issue, it has received special attention. Many different techniques

Table 1 Comparison among approaches used for the recognition of signals in hand gestures

Method	#Recognize d gestures	#Total gestures used for training and testing	Recognition percentage	Database used
[5]	26	1040	DP 98.8%	American sign language (ASL)
			MLP 98.7%	
[4]	6	60	Normal method 84%	Own database
			Scaling normalization method 95%	
[3]	26	208	92.78%	American sign language (ASL)
[6]	31	130 for testing	90.45%	Own database
[1]	6	60	100% for more than 4 gestures	Own database
[2]	20	200	100% for 14 gestures, and >90 for 15–20 gestures	Own database

Table 2 An invariant matrix of certain hand gesture detection systems and application domains

Method	Application area	Invariant factor
[5]	Sign recognition	Rotation
[4]	Sign language	Rotation/translation/scaling
[5]	Robot control application	Translation/rotation/scaling
[4]	Sign recognition	Rotation/translation/scaling

for detecting gestures using various sign languages have indeed been proposed. For example, border statistics are used to recognize American Sign Language ASL, and MLP neural networks with dynamic compatibility are used to detect American Sign Language ASL. The neural network is used in the Japanese JSL sign language, which has 42 alphabets and 10 words. The Arabic sign language (ASL) was recognized using two types of neural networks: partial and fully-recurrent [12].

Proposed Hand Gestures Algorithm

Hand Landmark Model

Following palm detection over the entire image, our next hand landmark model uses regression to accomplish exact key point localization of 21 3D hand-knuckle positions within the detected hand regions, i.e., direct coordinate prediction. Even with partially visible hands and self-occlusions, the model develops a consistent internal hand posture representation shown in Fig. 5 and Table 3.

We manually tagged 30 K real-world photographs with 21 coordinates to obtain depth data, as seen below. We additionally render a design and develop hand model on various backgrounds and map it to the associated 3D coordinates to better cover the available hand poses as well as provide superannuation on the nature of hand geometry.

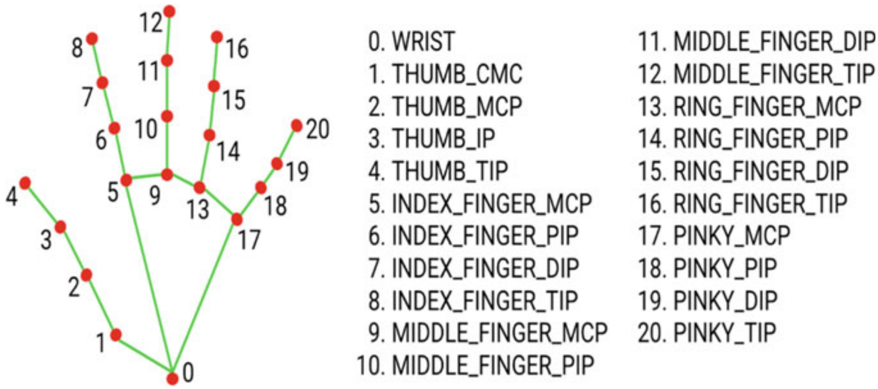


Fig. 5 Hand points capture

Table 3 MediaPipe algorithm and its invariant factor

Method	Application area	Invariant factor
MediaPipe algorithm	Gesture recognition	Mouse, swipe, static (Hand landmark model)

The capacity to recognize the shape and motion of hands can help improve the user experience across a wide variety of technical domains and platforms. It can, for example, serve as the foundation for comprehending sign language and controlling hand gestures, as well as enabling the overlaying of digital video content on top of physical environment in augmented reality. Because hands frequently occlude themselves or one other and lack high contrast patterns, robust real-time hand recognition is a difficult computer vision problem.

Hands by MediaPipe are a high-resolution hand and finger simultaneously managing. Machine learning (ML) is used to deduce 21 3D features of a hand from a single shot. Our solution delivers real-time performance on a cell phone, and even scales to several hands, whereas existing state-of-the-art systems rely mostly on strong desktop environments for inference. We anticipate that by making these hand perception capabilities available to the broader research and development community, new applications, and research pathways will arise, inspiring new applications and study.

Palm Model Detection: We established a special detector model for initial hand detection, comparable to a motion detection prototype in MediaPipe Edge Mesh. Hand detection is a difficult task because our model must detect obscured and self-occluded hands across a wide range of hand sizes (20x). Faces possess high contrast patterns around the eyes and mouth, but hands lack such features, making visual recognition difficult. Instead, adding context such as arm, body, as well as person features helps hand location.

Our approach tackles the issues in a unique way. First, humans train a palm sensor rather than a hand detector because detecting rigid objects such as palms as well as fists is much easier. Given that palms are smaller objects than hands, the non-maximum suppression enhancing effective well for handshakes. By ignoring aspect ratios, square bounding boxes can model palms, decreasing the number of supports by 3–5. In order to extract features from larger scenes, an encoder–decoder feature selection method is used (equal to the RetinaNet approach). Finally, we reduce loose connections during training to promote a large number of anchors. Using the above methods, we achieve a 95.7% accuracy in palm detection. A frequent cross-entropy loss with no decoder gives 86.22%.

Data-driven media pipelines can be built with MediaPipe. Using the MediaPipe framework, we can create stunning media processing pipelines.

Some of MediaPipe's most popular uses are

- Multi-hand tracking
- Face detection
- Object detection and tracking
- Objectron: 3D Object detection and tracking
- AutoFlip: Automatic video cropping pipeline etc.
- Hand landmark Model
- Hand tracking landmark model.

A separate palm detection model is used by the MediaPipe, which then performs 21 3D palm coordinates precise key point localization.

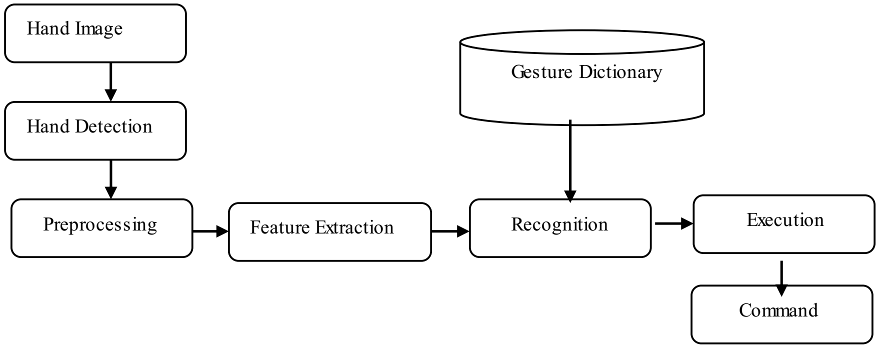


Fig. 6 Flowchart for recognization using MediaPipe

For example, a palm detection model comes back an oriented hand frame from the full image. The palm detector feeds the cropped image to a hand landmark model that needs to return elevated 3D hand key points.

Process Flowchart for Recognization Using MediaPipe

Figure 6 shows that the process of flowchart for recognization using MediaPipe.

Results

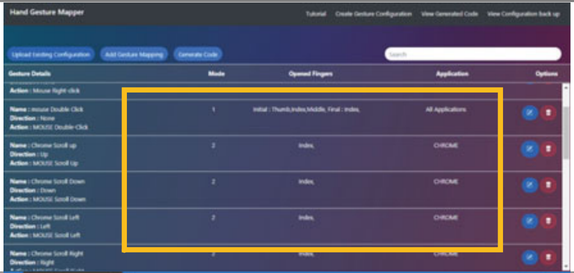
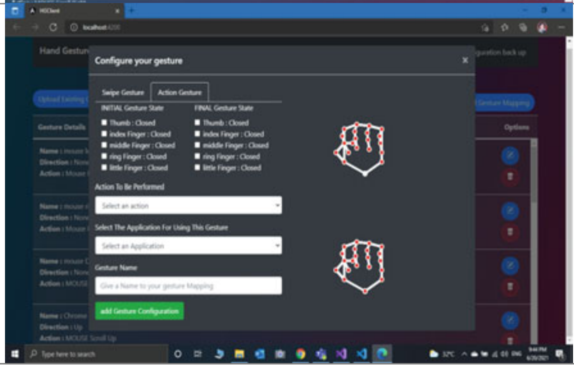
Table 4 shows that the result screen and its description.

Conclusion

Over the previous few years, hand motion detection technology has advanced dramatically. Though it has evolved, this technology is still only available to those who have a good understanding of programming and hand gesture recognition technologies. As a result, the great majority of PC users cannot benefit from this technology.

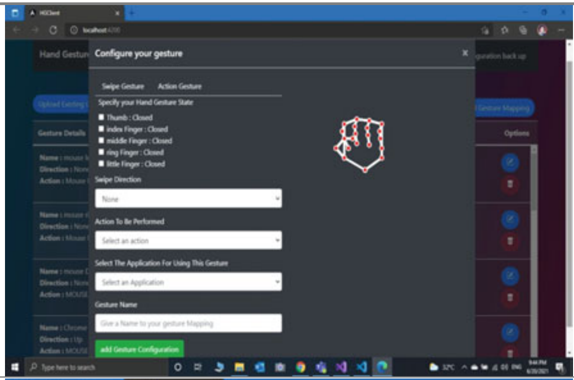
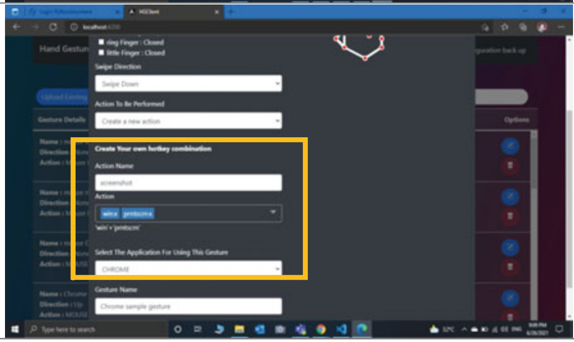
Hand gesture mapper is a paper that serves as a bridge between typical windows users with hand recognition technologies. It lets users to customize the hand gestures without needing to type any coding; instead, they may construct a gesture configuration by picking a few parameters from a Web-based interface. The server then generates code based on the user's gesture setup, which can be run in any Python IDE. When the generated code is run, the user will be able to use the hand gestures that have been configured. As a result, this initiative acts as a link between the terminal user communities with hand gesture technology by making it accessible to everyone.

Table 4 Result screen and its description

S. no.	Result screens	Principle
1		Gesture configurations
2		Adding an action gesture

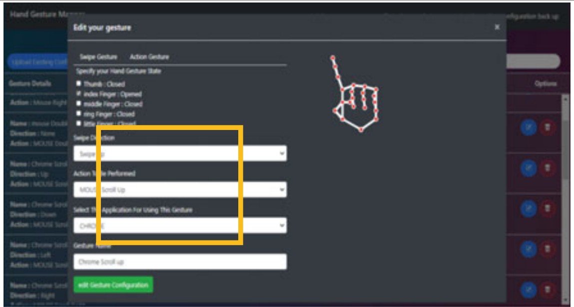
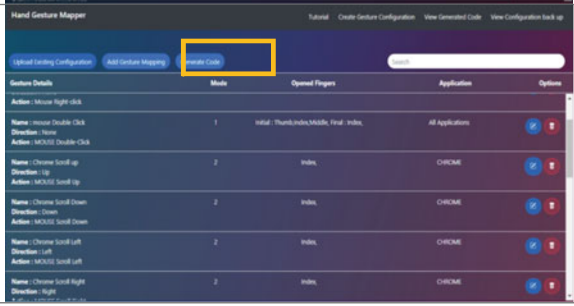
(continued)

Table 4 (continued)

S. no.	Result screens	Principle
3		Adding a swipe gesture
4		Adding new shortcut

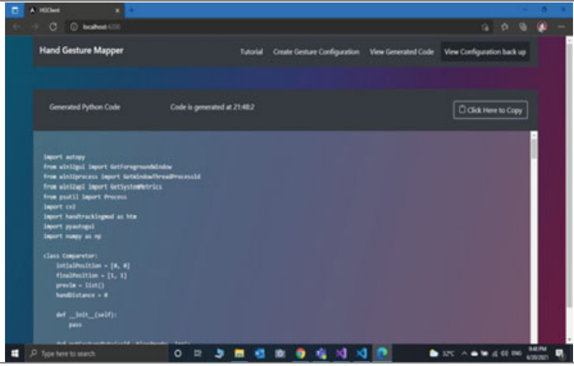
(continued)

Table 4 (continued)

S. no.	Result screens	Principle
5		Editing a gesture configuration
6		Code generation

(continued)

Table 4 (continued)

S. no.	Result screens	Principle
7		View configuration backup

References

1. Hasan MM, Mishra PK (2012) Features fitting using multivariate gaussian distribution for hand gesture recognition. Int J Comput Sci Emer Technol IJCSET 3(2)
2. Hasan MM, Mishra PK (2012) Robust gesture recognition using gaussian distribution for features fitting. Int J Mach Learn Comput 2(3)
3. Kulkarni VS, Lokhande SD (2010) Appearance based recognition of American sign language using gesture segmentation. Int J Comput Sci Eng (IJCSE) 2(3):560–565
4. Hasan MM, Misra PK (2011) Brightness factor matching for gesture recognition system using scaled normalization. Int J Comput Sci Inf Technol (IJCSIT) 3(2)
5. Wysoski SG, Lamar MV, Kuroyanagi S, Iwata A (2002) A rotation invariant approach on static-gesture recognition using boundary histograms and neural networks. IEEE Proceedings of the 9th international conference on neural information processing. Singapura
6. Stergiopoulou E, Papamarkos N (2009) Hand gesture recognition using a neural network shape fitting technique. Elsevier Eng Appl Artif Intell 22(8):1141–1158. https://doi.org/10.1016/j.engappai.2009.03.008
7. Lv Z, Xiao F, Wu Z, Liu Z, Wang Y (2021) Hand gestures recognition from surface electromyogram signal based on self-organizing mapping and radial basis function network. Biomed Signal Proc Control
8. Gao R, Zhang M, Zhang J, Wang L, Zhang D (2021) Towards position-independent sensing for gesture recognition with Wi-Fi. Proc ACM on Interact, Mobile, Wearable UbiquitTechnol 5(2): 3463504
9. Thakur A, Budhathoki P, Upreti S, Shrestha S, Shakya S (2020) Real time sign language recognition and speech generation. J Innovative Image Proc 2(2):65–76
10. Vivekanandam B (2020) Evaluation of activity monitoring algorithm based on smart approaches. J Electron 2(03):175–181
11. Jeon M-J, Lee SW, Bien Z (July 2011) Hand gesture recognition using multivariate fuzzy decision tree and user adaptation. J Korea Rob Soc
12. Richarz J, Fink GA (2011) Visual recognition of 3D emblematic gestures in an HMM framework. J Ambient Intell Smart Environ 3(3):193–211
13. Bautista MÁ, Hernández-Vela A, Ponce V, Perez-Sala X, Baro X, Pujol O, Angulo C, Escalera S, Probability-based dynamic time warping for gesture recognition on RGB-D data, international workshop on depth image analysis and applications
14. Tang D, Jin Chang H, Tejani A, Latent regression forest: structured estimation of 3D articulated hand posture 2014 IEEE conference on computer vision and pattern recognition (CVPR)
15. Stern H, Shmueli M, Berman S, Most discriminating segment—longest common subsequence (MDSLCS) algorithm for dynamic hand gesture classification November 2013 pattern recognition letters
16. Levy E, David OE, Netanyahu NS, Genetic algorithms and deep learning for automatic painter classification. Proceedings of the 2014 annual conference on genetic and evolutionary computation

An Energy-Competent Enhanced Memetic Artificial Bee Colony-Based Optimization in WSN

S. Sowndeswari and E. Kavitha

Abstract Wireless sensor networks (WSNs) are a significant technology for the twenty-first century because of its wide range of applications in various fields. Energy consumption and network security are the major challenges among other challenges in WSN because of existence of various hard problems in wireless sensor networks. Those hard problems cause the reduction of energy in each node of the network and also cause security threat which in turn decreases the packet delivery ratio and lifetime of the entire network. Some of the hard problems include routing, clustering, localization of the nodes, etc. These hard problems cannot be best solved using deterministic methods. Optimization methods are the best alternate to deterministic methods to address the hard problems in WSN. Mostly, the research involves multiple objectives which can be achieved by metaheuristic algorithms. Population-based metaheuristic algorithm is preferred than single solution-based metaheuristic algorithm because of its wide exploration to find the new good solution. In this research work, an energy-competent clustering and secure routing algorithm is proposed using artificial bee colony (ABC) metaheuristics with memetic technique which achieves the desired performance, and the results can be simulated using NS2/MATLAB.

Keywords Wireless sensor networks · Energy consumption · Hard problems · Optimization · Metaheuristics · Memetic · Secure routing · ABC algorithm

Introduction

A wireless sensor network (WSN) is an infrastructure less network does not possess static topology because of random movement of sensor nodes. Unpredictability in topology is caused by a simple change in node position. These sensor nodes can

S. Sowndeswari
Department of ECE, Sambhram Institute of Technology, Sir M VIT, Bengaluru, India

E. Kavitha (✉)
Department of Electronics and Telecommunication Engineering, Sir M Visveswaraya Institute of Technology, Bengaluru, India

V. Bindhu et al. (eds.), *Proceedings of Third International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 844,
https://doi.org/10.1007/978-981-16-8862-1_40

detect, process, and transfer data to the next node until the sink node is reached. The WSN sensor nodes are considerably cheap compared to MANET networks and require much less maintenance once they are deployed as a network in a big geographic area. In case of failure of single node, the entire network will not be functional.

Maintaining energy consumption by the nodes is a biggest challenge and ongoing global research problem in wireless sensor nodes due to limitations in memory, computation, processing, and power [1]. As wireless sensor network is used for wide range of applications mainly for continuous monitoring, controlling the energy usage is a herculean task. Different techniques involved in controlling energy consumption include keeping the radio transceiver in sleep mode in the absence of data, using structure-based hierarchical routing protocol because of good scalability, proper routing, efficient communication etc., Hierarchical routing protocols use a two-layer approach, with one layer used for sensing and the other for efficient routing. Because of its high scalability and efficient communication with low energy consumption, the cluster-based hierarchical routing protocol is preferred over the grid-based approach [2] (Fig. 1).

Wireless sensor network consists of several nodes ranging from few to several millions of nodes, and clustering of such millions of nodes and identifying the best path between the nodes or route is a toughest task. Clustering and identifying best available path for data transfer are the major requirements in wireless sensor network to decrease the energy consumption and increase the packet delivery ratio. Thus, the overall network lifetime is improved, and suspicious nodes can be avoided in data transfer. Compared to conventional clustering algorithms, an improved clustering based on population-based optimization technique is required to achieve the enhanced performance of the wireless sensor network. One such technique is an enhanced memetic artificial bee colony optimization technique which is a population-based metaheuristic approach for obtaining optimization in wireless sensor networks.

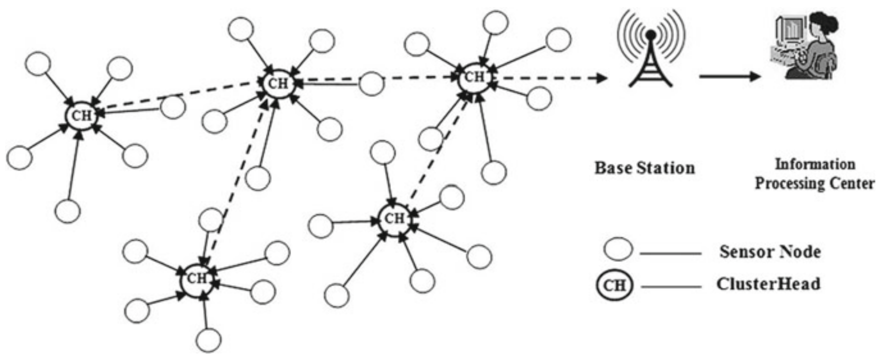


Fig. 1 Clustering in wireless sensor network

Related Works

Ant colony optimization (ACO) is swarm intelligence-based routing algorithm based on behavior of foraging behavior of ant colonies to solve the hard problems in wireless sensor networks [3]. Pareto multi-objective optimization strategy is included in conventional ACO to solve the security issues and issues related to limited resources. In multi-objective ACO algorithm, only two objectives are considered, and constraints such as maximizing network reliability and reducing failure rate of the network are not considered [4]. ACO uses pheromone concentration for finding the shortest feasible path in wireless sensor networks for the fastest packet transmission. This optimization provides reliable packet transmission and faster convergence but fails to consider the energy optimization.

Particle swarm optimization (PCO) [4] uses particle position and velocity for finding the optimal path which helps in faster packet transmission and convergence. This optimization considers the energy parameter of each sensor node, thus provides better energy optimization compared to ACO. In PCO, the number of hops is larger which causes network overhead. PCO works well when combined with other techniques.

Artificial bee colony (ABC) algorithm [4, 5] uses nectar amount to find the shortest path for packet transmission in wireless sensor networks by considering the energy parameter of the sensor nodes. The ABC algorithm achieves energy optimization but fails to meet the faster convergence rate [6]. ABC algorithm can be used to solve the complex problems in day-to-day life. But ABC algorithm attempts premature convergence and achieves slow convergence rate in wireless sensor networks. In order to overcome this, limitation of conventional ABC algorithm, memetic search is used with conventional ABC algorithm to explore the large search space. Thus, the premature convergence is reduced and faster convergence is achieved by using memetic search process.

Background

Memetic Algorithm

It is a problem-specific local search evolutionary algorithm that balances algorithm exploration and exploitation to improve quality. To limit the likelihood of premature convergence, the algorithm employs a local search approach [7]. In a memetic algorithm (MA), the fitness function is a function that takes a potential solution to the problem as an input and outputs how fit the answer is to the problem. To reach better milestone in finding solutions, memetic computation (MC) is evolved from memetic algorithms. Memetic computation uses a combination of memes which indicates group of information encrypted in complex structures which further interacts with each other for solving the problem.

Artificial Bee Colony Algorithm

In artificial bee colony (ABC) approach, each food source indicates viable path between the source node and sink node. Location of food sources are altered by artificial bees from colonies. Each colony is proficient of flourishing an intelligence which can be used for seeking the food. The food seeking operation [8] is divided into three major parts: employed food seekers for creating new food sources, onlooker food seekers for renewing the food sources depending on the nectar quantity, and scout food seekers for locating the new food sources in place of rejected food sources. In ABC, stair size is a combination of arbitrary number φ_{ij}, present solution and arbitrarily selected solution. The stair size decides the quality of the upgraded solution. If stair size is too big, the upgraded solution transcends the true solution, and if stair size is too small, then convergence rate of ABC may decrease significantly. Therefore, incorporating the memetic computation into standard ABC averts the situation of transcending the true solution while simultaneously maintaining the convergence speed.

Proposed Enhanced Memetic Artificial Bee Colony (EMABC) Algorithm

In enhanced memetic artificial bee colony (EMABC) algorithm, a safe cluster-based routing is developed for improving the performances of the WSN. This research comprises three stages such as clustering, secure cluster head (CH) selection, and secure routing path generation. Initially, the K-means clustering algorithm is used to divide the network into clusters. Subsequently, the near optimal secure CH is selected by using the enhanced memetic artificial bee colony (EMABC) algorithm. The originality in EMABC algorithm is considering all the four different fitness values such as trust value of the nodes, residual energy, distance, and node degree. The trust value contemplation in the EMABC is used to avert the black hole attacks during the CH selection. The black hole assault is contemplated as one of the wide active attacks which devalues the performance and reliability of the network as a result of dropping all incoming packets by the malevolent node. Moreover, this EMABC algorithm is used to generate the secure routing path between the source and destination nodes. Similar to the ad-hoc on-demand distance vector (AODV), the EMABC routing uses the route request and route reply messages to generate the routing path. In the route discovery process, the source node telecasts the route request messages to the neighbor nodes. Then, the node which has better fitness transmits the route reply message to the source node. Likewise, the secure data transmission path is generated amid the source and destination nodes. After identifying the transmission path, the data packets are sent to the destination node. This research considers all the four optimal fitness values, and performance is achieved using EMABC algorithm. Hence, the performance of enhanced memetic artificial bee colony (EMABC) algorithm is better than any other population-based heuristic algorithms.

Steps of Enhanced Memetic Artificial Bee Colony Algorithm

The original artificial bee colony algorithm works in three different stages such as employed food seekers, onlooker food seekers, and scout food seekers. Food availability is found out by only one employed food seekers. Using this information, onlooker food seekers makes a decision that which food source to visit. When availability of food is drained, then the unused food seekers become scout food seekers. In this working protocol, every food source represents one of the N feasible paths between each pair of nodes in wireless sensor networks. The location of a food source indicates possible solution for routing, and the nectar quantity of a food source is comparable to the strength of correlated solution measured by a fitness function. And each node stores two information in the neighbor table, first one is total time required for a packet to reach the next node, and second one is unused node energy. In order to get the accurate path in the search process, a local search space should be improved which is lagging in original ABC algorithm. Thus, by incorporating memetic computation [9] into original ABC algorithm, local search space is improvised which is done by adding one extra step to original ABC algorithm.

EMABC is population-based algorithm of size N where the candidate solutions a_i are vectors of M design variables within a decision space S. The initial population is generated randomly as follows:

$$a_{i,j} = \mathrm{rand}(0,1).(u_b - l_b) + l_b \text{ for } \quad j = 1, 2 \ldots M \tag{1}$$

where rand(0,1) generates a random value between 0 and 1, l_b, u_b represents lower and upper bounds of candidate solution a_i.

The exploration tasks in EMABC is given as follows:

1. Stochastic long distance type (comparable to employed food seekers in ABC)
2. Stochastic moderate distance type (comparable to onlooker food seekers in ABC)
3. Deterministic long distance type
4. Random long distance type (comparable to scout food seekers in ABC).

A. Stochastic long distance type

To get a trial solution, three operations such as mutation, crossover, and selection are performed on each of the candidate solution a_i. Mutation operation is performed as the first step to produce the trial solutions which can be calculated using the Eq. (2)

$$u_i^{(t)} = a_{r_1}^{(t)} + \left(a_{r_2}^{(t)} - a_{r_3}^{(t)}\right) \tag{2}$$

where $a_{r_1}^{(t)}$, $a_{r_2}^{(t)}$, $a_{r_3}^{(t)}$ are randomly picked candidate solutions, t is iteration number.

The number of modifications permitted to the trial solution is calculated using Eq. (3) which comes under crossover operation.

$$v_{i,j}^{(t)} = \begin{Bmatrix} u_{i,j}^{(t)} \text{ if}\big(\text{rand}_j(0,1) \le CR \| j == j_{r4}\big), \\ a_{i,j}^{(t)} \text{ otherwise} \end{Bmatrix} \tag{3}$$

where CR is crossover probability and j_{r4} is random dimension in trial solution.

The supreme solution between candidate and trial solution is calculated using Eq. (4)

$$a_i^{(t+1)'} = \begin{Bmatrix} v_i^{(t)} \text{ if } \mathrm{f}(v_i^{(t)}) \le f(a_i^{(t)}) \\ a_{i,j}^{(t)} \text{ otherwise} \end{Bmatrix} \tag{4}$$

where f represents the fitness function.

As this exploration is similar to employed food seekers in conventional ABC, a supreme solution is obtained using the fitness function. After the entire employed bee finishes the search process, they share that information to onlooker food seekers.

B. Stochastic moderate distance type

This exploration is similar to onlooker food seekers in conventional ABC, and the operations in obtaining the trial solution involve mutation, crossover, and selection which is similar to stochastic long distance type using Eq. (5)

$$u_i^{(t)} = a_i^{(t)} + \left(a_{\text{sup}}^{(t)} - a_i^{(t)}\right) + \left(a_{r_1}^{(t)} - a_{r_2}^{(t)}\right) \tag{5}$$

where $a_{\text{sup}}^{(t)}$ represents the supreme solution.

The crossover and selection operations are performed using Eq. (3) & (4)

At the end of this stage, best food path is obtained based on the quality of the nectar amount calculated by the onlooker food seekers.

III. Deterministic short distance type

This exploration tries to bring the candidate solution into local optima which are mainly concerned for maintaining the diversity and averting the situation of transcending the true solution. The success of any optimization algorithm depends on maintaining the diversity. The fitness on diversity can be expressed as

$$\varphi = 1 - \left| \frac{f_{\text{avg}} - f_{\text{bst}}}{f_{\text{wrst}} - f_{\text{bst}}} \right| \tag{6}$$

where f_{avg}, f_{bst}, and f_{wrst} are the average, best, and worst fitness values of food paths in the population.

IV. Random long distance type

When the food source is depleted, the food seekers become scouts, and the search process is repeated as described above in three different exploration tasks.

EMABC Algorithm

Step 1: Initial population is randomly created using Eq. (1)
Step 2: For each employed food seeker, the supreme solution is obtained using three operations such as crossover, mutation, and selection.
Step 3: Each onlooker food seeker updates the solution using the same three operations mentioned in step 2, and best solution is obtained using the nectar amount.
Step 4: Local search space is obtained by maintaining the diversity based on fitness function using the Eq. (6).
Step 5: In place of abandoned food sources, scout food seekers discovers the new food sources using the step 2, and best solution is obtained using step 3.
Step 6: The best solution or best food source so far found is memorized.

Experimentation and Result Analysis

The simulations are accomplished in ns 2.34 platform and organized in a square area of 1200 m × 1200 m with 100 sensor nodes which are deployed randomly. The simulation is executed to show the energy expenditure of the nodes using EMABC algorithm by changing the number of black hole nodes in each simulation step. The proposed algorithm EMABC achieves good performance than other population-based heuristic algorithms in terms of reduced energy consumption (Fig. 2 and Table 1).

The below figure shows the xgraph of black hole nodes versus the energy consumption. The energy consumption of the nodes is less in the proposed algorithm compared to the conventional population-based algorithms. The simulation is repeated by changing the number of black hole nodes while keeping the number of deployed nodes constant (Fig. 3).

The below figure shows the xgraph of black hole nodes versus routing load. The routing load is considerably less in the proposed algorithm compared to the conventional population-based algorithms. The simulation is repeated by changing the number of black hole nodes while keeping the number of deployed nodes constant (Table 2).

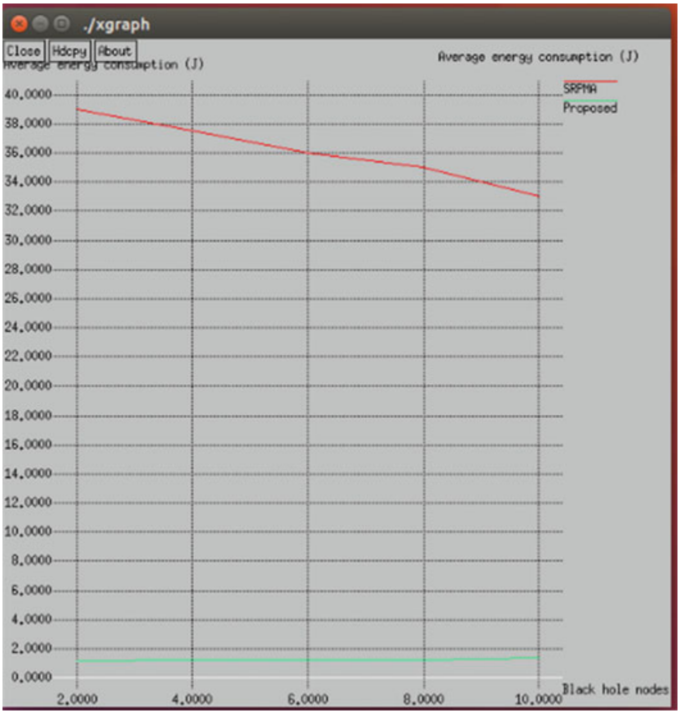


Fig. 2 Energy expenditure versus black hole nodes

Table 1 Simulation settings and parameters

Parameters	Values
Wireless propagation model	Two ray ground
MAC protocol	Mac/802.11
Network interface	WirelessPhy
Antenna	OmniAntenna
Queue	DropTail/PriQueue
Routing	AODV
Topography	1200 m x 1200 m
Initial energy	50 J
Number of nodes	100
Interface queue length	100
Simulation time	100 s

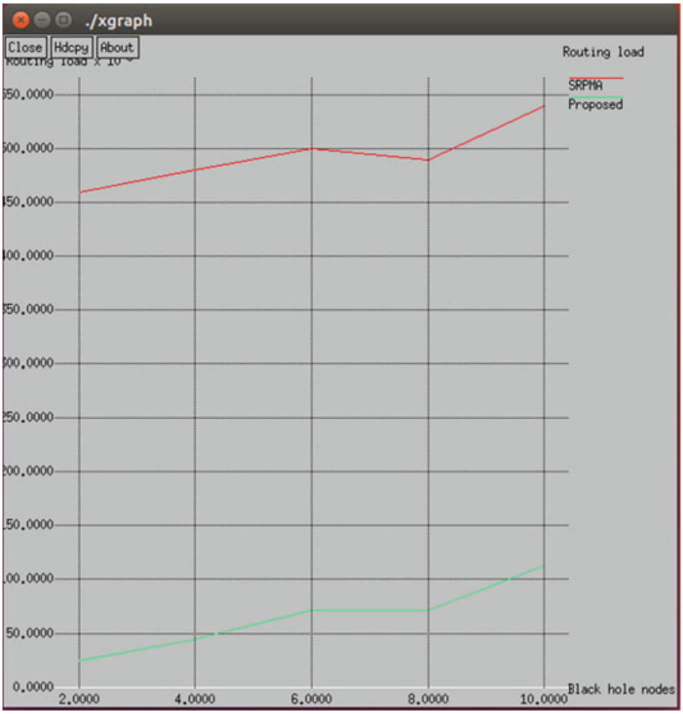


Fig. 3 Routing load versus black hole nodes

Table 2 Routing load, total energy consumption, and number of packets sent and received in the presence of different number of black hole nodes

Black hole nodes	No. of packets sent	No. of packets received	Total energy consumption	Routing load
2	392	388	1.1359	0.0242784
4	392	386	1.23102	0.0447927
6	392	384	1.19998	0.07125
8	392	384	1.19998	0.07125
10	392	377	1.35849	0.112626

Number of nodes deployed = 100

Conclusion

This research work uses memetic computation with existing artificial bee colony algorithm to give the solution to energy constraints in WSN. The proposed algorithm EMABC explores large search space by using memetic search to reduce premature convergence and improves the faster convergence rate. Thus, the algorithm proves better performance in view of achieving one of the fitness functions in this research such as reduced energy consumption than the other popular population-based algorithms. Future work includes the achievement of other fitness functions such as trust value, distance, and node degree.

References

1. Gao W (2016) Improved ant colony clustering algorithm and its performance study. Comput Intell Neurosci 2016. https://doi.org/10.1155/2016/4835932
2. Trejos J, Murillo A, Piza E (Jan 2004) Classification, clustering, and data mining applications. https://doi.org/10.1007/978-3-642-17103-1
3. Mamta (2014) Various clustering techniques in wireless sensor algorithms. 3(6):381–384
4. Jan B, Farman H, Javed H, Montrucchio B, Khan M, Ali S (2017) Energy efficient hierarchical clustering approaches in wireless sensor networks: a survey. Wirel Commun Mob Comput 2017 https://doi.org/10.1155/2017/6457942
5. Sun Z, Wei M, Zhang Z, Qu G (2019) Secure routing protocol based on multi-objective ant-colony-optimization for wireless sensor networks. Appl Soft Comput J 77:366–375. https://doi.org/10.1016/j.asoc.2019.01.034
6. Al Aghbari Z, Khedr AM, Osamy W, Arif I, Agrawal DP (2020) Routing in wireless sensor networks using optimization techniques: a survey. Wirel Pers Commun 111(4):2407–2434. https://doi.org/10.1007/s11277-019-06993-9
7. Adamou A, Ari A (2016) Energy efficient clustering algorithm for wireless sensor networks using the ABC metaheuristic
8. Problems HD (2020) A memetic artificial bee colony algorithm for. 2050008:1–14. https://doi.org/10.1142/S146902682050008X
9. Fister I, Brest J, Zumer V (2012) Memetic artificial bee colony algorithm for large-scale global optimization. pp 10–15
10. Bansal JC, Sharma H, Arya KV, Nagar A (2013) Memetic search in artificial bee colony algorithm. pp 1911–1928. https://doi.org/10.1007/s00500-013-1032-8
11. Karaboga D (Dec 2015) A powerful and efficient algorithm for numerical function optimization: artificial bee colony (ABC) algorithm—a powerful and efficient algorithm. https://doi.org/10.1007/s10898-007-9149-x
12. Raquel CR, Naval PC (2005) An effective use of crowding distance in multiobjective particle swarm optimization. GECCO 2005—Genet Evol Comput Conf 257–264. https://doi.org/10.1145/1068009.1068047
13. Singh N, Singh DP, Pant B (2019) ACOCA: ant colony optimization based clustering algorithm for big data preprocessing. Int J Math Eng Manag Sci 4(5):1239–1250. https://doi.org/10.33889/IJMEMS.2019.4.5-098
14. Ahmed ZE, Saeed RA, Mukherjee A, Ghorpade SN (2020) Energy optimization in low-power wide area networks by using heuristic techniques. INC
15. Hu K, Tsai C, Chiang M, Yang C (2015) Optimization for clustering. Hindawi Publ Corp Math Probl Eng 2015:1–11
16. G. Wang, Y. Wang, X. Tao (2009) An ant colony clustering routing algorithm for wireless sensor networks. 3rd international conference on genetic and evolutionary computing WGEC 2009. pp 670–673. https://doi.org/10.1109/WGEC.2009.22
17. Sowndeswari S, Kavitha E (2020) Improving energy conservation and securing wireless sensor networks—a comprehensive survey. 7(12):907–910
18. Ennajari H, Ben Maissa Y, Mouline S (2017) Energy efficient in-network aggregation algorithms in wireless sensor networks: a survey. Lect Notes Electr Eng 397:135–148. https://doi.org/10.1007/978-981-10-1627-1_11

19. Singh SP, Sharma SC (2015) A survey on cluster based routing protocols in wireless sensor networks. Procedia Comput Sci 45(C):687–695. https://doi.org/10.1016/j.procs.2015.03.133
20. Parwekar P, Rodda S, Kalla N (Oct 2018) A study of the optimization techniques for wireless sensor networks (WSNs). https://doi.org/10.1007/978-981-10-7512-4
21. Mao W, Lan H, Li H (2016) A new modified artificial bee colony algorithm with exponential function adaptive steps. 2016(i)
22. Dhaya R, Kanthavel R (2021) Bus-based VANET using ACO multipath routing algorithm. J Trends Comput Sci Smart Technol (TCSST) 3(01):40–48
23. Smys S, Wang H (2021) A secure optimization algorithm for quality-of service improvement in hybrid wireless networks. IRO J Sustain Wireless Syst 3(1):1–10
24. Manoharan JS (2020) Population based metaheuristics algorithm for performance improvement of feed forward neural network. J Soft Comput Paradigm 2(1):36–46

Smart Helmet for Motorcyclist

S. V. Viraktamath, Vidya, Rohan Bhadavankar, Siddhartha S. Nabhapur, and Nikhil Rati

Abstract Motorcycle accidents are occurring at a great phase, and the number of deaths caused by this also is at its peak. Currently, no technical innovations are being used to reduce the number of injuries and deaths caused by them. According to a study, almost half of people die as a result of a delay in receiving care after an accident. The major cause of this delay could be a lack of knowledge, traffic challenges, or other factors. This delay may be caused by inadequate information regarding the accidents, traffic problems, or other reasons. To tackle these current issues, the smart helmet is designed that provides the best solution to all of the issues mentioned above. The biker wears a smart helmet, which is a type of protective headgear that makes biking safer than before. The primary purpose of this smart helmet is to ensure the rider's safety. Advanced functions such as alcohol detection, accident identification, location monitoring, use as a hands-free device, and fall detection can be used to do this. This is not only a feature of the smart helmet but also a feature of the smart bike.

Keywords Arduino · Accident detection · Alcohol detection · GSM and GPS · Microcontroller

Introduction

Two-wheelers are the foremost sold vehicles all through India. The Indian two-wheeler industry enrolled a deals volume of 13.7 million units in 2012–13, a development of 2.9% over the year FY 2011–12. As the number of bikers in our nation is expanding, the street accidents are moreover expanding day by day, due to which numerous deaths are happening. The most common kind of negligence is failure to wear a helmet, which accounts for the majority of death cases. In addition, many deaths occur as a result of the injured person's inability to receive timely medical assistance. According to Section 129 of the Motor Vehicles Act, 1988, everyone riding a two-wheeler must wear protective headgear that meets Bureau of Indian

S. V. Viraktamath · Vidya · R. Bhadavankar (✉) · S. S. Nabhapur · N. Rati
Department of Electronics and Communication Engineering, SDM College of Engineering and Technology, Dharwad, India

V. Bindhu et al. (eds.), *Proceedings of Third International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 844,
https://doi.org/10.1007/978-981-16-8862-1_41

Benchmarks (BIS) requirements. Furthermore, intoxicated driving under the influence (DUI) may be considered a criminal offense under the Motor Vehicle Act of 1939, which stipulates that a bike rider who breaches the act would be penalized [1]. Currently, bike riders are able to easily escape from the law. Safety and security are the most crucial zones in all area of our lives. This inspires us to consider creating a framework that ensures the safety of bikers by making it mandatory to wear a helmet in accordance with government regulations, as well as inducing proper and fast medical attention in the event of an accident. The system intends to improve cyclist security and safety in the face of street accidents [2]. The development of this system has been fueled by social responsibility.

Literature Survey

In all the related work discussed below, it is observed that the rider needs to wear the helmet compulsory; otherwise, the ignition is not turned ON and usage of an RF module for wireless communication between the transmitter (Tx) and receiver (Rx). In some cases, if the bike rider is found alcoholic, the ignition is immediately locked, and in the event of an accident, a message is sent by GSM, along with the rider's location via GPS. In [2], a simple approach is used to make helmet wearing compulsory, where without the helmet bike cannot start. In [3], two Arduinos are used: one for the Tx unit and another for the Rx unit. The force sensing resistor (FSR) detects whether or not the helmet is worn, and both accelerometer and vibration sensor for detection of an accident. In [4], FSR is used to verify whether or not the helmet is worn, BLDC fan is used to detect speed and gives a warning if speed exceeds 100 km/h, and PIC16F84a microcontroller is used. A proximity sensor is used in [5] to detect whether or not the helmet is worn. The Internet of things (IoT) is employed in [6, 7], so that the user's data can be transferred to the cloud for monitoring of activities such as helmet use, alcoholic consumption, and accident conditions. In [8], The IoT is used and the helmet works on the Raspberry Pi 3 module, which is interfaced with both the vehicle and the cloud in which the image can be accessed and sent to the receiver. The technology in [9] includes a speech module that controls the visor, turn indication, headlights, horn, and ignition system, giving the user a lot more control over the vehicle. In [10], two Arduinos are used, and in addition to other sensors, the temperature sensor is used. The solar panel is used to power up. In [11], the vibration sensor, pressure sensor, and accelerometer are used to detect an accident. The helmet is fitted with a cooling unit in [12], which uses an external power source such as solar energy to lower the temperature. An emergency alert system might be used to request assistance from cops or family members in the event of additional situations such as harassment, abduction, or molestation. In [13–15], authors have proposed a review and comparative studies of different types of helmets.

From the works discussed in this section, it is seen that some of the work carried is very simple, like limited to detection of helmet worn or not, and some are complex which involves IoT, Raspberry Pi, and many other things which may result in increase in the cost. Hence, the proposed system in this paper has most of the things required in one single unit like alcohol detection, accident detection, and helmet detection which are essential for a safe drive.

Technical Studies

Microcontroller

A microcontroller board based on the ATmega328 is the core component of the helmet's hardware part. It has 28 pins. ATmega328 microcontroller is the main component of Arduino UNO. As soon as the microcontroller is powered, the peripherals such as GPIO and TIMER are initialized. This microcontroller is mainly used to detect the RFID whether it is similar to what is required or not, which helps in avoiding theft. ATmega32 is used in vehicle unit. It has 32×8 general working purpose registers, 32 K bytes of flash program memory, 2 K bytes of internal SRAM, 1024 bytes EEPROM. It has 32 pins (4ports $\times$ 8pins) configurable as digital I/O pins. It has 3 inbuilt timer/counters, two 8 bits (timer0 and timer2) and one 16 bits (timer1). Also, it has 3 external interrupts. Operating voltage is in range of +4.5 to +5.5 V and has an 8 MHz calibrated internal oscillator.

Alcohol Sensor

In the MQ sensor series, the MQ3 sensor is one of the most extensively utilized alcohol sensors. It is a sensor that detects the presence of alcohol gas in the air and outputs an analogue voltage reading. It can detect alcohol concentrations as low as 25 parts per million and as high as 500 parts per million (ppm). In this model, alcohol sensor is present near the front end of the helmet unit for sensing the presence of alcohol content in the riders' breath. These sensors are highly sensitive to alcohol and less sensitive to benzine. With a sensing range of 0.04 mg/L to 4 mg/L, it is perfect for breathalyzers. The voltage range of the MQ-3 sensor is 2.5–5.0 V.

RF Encoder and Decoder

A pair of encoder and decoder boards composes the module. The HT12E IC can only be used in combination with the HT12D IC. These two ICs work together to produce an encoder and decoder pair. 12-bit encoders/decoders are capable of sending and receiving 12-bit data. An 8-bit data address will be shared by the encoder and decoder

IC pair. These ICs are frequently seen in radio frequency (RF) pairs. They both use a voltage range of 2.4–5 V to operate.

LCD

LCD is an electronic display module, which is a commonly used component in most of the embedded applications. In this model, a 16 * 2 LCD is used, which can display 16 alphanumerical characters on the two rows. Each of the characters is made up of 5×8 pixels. This can operate at a voltage range of 4.7–5.3 V.

GSM Module

A GSM or GPRS module is a circuit that connects a computer or mobile device to a GSM or GPRS network. A GSM modem might be a separate device having a serial, USB, or Bluetooth interface, or it can be a mobile phone with GSM modem capabilities. Its dimensions are $24 \times 24 \times 3$mm.

GPS Module

The NEO6MV2 GPS module makes it possible to receive GPS signals. This allows the model to establish its actual location on earth by supplying the longitude and latitude of its position, as well as the speed and direction of its movement and the number of satellites visible. A serial Tx/Rx connection is used for communication (just two I/O are required). NEO6 modules are suited for battery-operated mobile devices with severe cost and space requirements because of their tiny architecture, power, and memory options.

Vibration Sensor

A structure's vibration or acceleration of motion is detected and measured using vibration sensor modules. They have a piezoelectric transducer that transforms mechanical force from vibration or movement into an electrical current.

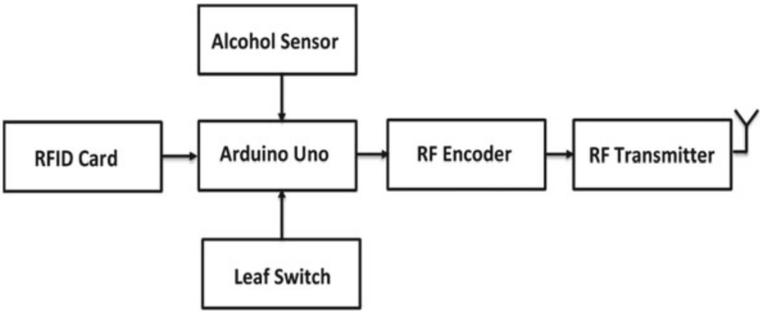


Fig. 1 Helmet unit

Leaf Switch

A leaf switch is a switch that is installed in a helmet model to detect whether or not the wearer is wearing one. The microcontroller gets a signal indicating whether or not the user is wearing a helmet when the switch is pressed.

Methodology

Block Diagram

The entire system of smart helmet is mainly divided into two units: vehicle unit and helmet unit. The below block diagram represents the clear idea of each unit. Figure 1 shows the helmet unit where it consists of alcohol sensor to detect the alcohol content in the rider's breath, a leaf switch to detect the helmet, and RF encoder to encode the data and it is transmitted using RF transmitter; the whole helmet is protected with RFID card reader against theft.

Figure 2 indicates the vehicle unit where all the inputs from the different sensors are applied to the microcontroller. The microcontroller processes all the inputs from the sensor and takes the required actions. The microcontroller is the main head of the unit. The data received by the helmet unit will be processed and gives the output; the GSM and GPS are used to send the SMS to the family members whenever the rider met with an accident with live location.

Flowchart and Working

In this section, the design methodology of the system is indicated in flow diagram. The sensors used here sense all the parameters, check all the conditions, and give

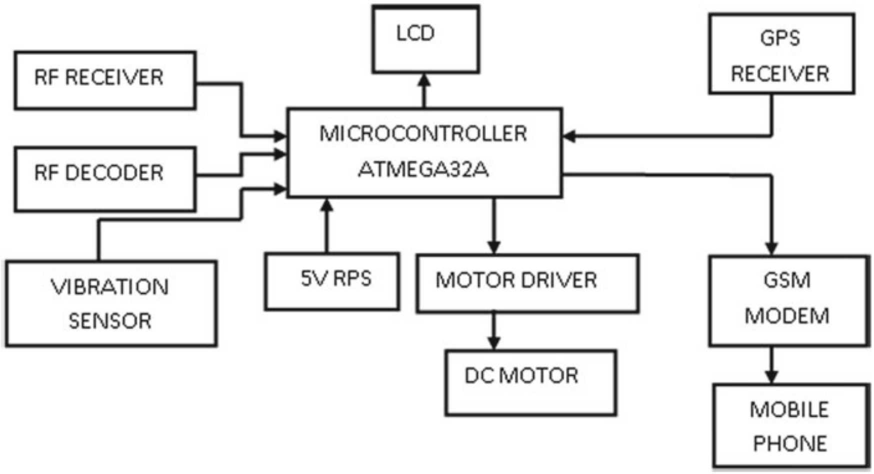


Fig. 2 Vehicle unit

the output. If the necessary conditions are not met, it checks again all the parameters and gives required result. As soon as the microcontroller is powered, the peripherals such as GPIO and TIMER are initialized. The first step in the flowchart involves the initialization of all the ports and then setting the RF communication between the bike and the helmet unit. The RF decoder and the vibration sensor are monitored by the microcontroller. The next step involves the arrival of the signal from the leaf switch where it checks whether the helmet is worn or not and the alcohol sensor placed at the front end of the helmet senses the alcohol content in the riders' breath, and if it is between the range of 25 and 500 ppm, it sends the signal as the rider is drunken. The ignition will be locked until the rider wears the helmet, and if the rider is drunk, a message will be displayed on the LCD. If both the signals are valid, the user will be able to access the vehicle. The vibration sensor present in the vehicle unit senses the vibration, if any accidents occurred. In case of any accidents occurred, a message is sent to a mobile number of the authorized person with the location. Figure 3 shows the flow of the system.

Results

The smart helmet system is successfully tested, and the results obtained are analyzed. After the testing, it can be integrated into a helmet. Figure 4 gives the schematics of the vehicle unit, and Fig. 5 shows the helmet unit. The driver is not allowed to start a bike if he is not wearing a helmet, also displays the message as "PLEASE WEAR THE HELMET," and the message is also sent to the given number. If the accident is detected or rider has consumed alcohol, it is displayed in LCD as shown in Figs. 6

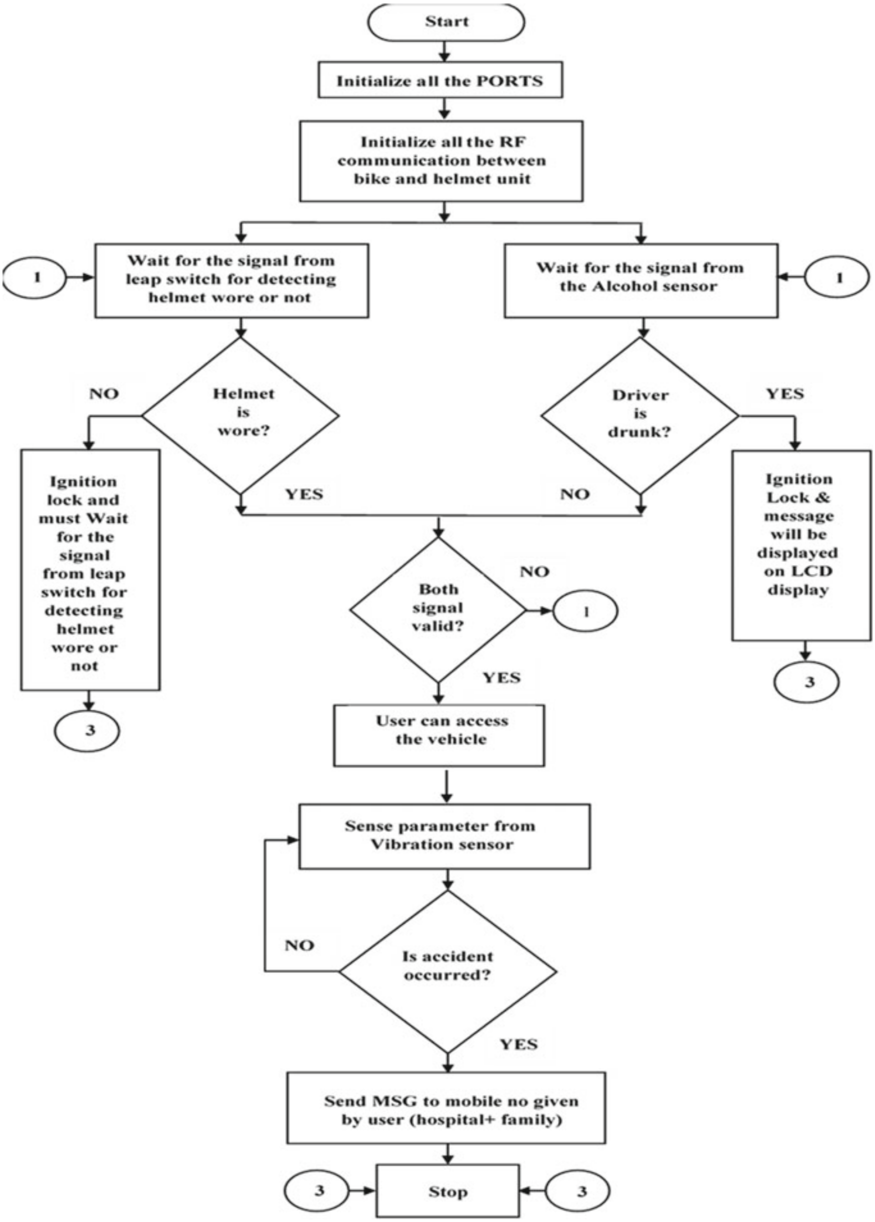


Fig. 3 Flowchart of the model

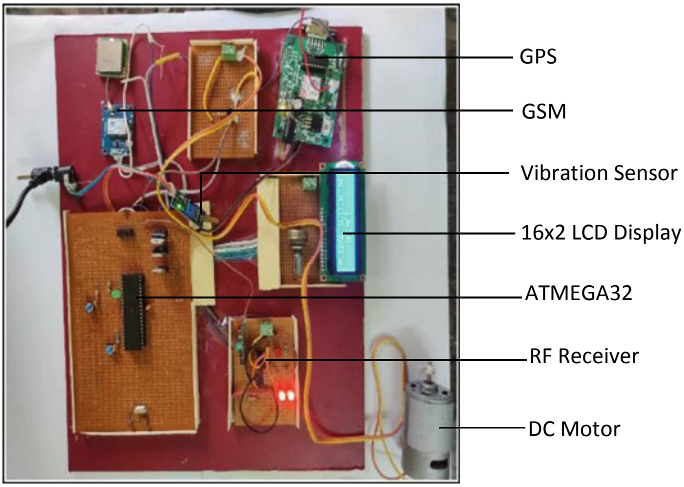


Fig. 4 Vehicle unit

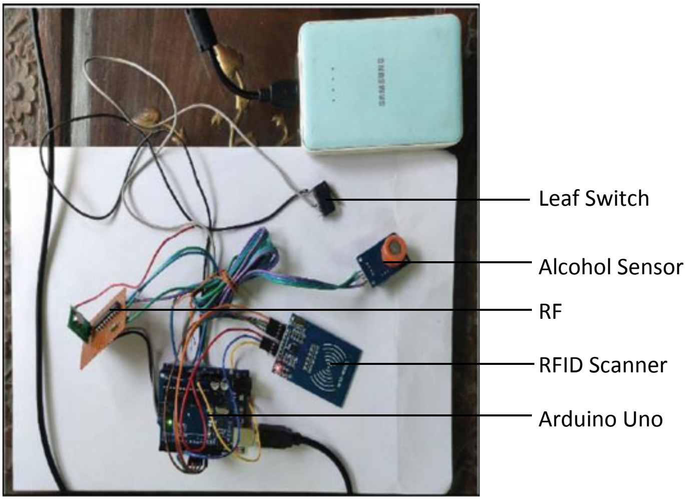


Fig. 5 Helmet unit

and 7. When the rider meets with an accident, the vibration sensor in vehicle unit detects it and a message is sent with the help of GSM to the designated cell number and also the location of accident spot which is obtained from GPS. This is depicted in Figs. 8 and 9.

Fig. 6 Accident detection

Fig. 7 Alcohol detection

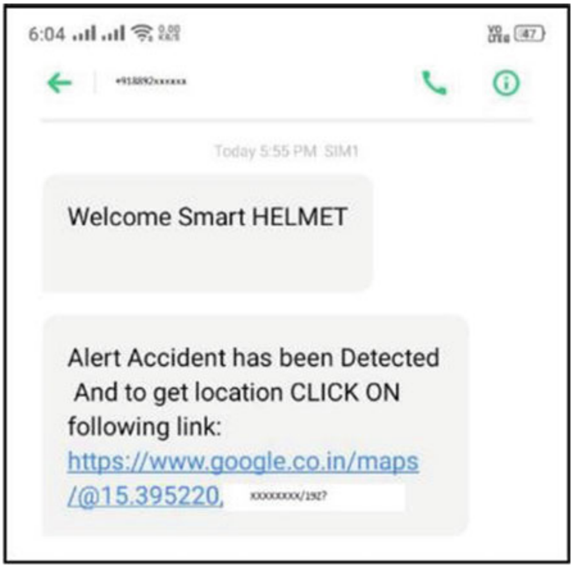


Fig. 8 Message received on registered mobile number when accident detected

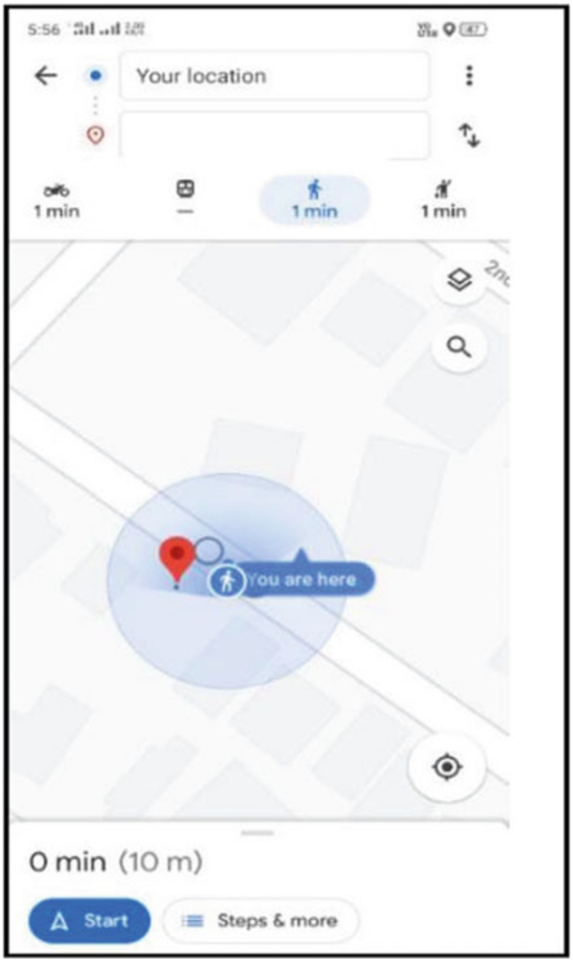


Fig. 9 Location of the accident

Future Scope

- The use of various bioelectric sensors on a helmet, which can track diverse actions, can improve this safety system technology.
- The entire approach can be adopted in cars simply substituting a seat belt for the helmet.
- A small camera can also be placed in the helmet for recording the activities of the driver. It also gives a message about the passing vehicles.
- A solar panel can also be inserted at the helmet top, by which the helmet unit gets a power supply. This power can be used for charging the mobiles.
- The model as a whole can be employed in a real-time safety system.

Conclusion

The smart helmet acts as a rider's protective headgear that helps to reduce the number of people dying from road accidents. If not treated as soon as possible, a person may die. Violations of traffic laws can also be reduced. The suggested system of smart helmet stipulates that the user must wear a helmet when riding the bike, without which the bike will not start. The alcohol sensor detects alcohol in the breath of the rider. The bike will not start if found to be alcoholic. Finally, in the event of an accident, a message containing the accident location is sent to a family member or any registered number via GSM. This aids in the provision of timely assistance to the rider and the prevention of casualties.

References

1. Chitte PP, Salunke Akshay S, Thorat Aniruddha N, Bhosale Nilesh T (2016) Smart helmet & intelligent bike system. Int Res J Eng Technol 03(05):483–487; Clerk Maxwell J (1892) A treatise on electricity and magnetism, 3rd edn, vol 2. Clarendon, Oxford, pp 68–73
2. Agarwal N, Kumar A, Pratap Singh P, Sahani R (2015) Smart helmet. Int Res J Eng Technol 02(02):19–22
3. Shravya K, Mandapati Y, Keerthi D, Harika K, Senapati R (2019) Smart helmet for safe driving. In: E3S web of conferences, vol 87, p 01023. https://doi.org/10.1051/e3sconf/20198701023
4. Mohd Rasli MKA, Madzhi NK, Johari J (2013) Smart helmet with sensors for accident prevention. In: 2013 International conference on electrical, electronics and system engineering (ICEESE), pp 21–26. https://doi.org/10.1109/ICEESE.2013.6895036
5. Nandu R, Singh K (2014) Smart helmet for two-wheelers. Adv Automob Eng 3:110. https://doi.org/10.4172/2167-7670.1000110
6. Divyasudha N, Arulmozhivarman P, Rajkumar ER (2019) Analysis of smart helmets and designing an IoT based smart helmet: a cost-effective solution for riders. In: 1st International conference on ınnovations in ınformation and communication technology (ICIICT), pp 1–4. https://doi.org/10.1109/ICIICT1.2019.8741415
7. Rao PK, Sai PT, Kumar NV, Sagar SK, Vidya Y (2020) Design and ımplementation of smart helmet using IoT. In: International conference of advance research & ınnovation (ICARI)
8. Rao S, Vishnupriya SM, Mirnalini Y, Padmapriya RS (2018) The high security smart helmet using internet of things. Int J Pure Appl Math 119:14439–14450
9. Prashanna R, Sangameshwaran M, Poovendan V, Pavanptanesh G, Naveen C (2018) Voice controlled smart helmet. Asian Rev Mech Eng 7(2):1–5
10. Vinod GV (2018) Smart helmet. https://doi.org/10.13140/RG.2.2.11068.69762
11. Manzoor F, Bashir S, Manzoor A, Wani Z, Mohi Ud Din S (2017) Fazz smart helmet. Int J Comput Sci Mob Comput 6(6):332–335
12. Kodanda Ramaiah GN, Hunagund PV, Panjagal SB (2015) Smart helmet: the next geneartion solar gadget. Int J Adv Innov Thoughts Ideas 03(02). https://doi.org/10.4172/2277-1891.1000159
13. Baburaj A, Thasni VT, Reshma NS, Yadhu Krishnan P, Deepak KN (2020) Intelligent smart helmet system: a review. Int J Adv Res Comput Commun Eng 9(1):18–23. https://doi.org/10.17148/IJARCCE.2020.9103
14. Korade P, Gupta M, Shaikh A, Jare S, Thakur Y (2018) Smart helmet: a review paper. JSDR 3(11):170–171
15. Sharma S, Vaideki M, Ashok M (2017) A comparative study of smart helmets in IoT. Int J Adv Res Manage Arch Technol Eng (IJRMATE) 3(5):53–57

Review of Cyber-Attacks on Smart Grid System

Shreyas Mavale, Jayesh Katade, Nachiket Dunbray, Sparsh Nimje, and Balaji Patil

Abstract With increase in use of technology along with new inventions, consumption of energy required is increased in exponential manner. In recent development of communication infrastructure in smart grids, it leads to new issues in physical power system related to cyber security. In traditional ways to handle cyber-attacks on power system, it generally involves separation of cyber domain and physical domain. Therefore, it is very important to have unification of cyber and physical power system. So to handle these issues, cyber physical power system (CPPS) is introduced. The CPPS mainly consists of core physical power system tightly integrated with cyber system. The purpose of CPPS is to monitor and control the smart grids efficiently and reliably. CPPS consists of various phases such as power transmission, power generation, power distribution, utilization of power, supervisory control and data acquisition system (SCADA), and utilization of power. These phases in smart grids are prone to cyber-attacks so it is important to summarize, analyze, and monitor cyber-attack methods on CPPS for the defense against different cyber-attacks. In this paper, we provide a comprehensive review of different kinds of cyber-attacks on physical power system phases. Also, paper analyzes various scenarios of cyber-attacks on systems and equipment such as SCADA system, smart meters, and communication system. Finally, according to several characteristics of cyber-attack methods, some preventions and detection methods are presented.

Keywords Cyber physical system (CPS) · Cyber physical power system (CPPS) · Cyber security · Cyber-attack · Smart grid · Supervisory control and data acquisition (SCADA) · Distributed denial of service (DDoS) · False data injection (FDI)

S. Mavale (✉) · J. Katade · N. Dunbray · S. Nimje · B. Patil
School of Computer Engineering and Technology, MIT World Peace University, Pune, India

B. Patil
e-mail: balaji.patil@mitwpu.edu.in

V. Bindhu et al. (eds.), *Proceedings of Third International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 844,
https://doi.org/10.1007/978-981-16-8862-1_42

Introduction

Nowadays with increase in rate of natural disasters such as floods, hurricanes, tornados are compromising the reliability and performance of electricity grid. With ongoing industrial 4.0 revolution, use of Internet gets exponentially increased which leads to whole new sector known as cyber security and further which leads to soaring risk of intentional physical cyber-attack on smart grids. The advancements in smart grid technologies are also expected to expand cyber flaws and vulnerabilities of power grid system through remote access points. Cyber-attacks on Ukraine's physical power system in 2015 are great example of how cyber security compromisation can affect whole power system. More than thousands of homes and other facilities experienced a power outage for days. This cyber-attack was done by malware known as "BlackEnergy" which was installed on central control center [1]. Also, attacks on substation transformers in California are recent examples in cyber physical attacks on smart grid system.

Compared to physical intrusions in power system, cyber-attacks are hard to locate as attackers can be anywhere with network access. Cyber-attacks mostly depend on configurations of communication network in power grid. As we cannot perform different kinds of scenarios and testing on physical power system directly, different types of testbeds have been developed for power grids which serve the purpose of analyzing, modeling, testing, and impacts on other subsystems.

With integration of cyber system and physical power system, information sharing and communication network between them become important sector. Because of large-scale usage of communication technologies in physical power system, automation within it and remote access control of system have been gradually increased. Because of installation of intelligent electronic devices (IEDs) on physical power grid system, physical power system operators are able to control and monitor whole system from a remote control center. All of these control and monitoring centers are based on information and communication technology (ICT).

The remainder of this paper is organized as follows: The literature survey is represented in Section "Literature Survey". Goals of in smart grid system are presented along with overview of whole cyber physical system (CPS) in Section "Smart Grid". Section "System Architecture of CPPS" presents CPPS system architecture along with various types of levels in CPPS model. Section "Scenarios of Different Cyber-Attacks in CPPS" gives an outlook of the applicable scenarios of different cyber-attack in CPPS. Finally, conclusion is in Section "Conclusion".

Literature Survey

In cyber physical attacks, detection and prevention are being the subjects of comprehensive research in recent years [2] and various attacks which specially targeting stability of power systems have been studied [3]. Cyber-attack mainly targets the information principles which are confidentiality, integrity, and availability.

Liu et al. modeled local FDI attacks, which made use of reduced network information and passed the examination of the state estimator [4]. Zhang et al. presented a data integrity attack detection method based on a gray relational analysis method, which evaluated the correlation between measurements and control variables [5]. According to ISA 84/IEC61511 [6], functional safety is aimed toward protecting and monitoring devices from accidental failures or failings so as to realize or maintain a secure state of the system. Security refers to cyber security. Consistent with ISA99, cyber security attempts to guard the cyber environment of the authorized users or organization, including networks, information in storage or transit, devices, processes, all software, etc.

One of the most important studies related to cyber security in power grids is false data injection (FDI), work of Dán and Sandberg [7], who study the problem of identifying the best k sensors to protect in order to minimize the impact of attacks, and Kosut et al. [8], who consider attackers trying to minimize the error introduced in the estimate, and defenders with a new detection algorithm that attempts to detect false data injection attacks.

In 2020 Yohanandhan, R. V. and his team, explained different cyber-attacks and analysis, modeling and simulation of those attacks [9]. North American Electric Reliability Corporation (NERC) states various assets as well as rules and regulations for protection and smooth operation of physical power system [10]. Jahromi, A. A. and team in 2020 address the vulnerabilities which are present in communication-assisted protection scheme. DDoS, FDI, etc. These kinds of vulnerabilities can be found in this scheme [11]. In 2016, Jokar P. and his team used a consumption pattern-based energy theft detector which does not invade customer's privacy and use to detect malicious consumption patterns. They also used anomaly detectors and transformer meters to make algorithm robust against malicious changes [13]. Amin, S. and his team gave elaboration about how parameter tampering in meters can affect whole power consumption of customer also about consumption test, the nonparametric cumulative sum (CUSUM) algorithm to evaluate electricity theft detection system [14]. In 2008, Ten C. W. and his team emphasizes on three-level substation model for cyber system. Ten C. W. also explained about reducing password threshold for lower probability of intrusion attempts [19].

Smart Grid

With advancements in technologies integrating with physical power system, new term is introduced as smart grids. Smart grid is same as electrical grid which consists of several of components such as smart distribution system, advanced metering system for power distribution and integration of different circuit breakers (Fig. 1).

Nowadays, cyber physical system (CPS) is widely used in various sectors. As CPS is mainly integration of sections such as physical systems and cyber system. CPS consists of embedded systems which uses various wireless network protocols for real-time monitoring. CPS has not only applications in power grids sector but

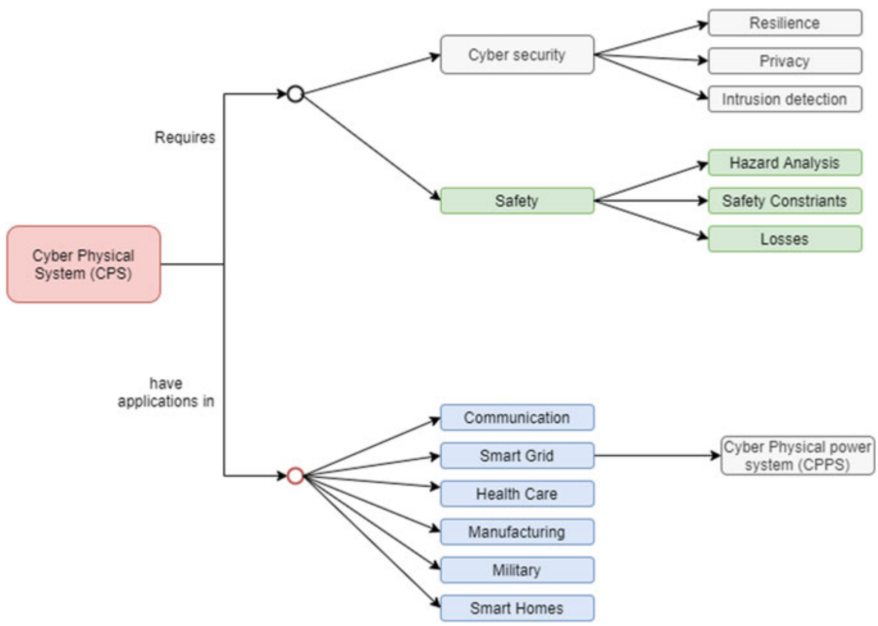


Fig. 1 Overview of cyber physical system

also in public health sector, safety, and precautions sector, water system, military, etc. There are many reasons for updating in electric power grid system to smart grid system. Some of them are as follows:

i. Efficiency: With increase in demand for electricity every year, it is necessary to minimize power losses and optimize power transmission. Efficiency also depends upon several factors such as implementation of new technologies and renewable sources for power generation.
ii. Reliability: Power grid system should be reliable and able to provide required power supply to given areas. With the help of newly developed sensors and actuators, power grid system can receive real-time data of operations and detection of attacks can be done in better way. For example, smart meters are installed to monitor consumption of energy not only inside of house but also of outside.
iii. Consumer choice: There should be transparency between consumer and power provider, as most of the consumers only receive monthly bills and updates related to their energy usage. Customers are also not aware about their energy consumption, and prices they are paying at different times in a day. Also, they do not know about how much energy is generated through renewable resources.

Smart grid systems are designed to overcome the above problems, but while integrating traditional physical power systems with modern technologies it also gives birth to new problems in cyber sector such as loss of energy and false billing to customers for energy usage.

System Architecture of CPPS

In smart grids, it is very efficient to manage system remotely as well as organizations can easily track the power consumptions data. To improve efficiency and stability in smart grid system, demand–response program provides a mechanism to control energy usage by providing incentives to consumers for reducing energy consumptions during peak hours. At present, these demand–response programs are mainly used by government and commercial consumers for large buildings and areas. And functioning of these programs is practically based on sending incentive signals vi calls or messages. For example, a company can send signals to consumers that to lower their consumption of energy in peak times.

Consumers with energy generation and storage capabilities are known as prosumers. Demand–response program can be extremely useful for prosumers as they can exchange energy with each other and can maintain their own security protocols. Many vulnerabilities and problems occur while controlling physical power system with remote technologies, it leads to a term called cyber physical power system (CPPS).

It is integration of cyber security and physical power system. CPPS mainly covers various areas such as transmission, distribution, and power generation. CPPS is integration of computation, networking, and physical power system which includes different types modeling, innovations, and creations. CPPS uses embedded computer network for communication, computation and organization of physical power system. In CPPS, there are three levels of interaction (Fig. 2).

(1) Level 1:

First level of CPPS is between transformer, generator and transmission line, etc., with controller of physical power system. Controller of physical power system calculates the control signal and also fetches information from other system components. This information is further used optimize the operations of power system.

(2) Level 2:

Second level of CPPS is between communication infrastructure and physical power system. Communication infrastructure works with coordination of sensors, actuators, controlling units, and communication units. There are some challenges in communication infrastructure such as time delay, bad data, and loss of data which may affect operations of physical power system.

(3) Level 3:

Third level of CPPS is between cyber security system and communication infrastructure. Components of cyber security system are communication servers, master servers, communication structures, computing stations, application softwares, cyber-attack, and defense models. This layer performs actions like operation planning, analysis of power system, optimization of voltage and power supply, monitoring of system.

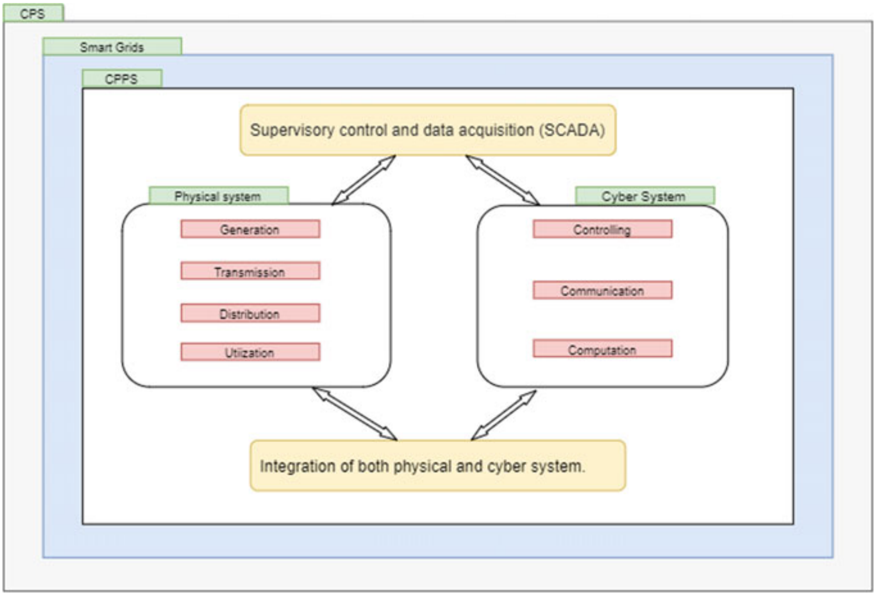


Fig. 2 Overview of smart grid system

In CPPS, distribution system and network transmission are very huge, complex, and heterogeneous which generate possibility of cyber-attacks. Various components of cyber security system like operating systems, remote access points, and connection ports are prone to cyber threats. It is extremely important to calculate the impact caused by cyber-attack on physical power system as attack does not directly impact physical power system but causes instability in whole system.

Traditional methods used in power system are totally based on physical parts of power grids. Optimization of CPPS is needed for ensuring secure and safe operation of power grid.

With the help of simulation tools and testbeds, we can predict the impact of cyber-attack and also analyze it. The risk factor of cyber-attack by vulnerabilities and their impact is shown as [9]

$$R = [T] * [V] * [I]$$

where

R stands for risk.

T stands for threat to power system.

V stands for vulnerability.

I stands for impact.

Here vulnerability can be of various types and threat can evaluate as motivation of attack and resources available for attack, etc. Developing an integrated risk assessment modeling framework is main motive of research in cyber security.

Scenarios of Different Cyber-Attacks in CPPS

CPPS consists of different phases as generation, transmission, distribution, communication, computation, and controlling. Many phases are vulnerable to cyber-attacks such as DDOS and false data injection which can compromise whole system. Cyber-attacks on smart grids are mainly on following phases.

A. Intelligent electronic devices (IEDs)
B. Advanced metering system
C. Supervisory control and data acquisition (SCADA)
D. Communication channel.

In cyber-attacks on the above phases, it causes compromisation of communication networks, malfunctioning of sensors and actuators, compromisation of whole system (Fig. 3).

Intelligent Electronic Devices (IED)

In recent years, traditional devices which were used in power systems have been replaced by intelligent electronic devices (IEDs) and intelligent controllers. IED provides remote control center and digital communication. North American Electric Reliability Corporation (NERC) succeeds to develop a system called critical infrastructure protection (CIP) for "identification and protection of important cyber assets for reliable operation of physical power system" [10].

IEDs are connected with phases of physical power system which are power generation, power transmission, power distribution which can be accessed remotely. Generally, most of the power substations are unmanned and operated via remote control technologies; there are high chances of cyber-attack on substation communication networks (SCN). Once an attacker exploits the system via methods like password cracking, malware installation, they gain access to important data of substations such as maintenance records, status of operations, system topology, measurements and operating plans of system). If there is vulnerability in communication system, then attackers can access multiple substations at a time which may lead to triggering number of cascading events on physical power system causing a blackout.

DDoS or false data injection attack extends the duration of fault clearing time at critical transmission lines by different methods such as packet flooding or by alternating signals which may lead to blackout and instability in smart grid system. This kind of attacks could be occurring with the help of malware injection or adding new

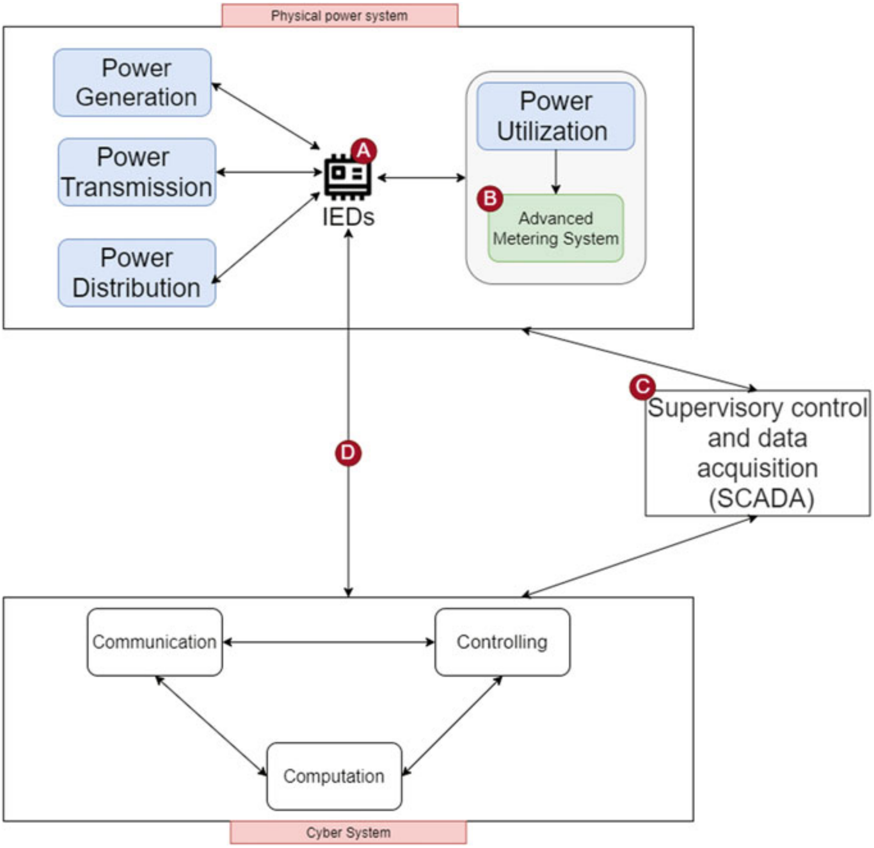


Fig. 3 Attack-prone systems or devices in CPPS

communication device which can access the network channel [11]. DDoS corrupts network with extensive packet flooding which makes internal communication of power stations impossible.

Also, there was certain type of attack where attacker used electromagnetic interference to cause an actuator to follow his commands. Traditionally most of the attacks on CPPS are software-based attacks, but system can be compromised by physical way by injecting false signals or injecting malware into system. These types of attacks are known as transduction attacks. Attacker can manipulate system and its environment by altering sensor data and injecting false data.

Advanced Metering System (AMI)

AMI is a smart metering system for customer-side which helps to build relationship between power consumer and providers directly. Traditionally used meters are mainly used for recording usage of power by consumer but smart meters are able to record not only energy flow inside of house but also out of house. So with the help of smart meters, consumers can become producers by installing solar panels or wind generators.

Smart meter can also work as a controller, and home appliances can be controlled by mobile phones with the help of Internet. It can be also serves as router in home area network (HANs). Nowadays, most of the smart meters are developed on basis of ZigBee protocol which is defined in IEEE 802.15.4 [12]. ZigBee is mainly developed for low power consumption devices, so it has limit for communication distance.

As smart meters are connected with Internet, cyber-attacks are most likely to happen on that certain point. ZigBee end-device sabotage attack is possible by activating sensors to send messages every time when device turn on from sleep mode. So by this method a sensor which is under attack is forced down to reply the malicious user and compromise system. Also, attacks such as DoS, DDOS, sinkhole, and wormhole are executed to disconnect nodes from connected networks. In sinkhole, attack node which is compromised will attract other network packets to create confusion in routing phase while in wormhole attack node which is compromised will receive malicious code and forward malicious packets to all connected network [13].

Smart meters can be compromised while attacker can modify the values for reading of energy consumptions. Different detection systems have been developed for energy theft [14–16]. It causes increase or decrease in billing of energy consumption, cutting down power of different locality remotely, etc. One of the solutions for energy theft is by monitoring load profiles and recognizes changes such as drastic change in power usage at certain time or unusual power usage readings by meters.

Supervisory Control and Data Acquisition (SCADA)

For online operation and monitoring of the critical infrastructures, SCADA systems are deployed in various industries, like power, oil and gas, transportation and manufacturing. Abnormal operating conditions of an influence system are often detected from a foreign location through a SCADA system. Thus, the reaction time to correct an abnormality is reduced. Additionally, utilities can reduce routine and emergency visits of field crews to remote sites.

In SCADA systems, many critical vulnerabilities may occur such as disrupts in communication, leakage of critical information, injection of malicious codes and commands, false data injection in control center of system. In any cyber-attacks on physical power system vulnerability index is calculated based on effect caused by the intrusion on power grids.

Most crucial attack in power system is when attacker gains access of SCADA system and launches different attacks which may cause catastrophic damages. Inclination of new technologies toward Internet protocol (IP)-based systems results in new vulnerabilities associated with SCADA system. Awareness and personnel trainings for SCADA system are very crucial as the exponential increasing dependency on communications over Internet protocols [17, 18]. As CPPS testbeds, SCADA testbeds are also effective solution for detecting vulnerabilities and analysis of them [19]. But there is no systematic method for modeling and analysis of critical assets in power system infrastructure.

Remote access systems have capability of controlling other machines or systems via network. An intrusion in substation enables an attacker to create false data to cause unnecessary operations on internal devices. An attacker uses some of the following methods to get entry in network layer of power systems:

- Port Scanning: With the help of port scanning admins can determine the ports on machine which are in listening state for potential access.
- War dialing: It executes different scripts in surrounding networks to detect connection, which can be threat to power system
- Traffic sniffing: Tools such as wireshark are used as network analyzer to capture network traffic.
- Password cracking: In this method, attacker uses randomly generated password list to gain unauthorized access.

Different types of attacks on SCADA system are as follows

Directed attacks. Attacks which have short-term effects on control system are known as directed attacks. This type of attack leads to shutting down of SCADA system for certain amount of time or deletion of system files. DoS and DDoS attacks are examples of directed attacks.

Intelligent attacks. Attacks which are well-planned and executed with proper management are known as intelligent attacks. In this type of attacks, attacker must have in depth knowledge about internal working of system. This type of attacks can cause major power outage which may leads to catastrophic.

Access point vulnerability. An access point provides the port services to determine a connection for an intruder to penetrate the SCADA computer systems. The vulnerability of a scenario through an access point is evaluated to work out its potential impact. The impact factor represents the extent of impact on a system when a substation is removed, i.e., electrically disconnected, by switching actions because of the attack [20].

Most crucial element of cyber security is the software on which system is running. The number of vulnerabilities increases every year, so control system should be updated and upgraded. Vulnerabilities of operating system are also most important as they can establish a malicious connection and can be easily compromised. In OS well-known ports are from 0 to 1024 for establishing connection. Attacker generally uses unused ports and services to gain the access of control system. Cyber-attack can cause the following consequences:

- Loss of information and loss of load
- Economic losses
- Damage of equipment
- Line failures
- Instabilities in frequency
- Increase in cost of operations
- Blackout for great amount of time.

SCADA system security model. SCADA systems specifically have password policies and some firewall rules for security purpose. There are two sub-models present as (1) Firewall model and (2) Password model.

Firewall model. In firewall model, it mainly monitors packets between network layer and uses different types of filters according to security levels. Firewall rules are mainly configured for filtering out unwanted traffic.

Some of the rules are as follows:

(a) Ports which are currently used in service.
(b) IP address or range of IP addressed which are connected to system.
(c) Types of protocol used in system.

Because of high volume of traffic, it is impossible to monitor or administer every detail in network. Hence, firewall analyzer is installed to detect abnormal traffic in network.

Password Model. It is used to locate previous attempts of breaking into system and analyze the log files which will help control system to develop a new model for prevention. Password model detect the response rate of machine with the help of central processing unit (CPU) and then study the behavior of attacks that occur over a certain period of time.

Cyber-Net is a combination of firewall and password model. In the SCADA system, there is already preinstalled algorithm present for minimizing the random error generated by sensors, and actuators known as bad data detection. False data injection can cause malfunctioning of system by evading that detection algorithm and sending false data.

Communication Channel

Attacker can compromise communication paths between controllers, sensors, and actuators with the help of DoS, false data injection attacks. Also, attackers can block or delay the controlling signals which further delay the future operations. Traditional methods in CPPS focus on separate parts of physical power system. Therefore, it is important to have integrated system for monitoring transmission, generation, and distribution combinely.

In physical power system, firewall works as a frontline defense for protection of system as it filters packets such that a packet can pass firewall if it fulfills the rules defined by the user of system. If attacker is trying to do IP scanning or port scanning, these events are recorded as log file in system. But firewall only examine and detect anomalies in lower layer of communication, i.e., network layer, so attacks which are on application layer or transport layer cannot be detected easily. Different types of IDEs have been proposed for communication system [21].

Types of Intrusion detection system (IDS) in communication channel. Physical power system is mainly consisting of substations, power generation units, distribution systems, and transmission systems while cyber systems consist of digital communication of data and SCADA system. In smart grids for different protection range, different types of IDSs are used such as for substations:

- Network-Based IDSs:

Network-based IDSs mainly monitor and analyze network traffic in local area network. It mainly checks header information of packets and content of packets which are passing through network layer.

- Host-Based IDSs:

Host-Based IDS is installed individually on more than one data servers. Host-based IDS mainly detects interruptions in measurements and status of physical devices. Host-based IDS (HIDS) can utilize log files to detect anomaly in power system.

All of the above cyber-attacks on smart grid phases, their prevention methods and impacts are summarized in Table 1.

Conclusion

Cyber security plays a very vital role in cyber physical power system (CPPS) and has wide attention from governments and academics. The proposed paper is focused on different kinds of cyber-attacks possible on smart grid system along with preventions on them and impacts of those attacks. During our research work, we found vulnerabilities in smart grid phases: With compromising sensors and actuators, exploiting a particular type of packet, attack on IEDs, injecting malware or Trojan in system, remotely reconfigure/program network nodes as the attacker wishes, hence compromising data communication security of the network. After analyzing various attacks on different phases of smart grids, we conclude that most of the phases are vulnerable to mainly DoS and DDOS type of attacks. This paper also discusses prevention and detection methods against cyber-attacks.

It is necessary to understand and study the principles of different attacks and develop effective security. Research on security and safety integration in CPPS is

Table 1 Attacks on CPPS phases and impact

S. No.	Phases	Attacks	Preventions and detection	Impacts
1	Generation Transmission Distribution (IEDs)	1. DDoS 2. DoS 3. FDI [4] 4. Alternating frequency 5. Malware injection [8] 6. Electromagnetic interference 7. Transduction attacks	1. Isolation of system 2. Bump-in-the-wire [9] 3. Cryptographic algorithm 4. Firewall installation 5. Use of blockchain technology 6. CPPS testbeds [18]	1. Instabilities in frequency 2. Major economic losses 3. Increase in case of operations 4. Exploitation of maintenance records, status of operations, system topologies 5. Blackout and instabilities [1]
2	Power Utilization (AMI)	1. DDoD 2. Sinkhole 3. Wormhole [7] 4. Absolute slot number 5. Time synchronization tree attack 6. DoS	1. Monitoring of load profiles and recognize drastic changes 2. Detect unusual power usage readings	1. Modification of meter reading values [13] 2. Cutting down of power in locality 3. Modification of readings of power generation
3	SCADA	1. DDoS 2. DoS 3. Injection of malware or Trojan on system software 4. FDI 5. Attack on system software	1. Firewall model 2. Password model 3. Cyber-net model 4. Bad data detection algorithm	1. Loss of load 2. Line failure [10] 3. Compromisation of whole system 4. Blackout for great amount of time
4	Communication network	1. DDoS 2. DoS 3. Man-in-the-middle attack 4. Identity spoofing 5. FDI	1. Intrusion detection system such as network-based IDS and host-based IDS y[21]	1. Compromise communication paths between controllers, sensors and actuators 2. Block or delay the controlling signals

still in progress and also needs future improvements and supplementation. A gap identifies in existing methodologies on security and safeties of physical power system are as follows:

i. Lack of security and safety measurements in cyber physical system. More research is needed in CPPS modeling and analysis.
ii. No unified method for detection and prevention of attack on physical system.
iii. Most of the existing technologies for risk assessment are not able to find potential threats and cannot predict attack scenarios and evaluate probability of attack.
iv. Most of the previous works in CPPS are not able to differentiate between errors of risk caused by incidents and malicious attacks.
v. Insufficiency in statistical information on intrusion detection attempts to invade the energy infrastructure. This type of limitations can be removed with the help of test beds development. As test beds are powerful evaluation and development tools in field of cyber physical power system.

To achieve high-performance rate and high efficiency in physical power system CPPS is used. CPPS has gained considerable amount of attention in recent years for modeling, analyzing, and simulation. Highly complex designs on testbeds can reduce performance of power system. So it is important to detect and prevent cyber-attacks against smart grids.

References

1. Assante MJ (2016) https://ics.sans.org/blog/2016/01/09/confirmation-of-a-coordinatedattack-on-the-ukrainian-power-grid
2. Khaitan SK, Mccalley JD, Liu CC (2015) Cyberphysical systems appraoch to smart electric power grid. Springer, Berlin Heidelberg
3. Kundur D, Feng X, Mashayekh S, Liu S, Zourntos T, Purry KB (2011) Towards modelling the impact of cyber attacks on a smart grid. Int J Secur Networks 6(1):2–2 https://doi.org/10.1504/ijsn.2011.039629
4. Liu X, Bao Z, Lu D, Li Z (2015) Modeling of local false data injection attacks with reduced network information. IEEE Trans Smart Grid 6(4):1686–1696
5. Zhang Z, Wang Y, Xie L (2018) A novel data integrity attack detection algorithm based on improved grey relational analysis. IEEE Access 6:73423–73433
6. IEC 61511 (2016) Functional safety—safety instrumented systems for the process industry sector
7. Dán G, Sandberg H (2010) Stealth attacks and protection schemes for state estimators in power systems. In: First IEEE smart grid commnunications conference (SmartGridComm)
8. Kosut O, Jia L, Thomas R, Tong L (2010) Malicious data attacks on smart grid state estimation: attack strategies and countermeasures. In: First IEEE smart grid commnunications conference (SmartGrid- Comm)
9. Yohanandhan RV, Elavarasan RM, Manoharan P, Mihet-Popa L (2020) Cyber-physical power system (CPPS): a review on modeling, simulation, and analysis with cyber security applications. IEEE Access 8(151):019–151064. https://doi.org/10.1109/access.2020.3016826

10. North American Electric Reliability Corporation (NERC) CIP Standard. http://www.nerc.com/fileUploads/File/Standards/Revised_Implementation_Plan_CIP-002-009.pdf Accessed on 2 May 2006
11. Jahromi AA, Kemmeugne A, Kundur D, Haddadi A (2020) Cyber-physical attacks targeting communication-assisted protection schemes. IEEE Trans Power Syst 35(1):440–450. https://doi.org/10.1109/tpwrs.2019.2924441
12. (2007) Wireless Medium Access Control (MAC) and PHY specifications for low rate wireless personal area networks (WPANs). IEEE Standard 802:30–30
13. Jokar P, Arianpoo N, Leung VCM (2016) Electricity theft detection in AMI using customers' consumption patterns. IEEE Trans Smart Grid 7(1):216–226. https://doi.org/10.1109/tsg.2015.2425222
14. Amin S, Schwartz GA, Cardenas AA, Sastry SS (2015) Game-theoretic models of electricity theft detection in smart utility networks: providing new capabilities with advanced metering infrastructure. IEEE Trans Smart Grid 35:66–81
15. Liu Y, Hu S (2015) Cyberthreat analysis and detection for energy theft in social networking of smart homes. IEEE Trans Smart Grid 2:148–158
16. Amin M (2002) Security challenges for the electricity infrastructure. Computer 35(4):supl8–supl10. https://doi.org/10.1109/mc.2002.1012423
17. http://www.tswg.org/tswg/ip/21StepsSCADA.pdf
18. Davis CM, Tate JE, Okhravl H, Grier C, Overbye TJ, Nicol D (2006) SCADA cybersecurity test bed development. In: Proceedings of the 38th North American power symposium, pp 483–488
19. Ten CW, Liu CC, Manimaran G (2008) Vulnerability assessment of cybersecurity for SCADA systems. IEEE Trans Power Syst 23(4):1836–1846. https://doi.org/10.1109/tpwrs.2008.2002298
20. Sun CC, Liu CC, Xie J (2016)
21. Khraisat A, Gondal I, Vamplew P, Kamruzzaman J, Alazab A (2020) Hybrid intrusion detection system based on the stacking ensemble of C5 decision tree classifier and one class support vector machine. Electronics 9(1):173

Cyber Security Challenges in the Latest Technology

Manjot Kaur

Abstract In the world of information management, data protection plays an significant position. Protecting knowledge has been one of today's greatest problems. Once we hear of privacy protection, the first topic that comes to mind is "Internet crime," that is rising exponentially each single day. Data protection, cyber security or information technology security are the defense of computer systems and networks from misuse or harm to their equipment, software or electronic records, as well as from interruption or mismanagement of the services they offer. Cyber security is a subset of computer security that deals with the Internet. The major security goal is to project the device using various rules and to set up various safeguards to protect it from online attacks. Cyber security is a practice that protects systems, networks and data breaches. These cyber attacks are typically targeted at obtaining, modifying or damaging private information; extorting users' money; or interrupting regular businesses processes. Different governments and corporations are taking many steps to combat these cyber-crimes. To avoid online assaults and improve Internet security, a variety of approaches are employed. The execution of successful information protection policies is especially difficult today because there are more computers than people and criminals are getting increasingly creative. Despite response to numerous initiatives, information defense is also a big concern. This article mainly addresses topics pertaining to the security of information, particularly emerging technology. It reports on new approaches, policies and innovations in information security which change the profile of information security.

Keywords Cyber security · Cyber-crime · Cyber ethics · Social media · Cloud computing · Cyber stalking · Malware · Encryption

---

M. Kaur (✉)
Chandigarh University, Gharuan, India
e-mail: manjot.cse@cumail.in

V. Bindhu et al. (eds.), *Proceedings of Third International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 844,
https://doi.org/10.1007/978-981-16-8862-1_43

Introduction

People can send and receive any kind of data simply by clicking a button, may be an e-mail, audio or video, but have they realized how simple their ID is to upload or upload securely without a leak of information to another person? Cyber protection is the key. The Internet is the rising network on a day-to-day basis. Some technological technologies turn the face of man into the urban world of today. But because of these new technology, we cannot secure our private information very well, cyber-crime arises day by day. Upward of 60% of overall business transactions are already carried out electronically, and a strong degree of security is required in this field for safe and efficient transactions [1]. Cyber attacks are also a worldwide epidemic that impacts several aspects of human life. The global security challenge continues to grow at a fast rate, with an growing amount of data breaches per year. Risk-based safety studies have shown that only in the first nine months of 2019 were an astounding 7.9 billion documents exposed to data violations. This estimate is more than twice (112%) the amount of documents released during the same time in 2018. Health facilities, distributors and governmental bodies have witnessed the most serious infringements, with violent offenders accounting for most of the cases. Many of these industries become more appealing to cyber attackers as they gather financial and medical data, but all companies that utilize networks may be aimed at consumer data, industrial theft or user assaults [2].

Cyber protection has therefore become a most recent problem. In addition to protecting information in the IT sector, the application of cyber security is restricted to many other areas such as cyber space, etc. Emerging innovations such as cloud storage, Internet networking, e-commerce, online banking and so on often need a high standard of security. Since systems have certain basic individual expertise, they have become a must for their safety. In order the defense and economic well-being of a country, improving data protection and securing sensitive knowledge infrastructure are important. The Internet has been an important part of both the creation of digital technologies and public policy, making us healthier (and defending Internet consumers). Combating cyber-crime needs a more systematic and safe strategy because technical solutions alone cannot deter any crime; it is essential that law enforcement authorities be allowed to successfully investigate and prosecute cyber-crime. Today several nations and governments enforce stringent electronic security regulations to prevent the destruction of any critical information. It must now be prepared to defend itself from this cyber security and from these rising cyber-crimes. With the magnitude of the cyber challenge expected to continue to increase, the International Data Corporation estimates that global spending on information protection technologies would hit a whopping $133.7 billion by 2022. Governments all over the world have reacted to the growing cyber challenge by offering guidelines to help companies adopt successful information protection activities. A system for information defense has been developed in the United States by the National Institute of Standards and Technology (NIST). In order to counter the dissemination of malicious code and early warning assistance, the system advises constant, real-time surveillance of all electronic infrastructure [2].

Cyber-Crime

The word cyber-crime is the primary means of commissioning and theft, like any illegal activity using a computer. The U.S. Department of Justice is extending cyber-crime concepts to cover any unauthorized operation, which saves information using a device [3]. Cyber frustration is a particular type of cyber-crime stalking. Various forms of violence taking place in cyberspace or use cyberspace to commit a criminal offense. Harassment may be physical, political, racist or otherwise. Those who are interested in perpetuating these attacks are accused of computer crimes. Cyber stalking takes us to another relevant field of infringement on netizens' privacy. The list of cyber-criminals involves Internet-enabled criminals such as network intrusions and the spread of computer viruses and web-based versions of current crimes, including data stealing, abuse, intimidation and extremism, which have become an significant concern for individuals and countries. Normally, cyber-crime in the common man's language can be defined by the computer and the Internet as a crime to steal a person's identity or sell trafficking or stalk victims or disrupt malevolent program operation [4]. New strategies that are proliferating through the usage of internet operation lead in manipulation, insecurity allowing an acceptable route for the transmission of sensitive data to conduct an offense by criminal behavior. Operations include assaults on the Information Center Computer Network, extortion, child trafficking, electronic purchase fraud, Internet advertising fraud and the introduction of harmful Internet activities such as malware, worms and third-party manipulation such as phishing, spam scams, etc. The universal method of a network such as the Internet at all stages of the network wants to rebound from conducting illegal activity in the world and to avoid malicious activity by preventing illegal activity by enforcing various types of firewalls inside its offline jurisdiction for each country in order to track and deter crimes committed in cyberspace [5]. Network protection mechanisms are used to deter users from breaching networks that contain firewalls, virtual private networks(VPN) and encryption algorithms. Because of these, the VPN performs a crucial function in stopping criminals from breaching the network. As technology plays an important role in a person's life, cyber-crimes will increase as well as technology developments. Cyber-crime covers a wide array of crimes like abuse of identities of individuals, fraud, criminal crimes, prostitution, selling contraband items, illegal downloading records, etc. According to the authors, some crimes involving a machine and that Internet is cyber fraud [6]. The following are some of the common and disturbing cyber crimes.

Financial Crimes

With the Internet market mounting, banking, financial offenses have turned into really unsettling. Economic violations require loans card theft, and electronic robbery of money banking, etc. Payment card scam offenders are often advised to seek details about their accounts from a government agent or from people from financial institutions. The plaintiffs had no formal investigation, and no credit card records for these criminals have been released. Criminals will then strip them of their identities, and social consequences are the most critical ramifications.

Drug Trafficking

Drug traffickers make a major contribution using the new cyber-crime to market drugs software to encrypt mails [6]. They decide when and how to get the trade done, usage primarily by couriers. Because there is no immediate contact between buyer and dealer, these transactions are more convenient to purchase illicit drugs and even other things for frightened people.

Cyber Pornography

Pornographic websites allow pornographic films to download videos and photos, pornography magazines online (photographs, scripts, etc.), all come under this category [7]. The British Home Research Computer Relations Committee report Pornography (Maison des Commons, 1994) states the "Old Horror is machine pornography" (Maison des Commons, 1994: 5). The USA Carnegie Mellon Institute is another one of these institutes and has created a broad variety of tests and gathered sexual data for youngsters and computers.

Cyber Terrorism

Act of Crime conducted in Cyberspace is termed cyberterrorism. Cyberterrorism may be described as the malicious usage of machines, networks and the public Web to disrupt and damage personal objectives. Electronics Terrorism may involve pure broadcasting Bombing Data on the Web what might happen at a certain moment in looks away [4]. Cyber criminals are those individuals that are bullying and coercing a individual, arranging or even controlling a program. They are targeted by computers and networks for financial, cultural or political objectives. Experienced cyber-terrorists that are extremely experienced at hacking may inflict significant disruption to government networks, health documents, and national security services that could leave a society, population or institution at chaos in fear of more attacks [5]. Cyber terrorism is viewed in the same manner as conventional terrorism, which thus involves only attacks that endanger property or life, which can be described as the manipulation of target computers which knowledge, especially through the Web, to trigger actual, real-world harm or serious disruption of infrastructure.

Online Gambling

Online games on thousands of websites running servers worldwide. These websites are among the most frequented relevant Money Launderers locations [8]. Online gamblers were smaller relative to non-Internet gamblers, participated in a higher variety of gaming practices, and they were more inclined to bet on soccer. Both inequalities were far greater than moderate-risk participants at question [9]. Non-Internet gamblers were likely to accomplish on online game consoles, and this gambling practice was undertaken by a slightly larger percentage of problem gamblers.

Cyber Stalking

"Stalking," as set in Oxford Dictionary, means "the stealthy search" Electronics Stalking follows that of an individual or the location of the organization, on the Website. Harassment and stalking are problems faced in real life in several people, particularly women. This may mean sending dangerous or non-dangerous messages via social networking blogs or even e-mails on victims' newsletters [10]. Similar to the David Wall [6], cyber-crime is one of the rising types. In fact, this is a crime in which the victim remains consistently harassed by yet another case, trying to send frequent mails to any individual with objectionable content and notifications of threat.

E-Mailing Spoofing and Phishing Scams

Cyber criminals hack known and unknown e-mails as well. Spoofing e-mails in reality, e-mail from a while another e-mail message seems to have sent the source. Spoofing via e-mail seems to be a very secure strategy. Typical cause of monetary loss. Phishing is called the act which attempts to obtain information that is confidential, such as passwords and credit card information by claiming to be a confident individual in an e-commerce company. Phishing e-mails definitely would include access to the malware sites [11].

Cyber Security

Data security and secrecy would therefore be the highest degree of enforcement standards that are taken care of for a organization. Living in a environment where all the material is literally or digitally processed. Social networking sites provide a secure way for people to connect with their families and friends. Continues to use

Table 1 Cyber security incidents reported to Cyber999 in Malaysia from January–June 2012 and 2013

Incidents	January–June 2012	January–June 2013	% Increase/(decrease)
Fraud	2439	2490	2
Intrusion	2203	1726	(22)
Spam	291	614	111
Malicious code	353	442	25
Cyber harassment	173	233	35
Content related	10	42	320
Denial of services	12	10	(17)
Vulnerability reports	45	11	(76)

Web pages on social media for the intent of collecting sensitive details in the case of domestic consumers. NIt is not just social networking that a single person will take all appropriate security measures during bank transactions.

The information security risks are apparent from January to June 2012, as seen in Table 1 contrast of cyber security incidents published to Cyber999 in Malaysia. Also as violence rises, protective programs are growing. The Silicon Valley Bank report reveals that businesses think Cyber threats pose a significant challenge both to their data and their sustainability, according to a research by U.S. it and Healthcare Administrator nationally.

- Ninety-eight% of companies maintain or improve the information protection capabilities, half of which are increasing online assault resources this year.
- Many businesses plan for when cyber attacks happen, if not.
- Only one-third are entirely assured in their cyber security, and far less assured in their company partners' security measures.

There's going to be new attacks on smartphones focused on Android, but they're not going to be that. The assumption that notebooks work in the same manner as mobile phones implies that the same malware as smartphones is likely to damage them. Macs will tend to increase the number of malware cases, although much fewer than PCs. Windows 8 would allow users to create apps on almost any computer (PCs, tablets and smartphones) running Windows 8, and malicious applications such as those for Android can be developed, which is why these are some of the expected cyber security patterns.

Trends Changing Cyber Security

Some developments that have a significant effect on information protection are described below.

Web Servers

Web apps continue to threaten the collection or spread of malicious code. Cyber criminals are utilizing legitimate, compromised Web servers to distribute their malicious code. However, threats on data-robbery are still a major concern to much in the newspapers. Further emphasis on securing database servers and software apps is necessary. In fact, web servers offer the perfect place to access the data for such cyber criminals. Therefore, must always use a safer browser, particularly during important purchases, to avoid being a target of such crimes.

Cloud Computing and Its Services

Today cloud technologies are increasingly being embraced by both small, medium and big businesses. It means that the earth is slowly rising toward the clouds. The new development raises a big information security problem as typical inspection points are wandering via traffic. In fact, with the number of cloud-based services increasing, protection protocols will also have to be updated to insure that essential details are not compromised with web apps and service providers. Throughout the development of cloud providers' own models, several security issues are already answered. Cloud can be enormous, but it is also important to remember that the cloud grows to improve its safety issues.

APT's and Targeted Attacks

APT is a radically modern method of warehousing for criminal organizations. For decades (most after the initial compromise) network protection technologies like site filtering or IPS were instrumental in detecting targeted attacks. As risks are employed for bolder and more nuanced methods, network safety as well as other security resources needs to be integrated to deter attacks. We will also improve our security mechanisms to deter more threats from coming in future.

Mobile Networks

Today we in every area of the world are able to communicate with others. Security is therefore a very major issue for these cell networks. Firewalls and other protection mechanisms are also porous, because users utilize gadgets like laptops, computers, PCs and so on which include more securities other than those in the applications used. The protection issues of such cell networks will also be addressed. In case of protection issues, certain mobile networks are especially susceptible to such cyber-attacks.

IPv6: Recent Internet Protocol

IPv6 is the current IPv4 standard and the foundation of our networks such as the Internet as a whole. IPv6 is a modern networking protocol. Encryption and credibility-checking is a basic IPv6 feature for all communications and is enabled by all compliant equipment and software. While IPv6 is an effective substitute for external IP addresses, certain essential modifications of the protocol in the protection policy are needed. Upgrading to IPv6 is also easier meaning that risks to information protection can be that as quickly as possible.

Code Encryption

Encryption is the method, in which documents can be stored (or information), hacks can't interpret. In order for those obtaining a hidden key (formally called a decryption key) or password to interpret the data, protection of data converts to certain types, or codes. The letter or text is protected in an encryption scheme using a cryptographic algorithm, rendering it an unreadable chip disk. Using an encryption key, which specifies how to decipher the letter, is usually done. Throughout very early phases, security preserves personal protection and integrity. The usage of cryptography, though, presents potential safety issues. Corrupting is used frequently for encrypting transport data (e.g., networks, e-commerce), mobile phones, wireless microphones, wireless intercoms, etc. Wireless data is shared. Therefore, by encrypting the code, one will know if there is any information leakage.

Role of Social Media in Cyber Security

The White Collar Crime Center (NW3C) white paper reports "Criminal usage of social media" [12], in recent years, social media has been on the rise and the media climate changes. In sites in social media, millions such as Facebook, Instagram and YouTube current users use these services to connect instantly happily with one another. Citizens use social media platforms to connect with each other, and the public sector to market and hire new workers. Statistics on usage of social media as of top-six social websites Shown at Fig. 1. As shown in Fig. 1 Facebook has above 2.2 billion monthly active users. Social media according to the NW3C report [12], the most popular online operation is networking; the Internet PC users operate for 1 h on social networking ranges for more than 12 min, while the Web mobile users invest on social media networks for about 18 min on one hour of activity (Fig. 2).

In an increasingly wired world, as we are connected, businesses need to find innovative approaches to protect personal information. Throughout information protection, social media play an increasingly critical position and can make a huge

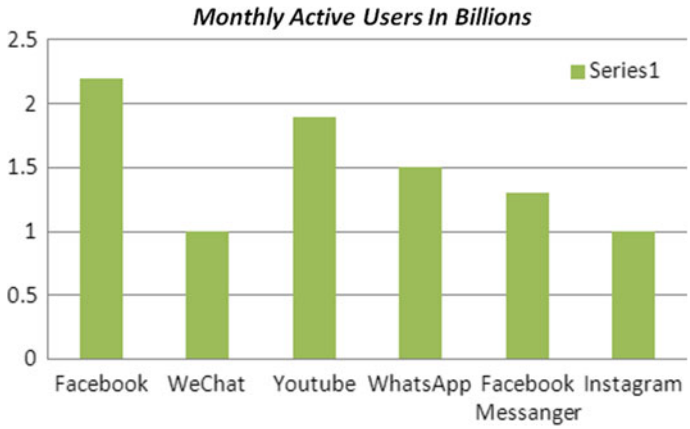


Fig. 1 Top-six: social website MAUs in billions

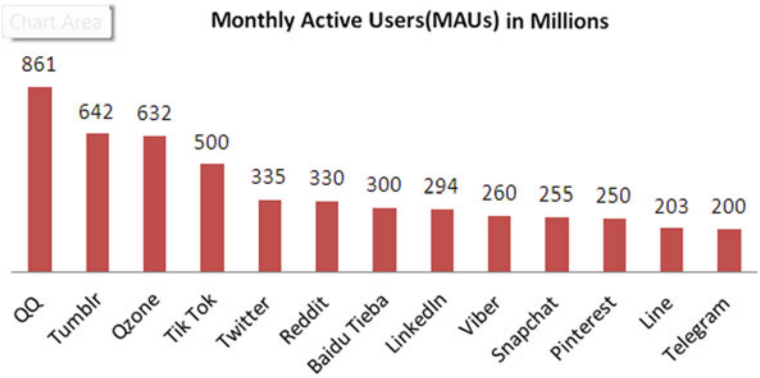


Fig. 2 Social website MAUs in millions

difference to specific cyber attacks. Workers' use to social media is in instant and so there is the chance of an attack. This seems to be an enormous forum for cyber criminals to exploit confidential data and steal sensitive information as most of them utilize social media or social networking platforms on a regular basis. In a environment where we tend to surrender our personal information , organizations have to insure that they are equally successful in identifying threats, reacting in real time and stopping hacking. Since this social forum attracts users quickly, hackers use it as a trap for obtaining the details and the required data. Therefore, people ought to take appropriate steps to avoid loss of knowledge, especially when engaging with social media.

The desire of people to exchange knowledge with millions of viewers is essential to the social networking problem confronting businesses. In addition to providing us the capacity to transmit economically valuable details, the social network often provides the same power to spread misleading and malicious information. The rapid spread of false information via social networking has been among the new risks identified in the 2013 Global Risks Report. While social networking may be used for cyber theft, other businesses cannot afford to avoid using social media, as it plays a major role in the advertising of a business. While social networking may be used for cyber theft, other businesses cannot afford to avoid using social networks, because it plays a major role in the advertising of a business. The firms, however, should recognize this and accept the value of research and have adequate safeguard solutions to avoid risks, especially in social conversations. Many strategies and technology are to be used in the regulation of social media. Six offenses involving social media are addressed in the study by NW3C:

1. Media and phishing Apps
2. Cyber-stalking
3. Cyber-casing
4. Identity theft.

Social Engineering and Phishing

Information engineering takes advantage of social conditioning to get personally identifiable information. People who use social networking sites get updates from their peers calling for urgent financial assistance. Such notes were not directly received by their parents, but by the thief who took the addresses and passwords from their families. The device security is due to its simplicity of existence Trend Micro business names Facebook "scam minefield" [12]. The report by Symantec Corporation describes phishing as one of its several Behavioral Science approaches Cyber attackers employ different techniques to use social media tricks and strategies to get future target knowledge. Phishing e-mails can look like from the manager asking workers to access passwords, or from the individual's bank. Instead of thought rationally, cyber-criminals ensure sure they care about their goals [13]. Criminal usage will send million e-mails anticipating useful details to be gathered. A website such as Facebook or Bank is the most common phishing tool [14].

Cyber-Stalking

Within the Internet realm, bullying using social media or any other electronic medium that can induce sensations of cyber-stalking is frustration, violence and social distress to the victim. NW3C 's report on cyber-stalking [15] further stresses the cyber-stalker's intentional motives by distinguishing it from theft. The study reports that the effect of their behavior on the target is not troubling to detect offenders, though cyber stalkers are well informed and actively doing so. E-mails, instant messages, telephone calls and other contact methods may be used by cyberstalkers to threaten individuals. Cyberstalking may assume the form of sexual assault, unwanted interaction or inappropriate access to family life and activities [15].

Cyber-Casing

National research on the unauthorized usage of the concept in cybercasing in social network white color crime centers is a tool used to create a modern environment. Geotagging is one of the popular features social media platforms have provided over the last years [12]. IoS apps have beyond any legitimate reason played an significant part in supporting this phenomenon. Geographical awareness is the primary component in the cybercasing cycle that let offenders prepare and execute their evil plans [16].

Identity Theft

ID fraud is identified by researchers in [17] as a way to get a suspicious service to identify a individual confidentially. Work perceives identity fraud as the deliberate use, with malicious intent, of victim's personal knowledge, without any lawful authority [18]. According to the FBI 's Internet Crime Complaint Center (IC3) cybercrime survey of 2016, identity crime rated seventh in the United States alone, with 16 878 perpetrators and settlements of USD 58,917,398 [17].

Cyber Security Techniques

In view of the increasing number of illicit attempts to hold private data in order to exploit it in order to intimidate or compel users into pursuing the content, cyber protection is becoming more common.

Authentication

This attempts to test the user identification on the basis of the passwords stored into the network protection domain using this basic information security technique. A most common form of control is password processing, although several other technologies, such as the SIM card in everybody's mobile phone are possible. In order to recognize a particular mobile phone, SIM cards have different identification numbers, which are distributed through a secure contact chain. The greatest challenge confronting the authentication method is foiling the authentication document by unauthorized individuals. You can receive the password transmitted from an untrustworthy means from the wrong people who will use it to disguise it as their initial user. Encryption helps to combat this issue.

Encryption

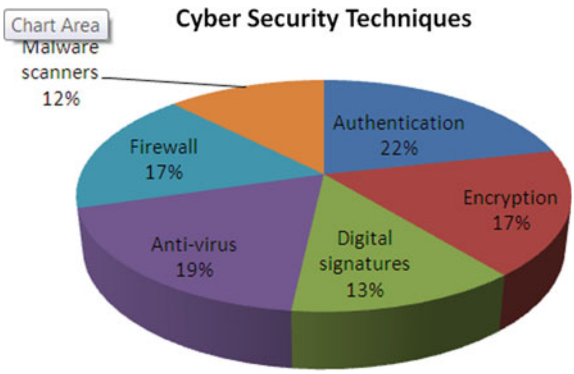
Cryptography renders data undecipherable, by having a special key. To counter encryption, you will attempt to solve complicated mathematical problems including factoring large primes involving computing resources and time. Encryption algorithm is protected by the same key and the degree of security is the same as that of the telephone. The distribution of the key can be exposed to possible security risks. A public key for encrypting the message and a private key are used to send the encrypted with an asymmetric encryption. Most authentication protocols currently use asymmetric public key encryption.

Digital Signatures

It is possible to erect digital signatures from the same mathematical algorithms used in asymmetric encryptions. By encoding some details, a consumer can verify that he possesses a private key. All may have the same public key decryption to test the credentials. It has the same reciprocity as the public key encryption system which often means that the identified user has just a private key (Fig. 3).

Threats of the computer virus or malicious applications triggering inappropriate instructions have reached massive dimensions without clear consumer permission. Two features in anti-virus applications avoid the virus from spreading in a device and tests the already developed virus programs. Most viruses have been developed to target the Windows OS, which is the most common mass desktop workstation.

Fig. 3 Cyber security techniques



Viruses designed specifically for such operating systems can also target Apple and Linux users.

Firewall

A firewall is a network security mechanism that tracks and manages incoming and outgoing network traffic based on default protection regulations. In computation, firewalls are popular for personal and business use, and many devices have one single firewall, like Apple, Windows and Linux. They are commonly recognized as an essential aspect of network defense. Firewall technology may be used to secure networks by dynamically deploying it in a single monitoring screen station where a private or the Intranet links to the public Web and enables traffic control and logging [19]. Firewalls avoid any effort at unauthorized access to a device by hackers on the Internet whether linked directly or by certain network connections. Many operating systems pack firewalls and are turned on by chance. Commercial firewalls will be supported if the default firewall 's protection level is not high enough or if legal network operations are messed with.

Malware Scanners

A firewall is a software or hardware part that allows hackers, malware and worms to be monitored across the Internet. The new firewall receives all messages that reach or depart the Internet, checking every message and rejecting every message that may not satisfy the establish protection requirements. Consequently, firewalls perform an significant part in malware protection. It's a difficult issue to spot malicious malware. To network providers and IT contractor, the huge and rising environment of malicious applications and resources poses an immense challenge. Antivirus program is among the most commonly utilized malware identification and stoppage. Nonetheless, the growing complexity of modern malware ensures for any single provider that signatures with any new attack is becoming more and more complicated [20].

Cyber Ethics

Cyber ethics is a philosophical study of computer-related ethics, which includes user behavior and the programs of computers and how this affects people and society. Cyber ethics is just the Internet code. There is a good opportunity for us to use the Internet in a safer and correct manner when we practice this cyber ethics. Information ethics work in cyber technology is focused on ethical, legal and social problems. Four physical, moral and legal structures are surrounded by computer technology after study effects. The rules, values and legal laws of social networking have been established. Development is responsible for the problems and crimes in cyber technology. New priority was introduced for growth of Internet e-commerce in the legal problems in the information technology sector. Inclusion, shuffling and exchange of information became simpler due to the Internet and emerging technology. Consumer records, data security and privacy should be safeguarded. Cyber ethics mainly deals with the following domains:

- Privacy
- Property
- Security.

Privacy

The manner started to live in the environment of Internet today is affecting our privacy and security. The biggest task at the moment is to increase the amount, duration, variety and quality of data inside social media networks, contributing to the growth of many parties involved, including privacy and safety; on the other hand, it is also a medium for crime reduction and detection as far as smart handling is involved [21]. Facebook is more about advertising; e-mail, instead of becoming a diversion in the way, consists in the very negative impact of breaking the capability of the inbox and the time it is consumed screwing hours. Often spam happens as tracking systems from the other sites check the Internet (IP) addresses of their user, code the hosting service and the name of the customer and then blast the e-mail names that could be connected with the identity that was signed to the account. Spam will also occur. This is why users collect updates in the contents of their previously accessed pages from parent corporations and associates.

Property

The theorist of the Renaissance, Immanuel Kant (1785), developed the concept of the social truth in reaction to the reforms, including the focus on the establishment of the rights of intellectual property in Western nations, possession of personal property and autonomy, and that ethics should always "accept circumstances under which anything that should always not take place" and "Act simply [so your actions] are considered to be a common rule for all moral beings." Categorical imperative. The philosophy of utilitarianism whereby all individuals will do what is best for a vast amount, was introduced by John Stuart Mill and Jeremy Bentham in the following half century. Consequently, cyberethics research explores new internet law standards in virtual space protection [21]. The principles should not be mistaken with the statute, but once rules have been enforced we will govern moral responsibilities. An individual is therefore allowed to participate in fair and equal exchange with others in the media, providing that it does not affect him or other individuals.

Security

Despite the increasing number of cyber-crimes, several social media websites offer their best services. For many nowadays, cyber security always is a major concern. Cyber protection is very relevant for social media applications without information fraud, but the challenge "to ensure 100% cyber security in the actual world" is very challenging [22]. The worldwide state relies mainly on information security, knowledge of how different new technologies and social media can be exploited. We have thousands of cyber-crimes daily, according to current surveys in this real world. Hundreds of millions of exposes to data were recorded in an analysis into security frauds, resulting in about 10 documents per second revealed. Internally, inside organizations, 27% of such attacks were carried out.

Discussion and Conclusion

The Net provides all we want and everything we can. E-mails also replaced e-mails, video conferences are the means for social contact in today's parties and events networking. The world's left. The social network is used to exchange ideas , feelings, habits, connect with friends and build new ones. Thanks to the success of social networking sites, its 2.22 billion UserBase in 2019 is projected to hit $3.02 billion in 2021. Social networks produce an avalanche of big data with their huge populations, or we can call them large social data. This can be seen as a chance for our data driven society, but it has its own constantly developing challenges, like volume, variety, speed, truthfulness, volatility, quality and volunteerism. In all aspects of life, the productive utilization of this enormous quantity of large data is endless. Social media is an ideal place for those seeking a potential buyer of products and services, a vacancy employee, and an excellent target for crime identity kidnapping and online bullying, for starters. Social media is, as with other professionals, a good tool both for law enforcement. In this study, different crimes with strong links to social media were discussed. If friend's social media account is compromised or the victim doesn't know, can easily become a victim of burglary or social engineering. If people of social networking don't respect the protection of their info, post personal details too frequently, and open connections that they use, ransomware and identity fraud are a little like that. In the case of cyber bullying and cybercasing, the details that the consumer of social media publicly communicates ought to be careful. Users can verify the validity of their e-mail or request for personal details in order to be protected from crunch cards and catastrophe scams. If it comes to privacy breaches, the first level of protection is to guarantee the monitoring devices are still enabled.

The present study discussed and analyzed all these types of crimes. The authors plan to design and develop features in the future of a mobile app that raises end users' awareness. This app helps smartphone users to learn about different vulnerabilities, including text, images and videos, of the phone, but also of the smartphone itself [23].

It is also really necessary that all are informed of these crimes and are alert to avoid any damages. In order to guarantee justice and punish the victims, some laws known as cyber laws have been drafted by the judiciary. It is therefore advisable for everyone to know these laws. In addition, cyber-crime cannot merely be referred to as a technical problem. Rather, it's an approach-based problem as it doesn't affect the organizations on the computers. The people who use technology instead are causing the damage. Rather, it's an approach problem because computers don't damage organizations and attack them. The people that exploit the technology are the ones who are causing the damage, instead. It is we who must be vigilant to understand the various approaches such criminals could take. Intellectual thinking is needed to feel such a situation, which can lead to such damage.

References

1. Jain N, Shrivastava V (2014) cyber crime changing everything—an empirical study
2. https://www.kaspersky.co.in/resource-center/definitions/what-is-cyber-security
3. Lyne J, A sophos 04.12v1.dna, eight trends changing network security
4. www.britannica.com/ebchecked/topic/130595/cybercrime
5. Hussain WS, Ibrahim NJ (2019) A survey of cybercrimes, investigations and penal laws imposed on the criminals 8(1):131–143
6. Wall D, Cyber crimes and internet, crime and the internet, ISBN 0-203-164504, ISBN 0-203-164504, p 1
7. Chatterjee BB, Last of the Rainmacs? Thinking about pornography in cyber space, crime and the internet, ISBN 0-203-164504, p 74
8. http://satheeshgnair.blogspot.in/2009/06/selected-case-studies-on-cybercrime.html#email_spoofing_329149830155074
9. Gainsbury S, Russell A, Hing N, Blaszczynski A (2013) The impact of internet gambling on gambling problems: a comparison of moderate-risk and problem internet and non-internet gamblers. Psychol Addict Behav?: J Soc Psychol Addict Behav 27:02
10. Singh G, Singh J (2013) Investigation tools for cybercrime. Int J Comput 4(3):141-154. ISSN 0974-2247
11. en.wikipedia.org/wiki/phishing
12. Nw3c, criminal use of social media (2013) nw3c
13. Symantec, what is social engineering? Symantec Corporation
14. Sauter M (2012) Nine major ways criminals use facebook. Fox business
15. Nw3c, cyberstalking (March 2015) nw3c
16. Saariluoma P, Sacha H (2014) How cyber breeds crime and criminals. In: The society of digital information and wireless communications (SDIWC)
17. Dadkhah MLM, Borchardt G (2018) Identity theft in the academic world leads to junk. Sci Eng Ethics 24(1):287–290
18. Irshad S, Soomro TR (2018) Identity theft and social media. Int J Comput Sci Network Secur 18(1):43–55

19. Abie H (2000) An overview of firewall technologies. Telektronikk 96(3):47–52. http://www.nr.no/~abie
20. Salam S, Maged D, Mahmoud D (2014) Malware Detection in cloud computing. Int J Adv Comput Sci Appl 5(4)
21. Soomro TR, Hussain M (2019) Social media-related cybercrimes and techniques for their prevention. Appl Comput Syst 24(1):9–17
22. Hedayati A (2012) An analysis of identity theft: motives, related frauds, techniques and prevention. J Law Conflict Resolution 4(1)
23. Tran PH, Tran KP, Huong T, Heuchenne C, Tran PH, Le H (2018) Real time data-driven approaches for credit card fraud detection

Network Adaptive Self-configuring Multi-mode LEACH Protocol for Hierarchical WSN

S. Kavitha Rani

Abstract Wireless sensor networks (WSNs), a distributed and infrastructure-free network, have expanded the horizons of the Internet of things (IoT) and machine-to-machine (M2M) communication systems. The rising demands with energy efficiency have been motivation for research on more efficient routing solutions. LEACH being most effective, but with less attention towards dynamics of topology and QOS requirements, has limited its ace. This paper proposes a network adaptive multi-mode LEACH protocol (M2M-LEACH) to meet QoS and energy efficiency requirements of IoT/M2M communication system. It performs a network adaptive grid partitioning, accommodating three different transmission modes—node to gateway (CNG), node to base station (CN-BS) and node to the cluster head (CH) and CH to base station (BS) while considering residual energy and inter-node distance for CH selection. Thus, the strategic implementation of the aforementioned multimode transmission with multi-parameter adaptive CH selection enabled timely and reliable transmission over large-scale dense WSN. The MATLAB-based simulation showed that the proposed M2M-LEACH protocol achieves better energy efficiency, throughput and network lifetime. It has also proven to achieve node alive rate then classical LEACH and its improved variants, resulting its suitability towards QOS and energy-efficient IoT/M2M communication systems.

Keywords WSN · Adaptive routing · LEACH · Cluster head (CH) · Gateway · Multi-mode transmission · Throughput · Cluster nodes (CN)

Introduction

The hierarchical wireless sensor network architecture is shown in Fig. 1. It basically consists of cluster of resource-constraint sensor nodes, which are the basic unit of WSN. These sensor nodes sense data post their deployment and forward it to cluster head, which aggregates all the data received from sensor nodes and transmits

S. Kavitha Rani (✉)
Department of E&C, Global Academy of Technology, Bengaluru, India

V. Bindhu et al. (eds.), *Proceedings of Third International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 844,
https://doi.org/10.1007/978-981-16-8862-1_44

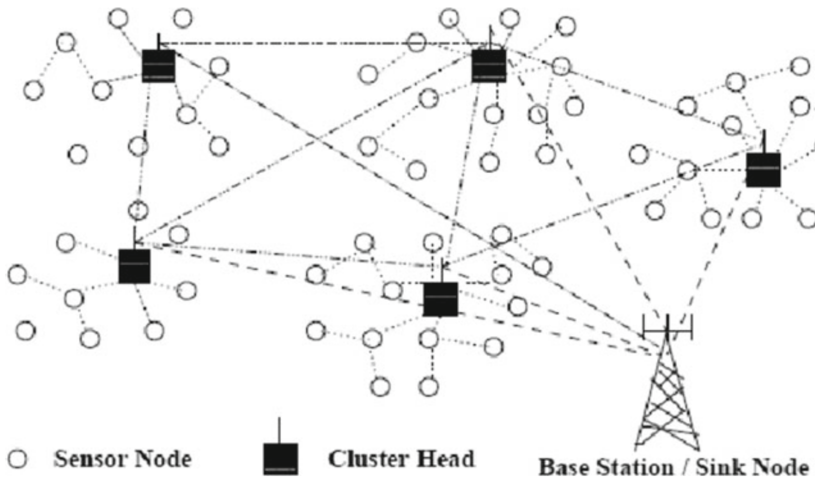


Fig. 1 HWSN architecture

to BaseStation through wireless communication links. BaseStation is a resource-rich node that performs all the required data processing. It acts is as a gateway node between physical world and WSN and has much larger communication range compared to the sensor nodes.

Increase in application demands has resulted in various wireless communication technologies. WSN is been widely used in health care, wartime field surveillance, industries, monitoring-related applications, etc. [1, 2]. WSN, being a distributed and infrastructure less technology, provides potential solution to meet communications needs including IoT and M2M communication system [1–3]. Since WSN is a resource-constrained technology, demands routing solutions along with energy efficiency and high network life time [3–5]. Most of the WSN applications like IoT & M2M system have a thirst for delay resilient transmission with afore-mentioned QoS with energy efficiency assurance [1–4, 7]. With respect to WSN, QoS services require maintaining negligible delay undeniable fault tolerant, higher throughput with the ability to reliable transmission.

Routing in dense heterogeneous WSN has become a challenge with multiple enhanced solutions to meet their above-stated demands [6, 7]. Many researches have shown that cleaving large-scale network into small sub-network regions result in a more reliable, computationally efficient transmission while addressing resource utilization problem. Also, introducing a gateway node within network can reduce multiple hop counts with less dependency on CH-based routing resulting in resource-efficient transmission in WSN. This proposed protocol considers multiple parameters like residual energy and inter-node distance to perform CH selection.

This paper is organized as follows: Section "Related work" consists of related work done. Section "System Model" presents detailed explanation of system model. Section "Simulation Results" explains simulation results followed by conclusion in Section "Conclusion".

Related Work

To address energy efficiency demands, Khediri et al. [8] proposed a new next-generation multi-weight LEACH (MW-LEACH) protocol for WSNs. The authors considered the residual energy and the distance between the inter-CH connected nodes between CH as a parameter for the CH selection. In each round, the authors simply applied residual energy information for the selection of CH (a node with the highest residual energy near the centroid); however, it failed to provide delay resilient system, which is a must for contemporary leading WSN-enabled systems. To meet the demands of energy efficiency, Behera et al. [9] proposed a hierarchical clustering model. The author's adapted threshold-based CH selection method which in addition to CH selection performs dynamic switching between nodes based on their power level to obtain energy efficiency. The highest observed yield was 67% with the last active node retained until the 1750 round. To enhance LEACH, Murugaanandam et al. [10] came up with a fuzzy clustering concept which used multi-criteria-based reliable-CH selection as parameter. To improve the calculation, the authors applied the classic LEACH protocol to perform a unique selection of CH in each pool using the RE-TOPSIS protocol. Undoubtedly, the residual energy use, the distance between the neighbouring nodes, the energy utilization rate, the availability of the neighbour node at a distance of a hop and the node between CH sinks resulted in achieving a reliable communication over the WSN. Ahmed Elsmany et al. [11] also where the authors designed an energy-efficient and scalable routing algorithm (EESRA) to improve network lifespan. The authors suggested that the adoption of a multi-layer hierarchy with random selection of CH may lead to better energy efficiency. The focus was on application of hybrid MAC-based inter-cluster communication for better energy efficiency.

Seyyedabbasi et al. [12] applied multiple measures like number of hops, residual energy, number of links to neighbouring nodes and energy at the neighbouring node to carry out the CH selection. Shen et al. proposed an "energy-efficient centroid-based routing protocol (EECRP)" [1] for IoT systems. EECRP described the self-organizing distributed clustering model with adaptation of the clustering and centroid-based CH rotation model to aid routing solution for energy efficiency. Residual energy information was considered by the authors to estimate the centroid. However, it was considered adequate for the small network. Gupta et al. [13] presented an enhanced remote energy-based LEACH protocol (IDE-LEACH). Traversal expanses (between the source and the sink) along with node energy is been applied for the CH selection. Rahmadhani et al. [14] put forward a model for CH selection based on residual energy and packet loss to analyse the LEACH-based DTN network. To improve transmission reliability over LEACH-based WSN, Alnawafa et al. [15] proposed improved multi-Hop LEACH (IMHT-LEACH) which applied allied task separation with multi-level node. To enhance network life span, Kirsan et al. [16] came up with a new LEACH protocol with BaseStation performing CH selection. In this proposal, the energy level of each cluster was considered as a measure for CH selection. Even after authors claim this proposal to reduce overhead across each normal node, it limits its suitability under dynamic WSN conditions due to efficiency which is merely based on energy clustering. Chen et al. [17] suggested using a cluster model based on multiple CHs for energy-efficient LEACH; however, the authors did not address the resulting overhead and inconclusive transmission scheduling resulted in greedy network conditions. However, the authors failed to optimize the number of clusters that could have allowed for greater energy efficiency. Recently, Chen et al. [18] proposed an improved inter-cluster multi-hop LEACH protocol for communication over WSN. As per this protocol, the aggregated data was partitioned into various levels of transmission in accordance with level of significance. In this proposal, when data is sent by the originating CH node, increasing data level results in increase in radio links, between CH and sink node. Data with higher level ensures reception of packets at sink node using redundant links with multi-hops, and for the data with lower level, low power consumption transmission strategy is adopted to transmit data packets, to achieve the balance between the energy consumption and reliable transmission. Many of the protocols based on traditional clustering are limited in their application across dense hierarchical WSN due to inconsistent cluster size.

In the above-stated literature survey, authors have used single constraints CH selection procedure mainly concentrating on energy efficiency which is the atmost requirement for large WSNs. But the QoS requirement with non-negotiable delay resilience has not been addressed. Also, most of the routing approaches in classical clustering in hierarchical WSNs are found to have limitations over inconsistencies related to cluster size. Overhead computations at CH result in energy depletion causing reduction in network lifetime. This research work is been carried to address these challenges, resulting in energy-efficient, QoS-based routing protocol for dense WSNs using multimode transmission along with multi-constraint CH selection.

System Model

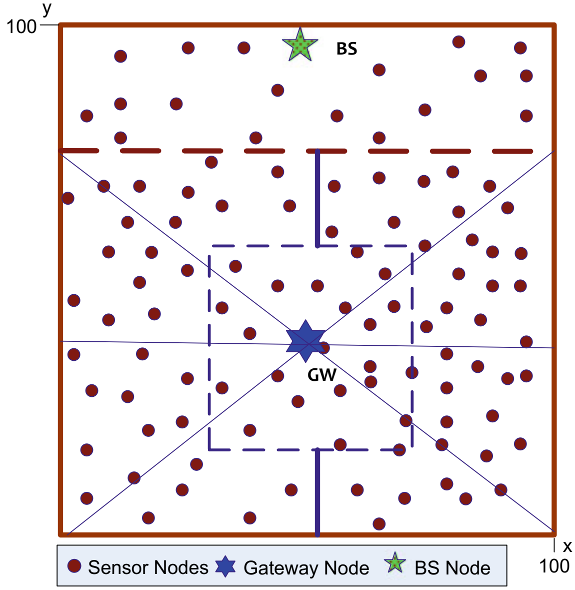
Even after being researched a lot, LEACH protocol has remained inferior to meet up surging demands due to rising complexities over time under dynamic network conditions. This research work emphasis on designing a new LEACH protocol to reserve QoS as well as energy efficiency in hierarchical WSN to meet varying demands of IoT/M2M communication systems. It focuses on improving cluster head (CH) selection along with delay resilient agent GateWay node-based multi-mode transmission to enhance energy efficiency with desired quality of service (QoS).

Proposed Model

Proposed model consists of.

- Dual mode, i.e. directly from cluster node to gateway (CNG/DTG) and Node-to-BaseStation (CN-BS) topology adaptive transmission along with classical node-to-CH and CH-CH-BS transmission as shown in Fig. 2, hence performing three different modes of transmission resulting in reduced CH exhaustion due to beacon overheads, data gathering and multi-hop transmission. This enables fast, resource-efficient transmission with high reliability.
- The concerned network is divided into four sub-regions as shown in Fig. 2, with two dedicated for cluster-based routing, while remaining for DTG and CN-BS. Partitioning is done based on inter-node distance and closeness to BaseStation and gateway node. This reduces multi-hop transmission dependency on CH increasing communication reliability.
- ClusterHead selection is done based on adaptive threshold condition for intra-/inter-cluster communication addition to residual energy while achieving required QoS and reliability. Also, M2M-LEACH protocol suggests to maintain region radio ranges between clusters to overcome redundant transmissions.

Fig. 2 Grid partitioning of network



Mathematical Model

Topology adaptive hierarchical routing M2M-LEACH protocol is developed on native LEACH model. In LEACH, large network is split into small clusters, inside which nodes (CN), CH and BS communicate cooperatively. Node with the highest potential to receive data from all nodes inside cluster and forward it to BS is selected as CH. CHs are elected for every round, so as to reduce energy exhaustion to increase network lifetime. The CH selection is done based on threshold value $T(n)$. Every node in WSN selects an arbitrary value between 1 and 0. Node whose values are less than $T(n)$ is selected as CH for that specific round. In classical LEACH protocol threshold value for CH selection is obtained as shown in Eq. (1)

$$T(n) = \begin{cases} \frac{p}{1-p\left[r \bmod \left(\frac{1}{p}\right)\right]} & n \in G \\ 0 \end{cases} \tag{1}$$

where p states probability of each node being selected as CH, r denotes number of rounds, and G states the set of nodes not selected as CHs in $1\backslash p$ round.

This work assumes heterogeneous network with static nodes along with BaseStation and gateway node being fixed which can be mobile in real time. BS is rich in resource and does not need any supplementary energy during simulation and can communicate with all nodes through broadcast message at particular power level. At nodes, data transmission is done using CSMA-CA MAC realization.

Energy is being consumed during transmission (E_{TX}), reception and during data fusion. The total transmission energy for K bit data over d distance can be obtained as per Eq. (2)

$$E_{TX} = \begin{cases} K * E_{\text{elec}} + K * \varepsilon f_s d^2 & d < d_0 \quad \text{(free space model)} \\ K * E_{\text{elec}} + K * \varepsilon f_{mp} d^4 & d \geq d_0 \quad \text{(multipath decay model)} \end{cases} \tag{2}$$

where E_{elec} represents the power loss due to transmission while $K * E_{\text{elec}}$ depicts total energy consumed by the node during reception of K bit data. f_s and f_{mp} are the energy required for power amplification in free space and multipath decay condition.

The distance threshold is defined as d_0 which is given as Eq. (3)

$$d_0 = \sqrt{\frac{\varepsilon f_s}{\varepsilon f_{mp}}} \tag{3}$$

Implementation of this protocol includes network initialization, nodes deployment (including BS and GW) as shown in Fig. 1 followed by CH selection, generation of cluster, scheduling, data acquisition and transmission. This mechanism repeats as per schedule over each execution round and is shown in Fig. 3.

Multi-parameter-based CH selection CH is elected in each region except one deploying DTG. Nodes with moderately high residual energy are elected as CH

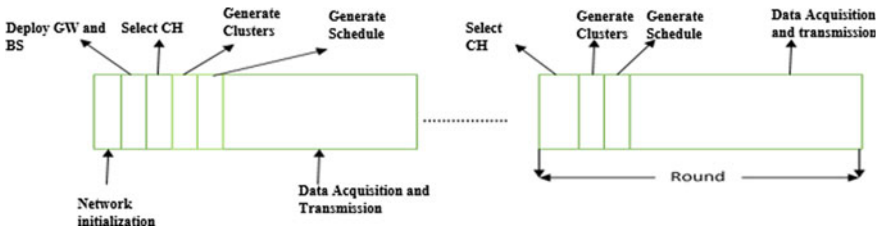


Fig. 3 Sequential implementation diagram

from set of candidate nodes for CH which increases network lifetime. Also, since nodes close to BS can directly communicate with BS, with distance less than d_0, can lower communication overhead. Nodes at distance greater than d_0 perform multi-hop transmissions to reduce energy exhaustion. Ace selection of CH impacts network life and hence is necessity for M2M LEACH-based WSN. Proposed work employs residual energy and distance factor for CH selection. Using these multiple constraint parameters, we have derived cost function for node p within a cluster q is given by Eq. (4)

$$\text{cost}_{p,q}(r) = E_p(r)^2 + \left(1 - E_p(r)^2\right).D_q(r)^2 \tag{4}$$

parameter r states current round(s) count, $E_p(r)$ is the ratio of residual energy of a specific node p to initial energy of CN Eq. (5) and $D_q(r)$ ratio of distance d_0 and distance between node q and BS Eq. (6).

$$E_p(r) = \frac{E_p}{E_0} \tag{5}$$

$$D_q(r) = \frac{d_0}{d_q} \tag{6}$$

The proposed model defines distinct threshold $T'(n)$ based on the above-stated cost function and is given as (7)

$$T'(n) = \begin{cases} \frac{p_{\text{norm}}}{1-p(r\text{mod}(1\backslash p_{\text{norm}}}X\text{cost}_{p.q}(r) & n \in G_1 \\ \frac{p_{\text{adv}}}{1-p(r\text{mod}(1\backslash p_{\text{adv}}}X\text{cost}_{p.q}(r) & n \in G_2 \\ 0 & \text{others} \end{cases} \tag{7}$$

G_1 states set nodes not selected as CH in previous round $\left(\frac{1}{p}\right)$.

G_2 states set of advanced nodes not selected as CH in previous round $\left(\frac{1}{p}\right)$.

p_{norm} states the probability of selection of normal node as CH.

CH broadcasts a multicast message once it is been elected during which CNs keep their radio ON to receive the beacon message and decide which cluster to join. In return, CNs send unicast acknowledgement to CH to join the respective cluster. On receiving acknowledgement from cluster nodes, CH schedules TDMA for data transmission from each node within the cluster. These connected nodes (CN) on schedule of TDMA gather sensed data and forward it to corresponding CH. If no data to transmit means, CN keeps its radio OFF. On reception of data from different CNs, CH performs data fusion and transmits it to BS. Thus, employing the aforementioned sequences explained in Fig. 3, M2M-LEACH protocol achieves QoS-centric, energy-efficient transmission delay resilient over dense WSN for IoT communication systems.

Simulation Results

Topology adaptive routing protocol is been developed on native LEACH protocol with simulation execution using MATLAB 2019b. For simulation, 100 × 100 m dimension consisting of 100 nodes has been considered being distributed randomly with BS placed at outer edge and GW node randomly deployed between the sub regions. For simulation, 3000 rounds are been considered with each data packet of 4000-bit size. Experimental set-up information is been stated in Table 1.

Unlike classical approaches, this proposed work considers multilayer hierarchical architecture consisting multiple clusters with GW node facility which could enable direct data transmission from CN to GW. Also, CNs were designed in autonomous manner to directly transmit to BS based on topologically favoured constraints. Hence, enabling multimode transmission, i.e. DTG, CH-BS,CH-CH-BS and DBS to meet delay resilient transmission with decreased dependency on CH resulting in energy-efficient protocol.

The simulated results are mainly compared with research work carried by Bahera et al. [3] which is similar to this research model. They developed a protocol called I-LEACH on hierarchical network which concentrates on energy efficiency but does not

Table 1 Simulation variables considered for execution

Variable	Description	Value
x_p	Node position or distance at *x*-axis	100 (m)
y_p	Node position or distance at *y*-axis	100 (m)
N	Total number of nodes	100
E_{Tx}, E_{Rx}	Total network energy	0.5 J
ε_{mp}	Energy consumption at the receiver side	0.0013 /pJ/bit/m^4
ε_{fs}	Energy consumption in free space model	10 /pJ/bit/m^2
ε_{amp}	Energy consumption in power amplifier	100 /pJ/bit/m^2
E_{DA}	Energy consumption during power amplification	5 /nJ/bit
d_0	Distance threshold or reference distance	87.7 m

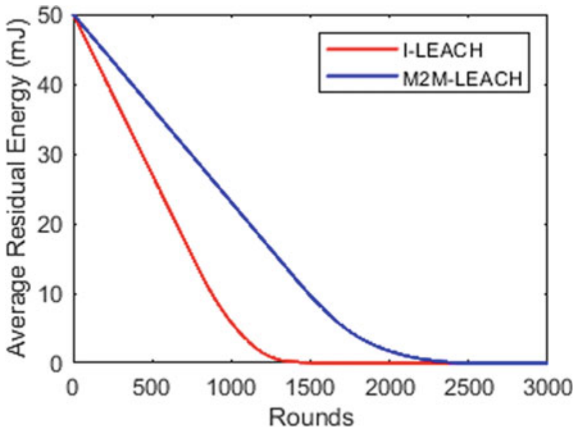


Fig. 4 Residual energy over rounds

consider multimode transmissions like direct to GW or BaseStation. Also, used only single parameter, i.e. residual energy for CH selection. Authors explained switching On or Off the radio of nodes can result in energy efficiency but did not address overburden on CH and latency. M2M-LEACH protocol out performed with respect to performance and reliability required by IOT/M2M systems. Along with I-LEACH, other variants like EECRP, LHC, EESRA, EECS are also been compared with M2M-LEACH in the graphs shown below.

To evaluate the performance, M2M-LEACH protocol is been simulated to retrieve the outputs in terms of residual energy, throughput, network lifetime and on death of first and last nodes.

Implementation of DTG and DBS resulted in reduced dependency on CH-CH transmission. Since the CH dependency is been reduced during transmission, there is reduction in redundant data and energy consumption at every participating nodes during cluster formation. Hence, higher residual energy is been retained at each node over larger time period. This results in higher energy efficiency which is shown in graph Fig. 4 in comparison with I-LEACH. M2M-LEACH protocol has its average energy retained till ~2480 round which is comparatively high with I-LEACH protocol where energy drop can be seen around 1500th round.

Figure 5 shows the throughput performance over rounds, i.e. number of packets delivered to BS by considering recent enhanced LEACH protocols. The M2M-LEACH protocol outperforms other improved variants of LEACH protocols which assures higher packet delivery efficiency. This in turn shows that desired QoS is being achieved. The M2M-LEACH protocol has significantly achieved higher throughput, i.e. 1.2×10^5 packets transmitted to BS while I-LEACH transmitting is 2×10^4 packets.

M2M-LEACH protocol performs continuous CH update and replacement resulting in energy exhaustion during this iterative process. Also, there is energy drop due to data transmission, reception and data fusion. This causes energy depletion

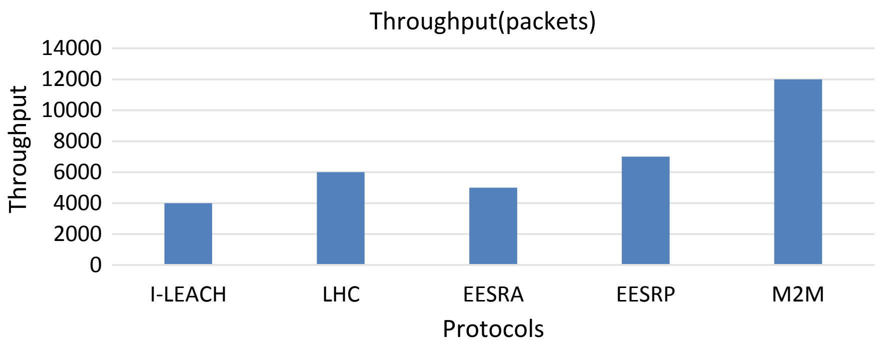
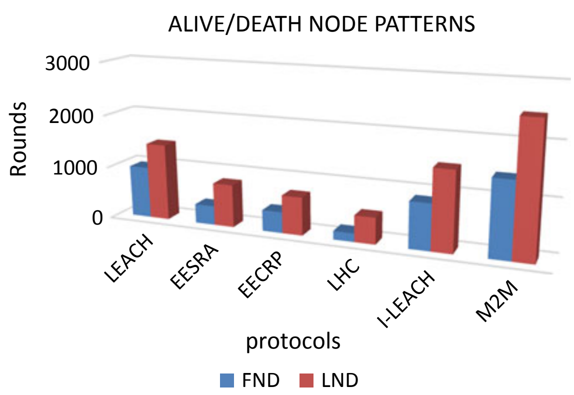


Fig. 5 M2M-LEACH protocol showing highest throughput

across each node and hence results in node death situation. Inappropriate routing decisions or allied scheduling might also result in early node death. Hence, a protocol with lesser death node or higher alive node pattern over increased number of rounds shows energy-efficient routing implementation. Since M2M-LEACH protocol considers resource-rich GW node, resulting reduction in overhead computations across CH, energy is been retained for comparatively longer time period than I-LEACH. Figure 6 represents relative performance in terms of first node death (FND) and last node death (LND) which signifies improvement in overall network lifetime and hence sustainability of M2M-LEACH protocol. FND occurred in M2M-LEACH protocol at 1424 round with LND occurring at 2480 round. Whereas LND in I-LEACH has been observed around 1500 round which is very low compared to proposed protocol.

Figure 7 explains the performance in terms of network lifetime which is dependent on the nodes alive over rounds. M2M-LEACH protocol shows superior performance compared to other existing LEACH protocols. M2M protocol has survived till 2500 round which comparatively high than other protocols.

Fig. 6 Death pattern over rounds for different protocols



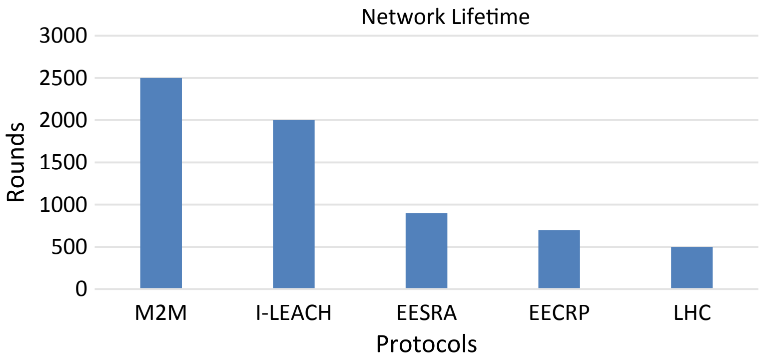


Fig. 7 Different protocols over rounds

Conclusion

Among the major LEACH routing protocols in WSN, M2M-LEACH protocol has performed better; however, many existing protocols have considered CH selection and allied enhancement. Even though CH selection is of main concern, considering only CH selection would result in QoS with non-negotiable delay resilience demand in large WSN unaddressed. The existing routing approaches incorporated CH selection based on either multiple node parameters, heuristic-assisted and path planning or higher CH implanting to meet energy efficiency and throughput. However, computational overhead and latency problems were not addressed. This research work emphasizes on achieving energy-efficient, delay resilient adaptive protocol for dense WSNs. The transmission of data through multimode based on topology information while adapting multi-parameter-based CH selection resulted in superior robust routing protocol with higher residual energy, throughput and network lifetime, guaranteeing reliable transmission with QoS and energy efficiency as per the requirements of IoT/M2M communication systems. Multi-mode transmission has resulted in reduction of latency and also computation over cluster head. M2M-LEACH protocol has achieved nearly 30% higher active nodes over 2500 rounds than other existing improved LEACH variants. Reduction of redundant transmissions between inter-cluster (CHs) is not been considered in this research work. It can be considered as topic for further enhancements.

References

1. Shen J, Wang A, Wang C, Hung PCK, Lai C (2017) An efficient centroid-based routing protocol for energy management in WSN-assisted IoT. IEEE Access 5:18469–18479
2. Huo J, Yang J, Al-Neshmi HMM (2020) Design of layered and heterogeneous network routing algorithm for field observation instruments. IEEE Access 8:135866–135882

3. Bahera TM, Samal UC, Mohapatra SK (2018) Energy-efficient modified LEACH protocol for IoT application. IET Wireless Sens Syst, pp 1–6
4. Othman JB, Yahya B (2010) Energy efficient and QoS based routing protocol for wireless sensor networks. J Parallel Distrib Comput 70(8):849–857
5. Wang XM, Li CB, Lu JL, Liu D (2012) QoSPMR: QoS-aware and priority-based multipath routing algorithm for WMSNs. In: Engineering and technology (S-CET), 27–30 May 2012, pp 1–5
6. Zanella A, Bui N, Castellani A, Vangelista L, Zorzi M (2014) Internet of things for smart cities. IEEE Internet Things J 1(1):22–32
7. Datta SK, Bonnet C (2015) Internet of things and M2M communications as enablers of smart city initiatives. In: 2015 9th International conference on next generation mobile applications, services and technologies, Cambridge, pp 393–398
8. El Khediri S, Khan RU, Nasri N, Kachouri A (2020) MW-LEACH: low energy adaptive clustering hierarchy approach for WSN. IET Wireless Sens Syst 10(3):126–129
9. Behera TM, Samal UC, Mohapatra SK (2018) Energy-efficient modified LEACH protocol for IoT application. IET Wireless Sensor Syst 8(5):223–228
10. Murugaanandam S, Ganapathy V (2019) Reliability-based cluster head selection methodology using fuzzy logic for performance improvement in WSNs. IEEE Access 7:87357–87368
11. Ahmed Elsmany EF, Omar MA, Wan T, Altahir AA (2019) EESRA: energy efficient scalable routing algorithm for wireless sensor networks. IEEE Access 7:96974–96983
12. Seyyedabbasi A, Dogan G, Kiani F (2020) HEEL: a new clustering method to improve wireless sensor network lifetime. IET Wireless Sensor Syst 10(3):130–136
13. Gupta S, Marriwala N (2017) Improved distance energy-based LEACH protocol for cluster head election in wireless sensor networks. In: 2017 4th International conference on signal processing, computing and control (ISPCC), Solan, pp 91–96
14. Rahmadhani MA, Yovita LV, Mayasari R (2018) Energy consumption and packet loss analysis of LEACH routing protocol on WSN Over DTN. In: 2018 4th International conference on wireless and telematics (ICWT), Nusa Dua, pp 1–5
15. Alnawafa E, Marghescu I (2017) IMHT: improved MHT-LEACH protocol for wireless sensor networks. In: 2017 8th International conference on information and communication systems (ICICS), Irbid, pp 246–251
16. Kirsan S, Udin Harun Al Rasyid M, Syarif I (2019)Efficient energy for cluster head selection using new LEACH-based routing protocol in wireless sensor network. In: 2019 International electronics symposium (IES), Surabaya, Indonesia, pp 70–75
17. Chen P, Zhang Y, Dai W (2018)LEACH protocol based on clustering and multi-leader selecting in wireless sensor network. In: 2018 37th Chinese control conference (CCC), Wuhan, pp 7298–7303
18. Chen H, Wang F, Zhang M (2017) An improved inter-cluster multi-hop LEACH protocol oriented reliable transmission in WSNs. In: 2017 3rd IEEE International conference on computer and communications (ICCC), Chengdu, pp 193–197

Smart Anomaly Detection Using Data-Driven Techniques in IoT Edge: A Survey

J. Manokaran and G. Vairavel

Abstract In this article, we survey the different data-driven techniques for smart anomaly detection at the Internet of Things (IoT) edge. Anomaly detection is a significant study issue because of its wide arrangement of use case, from data analytics to network protection, national security, low-cost solutions, and industrial automation. Since IoT sensors are dynamic in nature, they produce huge volume and multi-dimensional data, it is a complex process to detect anomaly. Machine learning-based edge computing can rectify the difficulties in IoT like network traffic, latency, and security, by automated response and shifting calculation physically nearer to the device edge where the information are generated. Our survey discovered several problems in research and conflicts using data-driven-based abnormality detection techniques for limited gadgets in genuine problems of IoT. Based on our findings, researchers may acquaint themselves with current methods, apply them to real-world issues, and grant to the growth of smart anomaly detection techniques.

Keywords Internet of Things · Edge computing · Cloud computing · Fog computing · Machine learning · Deep learning · Anomaly detection

Introduction

The Internet of Things (IoT) is expected to have a universal financial shock of up to $11.1 trillion per year by 2023 [1]. More than 125 billion smart gadget will be linked to the Internet by 2030 [2, 3]. They can collect a huge volume of information that is normally transmitted to the central cloud for analysis, decision making, and implementation. However, these consequences improved network traffic and latencies. This has prompted a new part of the IoT model, known as edge computing

J. Manokaran (✉) · G. Vairavel
SRM Institute of Science and Technology, Chennai, India
e-mail: mj8571@srmist.edu.in

G. Vairavel
e-mail: vairaveg@srmist.edu.in

V. Bindhu et al. (eds.), *Proceedings of Third International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 844,
https://doi.org/10.1007/978-981-16-8862-1_45

which performs calculations at the edge of the device [4]. The benefits of the IoT with edge computing are low latency, location awareness, resilience, scalability, and real-time processing.

We aim to survey the different data-driven techniques like machine learning (ML) and deep learning (DL) for smart anomaly detection at cloud and edge environments in various IoT domains. We discovered that there are a lot of issues and challenges for smart anomaly detection in an IoT edge environment.

The anomaly detection is a vast field with many unsolved problems for several years. When looking at IoT statistics, it is normal to discover abnormal states in a system that's being tracked by sensors installed near the unit. This type of investigation can be used in a variety of situations, including smart city, healthcare, autonomous transportation, credit card fraud detection, and automated commercial processes.

Time-sensitive applications need to detect anomalies in real time, even a small delay can lead to bad outcomes. Example a fault in a device that continuously monitors the heart rate of a cardiac patient could cause a heart attack [5]. It is far preferable to detect such anomalies minutes in advance rather than seconds later or after the occurrence has taken place. For the last decade, the volume of the data is less so the statistical method was used to find anomalies. Due to 5G, the volume and dimensional of data has expanded, so we move toward data-driven techniques.

Data-driven is a synthesis of many scientific disciplines that employ linear algebra, machine learning algorithms, and other techniques to extract models and novel insights from large amount of data. These techniques have been used generally for a lot of function task including pattern recognition, knowledge extraction, online fraud detection, smart transportation, and anomaly detection. In IoT, multiple sensors are used and each sensor produces different forms of output like image, audio, video, and numerical values so the detection of anomalies is a challenging task. Today's IoT has enabled the use of a deluge of electronic data necessitates an automated data analysis technologies. These are provided by machine learning and deep learning, which develops a way for automatic detection patterns in information and then uses the found patterns to predict future information [6, 7].

Erhan et al. [8] have done a multiperspective survey of anomaly detection techniques, particularly for sensors systems. Andrew cook et al. performed a survey mainly focus on IoT time series data [9]. Ashkan Yousefpour et al. done a complete survey mainly focus on Fog computing and future scope in Fog computing [10]. Most of the previous surveys deal with anomaly detection of particular application in the cloud environment, but we are mainly analyzing the various domains of anomaly detection problems in IoT edge. In this paper, survey of different data-driven techniques for smart anomaly detection are described for IoT edge applications, and we have also discussed the key issue and challenges found in the detection of anomalies in IoT Edge.

We hope to present an up to date survey of anomaly detection approaches, with a particular focus on methodologies appropriate for different IoT systems. The background of our study is depicted in Fig. 1, which includes a taxonomy of IoT edge techniques on top and software architecture algorithm on bottom.

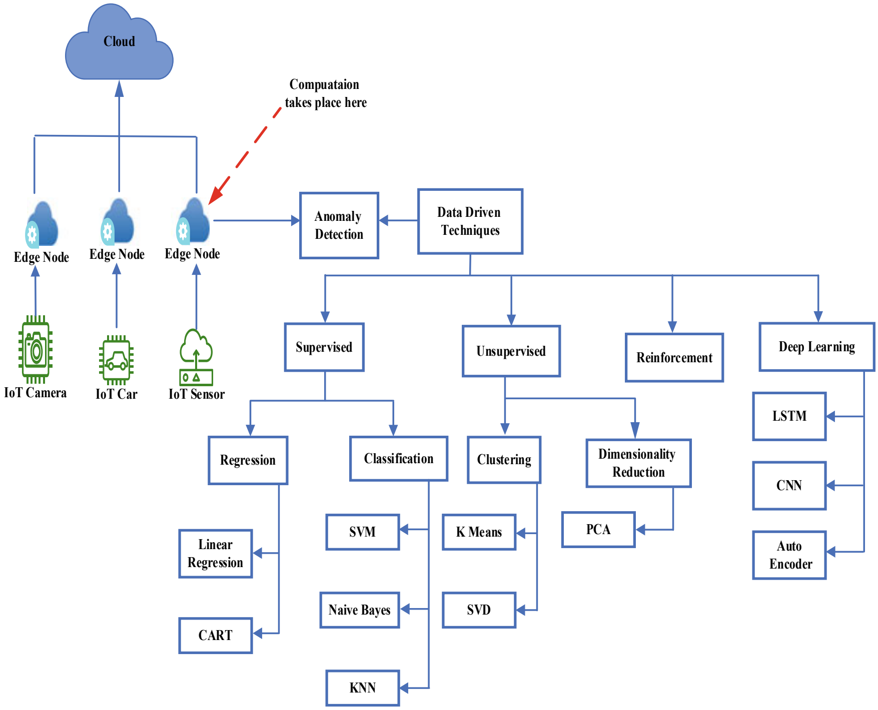


Fig. 1 IoT anomaly detection techniques using different data-driven techniques in edge

Anomalies

This segment explains a theoretical background about outlier detection for the benefit of inexpert persons.

Idea of Anomaly

Anomalies or outliers are a significant deviation from the benchmark. Hawkins defines an outlier as: "an observation which deviates so significantly from other measurements as to arouse suspicion that it was generated by a different mechanism" [11]. Another essential thing to be considered in abnormality is the category of anomaly. There are three types of anomaly detection [12].

Point anomalies: A single data point deviates from the balance data. In Fig. 2, a single data point is difference from rest of the group.

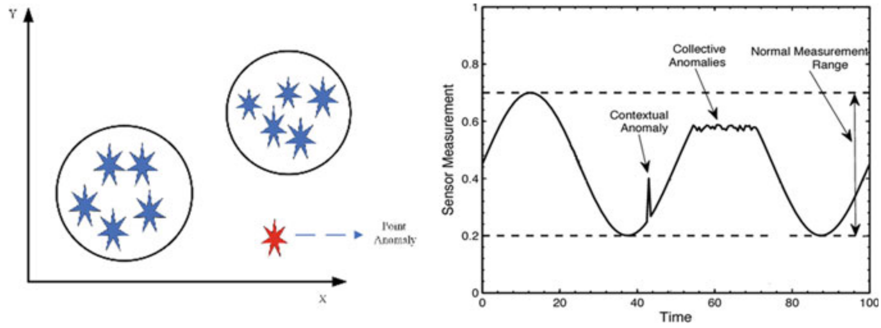


Fig. 2 Point anomaly, collective and contextual anomaly [12]

Contextual anomalies: A contextual anomaly happens when data is atypical in one setting yet not in another. A contextual anomaly is a divergence from the usual when seen in the context of the surrounding observations as shown in Fig. 2.

Collective anomalies: A cluster of closely connected data values deviates from the rest of the datasets as shown in Fig. 2.

Anomaly Detection in IoT

The basic data-driven-based anomaly detection steps in IoT are shown in Fig. 3; the goal of anomaly detection is to not only correctly identify abnormal observations, but also to limit the number of false positives by swiftly modifying the current patterns in the observed data. These abnormalities' detection techniques may be used to prevent resource waste in an industrial environment, to avoid hazardous situations on an avionics network, or to detect suspicious behavior in medical instruments. As a result, the ability to spot deviations can have a huge effect on the overall performance of any system being studied.

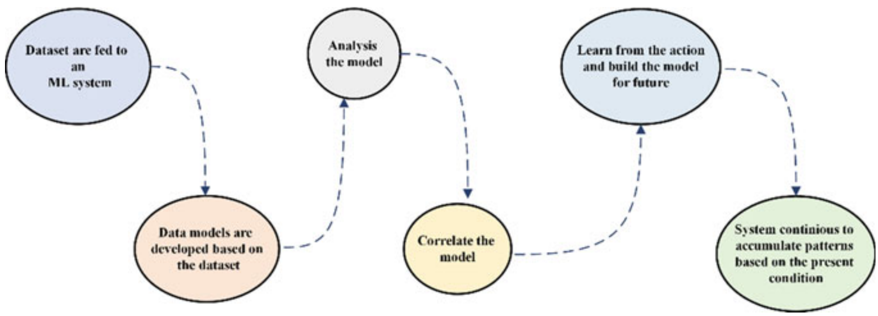


Fig. 3 Machine learning-based anomaly detection steps

In this, real-world probability of normal data is high compared to anomaly data. For example to analyze the heart pulse of humans, the anomaly data is obtained only when their abnormality is present so it is rare to get such anomaly data than normal data. This leads to a data imbalance problem which decreases the efficiency of the system. Most peculiarity identification models are prepared on typical conduct time arrangement designs. Abnormal conduct is portrayed as an event that doesn't observe the guidelines. Analysts have effectively evolved strong techniques for recognizing inconsistencies in chronicled information, ongoing examination, and the expectation of uncommon practices in IoT conditions. Our primary objective is to develop a smart anomaly detection system based on data-driven techniques [12].

Abnormal detection methods are used in variety of IoT smart applications like smart power management, smart monitoring, smart energy, and smart city. All the applications are mostly used sensor as an input device. We need to design and identified a sensor with low power, small size, and environment awareness type.

Types of the Data

Another essential portion to consider is the type of input information. The IoT is a main resource of information streams. A data stream is a collection of information records that are arranged in a logical order and are separated by implicit or explicit time stamps [13]. The data generated in IoT smart device have three nature; first is limited or finite data stream, second unlimited or infinite data stream third adaptive nature data stream i.e., evolve [14]. Anomalies in data streams can be recognized across a wide range of areas, with practical and critical implications. By assuring precision and immediacy, the relevance of identifying irregularities in the data stream grows in real implementations [15].

A variety of approaches have been presented for learning algorithms for abnormality detection in data streams such us IForestASD [16], Evolvable Takagi–Sugeno (T-S)models [17], SAND: Streaming subsequence anomaly detection [18], energy-based Online Sequential Extreme Learning Machine (e-b OSELM) [19], Auto cloud [20], Block DBSCAN [21], Local Outlier Factor (LOF) [22], MOA (Massive Online Analysis) [23], TEDA Clustering [24], Hierarchical Temporal Memory (HTM) [25], Long Short-Term Memory [26], and others. G. Ranganathan performed a real-time anomaly detection using the PySpark framework. Compared to the previous techniques, it worked well in terms of latencies, time consumption, and exactness [27].

Anomaly Detection at the Edge

In centralized infrastructure systems data is stored, business rules are executed, and data analytics tasks are performed distant from users and information sources. They

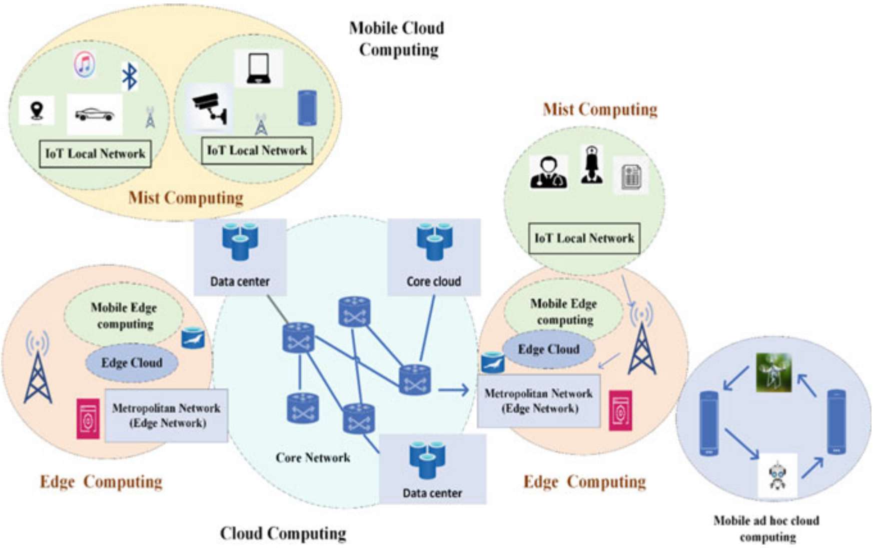


Fig. 4 Comparison of edge computing and its associated computing paradigms [10]

have a lot of processing power, huge storage and more reliable. However, with zettabyte (10^{21}) of trade and an explosion of interconnected gadgets, sending all information to the central cloud for computation is neither practical nor feasible. By pushing manipulation to the network's edges and data generators, edge computing has emerged as a means of addressing these issues [28].

Edge computing is particularly well suited to applications requiring ultra-low latency, as well as content delivery and caching [29]. The comparison of edge computing and its associated computing paradigms is shown in Fig. 4. In starting stage, there are lag software tools and package to support edge computing. So the execution of edge on real time and optimum usage data-driven techniques is so complicated nowadays software platforms like PyTorch, Anaconda, TensorFlow have toolboxes that enable simple programming and better performance anomaly detection processes that are suitable for edge nodes.

D. Sivaganesan, presents a life time improved data-driven trust apparatus based on blockchain as a decentralized and energy efficient approach for detecting gray and black hole attack in IoT-enabled sensor nodes [30].

Data-Driven Techniques for Anomaly Detection

A few literature works are reviewed here for discovering anomalies in various sectors by utilizing various machine learning and deep learning techniques.

Dong Yul Oh et al. have proposed sound anomaly detection in semiconductor assembly equipment by using the auto-encoder algorithm. Noise, data imbalance, cost are the major challenges faced here [31]. Mahmudul et al. have done a performance comparison of several classical ML algorithms for predicting attacks and anomalies in the kaggle dataset. According to this paper, the random forest (RF) algorithm is suitable for detecting cyber-attacks on IoT networks. The major problems faced here are real time, high volume, attack detection, intrusion detection [32].

Yu Liu et al. have done deep learning-based irregularity location techniques at vertical plant divider frameworks to improve automatic technology and awareness to accomplish prescient indoor environment maintenance. The major challenges which are addressed here are noise, sensor fusion, data missing, resource constraints [33]. Tharindu Lakshan Yasarathna et al. propose a network anomaly detection using One Class Support Vector Machine (OCSVM) and auto-encoder algorithms. The neural network-based method is mostly suitable for YAHOO, UNSW-NB15 network dataset. Class imbalance is the major problem here [34].

Pradeep Bedi et al. propose performance comparison of different classical ML algorithms which was done by analyzing the prediction attack and anomalies on open source datasets. Random forest algorithm is appropriate in finding digital assaults on IoT open-source datasets. Real-time, high volume, attack detection, intrusion detection are the major challenges faced here [35]. The measurement intrusion detection system (MIDS) allows us to find any unusual attack in the industrial system. Sohrab Mokhtari et al. find the stealthy attacks in the HIL-based augmented ICS dataset using random forest with a detection accuracy of 99.76%. The important challenges that are noted here are data imbalance, intrusion detection, attack detection, congestion [36].

Selim et al. have done detecting anomalies, malicious activities, and cyber-attacks in a cyber-physical of Industrial Internet of Things (IIoT). In the water system, the dataset Classification and Regression Tree (CART) algorithm has the best ML evaluation metrics. Data volume, real-time detection, sensor fusion, congestion, attack detection are the major challenges faced here [37]. Pierpaolo Dini et al. present ML-based intrusion detection in local area networks. In the classification of anomaly class and normal class, the K-nearest neighbor (KNN) algorithm performs better than artificial neural networks (ANNs). The major challenges which are addressed here are real-time, congestion, attack detection, intrusion detection [38].

The paper by Joseph Schneible depicts an appropriate way to deal with abnormality detection that utilizes auto-encoder (AE) which are particularly deep learning neural networks conveyed on each edge gadget to conduct examination and discern atypical perceptions. The important challenges that are noted here are data volume, real times, fault detection, resource constraints [39]. A power-saving scheme based on anomaly detection using ELM-B in the course of the machine's lifespan in Industry 4.0 are explained by S. K. Bose. Real times, fault detection, resource constraints are the major challenges faced here [40].

Matia Antoni et al. present a planning system for smart sound sensors that can catch and handle crude sound streams. The extricated highlights are shipped off the edge layer, where oddity recognition calculations running as miniature administrations can identify abnormalities progressively by examining the highlights acquired [41].

Zaffar et al. have developed an edge-based smart audio infrequent event finding system using two-stage unsupervised ML algorithm. Noises, data volume, data imbalance, real time, resource constraints, are the major challenges faced here [42]. Mbarka Soualhia et al. propose a method for detecting and predicting multiple defects at the base level of edge clouds. The outcome shows that the system can detect and predict several faults in real time. Real times, fault detection, resource constraints are the major challenges faced here [43].

Ibrahim et al. developed an anomaly detection loT (AD-IoT) system, which is smart anomaly detection based on the random forest ML algorithm, to address the loT cyber security threats in a smart city. To demonstrate the model's accuracy, he used a modern dataset to test it. The major challenges which are addressed here are real time, congestion, attack detection, resource constraints [44]. Juan et al. proposed super-fast video anomaly detection for traffic monitoring based on edge computing and data-driven techniques. Less miss rate and less false positive rate lead to an intelligent monitoring network. Data volume, real time, congestion are the major challenges faced here [45].

Information from a CHRIST osmotron water sanitization gadget in both damaged and standard working conditions. Information mining and artificial intelligence are being utilized to plan information-driven and model-driven edge oddity shortcoming identifiers for modern water purging frameworks. The important challenges that are noted here are real-time, sensor fusion, fault detection, data missing and corruption, resource constraints [46]. Tsukada et al. proposed ONLAD and ONLAD core, an IP center that is exceptionally upgraded for performing a quick consecutive method of figuring out how to follow thought float in 1 ms. ONLAD core permits edge applications to perform low power on gadget learning without requiring information it moves between the edge and the worker. Real-time, resource constraints are the major challenges faced here [47].

Chunde Liu et al. produced a various number of sensor information based on abnormal detection systems in IoT edge. Edge computing is the solution for insufficient computing capabilities. The performance of the proposed EDAD algorithm is better than classification methods in terms of detection accuracy, delay, and energy consumption. Real-time, sensor fusion, data missing and corruption, resource constraints are the major challenges faced here [48]. Suresh K. Peddoju et al. developed a patient monitor with a rapid fast and low-power IoT wearable device for identifying anomalies. The important challenges that are noted here are real-time, resource constraints [49].

A hybrid system for detecting events of interest related to patient fitness and atmosphere, as well as network intrusions was explained by Abdel Mlak said. Two modules work well together in a single, unified system, resulting in ease of administration and cost savings in system management. Attack detection, intrusion detection, resource constraints are the major challenges faced here [50]. Edge-based attack detection system using Extreme Learning Machines (ELM) ML approach in real-world attacks dataset. The major challenges which are addressed here are real-time, congestion, attack detection, intrusion detection, resource constraints [51].

Hang et al. present a stable video anomaly finding experiment based on convolutional neural network. He also shows that the proposed SecureAD can detect video anomalies safely without threaten data privacy. Data volume, real-time detection, congestion, resource constraints are the major challenges faced here [52]. Power grid anomaly detection system at edge environment for smart meter dataset was done author used DL techniques to avoid nonlinear characteristics of the time series data. He obtained classification accuracy improved to 92%. Each sample had a 935 microsecond inference time, which is sufficient for edge-based anomaly detection. The important challenges that are noted here are data volume, data imbalance, real-time detection, data missing and corruption, resource constraints [53] (Table 1).

Issues and Challenges for Anomaly Detection in IoT Edge

- *Noise.* Noise must be removed from edge gadgets in IoT before sending it to the cloud. If it's not eliminated, the error will be redundant which may increase power, cost, and latency problem. The main issue is that the amount of noise in different datasets can vary dramatically, and noisy instances can be distributed irregularly throughout the data space [31, 33, 42].
- *Data volume.* This problem is partially solved by edge computing which minimize the quantity of data sent to the cloud center; however, long-term and continuous data collection from various sensors need an advanced learning algorithm to perform the system in a better way [32, 35, 37, 39, 42, 45].
- *Data imbalance.* Because anomalies are rare in most cases, an imbalance in the anomalous to normal data ratio can result in a high false positive rate, lowering the accurateness of a prediction model. A high false positive rate refers to the percentage of non-anomalous events or behavior that is incorrectly classified as anomalous, resulting in a false alarm [31, 34, 36, 42].
- *Real-time detection.* The system needs a faster response because it involves high-speed streaming data [32, 33, 35, 37–40, 42–49, 51–53].
- *Sensor fusion.* It is extremely hard to gather information from various sensors and afterward cumulated for the outcomes. For example, several qualities like heat, CO_2 level, wind speed can be recorded from independent sensors and incorporated for displaying on a worker for natural effect examination [37, 46, 48].
- *Congestion.* Anomaly detection is very much complex due to a large amount of data. Unnecessary high throughput, like numerous access at the time of result announcement, can cause an over-burden on the anomaly detection engine [36–38, 44, 45].
- *Attack detection.* Anomaly detection is a significant worry due to various attacks like denial of service, data type probing, malicious control, malicious operation, scan, spying and wrong setup, etc. [32, 35–38, 44].

Table 1 Summary of ML/DL algorithm in anomaly detection

References	Computing environment	Data-driven techniques	To detect	Major finding
[31]	Cloud	Auto-encoder	To propose a method for detecting anomaly functional sounds in a very complex machine along with minimizing the cost of data-driven annotation	The proposed method worked well in identifying anomalies, and it can be used in more general circumstances which help to solve the issue of data imbalance
[32]	Cloud	Logistic regression (LR), support vector machine (SVM), decision tree (DT), RF, ANN	To discover attacks and anomalies in DS2OS traffic lines	With a prediction test accuracy of 99.4%, the random forest algorithm performs better
[33]	Cloud	Auto-encoder, long short-term memory encoder–decoder	To detect point and contextual anomalies in IoT-based vertical plant walls for indoor climate control	The possibility of using artificial intelligence-based anomaly detection techniques on vertical plant wall frameworks to improve automation and improve awareness to recognize prescient indoor environment help
[34]	Cloud	Auto-encoder, OCSVM	To concentrate on data-driven techniques to analyze cloud network data in order to detect anomalies (YAHOO data)	Auto-encoder performs better than kernel-based methods in detecting anomalies in cloud meshwork data
[35]	Cloud	ANN, LR, RF, SVM, DT	To detect assault and peculiar location in IoT framework	Random forest algorithm performs comparatively better with the prediction test accuracy of 99.4%

(continued)

Table 1 (continued)

References	Computing environment	Data-driven techniques	To detect	Major finding
[36]	Cloud	RF, KNN, and DT	To create a measurement intrusion detection system (MIDS) that allows the system to find any unusual behavior in the system, even if the intruder uses cracks in the control layer to hide it	Random forest algorithm does relatively better with the fault detection accuracy of 99.76%
[37]	Cloud	LR, linear discriminant analysis, SVM. KNN, CART, Naive Bayes	To analyze machine learning-based attack detection in a cyber-physical of critical water infrastructure in the IIoT	Classification and regression tree algorithm performs comparatively better with an anomaly detection accuracy of 94%
[38]	Cloud	KNN, ANN	Ingress detection systems for US Air Force LAN traffic data	In the classification of anomaly class and normal class, KNN performs better than ANN
[39]	Edge	Auto-encoder	An advanced distributed anomaly detection system	Fewer amounts of data are transmitted to the cloud which reduces the complexity of cloud
[40]	Edge	Extreme learning machine-boundary (ELM-B)method	A power-saving scheme based on anomaly detection Using ELM-B in the course of the machine's lifespan in Industry 4.0	Low-energy machine health monitoring system with less convergence time
[41]	Edge	Isolating forest, elliptic envelope	To become aware of anomalous sound occasions occurring in an industry surroundings using smart audio sensors	In terms of average computing time interval and user center processing unit load on the AGILE instance, elliptic envelop, the algorithm works better than the isolation forest algorithm

(continued)

Table 1 (continued)

References	Computing environment	Data-driven techniques	To detect	Major finding
[42]	Edge	IRESE unsupervised learning	Intelligent audio rare-event detection (gunshot, glass break, scream, and siren)	To detect rare events this system produce above 90% precision and recall measures
[43]	Edge	SVM, LSTM, CNN, RF, neural network	Using time-driven and data-driven methods to propose a model for detecting and predicting multiple faults at the system level of edge clouds	Error propagation in a cloud is avoided by finding real-time faults using data-driven techniques
[44]	Edge	Random forest	To propose an Anomaly Detection loT (AD-IoT) system (Smart city)	AD-loT can adequately accomplish the most noteworthy characterization precision of 99.34% with the least false positive rate
[45]	Edge	CNN, SVM	To discover the abnormal in an smart monitoring network of a super highway	Less miss rate and Less false positive rate lead to an intelligent monitoring network
[46]	Edge	SVM	Construct an anomaly detection system for IIoT	The proposed method worked well for identifying anomalies, in water industry with the help of data-driven methods and smart sensors

(continued)

- *Intrusion detection*. To detect any type of misuse that falls out of normal system operation [32, 35, 36, 38].
- *Cloud computing*. This methodology has a few downsides, including idleness, privacy, and reliance on network availability between information makers and consumers. To defeat these requirements, calculation ought to be moved as close as

Table 1 (continued)

References	Computing environment	Data-driven techniques	To detect	Major finding
[47]	Edge	Back propagation neural networks	To propose data-driven-based anomaly detector which learns on board	To develop a highly optimized less cost, minimum power board ONLAD core by using the data-driven methods. Here the data transfers between edge and cloud are not needed
[48]	Edge	Edge computing Data Anomaly Detection (EDAD)	To perform an anomaly detection for underground mining sensors	EDAD performs well in terms of accuracy, delay, and energy consumption
[49]	Edge	Models based on a supervised and unsupervised algorithm	Using an offline machine learning model and anomaly detection method to create a care observer system with less-power smart devices	The proposed framework will decrease the communication dormancy and give quick activities. The significant advantages of the proposed framework are to recognize wellbeing hazard designs and disperse crisis warnings to guardians
[50]	Edge	Support vector machine	To develop an efficient anomaly detection system for smart hospital	The IDC and EDC modules work well together in a single, unified system, resulting in ease of administration and cost savings in system management

(continued)

Table 1 (continued)

References	Computing environment	Data-driven techniques	To detect	Major finding
[51]	Edge	Extreme learning machine	Edge-based attack detection system using a scalable distributed machine learning approach (real-world attacks dataset)	The usage of edge computing is that the traffic problem of cloud can be eliminated
[52]	Edge	CNN	Edge-based unattackable video abnormality recognition structure dependent on convolutional neural network	To identify video inconsistencies while looking after security. Moreover, the hypothetical exploration and commonsense reproductions uncovered that secure anomaly detection could be done productively with insignificant calculation error
[53]	Edge	Long short-term memory	Power grid anomaly detection system at edge environment (smart meter dataset)	Classification accuracy improved to 92%. Each sample had a 935 microsecond inference time, which is sufficient for edge-based anomaly detection

conceivable to the IoT edge, ideally on entryways as opposed to straightforwardly on data makers [31–38].

- *Fault detection.* At clouds, infrastructure failures can collect and propagate, causing severe system and function performance degradation. As a result, it's critical to identify and correct these faults as soon as possible, e.g., banking [39, 40, 43, 46].
- *Data missing & corruption.* Data loss is caused by the external climate, which is difficult to detect. Information may be ruined by external components or gadget failures, making it difficult to detect anomalies and corrupted data [33, 46, 48, 53].
- *Resource constraints of IoT devices.* In terms of power, cost, and size, IoT devices are considered to be constrained devices [33, 39, 40, 42–44, 46–53].
- *High-dimensional data.* Interacting with the high-dimensional data generated by IoT devices is becoming more challenging. The number of attributes required for effective data analysis in a given application determines this. Techniques for

reducing dimensionality play an important role in dealing with and minimizing features more precisely [32, 35, 37, 42, 45].

Research Gap

In the above section, we discussed key issues and challenges present in anomaly detection. From the survey, we have found a lot of research areas for future researchers they are as follows,

- There is a major gap in terms of formalizing a method for accessing data logs and sensory data sources, as well as constructing and validating models in real-world settings.
- Most of the recent researches have concentrated on anomaly detection. The area of anomaly prediction and prevention has yet to be fully explored. It can be extremely useful in terms of predicting anomalies. New methods for proactively avoiding system failures and conducting root cause analysis must be developed and adopted.
- We were unable to locate any work that used fusion techniques during our search. These techniques would lay a strong basis for fusing sensory data streams and aiding in the investigation of suspicious activities.
- Till now, real-time analysis of data is not explored much for malware analysis.
- In the domains of intelligent transportation, smart city, and smart grid, there is a need to apply new prototypes and determine their ability to detect anomalies.
- The majority of ML algorithms used for anomaly detection are single, but we need advanced algorithms and hybrid techniques to extract accurate results.
- *Blockchain:* Blockchain has number of benefits include delegation, persistency, anonymity, and auditability, to name a few. Anti-money laundering tracking system, real-time IoT operating systems, NFT marketplaces, and public and social services are just a few of the themes covered by blockchain applications. It is an interesting approach to resolve the edge computing security problems. Integrating blockchain concept to the edge anomaly detection reach the objective of reducing the authentication time, this led to ultra-low latency.
- *Federated Learning (FL):* Federated learning is a cooperatively autonomous privacy preserving tools to overcome problem of statistical heterogeneity and data receptivity. Heterogeneous ML can be done on the network edge using information collected by the smart device.

Conclusion

Anomaly detection has earned a lot of interest from researchers in modern days, due to the improvement of minimum cost, high-impact sensor technologies in different application fields. This paper has described several data-driven techniques for anomaly detection in various IoT fields like smart cities, industry, and smart hospitals. Researchers working in the domain of IoT edge anomaly detection could find what are the key issues and challenges present in the area using our review paper. The IoT comprises countless gadgets that produce a lot of data, which require more calculation. To improve calculation speed, feature extraction is required. Finally, we conclude that determining which ML technique is ideally suited for all IoT data is difficult. The ML technique for anomaly detection can be chosen based on the application and nature of the information.

Apart from that the survey has significant flaws that should be addressed in future research. We did not discuss about preprocessing techniques and windowing techniques which are used in various anomaly detection methods.

References

1. Manyika J, Chui M, Bisson P, Woetzel J, Dobbs R, Bughin J, Aharon D (2015) The internet of things. McKinsey Global Institute, Technical Report, 06
2. Evans D (2001) The internet of things: how the next evolution of the internet is changing everything. CISCO White Paper 1:1–11
3. Howell J (2018) Number of connected IoT Device Surge to 125 billion bt 2030
4. Merenda M, Porcaro C, Iero D (2020) Edge Machine learning for AI-enabled IoT devices: a review. Sensors 20:2533
5. Chaunhan S, Vig L, Ahmad S (2019) ECG anomaly class identification using LSTM and error profile modeling. Comput Bio Medical 109:14–21
6. Murphy KP (2012) Machine learning: a probabilistic perspective. MIT Press
7. Sharma B, Sharma L, Lal C (2019) Anomaly detection techniques using deep learning in IoT: a survey. In: International conference on computational ıntelligence and knowledge economy (ICCIKE), Dubai, United Arab Emirates, pp 146–149
8. Erhan L, Ndubuaku M, Mauro D, Song W, Chen W, Fortino G, Bagdasar O, Liotta A (2021) Smart anomaly detection in sensor systems: a multi-perspective review. Inf Fusion 67:64–79
9. Cook AA, Mısirli G, Fan Z (2020) Anomaly detection for IoT time-series data: a survey. IEEE Internet Things J 7(7):6481–6494
10. Yousefpour A (2019) All one needs to know about fog computing and related edge computing paradigms: a complete survey. J Syst Arch pp 289–330
11. Hawkins DM (1980) Identification of outliers, vol. 11. Springer, Berlin
12. Chandola V, Banerjee A, Kumar V (2009) Anomaly detection: a survey. ACM Comput Survey 41(3):1–58
13. Yu K, Shi W, Santoro N (2020) Designing a streaming algorithm for outlier detection in data mining an incrementa approach. Sensors 20(5):1261
14. Salehi M, Rashidi L (2018) A survey on anomaly detection in evolving data: with application to forest fire risk prediction. SIGKDD Explor Newsl pp 13–23
15. Gul H, Javaid N, Ullah I, Qamar AM, Afzal MK, Joshi GP (2020) Detection of non-technical losses using SOST link and bidirectional gated recurrent unit to secure smart meters. Appl Sci 10(9):3151
16. Togbe M, Barry M, Boly A, Chabchoub Y, Chiky R (2020) Anomaly detection for data streams based on ısolation forest using scikit-multiflow. In: The 20th International conference on computational science and its applications, Caligari, Italy, July (2020)
17. Severiano CA, de Lima e Silva PC, Cohen MW, Guimaraes FG (2021) Evolving fuzzy time series for spatio-temporal forecasting in renewable energy systems. Renew Energy 171:764–783

18. Boniol P, Paparrizos J, Palpanas T, Franklin MJ (2021) SAND streaming subsequence anomaly detection
19. Wang X, Tu S, Zhao W, Shi C (2021) A novel energy-based online sequential extreme learning machine to detect anomalies over real-time data streams. Neural Comput Appl, pp 1–9
20. Silva M, Signoretti G, Andrade P, Silva I, Ferrari P (2021) Towards a customized vehicular maintenance based on 2-layers data-stream application.In: IEEE International workshop on metrology for automotive. IEEE, New York, pp 193–198
21. Chen Y, Zhou L, Bouguila N, Wang C, Chen Y, Du J (2021) BLOCK-DBSCAN, Fast clustering for large scale data. Pattern Recogn 109
22. Wang G, Chen Y (2021) Robust feature matching using guided local outlier factor. Pattern Recogn 117
23. Madhuri GS (2021) Anomaly detection in data streams using MOA. Turkish J Comput Math Educ (TURCOMAT) 12(6)
24. Toma GA (2021) Performance analysis of state of art clustering approaches in data mining. Int J Inf Technol (IJIT) 7(3)
25. Li K, Shan S, Zhang Q, Cai X, Zhijun Z (2021) A computational method to solve for the heat conduction temperature field based on data-driven approach. Thermal Sci, pp 165–165
26. Nguyen HD, Tran KP, Thomassey S, Hamad M (2021) Forecasting and anomaly detection approaches using LSTM and LSTM autoencoder techniques with the applications in supply chain management. Int J Inf Manage 57
27. Ranganathan G (2020) Real time anomaly detection techniques using PYSPARK frame work. J Artif Intell 2(01):20–30
28. Ghosh AM, Grolinger K (2021) Edge cloud computing for ınternet of things data analytics: embedding ıntelligence in the edge with deep learning. IEEE Trans Ind Inf 17(3):2191–2200
29. Abbas N, Zhang A, Taherkordi Y, Skeie T (2018) Mobile edge computing: a survey. IEEE Internet Things J 5(1):450–465
30. Sivaganesan D (2021) A data driven trust mechanism based on blockchain in IoT sensor networks for detection and mitigation of attacks. J Trends Comput Sci Smart Technol (TCSST) 3(01):59–69
31. Oh DY, Yun ID (2018) Residual error based anomaly detection using auto-encoder in SMD machine sound. Sensors 18:1308
32. Mahmudul Hasan Md, Islam M, Ishrak Islam Zarif Md, Hashem MMA (2019) Attack and anomaly detection in IoT sensors in IoT sites using machine learning approaches. Internet Things 7:100059
33. Liu Y, Pang Z, Karlsson M, Gong S (2020) Anomaly detection based on machine learning in IoT-based vertical plant wall for indoor climate control. Build Environ 183:107212
34. Yasarathna TL, Munasinghe L (2020) Anomaly detection in cloud network data. In: International research conference on smart computing and systems engineering (SCSE), pp 62–67,Colombo, Sri Lanka
35. Bedi P, Mewada S, Vatti RA, Singh C, Dhindsa KS (2021) Detection of attacks in IoT sensors networks using machine learning algorithm. Microprocessors Microsyst 82:103814
36. Mokhtari S, Abbaspour A, Yen KK, Sargolzaei A (2021) A machine learning approach for anomaly detection in industrial control systems based on measurement data. Electronics 10:407
37. Elim GEI, Hemdan EED, Shehata AM et al (2021) Anomaly events classification and detection system in critical industrial internet of things infrastructure using machine learning algorithms. Multimedia Tools Appl 80:12619–12640
38. Dini P, Saponara S (2021) Analysis, design, and comparison of machine-learning techniques for networking ıntrusion detection. Designs 5
39. Schneible J, Lu A (2017) Anomaly detection on the edge. In: IEEE Military communication conference. IEEE, New York, pp 678–682
40. Bose SK, Kar B, Roy M, Gopalakrishnan PK, Basu A (2019) ADEPOS: anomaly detection based power saving for predictive maintenance using edge computing. In: 24th Asia and South Pacific design automation conference

41. Antonini M, Vecchio M, Antonelli F, Ducange P, Perera C (2018) Smart audio sensors in the ınternet of things edge for anomaly detection. IEEE Access 6:67594–67610
42. Janjua ZH, Vecchio M, Antonini M, Antonelli F (2019) IRESE: an intelligent rare-event detection system using unsupervised learning on the IoT edge. Eng Appl Artif Intell 84:41–50
43. Soualhia M, Fu C, Khomh F (2019) Infrastructure fault detection and prediction in edge cloud environments. In: 4th ACM/IEEE Symposium on edge computing association for computing machinery, pp 222–235. New York, NY, USA
44. Alrashdi I, Alqazzaz A, Aloufi E, Alharthi R, Zohdy M, Ming H (2019) AD-IoT: anomaly detection of IoT cyber-attacks in smart city using machine learning. In: IEEE 9th Annual computing and communication workshop and conference (CCWC), pp 0305–0310, Las Vegas, NV, USA
45. Wang J, Wang M, Liu Q et al (2020) Deep anomaly detection in expressway based on edge computing and deep learning. J Ambient Intell Hum Comput
46. Garmaroodi MSS, Farivar F, Haghighi MS, Shoorehdeli MA, Jolfaei A (2020) Detection of anomalies in ındustrial IoT systems by data mining: study of CHRIST osmotron water purification system. IEEE Internet Things J 3034311
47. Tsukada M, Kondo M, Matsutani H (2020) A neural network-based on-device learning anomaly detector for edge devices. IEEE Trans Comput 69(07):1027–1044
48. Liu C, Su X, Li C (2021) Edge computing for data anomaly detection of multi-sensors in underground mining. Electronics 10:302
49. Peddoju SK, Upadhyay H, Bhansali S (2019) Health monitoring with low power IoT devices using anomaly detection algorithm. In: Fourth ınternational conference on fog and mobile edge computing (FMEC), Rome, Italy
50. Said AM, Yahyaoui A, Abdellatif T (2021) Efficient anomaly detection for smart hospital IoT systems. Sensors 21:1026
51. Kozik R, Choras M, Ficco M, Palmieri F (2018) A scalable distributed machine learning approach for attack detection in edge computing environments. J Parallel Distrib Comput 119:18–26
52. Cheng H, Liu X, Wang H, Fang Y, Wang M, Zhao X (2020) SecureAD: a secure video anomaly detection framework on convolutional neural network in edge computing environment. IEEE Trans Cloud Comput
53. Utomo D, Hsiung P (2019) Anomaly detection at the IoT edge using deep learning. In: IEEE International conference on consumer electronics, pp 1–2. Taiwan (ICCE-TW)

Breast Cancer Detection and Classification: A Comparative Analysis Using Machine Learning Algorithms

Shadman Sakib, Nowrin Yasmin, Abyaz Kader Tanzeem, Fatema Shorna, Khan Md. Hasib, and Saadia Binte Alam

Abstract The risk of fatality from breast cancer is increasing exponentially as the population rises. It is the most typical kind of cancer and the major cause of death in women throughout the world. Early detection and treatment of breast cancer could significantly improve the prognosis and survival rate. It is an important phase in rehabilitation and medication since it can help patients receive prompt medical services. An automated disease detection technique that employs machine learning (ML) and deep learning (DL) techniques assist medical professionals in the diagnosis of diseases and provide a reliable, efficient, and faster response thereby minimizing the chance of death. This research aims to perform a comparison among ML and DL methods for breast cancer detection and diagnosis. The five most popular supervised ML techniques named support vector machine (SVM), decision tree (DT), logistic regression (LR), random forest (RF), K-nearest neighbor (KNN), and a DL technique were used for classification using cross-validation technique. The Breast Cancer Wisconsin (diagnostic) dataset has been used as a training set to evaluate and compare each algorithm's effectiveness and efficiency through classification accuracy, recall, specificity, precision, false-negative rate (FNR), false-positive rate (FPR), F1-score, and Matthews correlation coefficient (MCC). Experimental results show that random

S. Sakib (✉)
Department of CSE, Leading University, Sylhet 3110, Bangladesh

N. Yasmin
Samsung R&D Institute, Dhaka 1205, Bangladesh

A. K. Tanzeem
Department of EEE, BRAC University, Dhaka 1212, Bangladesh

F. Shorna
Department of Pharmacy, BRAC University, Dhaka 1212, Bangladesh

K. Md. Hasib
Department of CSE, Ahsanullah University of Science and Technology, Dhaka 1208, Bangladesh

S. B. Alam
Department of EEE, International University of Business Agriculture and Technology, Dhaka 1230, Bangladesh

V. Bindhu et al. (eds.), *Proceedings of Third International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 844,
https://doi.org/10.1007/978-981-16-8862-1_46

forest (tuned) outperformed all the other models with accuracy and $F1$-score of 96.66% and 0.963, respectively.

Keywords Breast cancer classification · Machine learning · Early diagnosis · Disease prediction · Medical data mining

Introduction

Among women, breast cancer is known to be the second-most significant reason of death worldwide [1]. According to data of 2019, around 268,600 invasive breast cancer cases and 62,930 non-invasive breast cancer cases were said to be found among women in the USA [2]. Early stage testing and detection are found to be the most appropriate approach to enhance the probability of quick treatment and reduce the mortality rate among women. Mammogram is widely used as an inexpensive technique to detect any lesions, tumors, or cancerous cells in breast region [3]. Many underlying factors can bring about a negative effect on the accuracy of the results. A few and most common errors can be due to distraction or fatigue of the radiologist himself/herself the breast structure itself can be complex so the radiologist might find it difficult to locate the cancerous cells and also the early phase characteristics of the disease can be hard to analyze too [4, 5].

Computer-aided diagnosis also known as (CAD) comes into play for diagnosing breast cancer in an early stage. The whole process comprises of three steps all together: Firstly, finding and locating the area of possible tumor or cancerous cells by the aid of preprocessed mammogram, and secondly, getting a hold of the characteristics of the tumor including its shape, size, density, possible weight, texture and finally stating whether it is a benign or malignant version of a tumor following the noticeable features obtained from the second step [6]. X-ray was the only known screening method used earlier to detect any sort of malignancy including cancer of breast [7]. And since then various new and advanced detection technologies have been introduced which have proven to be more efficient compared to X-ray or mammography detection methods. These technologies are named as neural networks, data mining, and artificial intelligence.

In cancer research, machine learning (ML) techniques can be successfully applied to come across various patterns within a set of data, and as a result, we can also identify and state about the malignancy or benignancy of cancerous cells. One of the ways to expose the ML into supervised learning is by classification of the datasets which will be automatically used for future predictions due to its set algorithms. As such techniques are putting the clinics, hospitals, and other medical fields into huge advantageous situation by saving time and also aiding for early detection by the help of classification models created by previous cases of cancer [8]. In this study, five popular ML methods which included support vector machine (SVM), decision tree (DT), logistic regression (LR), random forest (RF), K-nearest neighbor (KNN), and a DL technique applied to a Breast Cancer Wisconsin (diagnostic) dataset are investigated and compared to analyze the classification performance of different classifying models.

The rest of the paper is structured as follows: In Section "Literature Review", previous work has been discussed. Section "Methodology" comprises a detailed methodology describing the dataset used, the proposed methodology, and the algorithms used. Section "Experimental Results and Discussion" represents the experimental results and the discussion including evaluation metrics, performance evaluation, and the comparison of performance. Finally, Section "Conclusion and Future Work" concludes the paper with future work.

Literature Review

Since decades, more and more researchers are putting their efforts to carry out studies through which they are trying to save time and bring the best output in cancer research. They are doing this by incorporating data mining algorithms in various available patient datasets. In addition to these, they have also been successful to classify stages and types of cancers including breast cancer. With the help of existing datasets, doctors and researchers are being able to solve challenging cases of cancer in patients within a short time frame as the algorithm plays its role and give results according to the symptoms and conditions of the patient which are further to be applied in patients for mitigation of the cancer. However, the results obtained by applying data mining in various categories for detection of breast cancer were found to very favorable. Table 1 portrays the different algorithms used till date and their efficiency in each case.

With reference to the comparative analysis constituted by Table 1, we can observe that Khuriwal et al. [9] obtained the best performing CNN model on the 'Mammogram MIAS' database, achieved an accuracy of 98%. In addition, Fathy et al. [10] also obtained a similar accuracy on the 'Screening Mammography' dataset with an accuracy of 96%. In addition, Saabith et al. [11] implemented an ANN model called MLP which had an accuracy of 97.28 on the 'Breast Cancer' dataset. Among all the different ML models, Amrane et al. [12]'s KNN model had demonstrated the highest performance accuracy with a figure of 97.51%, trained and tested on the 'Wisconsin Breast Cancer' dataset. Moreover, for the naïve Bayes classifier, Amrane et al. [12]'s model achieved the highest performance accuracy with a value of 96.19%, which is quiet close to the accuracy obtained by Amrane et al. [12]'s KNN model.

All classifiers considered, it was observed that CNN-based classifiers demonstrated better performance in terms of efficiency and classification accuracy, as opposed to other families of classifiers. ANN and ML models also had, more or less, similar performances.

Table 1 Existing works using diverse algorithms and dataset on breast cancer detection

Author	Dataset	Efficiency (%)	Algorithm	Findings
Nawaz et al. [13]	BreakHis Dataset	95.4	Deep learning (DL) convolutional neural network (CNN)	In comparison with DenseCNN model and state-of-art models, it showed more efficiency
Khuriwal et al. [9]	Mammogram MIAS database	98	CNN	Around 98% efficiency was found with CNN
Fathy et al. [10]	Screening Mammography dataset	96	DL	This method was found to have a sensitivity of 99.8% sensitivity
Amrane et al. [12]	Wisconsin Breast Cancer	97.51	KNN (k-nearest neighbors)	More accuracy was found with KNN compared to NB
		96.19	NB (Naive Bayes classifier)	
		79.97	J48	
		75.35	MLP	
Saabith et al. [11]	Wisconsin Breast Cancer	97.2818	J48	In this case, exactly same accuracy was achieved by both algorithms
		97.28	MLP	
		95.61	SVM	
		95.75	Bi-clustering and Ada boost techniques	
		91.3	RCNN	
		82.50	bidirectional recurrent neural networks (HA-BiRNN)	

Methodology

Dataset

The Breast Cancer Wisconsin (diagnostic) dataset [14] has been used in this study. In this dataset, there are 569 instances of computed features and classification of a digitized image of Fine Needle Aspirate (FNA) of a breast mass. There are 32 features of the dataset—two of them being the ID number and the diagnosis. The instances are diagnosed as either malignant or benign. There are 63% of instances belonging to the 'benign' class and the other 37% belong to the 'malignant' class. The other 30 features denote the mean, standard error, and the worst (mean of the three largest) values of 10 real-valued features that describe the characteristics of the cell nuclei present in the image.

In this study, 20% of the total dataset was used as test data, while the rest was used as training data. To prepare the data, the ‘diagnosis’ column was converted into int64 data type. Furthermore, the correlation between the features was analyzed, and upon finding strong correlation, all of them were decided to be used.

Proposed Methodology

The dataset has been used to analyze the classification performance of six models, of which five are ML models and the other is a DL model. In essence, we obtain a comparison between several popular machine learning models and between machine learning and deep learning techniques. The models used as classifiers are—support vector machine (SVM), decision tree (DT), logistic regression (LR), random forest (RF), K-nearest neighbor (KNN), and neural network (NN).

For each of the machine learning models, three instances were analyzed—using default parameters, using tuned parameters, and using tenfold cross-validation. The parameter tuning is done using GridSearchCV function of the scikit-learn library where an exhaustive search is performed over the provided parameter values. It performs the default fivefold cross-validation using the dictionary of parameters provided for each model. Additionally, the DL model is analyzed using default parameters and using tenfold cross-validation. Figure 1 portrays the overall process for evaluating the model in order to predict and classify breast cancer.

For the ML models, first their performance is analyzed using default parameters. Secondly, tuned parameters are used and their performances are analyzed again. The parameters specified for tuning and the best set of parameters found for each model are listed in Table 2.

Whereas for the DL model, two instances were analyzed—the first using a simple neural network (NN) and the second using tenfold cross-validation.

In the first instance, there are three layers to the network. The first layer is the input layer having input size of 29 (since we have 30 features), an output size of 8, and ReLU as the activation function. The second layer also has an output size of 8

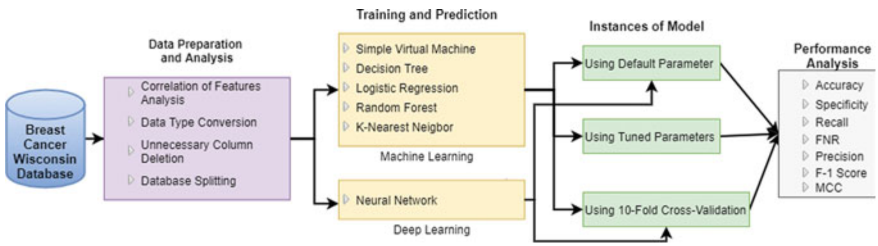


Fig. 1 Flowchart of proposed methodology

Table 2 Tuned and best set of parameters

Machine learning models	Tuned and best set of parameters
Support vector machine (SVM)	Gamma: 'scale' (*1 / (n_features * X.var())*), 'auto'(*1 / n_features*) Kernel: linear, radial basis function (RBF) Best parameters: gamma: 'scale', kernel: linear
Decision tree (DT)	Criterion: Gini, entropy Max depth: Between 1 and 10, excluding 10 Best parameters: criterion: entropy, max depth: 4
Logistic regression (LR)	Solver: 'newton-cg' (Newton method), 'liblinear' (library for large linear classification), 'sag' (stochastic average gradient descent), 'saga' (extension of 'sag') Max iteration: 10,000 Best parameters: solver: 'liblinear', max iteration: 10,000
Random forest (RF)	Number of estimators: 100, 200, 500 Max features: 'auto', 'sqrt', 'log2' Criterion: Gini, entropy Best parameters: criterion: entropy, max features: 'sqrt', number of estimators: 200
K-nearest neighbor (KNN)	Number of neighbors: 5, 7, 9, 11, 13, 17, 19, 21 Algorithm: 'auto', 'ball_tree', 'kd_tree', 'brute' Weights: uniform, distance Best parameters: algorithm: 'auto', number of neighbors: 13, weights: uniform

and ReLU activation function. And, the last layer is the output layer with output size of 1 and sigmoid activation function. The model is trained using an Adam optimizer, with a learning rate of 0.001, decay of 0.0001, and a clip value of 0.5. The model is fitted in batch sizes of 10 for 100 epochs. Prediction values above 0.5 are considered finally.

In the second instance, a NN of 5 layers is used. The first four layers are a repetition of the following sequence—a layer with output size of 8 and ReLU activation function and a dropout layer with dropout value of 0.2. The last layer is the output layer with output size of 1 and sigmoid activation function. It is trained and fitted using the same metrics as the first instance. Finally, the cross-validation scores are obtained using tenfold cross-validation. Figure 2 illustrates the layers of NN used in both case.

Classifier Description

Support Vector Machine (SVM)

Support vector machine (SVM) [15] is one of the most popular machine learning algorithms that provides sufficient accuracy using less computational power. This algorithm finds a hyper-plane to clearly classify data points in an N-dimensional space

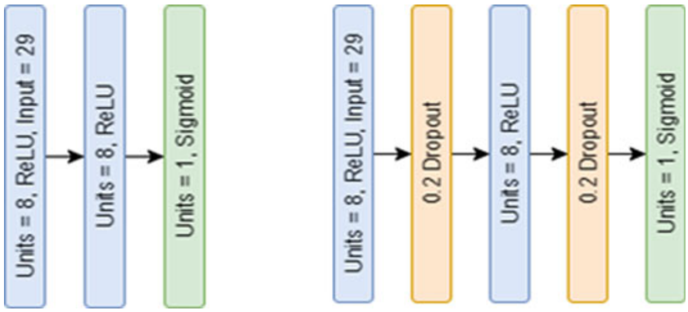


Fig. 2 Layers of neural network used: in the first instance (left) in the second instance (right)

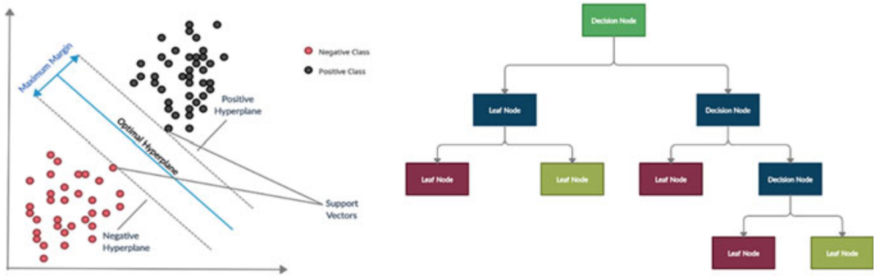


Fig. 3 General representation: hyper-plane derived using SVM (left) decision tree (right)

with maximum margin, i.e., space between data points of the different classes [16]. The hyper-plane acts as the decision boundary of the classification. The hyper-plane derived from SVM is depicted in Fig. 3 (left).

Decision Tree

Decision tree (DT) [17] is another popular supervised learning algorithm that can be applied and interpreted easily. It is essentially a tree diagram that contains internal nodes, leaf nodes, and branches. The internal nodes represent the independent variables of a classification or regression problem. The leaf nodes represent the outcome or the dependent variable of the problem [18]. And the branches represent the decision rules to move down the nodes. The decision making starts at the root of the tree moving downwards. Each internal node has a decision rule that is followed based on the comparison between the dataset's feature and the feature on the tree until a decision is reached, i.e., until a leaf node is reached. The general representation of a DT is shown in Fig. 3 (right).

Logistic Regression

Regression is a statistical process to predict the relation between a dependent variable and one or more independent variable. It is one of the categories of supervised learning. Logistic regression (LR) [19] is a class of regression thatpredicts categorical

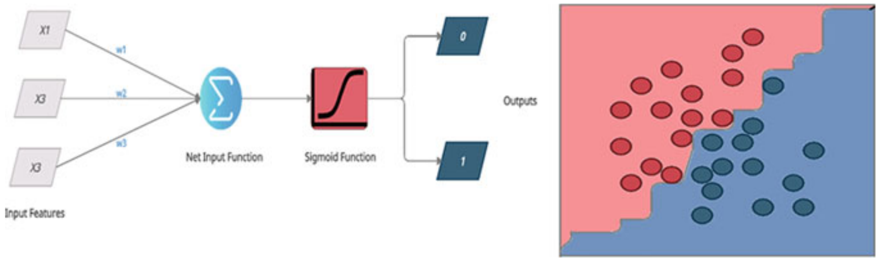


Fig. 4 Logistic regression (left) decision boundary of two classes generated by KNN (right)

variables using one or more independent variables. It is also known as logit regression or logit model. It provides the probability of the 'positive/true' class being decided. Based on that probability and a set threshold, the class can be categorized as 1 or 0. An illustration of logistic regression is shown in Fig. 4 (left). LR can be represented by the following equation:

$$\log\left(\frac{Y}{1-Y}\right) = \beta_0 + \beta_1 \mathrm{X}_1 + \beta_2 \mathrm{X}_2 + \cdots$$

Random Forest

Random forest (RF) is a decision tree-based ensemble method that is simple yet very powerful. It was first created by Tin Kam Ho in 1995 [20]. It creates several decision trees and classifies the data samples individually. Their decisions are then summed to find the greater number of votes to a class and thus reducing the overall error. This combining of decision trees is known as 'bagging.' In essence, if there is some correlation between the data samples and low correlation between predictions of individual trees, then the algorithm will converge to the right direction.

K-Nearest Neighbor

K-nearest neighbor [21] is another simple supervised learning algorithm used to model both classifications and regressions. It works on the basic idea that data points of the same kinds are in the same vicinity as each other and uses this idea to classify any given data point. For any given data point, K stands for the number of neighbors chosen to take the votes from; i.e., the closest K neighbors are enclosed in a perimeter, and the data sample is classified depending on the class of maximum number of neighbors. The parameter weight can be set to either uniform or with respect to distance. Figure 4 (right) demonstrates the decision boundary of two classes generated by KNN.

Neural Network

A neural network (NN) is composed of an input layer, an output layer, and layers in between called hidden layers [22]. The input layer consists of the individual values denoting the smallest unit of the input. The output layer consists of as many outputs as there are classes in the specific classification problem. Each hidden layer is responsible for recognizing a specific pattern. Each of the connections from all the neurons of one layer to the neurons of another layer has a weight assigned to them. The activation of a neuron can be found by calculating the weighted average of all the neurons connected to it from the previous layer along with a bias; this weighted average is then brought between 0 and 1 using a neural activation function such as the sigmoid function. The activation of a neuron can be found by calculating the weighted average of all the neurons connected to it from the previous layer along with a bias; this weighted average is then brought 0 and 1 using a neural activation function such as the sigmoid function.

Experimental Results and Discussion

In our work, we have employed the Breast Cancer Wisconsin Dataset that contains digitized FNA images of breast masses that depicts the characteristics of cell nuclei. The dataset constitutes a total of 569 images; 357 benign and 212 malignant. Thirty different features of each image were used to train our classifiers.

Performance Evaluation

In order to analyze the classification performance of different classifying models, we have employed five different models of ML and one DL model. The models were studied in order to see how accurately they were able to distinguish between benign tumors and malignant tumors (cancerous): malignant being the positive class and benign being the negative.

Moreover, we have made use of different performance measure parameters such as Specificity/TNR, Recall/TPR, FNR, FPR, Precision, F1-Score, and MCC, using the confusion matrices in order to provide more evaluations. The following segment discusses all the results obtained.

Performance Measure Parameters

Recall/True Positive Rate—Recall indicates what proportion of the total malignant samples were correctly predicted as malignant.

$$\text{Recall} = \text{TP}/(\text{TP} + \text{FN})$$

Specificity/True Negative Rate—Specificity indicates what proportion of the total benign samples were correctly predicted as benign.

$$\text{Specificity} = \text{TN}/(\text{TN} + \text{FP})$$

False-Negative Rate—It indicates what proportion of the total malignant samples were incorrectly classified as benign.

$$\mathrm{FNR} = \mathrm{FN}/(\mathrm{TP} + \mathrm{FN})$$

False-Positive Rate—It indicates what proportion of the total benign samples were incorrectly classified as malignant.

$$\mathrm{FPR} = \mathrm{FP}/(\mathrm{TN} + \mathrm{FP})$$

Precision—It indicates what proportion of the total positive predictions (all samples there were predicted as malignant) were actually positive.

$$\mathrm{Precision} = \mathrm{TP}\ /(\mathrm{TP} + \mathrm{FP})$$

F1-Measure—The $F1$-measure represents an overview of the overall performance of the model on the test dataset.

$$\mathrm{F1} - \mathrm{Measure} = (2 * \mathrm{Precision} * \mathrm{Recall})/(\mathrm{Precision} + \mathrm{Recall})$$

MCC (Matthews correlation coefficient)—The MCC is suitable for use in binary classification tasks. Its value ranges from $+1$ to -1. Closer the value of MCC is to $+1$, better the performance of the model, and further away it is, poorer the performance.

$$\mathrm{MCC} = \frac{(\mathrm{TP} * \mathrm{TN}) - (\mathrm{FP} * \mathrm{FN})}{\sqrt{(\mathrm{TP} + \mathrm{FP})\,(\mathrm{TP} + \mathrm{FN})\,(\mathrm{TN} + \mathrm{FP})\,(\mathrm{TN} + \mathrm{FN})}}$$

Confusion Matrices

For determining the different performance measure parameters, we will be using the confusion matrix of each of the classifier models in order to derive True Positive (TP), False Positive (FP), True Negative (TN), and False Negative (FN). Malignant is the positive class, while benign has been assumed to be the negative class. Table 3

Table 3 Confusion matrix parameter values for different models

Classifiers	TN	FP	TP	FN
SVM	70	2	36	6
Decision tree	70	2	37	5
Logistic regression	71	1	36	6
Random forest	72	0	39	3
KNN	72	0	33	9
Neural network	71	1	32	10

summarizes the parameters that can be derived using the confusion matrices of each of the classifiers.

True Positive (TP)—Indicates number of images that wer predicted as malignant and were correctly predicted.

False Positive (FP)—Indicates number of images that were predicted as malignant but were infact benign.

True Negative (TN)—Indicates number of images that were benign and were correctly predicted as being benign.

False Negative (FN)—Indicates number of images that were malignant but were predicted as benign.

Performance Comparative Analysis

The performances of the ML and NN models were measured by determining the accuracy of the total correct predictions made with respect to the total number of predictions made. Figure 5 presents the visual representation of the overall accuracies of the different models that were trained and tested on the dataset. The first conclusion that can be drawn from these values is that the random forest classifier had the best overall performance compared to all other models. It had achieved the highest performance accuracy of 97.37% when implemented with default parameters and the second-best performance accuracy of 96.66% when implemented with tuned parameters. Additionally, after implementing tenfold cross-validation on all the models, random forest had the highest cross-validation accuracy of 96.84% among all the models, which proves that the model can generalize most effectively on new dataset. On the other hand, the NN classifier had the minimal overall performance accuracy of 90.35% with split and also showed very low cross-validation accuracy of 85.61% in regard to the machine learning classifiers. Finally, we can also deduce that the performances of the ML models improved slightly after tuning the parameters.

From Table 4 and Fig. 6, we can observe that random forest (tuned) outperformed all the other models with an accuracy of 96.66%, outperforming all the other models. Logistic regression (tuned) was the second-best performing model on our dataset, with an accuracy of 95.08%. The DL network had the least classification accuracy—90.35%. In addition, random forest had the highest $F1$-measure. All parameters considered, random forest classifier has proven to be the best performing model.

Comparative Analysis of the Proposed Classifier with Existing Models

In order to measure the robustness and effectiveness of the proposed methodology, the best performing model has been compared with other state-of-the-art models which

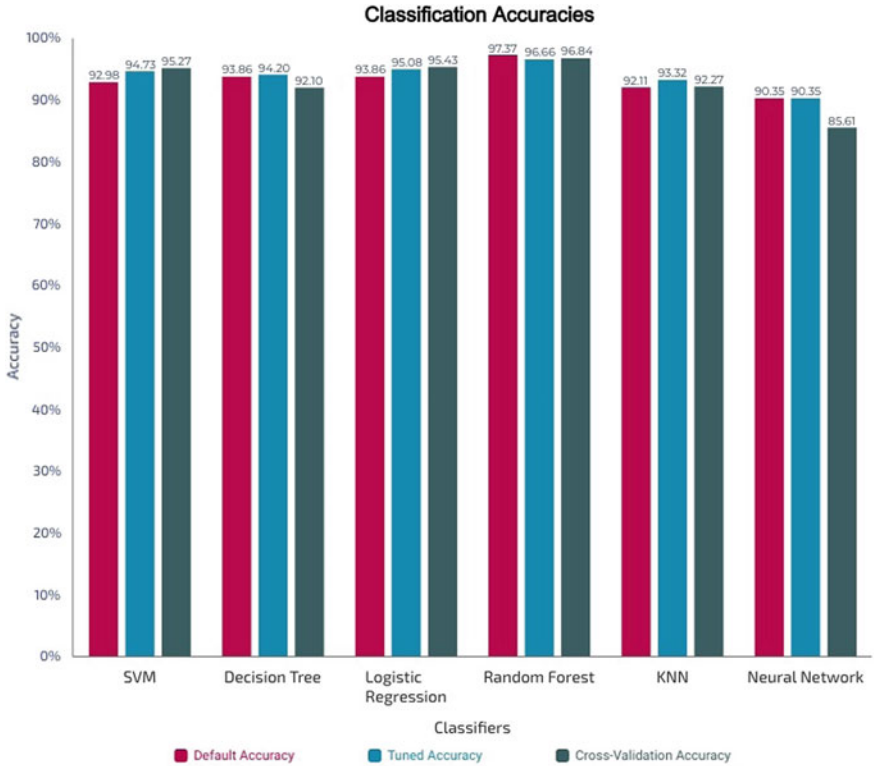


Fig. 5 Classification accuracies of the different machine learning models

Table 4 Performance measure parameter values for the different classifiers

Parameters	SVM	Decision tree	Logistic regression	Random forest	KNN	Neural network
Tuned accuracy (%)	94.73	94.20	95.08	96.66	93.32	90.35
Recall/TPR	0.857	0.881	0.857	0.929	0.786	0.762
Specificity/TNR	0.972	0.972	0.986	1	1	0.986
FNR	0.143	0.119	0.143	0.071	0.214	0.238
FPR	0.028	0.028	0.014	0	0	0.014
Precision	0.947	0.949	0.973	1	1	0.970
$F1$-Measure	0.900	0.914	0.911	0.963	0.880	0.854
MCC	0.849	0.900	0.869	0.944	0.836	0.796

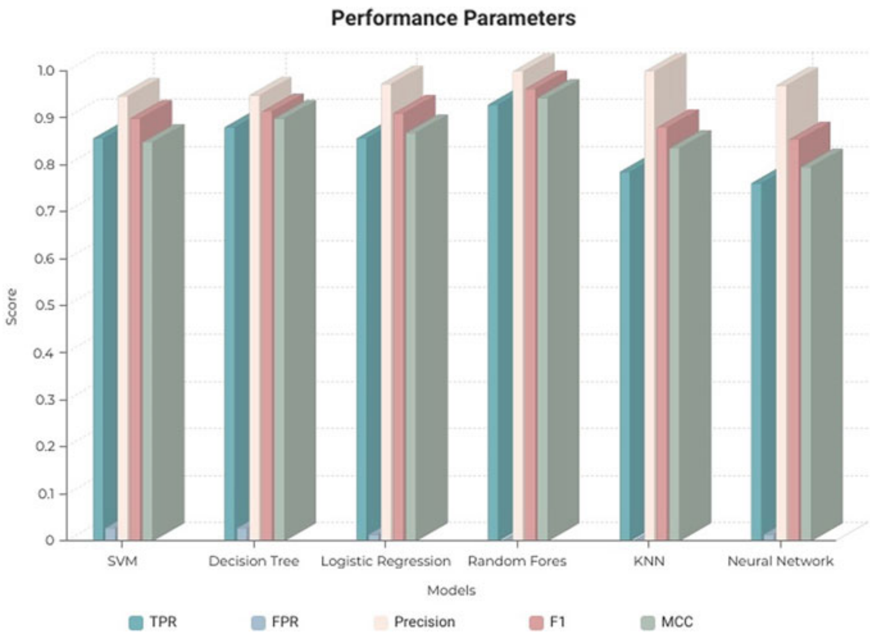


Fig. 6 Performance measure summary

similarly have been trained on the Wisconsin Diagnosis Breast Cancer (WDBC) dataset. In the proposed methodology, the random forest classifier had the best performance accuracy of 97.37%, and from Table 5, we can find that it outperforms classifiers such as naïve Bayes [23], classification and regression trees (CART) [24] and extra trees classifier [25] by a notable margin. On the other hand, classifiers such as artificial neural network (ANN) [26] and probabilistic neural network [7] outperformed the random forest classifier by a very small margin, having achieved

Table 5 Performance comparison of the proposed classifiers with other existing models

Classifiers	Performance accuracy (%)	Dataset ımplemented
Random forest [Proposed]	97.37	Wisconsin Diagnosis Breast Cancer (WDBC) 569 images; 357 benign and 212 malignant
Artifical neural network [26]	98.57	√
Probabilistic neural network [7]	98.15	√
Naive Bayes [23]	94.47	√
CART [24]	92.00	√
AdaBoost classifier [25]	98.23	√
Extra trees classifier [25]	97.34	√

98.57% and 98.15% performance accuracies, respectively. The better performance can be associated with the fact that these classifiers had been trained on a larger set of images than the random forest model. It is therefore, safe to state that the proposed random forest classifier is highly robust and is on par with the performances of most of the state-of-the-art models. However, it is noteworthy to mention that the AdaBoost classifier [25] can be said to have achieved the soundest performance accuracy when taking into the account the number of images it has been trained on. The AdaBoost classifier had been trained on 569 images, same as that of the random forest classifier and lower than other models; it achieved a performance accuracy of 98.23% which is higher but very marginal compared to the proposed classifier.

Conclusion and Future Work

Medical technology is constantly looking for ways to automate steps by either eliminating tedious steps or by providing a diagnosis to confirm. This can help the medical industry to provide healthcare efficiently and spontaneously. This literature is an attempt to contribute to this notion. In this study, we have used digitized images of FNA of breast masses to classify them as either malignant or benign. There were comparisons made using both machine learning and deep learning techniques. Furthermore, stability of the models was verified using cross-validation technique. The best performance was obtained by tuning the parameters of a random forest model with an accuracy of 96.66%. Logistic regression (with tuning) followed closely with an accuracy of 95.08%. The lowest performing model was the deep neural network (with cross-validation) with an accuracy of 85.61%. In future, we will be implementing models of DNN and transfer learning to increase the model's efficiency.

References

1. U. S. C. S. W. Group and others (2015) United States cancer statistics: 1999–2012 incidence and mortality web-based report. Atlanta Department of Health and Human Services, Centers for Disease Control Prevention, National Cancer İnstitute
2. Zhu H, Doğan BE (2021) American joint committee on cancer's staging system for breast cancer: summary for clinicians. Eur J Breast Heal 17(3):234
3. Lerman C et al (1993) Mammography adherence and psychological distress among women at risk for breast cancer. JNCI J Natl Cancer Inst 85(13):1074–1080
4. Huynh PT, Jarolimek AM, Daye S (1998) The false-negative mammogram. Radiographics 18(5):1137–1154
5. Ertosun MG, Rubin DL (2015) Probabilistic visual search for masses within mammography images using deep learning. In: 2015 IEEE ınternational conference on bioinformatics and biomedicine (BIBM), pp 1310–1315

6. Pereira DC, Ramos RP, Do Nascimento MZ (2014) Segmentation and detection of breast cancer in mammograms combining wavelet analysis and genetic algorithm. Comput Methods Programs Biomed 114(1):88–101
7. Übeyli ED (2007) Implementing automated diagnostic systems for breast cancer detection. Expert Syst Appl 33(4):1054–1062
8. Bazazeh D, Shubair R (2016) Comparative study of machine learning algorithms for breast cancer detection and diagnosis. In: 2016 5th International conference on electronic devices, systems and applications (ICEDSA), pp 1–4
9. Khuriwal N, Mishra N (2018) Breast cancer detection from histopathological images using deep learning. In: 2018 3rd İnternational conference and workshops on recent advances and innovations in engineering (ICRAIE), pp 1–4
10. Fathy WE, Ghoneim AS (2019) A deep learning approach for breast cancer mass detection. Int J Adv Comput Sci Appl 10(1):175–182
11. Saabith ALS, Sundararajan E, Bakar AA (2014) Comparative study on different classification techniques for breast cancer dataset. Int J Comput Sci Mob Comput 3(10):185–191
12. Amrane M, Oukid S, Gagaoua I, Ensari T (2018) Breast cancer classification using machine learning. In: 2018 Electric electronics, computer science, biomedical engineerings' meeting (EBBT), pp 1–4
13. Nawaz M, Sewissy AA, Soliman THA (2018) Multi-class breast cancer classification using deep learning convolutional neural network. Int J Adv Comput Sci Appl 9(6):316–332
14. Breast Cancer Wisconsin (Diagnostic) Data Set, UCI machine Learning Repository: Breast Cancer Wisconsin (diagnostic) data set. https://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+%28Diagnostic%29. Accessed 05 July 2021
15. Sakib S, Siddique MAB, Rahman MA (2020) Performance evaluation of t-SNE and MDS dimensionality reduction techniques with KNN, ENN and SVM classifiers. In: 2020 IEEE Region 10 symposium (TENSYMP), pp 5–8
16. Hasib KM et al (2020) A survey of methods for managing the classification and solution of data imbalance problem. arXiv Prepr. arXiv2012.11870
17. Cho H et al (2013) A similarity study of content-based image retrieval system for breast cancer using decision tree. Med Phys 40(1):12901
18. Hasib KM, Showrov MIH, Das A (2020) Accidental prone area detection in Bangladesh using machine learning model. In: 2020 3rd International conference on computer and ınformatics engineering (IC2IE), pp 58–62
19. Hwa H-L et al (2008) Prediction of breast cancer and lymph node metastatic status with tumour markers using logistic regression models. J Eval Clin Pract 14(2):275–280
20. Kam HT et al (1995) Random decision forest. In: Proceedings of the 3rd İnternational conference on document analysis and recognition, vol 1416, p 278282
21. Siddique M, Bakr A, Sakib S, Rahman M et al (2019) Performance analysis of deep autoencoder and NCA dimensionality reduction techniques with KNN, ENN and SVM classifiers. arXiv Prepr. arXiv1912.05912
22. Siddiaue MAB, Sakib S, Khan MMR, Tanzeem AK, Chowdhury M, Yasmin N (2020) Deep convolutional neural networks model-based brain tumor detection in brain MRI ımages. In: 2020 Fourth ınternational conference on I-SMAC (IoT in social, mobile, analytics and cloud) (I-SMAC), pp 909–914
23. Sharma S, Aggarwal A, Choudhury T (2018) Breast cancer detection using machine learning algorithms. In: 2018 International conference on computational techniques, electronics and mechanical systems (CTEMS), pp 114–118
24. Asri H, Mousannif H, Al Moatassime H, Noel T (2016) Using machine learning algorithms for breast cancer risk prediction and diagnosis. Proc Comput Sci 83:1064–1069
25. Dhahri H, Al Maghayreh E, Mahmood A, Elkilani W, Faisal Nagi M (2019) Automated breast cancer diagnosis based on machine learning algorithms. J Healthc Eng 2019
26. Islam MM, Haque MR, Iqbal H, Hasan MM, Hasan M, Kabir MN (2020) Breast cancer prediction: a comparative study using machine learning techniques. SN Comput Sci 1(5):1–14

Music Emotion Recognition Using Neural Networks

Divyasha Pahuja

Abstract This paper proposes usage of images to classify music according to genre. Musical characteristics are extracted from these images for music genre recognition and categorization. These textural or visual representations are created by converting the sound signal to spectrograms which are then utilized for feature extraction. By using the time–frequency graphs for data mining of musical characteristics we can do feature extraction to build a music genre recognizer. We use a multifarious dataset of 1000 music pieces, all divided into 10 genres and found that the classifier trained with data obtained from these spectrograms is very efficient at recognizing the genre of the clip. A CNN model hyperparameter tuned to give better results is used for classification. Our proposed model is tested with different optimizers and achieves a good accuracy rate for genre recognition with the GTZAN dataset.

Keywords Music genre · Feature extraction · CNN · Spectrograms · Keras · Classifier

Introduction

Music genre or emotion is a definitive category that is used to identify pieces and parts of music that belong to a certain predefined set of norms. It is said to be different from musical style, composition and form, although generally these ideologies are used interchangeably. Genres can be defined as categorical labels designed by us humans to identify the tone of the musical piece. Since human mind is subjective and can perceive each piece of sound/music differently, categorizing the musical pieces into genres is not a small task.

Music genre recognition comes under the umbrella of music ınformation retrieval (MIR). MIR deals with the understanding of musical pieces by looking at different aspects of music such as signal processing, music theory and machine learning itself. MIR can be used to peform various tasks like music generation, track recognition or

D. Pahuja (✉)
Bhagwan Parshuram Institute of Technology, Delhi, India

V. Bindhu et al. (eds.), *Proceedings of Third International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 844,
https://doi.org/10.1007/978-981-16-8862-1_47

even instrument recoginiton, beat checking, etc. So MIR is a field of study used to train and enable machines to understand, process and define music in an intelligent way.

Automated music genre recognition has been in works widely and serves as an important part of MIR as it can be and is in fact used for music recommendations, organizing and prioritizing musical databases and discovering new musical collections. Big companies (such as Spotify and SoundCloud) nowadays use music classification to be able to recommend music to the customer according to the customers taste or simply as a product to identify the piece of music (such as Shazam). Determining genre of music is the start of building an MIR. Machine learning fields like neural networks and deep learning have proved to be quite useful for extracting patterns and features from a wide pool of data.

In this paper, an interesting approach for music emotion recognition is discussed. Here audio signals are converted into spectrograms which form the basic input unit for our classifier. A spectrogram is a graphical representation of the range of frequencies over the time of an audio signal. A spectrogram usually looks like a heat map where the intensity is shown by the degree of brightness or varying color. The features extracted from the spectrogram are then fed to the model to train the model and build a reliable music genre recognizer.

The model is built with deep learning (DL) model which consists of various neural network algorithms that:

- Employ a multilayered nonlinear processing unit model for feature extraction and transformation. Each layer output becomes the input for the subsequent layer.
- Learn through supervised (e.g., classification) or unsupervised (e.g., clustering, pattern analysis) fashion.
- Learn from many levels of neuron layers that consist of distinct levels of density and abstraction.

This multilayer perceptron network is trained on the extracted features to correctly predict the music genres.

This convolutional neural network model uses best optimizer and learning rate tuned to give an improved recognition rate.

Related Work

Music genre and style recognition have been widely researched and performed over the time. The approaches considered for classifying music into genre relies on analysis of music by extracting feature information from the time–frequency visual representation generated from the audio and have been consulted from the following: Paper [1] presents the approach of spectrograms used for automatic audio signal conversion. The authors use zoning mechanism and do feature extraction locally on Latin Music Database (LMD) and then use support vector machine (SVM) classifier in

Table 1 Comparison of existing works using spectrograms or similar technique

Citation	Dataset	Model used	Recognition rate (%)
[1]	LMD	SVM	67.2
[2]	GTZAN	SVM + Random forest	71.3
[3]	LMD & ISMIR 2004	SVM (Gaussian kernel)	82.33 80.65
[4]	LMD	SVM	88.56
[5]	MagnaTagATune	CRNN	89.3
[6]	GTZAN & Ballroom	CNN-LSTM	94.2 93.8

ensemble with other classifiers to give better efficacy. Paper [2] also applies spectrogram or short-time Fourier transform (STFT) technique and uses combination of SVM and random forest classifier.

Papers [3, 4] use texture features based on local binary pattern (LBP) and identify harmonic structures and intensive beats from the LBPs for music classification (Table 1).

In paper [7], the authors apply Gaussian processes or GPs (Bayesian models) for emotion and genre classification of music. These are major tasks that are a part of MIR. Two types of recognizers are created—one for music emotion and another for music genre, and they used a combination of both SVM and GP models and computed and compared the differences in their efficiency and performances, both trained on similar databases. They found through experimentation that for both music emotion and music genre estimation the GP classifier performed better than SVM classifier.

Paper [8] forms the basis and initial research for automatic musical genre classification. The paper proposes in detail feature sets for presenting timbre texture, rhythmic content and pitch measures of the music. They use Gaussian mixture model (GMM) is combined with K-nearest neighbor (K-NN) classifier. Paper [9] proposes a new method for music genre recognition (MGR) that uses simplified chord sequences and uses random forest classifier combined with SVM classifier.

In paper [10], a semi-supervised classification algorithm is proposed which is trained on different labeled music tracks and many unlabeled tracks. Three features are then extracted from the each music piece, and manifold regularization method is used to develop the classifier.

Papers [11, 12] use genetic algorithm (GA) for feature selection.

In paper [13], they propose the use of time-constrained sequential patterns (TSPs) as effective features for music genre classification. It tokenizes each piece of music to a series of hidden Markov model indices and computes the frequencies of occurrence of TSPs for that piece.

Paper [14] states the problems that occur in evaluation of music emotion and concludes with recommendations for the same.

Paper [15] also demonstrates the problems that come with Music Emotion Recognition (MER) as human perception is subjective. The paper focuses on the importance of psychology and cognitive science. A new framework is created to utilize acoustic information.

Paper [16] addresses the absence of regional music dataset and proposes Brazilian Music Dataset (BMD) which covers our future prospects for classification of regional language.

Finally, paper [17] discusses about traditional classifiers and recent progress of deep learning and paper [5] introduces CRNN architecture that uses both CNN and RNN structures for extracting local features. Paper [6] calculates two-dimensional rhythmic and spectral features using the mix of CNN and CNN-LSTM models for music classification.

Dataset

Data plays a crucial role in any analysis. So the dataset collected is important for the analysis to run with high accuracy. The dataset used for this paper is GTZAN dataset, which is popularly used in the area of MGR; it consists of a 1000 tracks with each track belonging to one of the ten genres—country, classical, blues, jazz, hip-hop, metal, disco, pop, rock and reggae. All genres consist of 100 tracks each. All the tracks are in .wav format, each audio 30 s long sampled at 22,050 Hz. Data is handled using toolkits mentioned in paper [18].

Methodology

This paper explains the process step by step and the approach for achieving the desired results. Figure 1 demonstrates the proposed work flow. Firstly, the data is fetched and processed which is then suitable representing it in a time–frequency chart, i.e., spectrograms. Features are then extracted from the spectrogram created for each track and fed to the classifier. The classifier is then tested on different optimizers, and accuracy is calculated to determine the best optimization algorithm.

Data Preprocessing

Preprocessing of data is significant as it converts data into format suitable for processing. Preprocessing of data includes data loading and data cleaning.

Data Loading. Data loading is basically the procedure of loading of unclean data onto the environment where the analysis is to be performed. The GTZAN dataset of 1000 tracks is loaded to the IDE by using Librosa module in Python. Librosa is

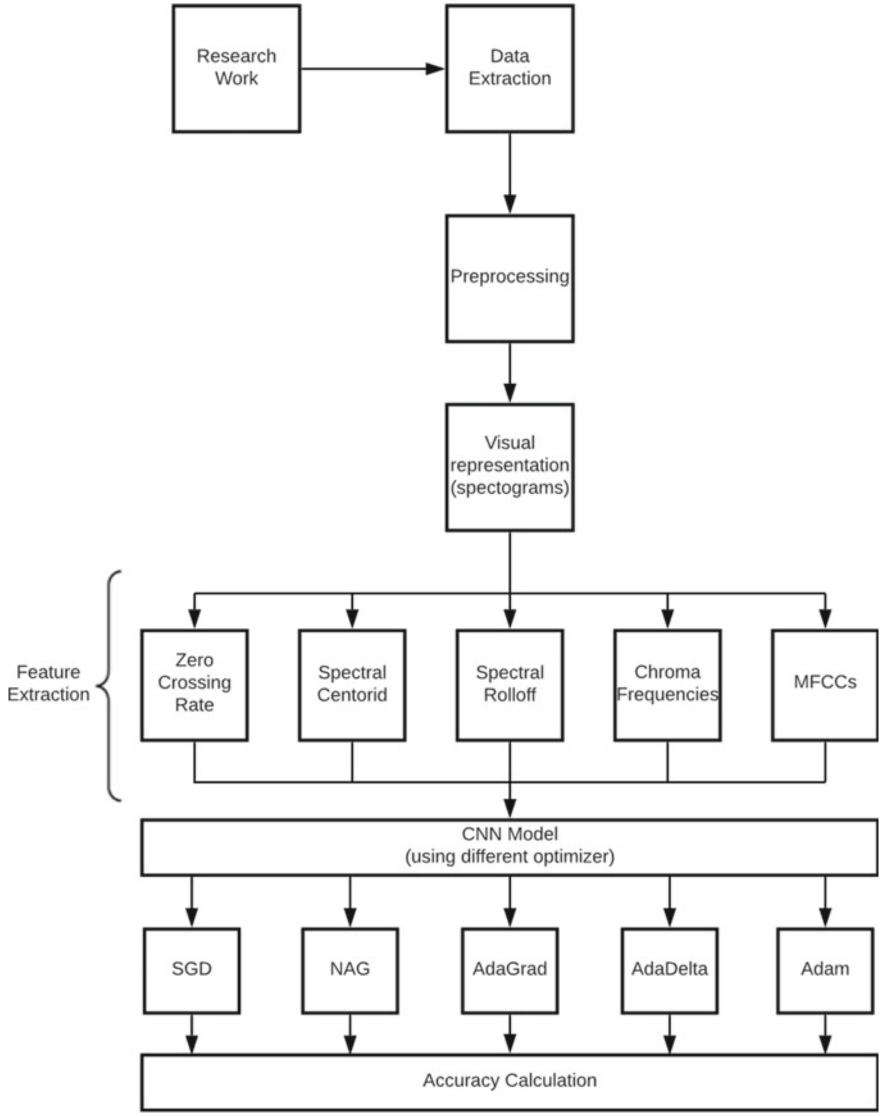


Fig. 1 Implementation process

Python package that handles audios, mainly music, and forms the core for creating a Music Information Retrieval (MIR) system.

Once it is loaded, the audio samples need to be converted from their raw form to format that is suitable for processing. When testing, the sound should be converted from .au to .wav format so the audio clips are compatible with Python modules and can be further processed into spectrograms.

Data Cleaning. Next preprocess task is cleaning of data, that is removing redundancies and inconsistencies from the initial data and to remove parts of the data which are not required for analysis.

Cleaning of data is done in the dataset formed from extracted features by:

- Removal of unnecessary columns (such as filename).
- Encoding of labels as the model can only intake data in numerals.
- Standardization of features, for normalized values throughout.

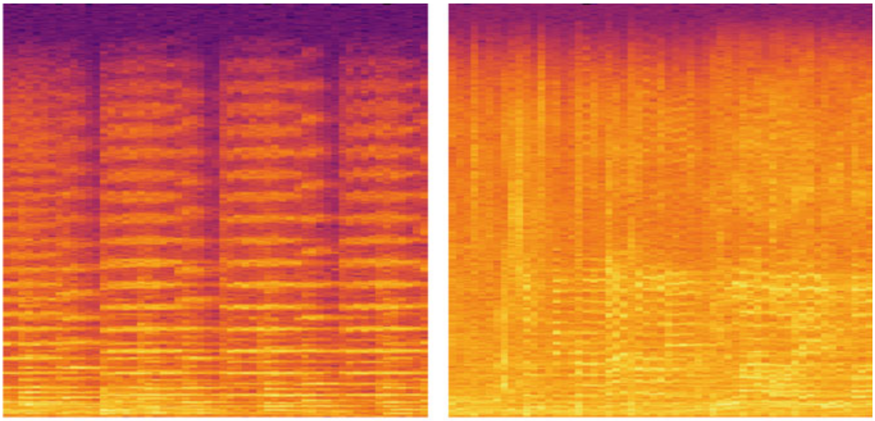
Feature Extraction

The feature extraction process is done from spectrograms. So, the first step of the workflow is to convert the preprocessed audio signal to time–frequency graphs, i.e., spectrograms. The spectrograms are created using Matplotlib in Python. Matplotlib provides multiple colormaps, and inferno colormap is used to create heatmaps of each track. The spectrograms are created with audio sample size 16 bits and a sampling rate of 22.5 kHz by default. We also tried sampling rate of 44.1 kHz for testing purposes.

All the sound clips are represented by the heatmap; an example of it is shown in Fig. 2. Figure 2 presents spectrogram of music piece of two different genres for comparison purposes.

After the creation of spectrograms, the next step in the process of genre recognition is extraction of features from these spectrograms.

Now, every sound signal consists of many characteristics and features that make it distinct. However, we have to extract only the features that are relevant to the



(a) (b)

Fig. 2 Spectrograms created for **a** classical audio and **b** rock audio clip problem at hand, that is genre recognition. This process of extracting characteristics and using them for detailed analysis is called feature extraction.

We extracted five meaningful features for our purpose from the spectrograms, those are Zero-Crossing Rate, Spectral Centroid, Spectral Roll-off, Chroma frequencies and Mel-frequency cepstral coefficients (MFCCs).

The significance of each feature is studied and explained in detail as follows:

Zero-Crossing Rate. The zero-crossing rate (ZCR) of an audio frame specifies the rate at which the mathematical sign of the audio changes; i.e., the signal changes its sign from positive to negative or vice versa, divided by the length of the frame. The graph of the signal shows the alternating values of the audio or sound crossing of the axis, which is the zero value. Zero crossing represents the point where there is no voltage present.

Zero crossing is a feature of importance as it shows most prominent frequency of an audio at a point of time and presents highs and lows in the audio. The ZCR therefore helps us in identifying the points in the audio where energy is high and makes easier to identify the genre of the clip as different genres have a change in high and low points differently. Hence, zero-crossing rate has and is being used for MIR as well as speech recognition.

Figure 3 shows the zero-crossing graph for the same classical audio used as an example for demonstrating spectrogram.

We can deduce from the graph that the sound track has seven zero crossings. Zero crossings can also be calculated by using the zero-corssings function of the librosa library which would also produce the same result as shown in Fig. 4.

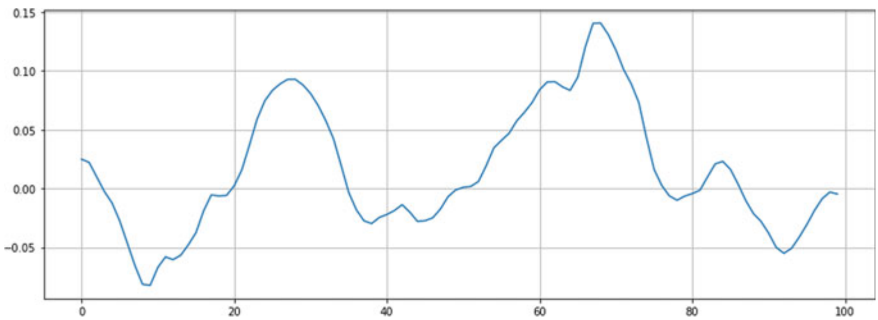


Fig. 3 Zero-crossing rate for classical audio clip

```
zero_crossings = librosa.zero_crossings(x[n0:n1], pad=False)
print(sum(zero_crossings))
```

7

Fig. 4 Calculation of zero crossings in Python

Spectral Centroid. It is the weighted mean of the frequencies present in an audio. So basically, it points out where the "gravity" or "center of mass" of a soundtrack resides. It is connected to the brightness of sounds. Higher values would mean brighter sounds.

Let us compare the two musical pieces, one from a classical genre and the other from rock genre. The classical song is the more or less same throughout its duration, whereas the rock song has more amplitude, i.e., energy especially toward the end of it. So, the mean or spectral centroid for the classical song will be somewhere in the middle of its frequency spectrum whereas the rock song would have high spectral centroid values, as a rock song has more energy or "brightness."

In Fig. 5, we show and compare the spectral centroid of classical and rock song, respectively, along with their audio signals.

Spectral Roll-off. The spectral roll-off feature calculates roll-off frequency, which is the threshold or cutoff frequency under which some part of the total energy resides.

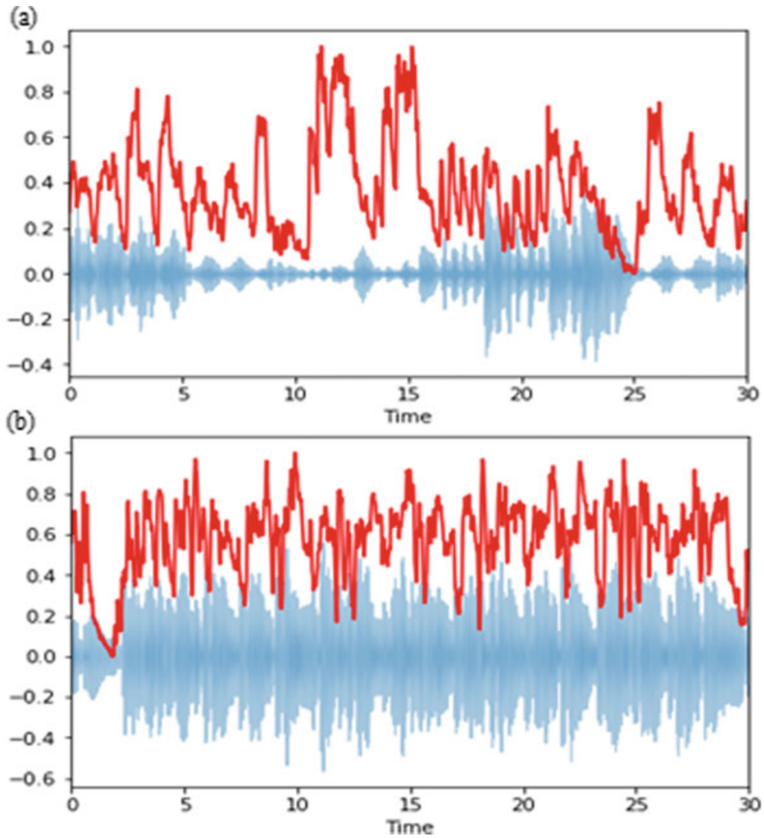


Fig. 5 Spectral centroid values over time for **a** classical song and **b** rock song

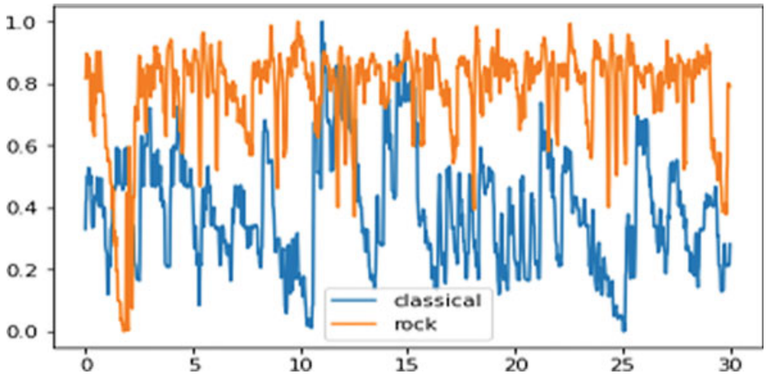


Fig. 6 Spectral roll-of values classical and rock song represented with different colors

So, the spectral roll-off represents the frequency which is under a specified percentage of overall spectral energy.

Spectral roll-off plays a significant role in genre recognition as the roll-off frequency can be used to distinguish the main audio from noisy sounds, which come above the roll-off frequency. They can also be used to ignore harmonic sounds which come below roll-off frequency.

Energy distribution below 85% of the total spectral energy magnitude is considered for the threshold roll-off.

Figure 6 shows the roll-offs of classical and rock music piece side by side.

Chroma Frequencies. The octave consists of 12 different musical chromas. The chroma frequencies are used to display the intensity of these 12 distinct semitones. So, the entire spectrum of an audio is mapped to 12 bins that constitute the 12 parts of the musical octave. Chroma features are related to the harmonic and symphonic

characteristics of music, pitch and timbre. Chromagrams form the twelve pitch classes and chroma-based features are also called as "pitch class profiles."

The chroma frequencies feature is important for genre recognition as it is powerful in analyzing and extracting information about music which can be distinctly categorized into those 12 bins or chromas.

Figure 7 represents the 12 chromas of the classical song audio.

Mel-Frequency Cepstral Coefficients (MFCCs). Mel-frequency cepstrum (MFC) is a linear cosine transform or LCT of the power spectrum based on a nonlinear mel scale of frequency. The magnitude of power spectrum is calculated from spectrograms and mapped onto mel scale. They imitate the characteristics of biological human sound.

MFC is made up of mel-frequency cepstral coefficients (MFCCs). They are computed from a kind of cepstral representation of the sound track. 20 coefficients are used to form the MFC.

Figure 8 shows the calculated 20 mel-frequency cepstral coefficients in the form of a spectral envelope over 1293 frames.

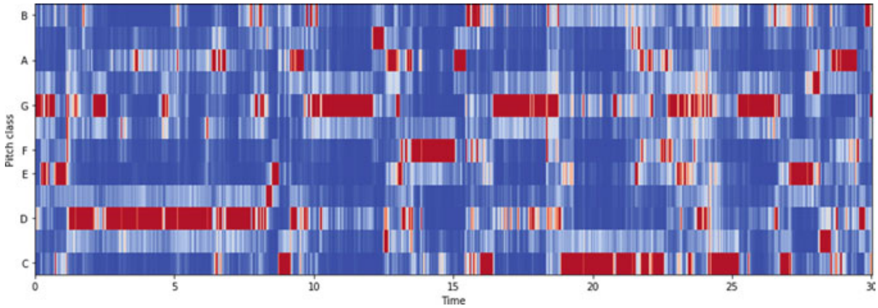


Fig. 7 Chroma frequencies of a classical song

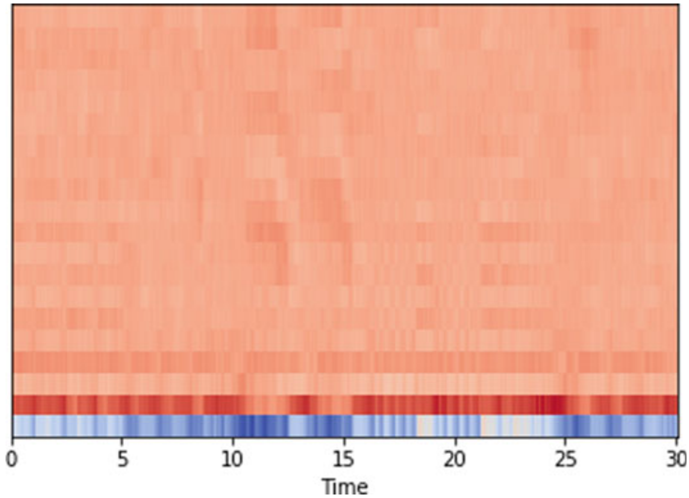


Fig. 8 MFCC values plotted on graph for classical song

Root mean square error, i.e., RMSE, values are also calculated for each audio. RMSE takes into consideration the errors while making predictions. It calculates squares of the distance of the regression points from the best fit. So, RMSE is good for estimating how spread out the data or how well the predicted functions fits the data.

The five features whose values are computed for each audio clip (also, including RMSE values and spectral bandwidth) are appended to a .csv file along with additional fields such as audio filename and the label (or genre) of that audio, which then forms the dataset that is used for training our model.

The feature values computed for audios belonging to 5 of the 10 genres is shown in Table 2 for reference. For convenience, only 2 MFCC coefficients out have 20 have been included.

Table 2 Feature values calculated for sound tracks

Genre/Feature	ZCR	Spectral centroid	Spectral roll-off	Chroma STFT	MFCC1	MFCC 20
Blues	0.127272	2116.943	4196.11	0.380260	−26.929	2.91899
Classical	0.087772	1360.408	2387.22	0.215064	−324.06	0.36627
Disco	0.188547	3290.414	6514.76	0.411989	−28.902	−0.1304
Jazz	0.059278	1266.406	2410.42	0.328119	−345.73	−0.7924
Rock	0.054982	1529.345	3271.66	0.365525	−111.14	−5.3326

Classifier Model

A convolutional neural network model is used in this work. A CNN is a part of deep learning study. CNNs are generally used for analysis of visual representation or images. CNNs work on matrices or grid-like topology. By using convolution, CNNs are able reduce the imagery into a form that is suitable for processing and analysis of that reduced data to make accurate predictions.

Our neural network model is built using Keras library from tensorflow module in Python. Using Keras, we created the building blocks of our neural networks which include the numerous layers of our CNN model, the adaption of activation functions and choosing the best optimizer for a good recognition rate.

First step is the division of dataset into train set, validate set and test set. The training of model is done on train set and then checked using validate set. The model is then used for predicting genres whose values are tested against the ones in test set. The dataset is split into training and test set in 70–30 ratio, respectively. Further, 200 records are set out from the training set for validation.

Next is creation of a multilayered perceptron network. A sequential model is created witch each subsequent layer less dense than the other, i.e., the number of nodes decreased with the layers. 2D (two-dimensional) CNN mode is used. The stages of the model are convolution—max pooling—convolution—max pooling—output as shown in Fig. 9. In convolution stage, kernels are applied to form feature

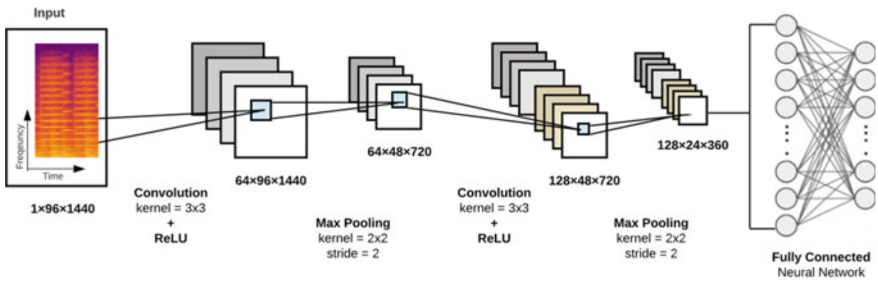


Fig. 9 CNN model architecture

maps from the input. In max pooling, an abstract form of the input is formed so that most important features from the feature map can be extracted.

The first 2D convolution layer is of size (64 × 3 × 3) and input size (1 × 96 × 1440). Features are then max pooled by strides of size (2 × 2) and so the input for second convolution layer becomes (64 × 48 × 720). Second convolution layer of size (128 × 3 × 3) max pooled with stride (2 × 2) to give (128 × 24 × 360) output size.

The hidden layers use ReLU activation function whose values are then passed on to the next hidden layer. ReLU activation function calculates the values using formula:

$$R(z) = \max(0, z) \tag{1}$$

where z is the input/weights from previous layer.

Softmax function is used just before the output layer for classification of data into multiple classes (genres in our case). Softmax function is calculated using:

$$\sigma(z)_i = \frac{e^{z_i}}{\sum_{j=1}^{K} e^{z_j}}. \tag{2}$$

where z = input;
e^{z_i} = standard exponential input function;
K = number of classes;
e^{z_j} = standard exponential output function;

The model is then fit to 50 epochs; i.e., the number of times full dataset is passed forward and backward through neural network, updating the internal model parameters each cycle, and batch size is kept 32. Loss function (difference between expected and current output) for multi-class classification is set to categorical cross-entropy. Typical loss function that is used in cross-entropy is computed by taking the average of all cross-entropies in the sample. The loss function [5] for N samples is given by:

$$L(w) = -\frac{1}{N}[y_n \log \hat{y}_n + (1 - y_n)\log(1 - \hat{y}_n)] \tag{3}$$

Most importantly, model is tuned to different learning rates (10^{-2}, 10^{-3}, 10^{-4}) to identify the optimal learning rate. A slow learning rate would lead to minuscule updates to weights each cycle. A very high learning rate can cause the divergence of loss function.

The momentum (β_1) which accelerates gradient descent in the right direction is not set as the default value of 0.9 serves our cause.

The model is compiled using different optimizers, and after experimenting with different optimizers for the different learning rates, Adam optimizer was found to have the highest accuracy rate (as shown in Table 3) as it reaches the global minimum faster than other optimizers.

Table 3 Comparison of performance of optimizers for various learning rates

Optimizer	LR 10^{-2} (%)	LR 10^{-3} (%)	LR 10^{-4} (%)
SGD	55.8	52.09	51.7
NAG	57.46	59.04	54.91
AdaGrad	66.8	69.7	69.37
AdaDelta	62.17	64.33	62.8
Adam	67.25	68.6	71.82

The accuracy is computed by determining the area under the ROC curve or AUC-ROC. ROC is a probability curve which plots true positive rate (TPR) opposite to false positive rate (FPR). The higher the AUC-ROC, the better is recognition rate. Its value ranges between 0.5 and 1. If value is under 0.5, it means there is discrepancy in data or the model has been setup wrongly.

$$\mathrm{TPR} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \tag{4}$$

$$\mathrm{FPR} = \frac{\mathrm{FP}}{\mathrm{FP} + \mathrm{TN}} \tag{5}$$

where TP = True positives
TN = True negatives
FP = False positives
FN = False negatives.

Results and Analysis

Stochastic gradient descent (SGD) uses only one batch to perform iteration and hence gives the lowest accuracy as it takes a long time to converge. Nesterov Accelerated Gradient (NAG) also gives similar results it is based on SGD with an additional momentum to guide the gradient.

From the values in Table 3, it can be deduced that in terms of recognition rate, the performance among adaptive learning optimizers (optimizers where learning rate or α is adapted automatically with each iteration), such as Adaptive moment estimation (Adam), Adaptive Gradient (AdaGrad) and AdaDelta is high, specially at lower learning rates as they require slower LR to converge. Of these, it can be seen that Adam performs best on learning rate of 10^{-4}. The difference in these adaptive optimizers is the disparity in calculation of learning rates. Adam performs best as it uses estimations of first moment and second moment of the gradient for adapting LR.

Conclusion and Future Work

In this work, we presented a different approach for music genre classification which is based on texture images. Firstly, visual representations are created by converting audio files into spectrograms. Features are then extracted from these spectrograms that define musical characteristics in a format that is compatible with our model. Different variants of CNN model are created and the best optimzer is chosen and hyperparameter tuned to give better results of classification of the GTZAN dataset.

The future scope of this project is vast and can be performed keeping the economic and time factors in mind. The future scope includes combining different techniques to see the result of combined approach of algorithms. The project now works only for songs present in some specific languages. This work can be extended for other languages. Also, live listening to the music and genre prediction is one of the major future targets along with an interactive and futuristic UI.

References

1. Costa YMG, Oliveira LS, Koerich AL, Gouyon F (2011) Music genre recognition using spectrograms. In: 18th International conference on systems, signals and image processing. IEEE, Bosnia
2. Elbir A, Ilhan HO, Serbes G, Aydin G (2018) Short time Fourier transform based music genre classification. In: Electric electronics, computer science, biomedical engineerings' meeting (EBBT). IEEE, Turkey
3. Costa YMG, Oliveira LS, Koerich AL, Gouyon F, Martins JG (2012) Music genre classification using LBP textural features. In: Signal processing, vol 92 (no 11), pp 2723–2737. Elsevier
4. Aguiar RL, Costa YMG, Nanni L (2016) Music genre recognition using spectrograms with harmonic-percussive Sound Separation. In: 35th International conference of the Chilean computer science society (SCCC). IEEE, Chile
5. Panwar S, Das A, Roopaei M, Rad P (2017) A deep learning approach for mapping music genres. In: 12th System of systems engineering conference (SoSE). IEEE, Waikoloa
6. Ghosal D, Kolekar MH (2018) Music genre and style recognition using deep neural networks and transfer learning. In: APSIPA Annual summit and conference, pp 1010–1018
7. Markov K, Matsui T (2014) Music genre and emotion recognition using Gaussian processes. IEEE Access 2:688–697
8. Tzanetakis G, Cook P (2002) Musical genre classification of audio signals. In: IEEE transactions on speech and audio processing, vol 10 (no 5), pp 293–302. IEEE
9. Pereira RM, Silla CN (2017) Using simplified chords sequences to classify songs genres. In: 18th International conference on systems, signals and image processing. IEEE, Hong Kong, pp 1446–1451
10. Song Y, Zhang C, Xiang S (2007) Semi-supervised music genre classification. In: IEEE International conference on acoustics, speech and signal processing—ICASSP '07, II-729–II-732. IEEE, Honolulu
11. Serwach M, Stasiak B (2016) GA-based parameterization and feature selection for automatic music genre recognition. In: 17th International conference computational problems of electrical engineering (CPEE). IEEE, Poland
12. Nayak S, Bhutani A (2011) Music genre classification using GA-induced minimal feature-set. In: Third national conference on computer vision, pattern recognition, image processing and graphics. IEEE, India, pp 33–36

13. Ren J, Jang JR (2012) Discovering time-constrained sequential patterns for music genre classification. In: IEEE transactions on audio, speech, and language processing. IEEE, Taiwan
14. Sturm BL (2013) Evaluating music emotion recognition: lessons from music genre recognition? In: IEEE international conference on multimedia and expo workshops (ICMEW). IEEE, San Jose
15. Lin C, Liu M, Hsiung W, Jhang J (2016) Music emotion recognition based on two-level support vector classification. In: International conference on machine learning and cybernetics (ICMLC). IEEE, Jeju, pp 375–389
16. Martins de Sousa J, Pereira ET, Veloso LR (2016) A robust music genre classification approach for global and regional music datasets evaluation. In: IEEE international conference on digital signal processing (DSP). IEEE, Beijing, pp 109–113
17. Zhong G, Wang LN, Ling X, Dong J (2016) An overview on data representation learning: from traditional feature learning to recent deep learning. J Fin Data Sci 2(4):265–278
18. Pandian AP (2019) Review of machine learning techniques for voluminous information management. J Soft Comput Paradigm (JSCP) 1(2):103–112
19. https://towardsdatascience.com/using-cnns-and-rnns-for-music-genre-recognition-2435fb2ed6af
20. https://librosa.org/doc/latest/index.html
21. https://keras.io/

Utilization of Machine Learning Techniques for Prediction of COVID-19 Epidemic

Tasmiah Rahman and Most. Nure Yesmin

Abstract COVID-19 pandemic is a deadly impact on the health and well-being of the world population. A developing country like Bangladesh has limited medical resources, and sometimes many people cannot get proper treatment in time. A continued increasing number of people tested positive for COVID-19 has caused a lot of strain on the governing bodies across the country, and they face difficulties to handle this situation. The aim of this work is to analyze the symptoms and predict the chances to get infected with COVID-19 disease. Five different machine learning algorithms are utilized to predict COVID-19 based on symptoms. Random forest, support vector machine, logistic regression, Gaussian Naive Bayes, and K-nearest neighbor algorithms have been used. We compare the performance before and after applying principal component analysis. The performance of K-nearest neighbor found the more accurate result before and after applying principal component analysis.

Keywords Machine learning · K-nearest neighbor · Principal component analysis · Bangladesh · COVID-19 · Treatment

Introduction

COVID-19 is one of the biggest challenges that the world is ever faced. It occurred in December 2019, in Wuhan, China, and spreading rapidly around the whole world. Novel coronavirus outbreak is announced as a worldwide pandemic by the World Health Organization (WHO) on March 2020. Severe acute respiratory syndrome coronavirus2 (SARS-CoV-2) is responsible for this disease. On August 17, 2021, the WHO reported that 207,784,507 people have been infected with COVID-19,

T. Rahman (✉) · Most. N. Yesmin
Department of Computer Science and Engineering, Daffodil International University, Dhaka, Bangladesh
e-mail: tasmiah15-8805@diu.edu.bd

Most. N. Yesmin
e-mail: nure15-9386@diu.edu.bd

V. Bindhu et al. (eds.), *Proceedings of Third International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 844,
https://doi.org/10.1007/978-981-16-8862-1_48

of which 4,370,424 have died globally [1]. Like other countries, this year COVID-19 has taken a terrible turn in Bangladesh. A developing country like Bangladesh has faced many difficulties to handle this terrible situation. In Bangladesh, the first COVID-19 infected patient was found on March 8, 2020. Since March 8, 2020, till August 17, 2021, 1,433,396 people have been affected and 24,547 have been died because of COVID-19 [2].

People are affected in COVID-19 various ways. COVID-19 affected people have a comprehensive scope of syndrome declare—going from gentle manifestations to serious disease. The virus enters a person's eyes, nose, mouth, and body through the hands of a person infected with the virus. The virus moves down our respiratory tract. Syndromes can seem 2–14 days' later openness to the infection. Anybody may have gentle for the sake of serious side effects. The virus then spreads to the lungs, infecting the small air sacs inside the lungs, resulting in shortness of breath in the patient. Signs and indications incorporate respiratory manifestations and include tiredness, fever, headache, dry cough, nausea, sore throat, runny nose, and diarrhea. The elder grown-up person with heart disease, such as lung disease or diabetes, is thought to be at high risk of being affected by coronavirus.

There have a large number of various kinds of variations of COVID coursing across the world. Gamma (P.1), delta (B.1.617.2), beta (B.1.351), and alpha (B.1.1.7) are some of the notable variants of COVID-19. Some important information among those variants:

Delta (B.1.617.2): This variation is presently the most well-known COVID-19 variation in the USA. It seems, by all accounts, to be spreading rapidly in numerous nations including the UK, where it has gotten the prevailing variation.

Alpha (B.1.1.7): It is a half expansion in transmission contrasted with past coursing variations. This variation additionally with extended risk of hospitalization and died.

Gamma (P.1): It decreases the viability of some monoclonal counteracting agent meds.

Beta (B.1.351): This variant is spread very faster than other variants of COVID-19 virus. It has an E484K mutation which changes the shape of a main part of the coronavirus spike protein and assists to elude antibodies useful against other variants.

Numerous vaccines have been developed to prevent coronavirus. Four types of COVID vaccines are approved. Those four types of COVID vaccines are Moderna, AstraZeneca/University of Oxford, Pfizer/BioNTech, and Sinopharm. A person needs to take two doses of the vaccine, 28 days apart, to complete the vaccination schedule. Defensive degrees of antibodies are by and large created fourteen days after receiving the second portion of the COVID-19 vaccine. These vaccines have not been able to completely control the virus but have been able to produce 60 to 90% of antibodies in the body. Artificial intelligence (AI) can help us in dealing with this pandemic.

In this epoch of technology different machine learning algorithms are used to develop the software to predict disease more accurately, so that doctors can give proper advice and proper treatment to the patients. In this work, several machine learning models are used to find the model which provides the best accuracy. By implementing this model, we will be able to reduce the death rate of the COVID-19 infected people.

Related Work

Many studies have already been conducted using various artificial intelligence to diagnose and foresee COVID-19 disease and recuperation. The related works that have been inspected so far show that ML strategies and other artificial knowledge methods played significant parts in the expectation, finding, and control of the COVID-19 pandemic, which can assist with diminishing the immense death on restricted medical care frameworks.

USA scientists Hoyt et al. published an article in 2020 where they used 197 United States patients' data. They used logistic regression algorithm which showed 12.58 diagnostic odds ratio that is higher for forecasting ventilation and practically triage COVID-19 affected people than a comparator quickly warning system, like Modified Early Warning Score (MEWS) that found the sensitivity is 0.78 where machine algorithm found the sensitivity is 0.90. The algorithm also showed ($p < 0.05$) higher specificity than MEWS [3]. Kolla et al. discussed with Gaussian Naïve Bayesian classifier, SVM, KNN + NCA, XGBoost classifier, random forest classifier, multilinear regression, and decision tree classifier; the machine learning models compare seven different algorithms. They find out high accuracy from using random forest regressor [4]. Lamiaa et al. conducted a study in Egypt by using 5000 COVID-19 patients' data. Supervised learning method are used for implementing this model. Then, they utilize regression analysis as a selected algorithm. Polynomial models like fourth-degree, sixth-degree, and fifth-degree are performed brilliant for finding the results. Specially, the fourth-degree model will assist the government to ready their process for one month [5]. Akib et al. in India expelled GitHub dataset where 1000 cases and 212 reports. In machine learning applications, they utilized supervised learning method for implementation. They applied eight different machine learning algorithms which are multinomial naive Bayes, support vector machine, bagging, logistic regression, adaboost, decision tree, stochastic gradient boosting, random forest and found the highest accuracy 96.2% from multinomial naive Bayes and logistic regression algorithms [6]. In another study in Italy, Luca et al. proposed a model where they used supervised learning method. They utilized K-nearest neighbor (KNN) classifier. The target of their proposed method was to detect the COVID-19 disease by using chest X-rays images. A total of 85 chest X-rays image is used for analyzing the model. The model shows the effectiveness and difference between other pulmonary diseases and COVID-19 disease [7]. R. Sujath et al. have introduced a model that could be help of to anticipate the spread of COVID-2019. Their dataset collects from Johan Hopkins University. They used linear regression, multilayer perceptron, and vector auto-regression methods where they can infer that the MLP technique is giving acceptable forecast results [8]. In another place, Maria et al. study to develop a machine learning (ML) which surveys taste disorders, the predictive value of smell, alongside different indications, COVID-19. Their data was collected from four different hospitals in the Spanish regions of Seville and Cadiz between March and April 2020. Support vector machine (SVM), logistic regression (LR), and random forest (RF) algorithms are utilized in their model. An average accuracy of 80%, a specificity of 78%, and sensitivity of 82% the results they find out by using those ML algorithms [9]. Yazzed et al. was used gradient boosting model. The dataset was collected from the Israel Ministry of health where 51,831 patients information is recorded. The model determined 0.90 auROC score with 95% CI:0.892–0.905 using test set where 71.98% specificity and 87.30% sensitivity or 79.18% specificity and 85.76% sensitivity. They also showed positive predictive value against sensitivity across auPRC of 0.66 with 95% CI:0.647–0.678 [10]. Aman et al. show that then 10 high population and high dense country for detect of patients where they applied nine different ML algorithm. Those algorithms were linear regressor polynomial (LRP), Bayesian ridge polynomial regressor (BRR), Holt-Winters exponential smoothing (HW), support vector regressor (SVR), auto-regressive integrated moving average (ARIMA), random forest regressor (RFR), XGBoost regressor (XGB), linear regression (LR), and auto-regressive moving average (ARMA). The model finds the highest accuracy 85%, which find by using the ARMA algorithm [11]. L. J. et al. proposed a model for COVID-19 infection where they applied five different machine learning (ML) algorithms like support vector machine, logistic regression, artificial neural network, naïve Bayes, and decision tree. They have collected a dataset from the General Directorate of Epidemiology, Secretariat of Health in Mexico. The model evaluates the highest accuracy of 94.99% which find by using the decision tree algorithm [12]. Dan et al. in Israel collected patients information from Sheba Medical Center that is utilized in that research as data. They utilized artificial neural network (ANN) algorithm's used APACHE II score, duration from symptoms to admission, white blood cell (WBC) counts, blood lymphocytes, and oxygen saturation count these variables for utilized the model. As compared to other techniques, the results show that machine learning algorithms found higher efficiency for the prediction of COVID-19 disease [13]. In Germany, Constantin et al. proposed a model which is built by neural network for analyzing data. They utilize 500 chest CT scans and 152 data of COVID-19 patients. They combined clinical software and machine learning to develop the platform early deployment, and lastly, they acquired the method for completely automated segmentation which is used in medical even in a complex case [14]. Chansik et al. have used five different ML algorithms; those are KNN, RBF-SVM, linear SVM, LASSO, and RF. The dataset was collected from The Korean National Health Insurance Service where patients are newly affected in Covid-19. Applying linear support vector machine and LASSO machine learning algorithms, they find out that the receiver operating characteristic curve (AUG (>0.9) which is high areas under sensitivities (>83% for 14- or 30-day mortality and > 90% for final mortality) specificities without compromising specificity (around 90%) and high balanced accuracies (>86% for 14- or 30-day mortality and > 91% for final mortality) [15]. Mr. C. Vijesh et al. have to make a paper where they introduce a recommender system which depends on location and orientation-aware that built with different work like ambiguity handling, profile making, item extracting, location detection, recommendation making, and orientation detection [16]. Dr. Joy et al. show a prognosis model that they used naïve Bayes and SVM machine learning algorithms. Applying those methods, they find out that the SVM method provides highest accuracy [17]. B. Thilaka et al. published an article where they find an approach to revoke a deliberate retirement scheme reck of a probability of adoption retirement desire. It cost due to declaring an intentional retirement plan and costs to the association because of one-time exceptional installments to the individuals who resign during the period examined [18].

Materials and Methods

In this work, we will follow some steps for finding our expected model. At first, we need to collect a dataset then preprocess the dataset by using different techniques. The selected algorithms are then trained and tested on the preprocessed dataset. The results are calculated and evaluated to identify the best model for COVID-19 prediction. The working process of this implementation is given in Fig. 1.

Dataset Information

Data collection is the most important task in any research work. In this research work, the dataset is collected from different hospitals in Bangladesh from the people who have COVID-19 symptoms and the people who are already infected in COVID-19 by using surveying and Internet. We are surveying 1933 people for collecting information. COVID-19 symptoms have been used to make this dataset. In our dataset, there are 21 features that are used to predict COVID-19 disease. There are 862 female and 1071 male information in our dataset. The COVID-19 target data are categorized into two groups. Group 1 represents the people who are infected by COVID-19 disease, and group 2 represents the people who are not infected by COVID-19. Early diagnosis and taking proper treatment of COVID-19 can save peoples life (Fig. 2; Table 1).

Dataset Preprocessing

It is very difficult to find an organized data in the real world. Real-life data are unorganized, incomplete, lack important behavior, and sometimes contain many errors. Data preprocessing is a proven technique to reduce this problem. It prepares the raw data for generating it applicable to build various machine learning algorithms. In our dataset, there are some missing values. We replace these missing values with the mean and median values of that column. Nominal missing values are replaced by the

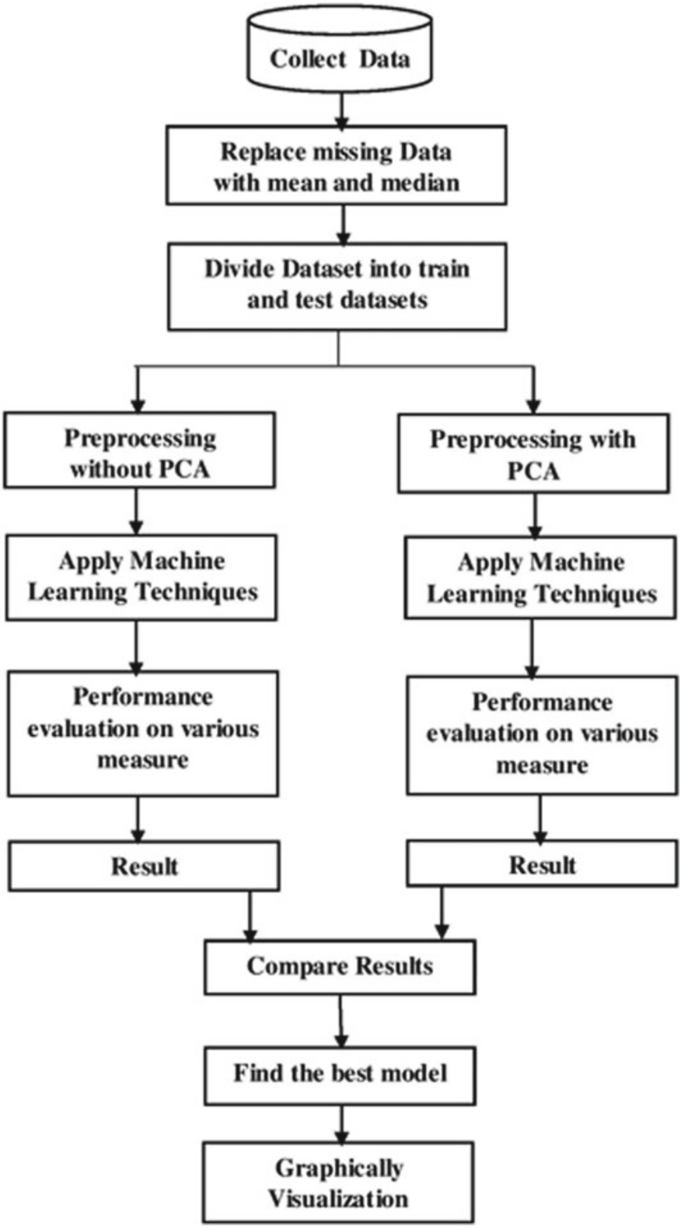


Fig. 1 Work flow of the activities

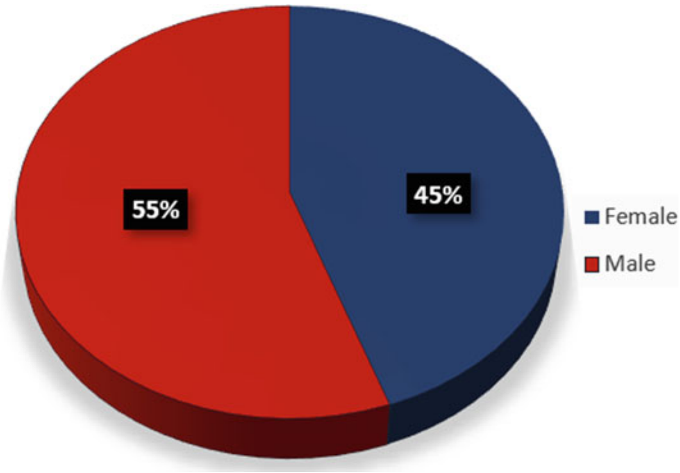


Fig. 2 Gender distribution of the dataset

Table 1 Description of dataset

S. No.	Attributes	Type	S. No.	Attributes	Type
1	Age	Numeric	12	Chills	Nominal
2	Gender	Nominal	13	Nausea	Nominal
3	Fever	Nominal	14	Conjunctivitis	Nominal
4	Oxygen Level	Numeric	15	Loss of movement	Nominal
5	Dry cough	Nominal	16	Chest Pain	Nominal
6	Sore throat	Nominal	17	Diabetes	Nominal
7	Tiredness	Nominal	18	High blood pressure	Nominal
8	Diarrhea	Nominal	19	Asthma	Nominal
9	Loss of taste	Nominal	20	Obesity	Nominal
10	Body aches	Nominal	21	Lung disease	Nominal
11	Runny nose	Nominal	22	COVID-19	Nominal

median value, and numeric missing values are replaced by the mean value of that columns.

Feature Selection Techniques

High-dimensional data analysis is very strenuous in the field of machine learning. When lots of features are used to predict the target value, sometimes model becomes confused and creates an overfitting problem. Solving this problem by removing immaterial and less important features by using efficient feature selection techniques, feature selection can reduce the computation time, help for understanding the model and data, and measure the outcome accurately. We can select features by using many different techniques. Automatic feature selection is a technique that removes the less important features not even changing anything. Dimensionally, reduction methodology is used to reduce the number of features without changing model performance. It makes a new combination of attributes. There are two dimensionally reduction techniques, the first one is recursive feature elimination (RFE), and the other is principal component analysis (PCA). In this research work, we use principal component analysis technique.

Principal Component Analysis

PCA is a dimensionality reduction technique that is utilized to shorten the dimension of a large dataset. It transforms the large set of attributes into the set that contains maximum information of the large dataset. Smaller dataset is easier to analyze. In PCA, N-dimensional of feature reduced into K dimension of feature where $K < N$. PCA works by converting highly correlated attributes to a shorter number of linearly uncorrelated attributes called principal components.

In our dataset, there are 1933 instances and 21 features. So, 21 is the maximum number of PCA. After applying PCA, our dataset is reduced to 10 principal components from 21 components.

Train Test Split

Train test split method is used for calculating the performance of machine learning algorithms. In these techniques, the dataset is divided into two parts, the large part is used to fit the model known as training dataset, and the other part is used to test the prediction of the model. In this work, we use 75% data for training the model and 25% data for testing the prediction.

Algorithms

To implement the result, we used five different machine learning-based algorithms that are used. They are K-nearest neighbor, logistic regression, support vector machine, Gaussian Naive Bayes, and random forest, and are compared these algorithms for a common dataset. These algorithms are compared based on different parameters.

Random Forest. A random forest (RF) classification was applied to ghastly just as mono- and multi-occasional textural highlights removed from Landsat TM symbolism to build the exactness of land cover characterization over a perplexing Mediterranean scene, with an enormous number of land cover classifications and low between class distinguishableness. Random forests make resolution trees on randomly selected data samples which get count from each tree and selects the best classification through casting voting. It likewise gives a very decent pointer of the element significance.

Support Vector Machine. Support vector machine (SVM) is a bunch of regulated learning techniques utilized for classification. By using SVM, linear and nonlinear problems can be solved. It shares a technique which is called the transformation of data of kernel trick. This kernel trick can transform the data based on transformation and get the optimal boundary between the possible outputs. SVM can maximize the margin by separating the hyperplane. It does some amazingly unpredictable data transformations, then, at that point sorts out some way to separate our data dependent on the marks or outputs we have characterized. Support vector machine algorithm's main benefit is that it can capture complex data and can make good relationships between data points without having performances on difficult transformation.

Logistic Regression. Logistic regression is a statistical analysis method. It has a place with the convention of linear classifiers and is somewhat like polynomial and linear regression. This is speedy and tolerably direct, and this logistic regression is beneficial for us to translate the results. Despite the fact that it is a technique for twofold plan, it can moreover be used to various issues. The outcome of the probability is measured by the odds of occurrence of an event. The logistic regression algorithm calculation numerical equation number (1) can be used to ascertain with the relationship among subordinate elements:

$$i = \ln\left(\frac{p}{1-p}\right) \tag{1}$$

Here, $\left(\frac{p}{1-p}\right)$ is the odd ratio. If P is the probability of an event, then $(1 - P)$ is the probability of it not occurring.

K-Nearest Neighbor. K-nearest neighbor is a simple, supervised learning algorithm that is used in machine learning for solving classification and regression problems. This algorithm basically uses data, and then, it classifies the new data points based on the same function. If the dataset is small, then KNN algorithm is a very handful to running in a shorter time. This algorithm's main target is to determine the value of a new data point based on which near it is to other data points that before exist in the dataset. This algorithm basically uses data, and then, it classifies the new data points based on same function. It can only store the complete data but cannot do any calculation at this point. It cannot make an assumption on underlying data, and for this reason, KNN is a nonparametric algorithm.

Gaussian Naive Bayes. Gaussian Naive Bayes is a machine learning algorithm which is used as conforming to Gaussian distribution, and it supports continuous valued feature. It is a powerful algorithm for imperative analysis. To find out a model correction or wrong output and accuracy, this algorithm can be used as machine learning algorithm. It is exceptionally adaptable with the quantity of indicators and information focuses. It is quick and can be utilized to make constant expectations. For multi-class prediction problems solution, Naive Bayes is very essential algorithm. At the same times, the probability is based on this theorem.

Experimental Result and Analysis

For the implemented result, we split our collected dataset into two sections 0.75% of data are utilized as training dataset and 25% data are utilized to test the model. By using this dataset, Gaussian Naive Bayes, support vector machine, logistic regression, K-nearest neighbor, and random forest algorithms are applied to predict the outcome. We also built confusion matrix for measure accuracy, precision, recall, $F1$-measure, and AUC. Confusion matrix has four parts: true positive, false negative, false positive, and true negative.

Evaluation Matrix

$$\text{Accuracy} = \left[(\text{No. of True Positive} + \text{No. of True Negative})/\text{Total}\right] \tag{2}$$

$$\text{Precision}(P) = \left[\text{No. of True positive}/\text{Predicted Yes}\right] \tag{3}$$

$$\text{Recall}(R) = [\text{No. of True Positive}/\text{Actual Yes}] \tag{4}$$

$$\text{F1} - \text{Measure} = [2 * (P * R)/(P + R)] \tag{5}$$

Area Under Curve (AUC). AUC is a graphical plot which is draw by the value of true positive rate and false positive rate. AUC is utilized for ascertaining the performance of a model. When the value of AUC is higher, it means the model is performed better.

Result Analysis

We use different machine algorithm for finding the best model which provides the highest accuracy. We compared these algorithms before and after applying feature selection techniques.

In Table 2, we see that before principal component analysis, we apply five machine learning algorithms which are Gaussian Naive Bayes, support vector machine, random forest, logistic regression, and K-nearest neighbor for assessing the performance of these models. We compare the accuracy and find the best accuracy in K-nearest neighbor model which is 0. 832. From Table 2, we see that the precision, recall, $F1$-measure, and AUC score which are 0.878,0.881, 0.880, and 0.829, respectively, also justify that K-nearest neighbor is the best algorithm among all algorithms which we applied for our dataset.

In Table 3 we see that after reducing our dataset by ten principal components, we apply five machine learning algorithms which are Gaussian Naive Bayes, support vector machine, random forest, logistic regression, and K-nearest neighbor for assessing the performance of these models. We compare the accuracy, and like Table 2, here we also find the best accuracy in K-nearest neighbor model which is 0. 834. From Table 2 we see that the precision, recall, $F1$-measure, and AUC score which are 0.881, 0.881, 0.881, and 0.825, respectively, also justify that K-nearest neighbor is the best algorithm among all algorithms which we applied for our dataset (Fig. 3).

Table 2 Performance of different model before PCA

	Precision	Recall	$F1$-measure	Accuracy	AUC
Random forest	0.874	0.848	0.858	0.826	0.824
Support vector machine	0.875	0.851	0.863	0.811	0.798
Logistic regression	0.882	0.848	0.865	0.816	0.807
K-nearest neighbor	0.878	0.881	0.880	0.832	0.829
Gaussian Naive Bayes	0.825	0.854	0.839	0.772	0.814

Table 3 Performance of different model after PCA

	Precision	Recall	$F1$-measure	Accuracy	AUC
Random forest	0.879	0.869	0.876	0.818	0.809
Support vector machine	0.867	0.854	0.860	0.807	0.806
Logistic regression	0.866	0.845	0.855	0.801	0.813
K-nearest neighbor	0.881	0.881	0.881	0.834	0.825
Gaussian Naive Bayes	0.841	0.866	0.853	0.793	0.819

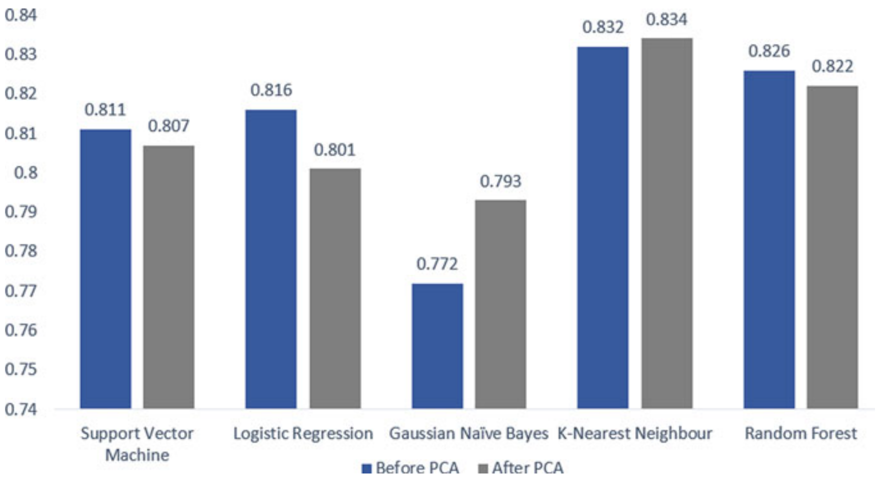


Fig. 3 Comparison of accuracy before and after applying PCA

Conclusion and Future Work

Now, COVID-19 is an ongoing global pandemic. Early detection of this disease can reduce the risk of health. In this study, five different machine learning algorithms are implemented which are Gaussian Naive Bayes, support vector machine, logistic regression, K-nearest neighbor, and random forest. We trained 75% data and tested 25% data in these models by using our collected dataset. These same algorithms are used for before principal component analysis and after principal component analysis where we use ten principal components. These results were compared by different statistical measures like accuracy, precision, recall, $F1$-measure, and AUC. The experimental results show that K-nearest neighbor algorithms provide the highest accuracy which is 0.832 before applying PCA and 0.834 after applying PCA. In future, we can improve this model by adding different essential features and advanced or combined algorithms.

References

1. WHO Coronavirus (COVID-19) Dashboard. WHO Coronavirus (COVID-19) Dashboard with vaccination data. https://covid19.who.int/. Last accessed 18 Aug 2021
2. 198 Covid-19 deaths, 7535 new cases reported in a day. The Daily Star. https://www.thedailystar.net/health/disease/coronavirus/events-who/deaths-infections/news/198-covid-19-deaths-7535-new-cases-reported-day-2154416. Last accessed 18 Aug 2021
3. Burdick H, Lam C, Mataraso S, Siefkas A, Braden G, Dellinger RP, McCoy A, Vincent JL, Green- A, Barnes G, Hoffman J, Calvert J, Pellegrini E, Das R (2020) Prediction of respiratory decompensation in Covid-19 patients using machine learning: the READY trial. Comput Biol Med 124:103949. https://doi.org/10.1016/j.compbiomed.2020.103949

4. Prakash KB, Imambi SS, Ismail M, Kumar TP, Pawan YVRN (2020) Analysis, prediction and evaluation of COVID-19 datasets. Int J Emerg Trends Eng Res 8:2199–2204
5. Amar LA, Taha AA, Mohamed MY (2020) Prediction of the final size for COVID-19 epidemic using machine learning: a case study of Egypt. Infect Dis Model 5:622–634. https://doi.org/10.1016/j.idm.2020.08.008
6. Khanday AMUD, Rabani ST, Khan QR, Rouf N, Mohi Ud Din M (2020) Machine learning based approaches for detecting COVID-19 using clinical text data. Int J Inf Technol 12:731–739. https://doi.org/10.1007/s41870-020-00495-9
7. Brunese L, Martinelli F, Mercaldo F, Santone A (2020) Machine learning for coronavirus covid-19 detection from chest x-rays. Proc Comput Sci 176:2212–2221. https://doi.org/10.1016/j.procs.2020.09.258
8. Sujath R, Chatterjee JM, Hassanien AE (2020) A machine learning forecasting model for COVID-19 pandemic in India. Stoch Environ Res Risk Assess 34:959–972. https://doi.org/10.1007/s00477-020-01827-8
9. Callejon-Leblic MA, Moreno-Luna R, Del Cuvillo A, Reyes- IM, Garcia- MA, Santos-Peña M, Maza JM, Martín-Jimenez DI, Palacios-Garcia JM, Fernandez-Velez C, Gonzalez-Garcia J, Sanchez- JM, Solanellas-Soler J, Sanchez-Gomez S (2021) Loss of smell and taste can accurately predict COVID-19 infection: a machine-learning approach. J Clin Med 10:570. https://doi.org/10.3390/jcm10040570
10. Zoabi Y, Deri-Rozov S, Shomron N (2021) Machine learning-based prediction of COVID-19 diagnosis based on symptoms. npj Digit Med 4:1–5. https://doi.org/10.1038/s41746-020-00372-6
11. Khakharia A, Shah V, Jain S, Shah J, Tiwari A, Daphal P, Warang M, Mehendale N (2021) Outbreak prediction of COVID-19 for dense and populated countries using machine learning. Ann Data Sci 8. https://doi.org/10.1007/s40745-020-00314-9
12. Muhammad LJ, Algehyne EA, Usman SS, Ahmad A, Chakraborty C, Mohammed IA (2021) Supervised machine learning models for prediction of COVID-19 infection using epidemiology dataset. SN Comput Sci 2. https://doi.org/10.1007/s42979-020-00394-7
13. Assaf D, Gutman Y, Neuman Y, Segal G, Amit S, Gefen-Halevi S, Shilo N, Epstein A, Mor R, Biber A, Rahav G, Levy I, Tirosh A (2020) Utilization of machine-learning models to accurately predict the risk for critical COVID-19. Int Emerg Med 15:1435–1443. https://doi.org/10.1007/s11739-020-02475-0
14. Anastasopoulos C, Weikert T, Yang S, Abdulkadir A, Schmülling L, Bühler C, Paciolla F, Sexauer R, Cyriac J, Nesic I, Twerenbold R, Bremerich J, Stieltjes B, Sauter AW, Sommer G (2020) Development and clinical implementation of tailored image analysis tools for COVID-19 in the midst of the pandemic: the synergetic effect of an open, clinically embedded software development platform and machine learning. Eur J Radiol 131:109233. https://doi.org/10.1016/j.ejrad.2020.109233
15. An C, Lim H, Kim DW, Chang JH, Choi YJ, Kim SW (2020) Machine learning prediction for mortality of patients diagnosed with COVID-19: a nationwide Korean cohort study. Sci Rep 10:1–11. https://doi.org/10.1038/s41598-020-75767-2
16. Joe MCV, Raj JS (2021) Location-based orientation context dependent recommender system for users. J Trends Comput Sci Smart Technol 3:14–23. https://doi.org/10.36548/jtcsst.2021.1.002
17. Chen JIZ, Hengjinda P (2021) Early prediction of coronary artery disease (CAD) by machine learning method—a comparative study. J Artif Intell Capsul Networks 3:17–33. https://doi.org/10.36548/jaicn.2021.1.002.
18. Thilaka B, Theetharappan N (2020) Optimal time for withdrawal of voluntary retirement scheme with a time-varying threshold. In: 2nd International conference on innovative mechanisms for industry applications. ICIMIA 2020—Conference Proceedings, vol 02, pp 598–602. https://doi.org/10.1109/ICIMIA48430.2020.9074885

Power-Efficient Bulk-Driven MCML D-Latch for High-Frequency Applications

Manikantha Vallabhaneni, Sreenidhi Balki, P. S. V. N. K. Mani Gupta, and Sonali Agrawal

Abstract This paper presents three different circuits for D-latch designed using bulk-driven MCML topology for high-frequency and low-power applications. The BD-MCML topology is the combination of the bulk-driven technology with MCML technology. The number of MOSFETs used in BD-MCML topology is less compared to MCML topology as the BD-MCML topology uses bulk terminal as one of the input. The three proposed circuits are designed using 45 nm technology and analyzed using Cadence Virtuoso with a supply voltage of 0.6 V. The performance of all the proposed D-latch circuits is evaluated and compared with the existing bulk-driven MCML D-latch circuit concerning metrics like power, delay, rise time, and fall time. This paper concludes that the proposed BD-MCML D-latch circuits are power efficient when compared with the existing BD-MCML D-latch.

Keywords MOS current mode logic (MCML) · Bulk-driven MCML (BD-MCML) · Body-bias effect · Sub-threshold region · D-latch · Current mirror circuit · CMOS · Current steering

Introduction

The exponential growth in the technology is demanding optimization in area, power, and delay of the circuits. The technology shrinking helps in designing the circuits with minimum area consumption. But, portable electronic circuits are demanding low-power and high-speed circuits. The increase in the bandwidth of the wireless communication networks is demanding circuits that can operate with constant power in the large frequency ranges [1]. To satisfy these demands, many new topologies are used to implement the logic circuits.

M. Vallabhaneni · S. Balki · P. S. V. N. K. M. Gupta · S. Agrawal (✉)
Department of Electronics and Communication Engineering, Amrita School of Engineering, Amrita Vishwa Vidyapeetham, Bengaluru, India
e-mail: a_sonali@blr.amrita.edu

V. Bindhu et al. (eds.), *Proceedings of Third International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 844,
https://doi.org/10.1007/978-981-16-8862-1_49

In VLSI, the most popular topology to design any circuit is CMOS. In conventional CMOS circuit, duality is an important property. The CMOS circuits have both pull-up networks designed using pMOS transistors and pull-down network designed using nMOS transistors. Based on the inputs, either pull-up network or pull-down network works to generate the output of the circuit [2]. In this CMOS topology, the power dissipation is directly proportional to the operating frequency of the circuit [3]. In communication and networking applications, the circuits operate in a wide range of frequencies so the power consumption of the conventional CMOS circuits in these applications varies with respect to the operating frequency and is very high at higher frequencies [4, 5].

To design the circuit which is independent of the frequency, a new topology called current mode logic (CML) is used. Unlike CMOS topology, the transistors in CML topology are always ON (fully or partially). In high-speed application, the CML topology is used [6]. In CML topology, bipolar junction transistors are used. If these BJTs are replaced by MOSFETs in CML topology, then it is called MCML topology. In MCML topology, the logic function is implemented using a pull-down network and two pMOS transistors, which act as load resistors, are used to form a pull-up network. As the constant current source is connected to the circuit, the current in the circuit is always constant. The current flowing through the circuit is always constant because of the current source [7]. This concludes that in MCML design the power consumption is independent of operating frequency and is always constant and static in nature.

The BD-MCML topology helps in designing circuits with low-power consumption. This topology uses low supply voltages which helps in designing low-power structures. In BD-MCML topology, MOSFETs are operated in the sub-threshold region of operation by using the bulk terminal as one of the input [8]. So, the number of MOSFETs in BD-MCML topology is less compared to MCML topology to design any particular circuit. The attributes of sub-threshold region are used in designing circuits with low supply voltage. The BD-MCML topology has good performance over MCML topology for high-speed applications. Similar to MCML topology, the nature of consumed power in the BD-MCML circuit is static [9].

In this paper, three novel designs of D-latch are proposed. These circuits are designed using BD-MCML topology. The proposed circuits have low power as the circuits work only for half clock cycle. The operating voltage of the proposed circuits is less as the transistors work in sub-threshold region. To compare the performance of the proposed and existing BD-MCML D-Latch, various parameters are considered like delay, power, fall time, and rise time. The performance analysis concludes that the proposed D-Latch designs are low-power and high-speed D-latch circuits when compared to the existing D-Latch design.

The rest of the contents of this paper are organized as follows: Section "Background" explains the background of BD-MCML topology. Section "Proposed BD-MCML D-Latch Circuits" provides explanation about all the three proposed BD-MCML D-latch circuits. Section "Performance Analysis" compares the performance of the proposed D-latch circuits with the existing circuit and Section "Conclusion" gives the conclusion about the chapter.

Background

The conventional CMOS circuits use pMOS and nMOS transistors in circuit designing. The input combinations to the circuit provide a path from voltage source to output through pMOS transistors or ground to output through nMOS transistors. Even though the CMOS topology has many advantages, the power consumed in conventional CMOS topology is directly proportional to frequency. So, the power consumption in CMOS designs at high frequency is very high and is dynamic in nature. This high-power consumption problem can be solved by using advanced circuit design topologies like MCML topology and BD-MCML topology. The MCML topology uses nMOS transistors to implement the logic and pMOS transistors work as load resistors. The MCML topology uses a constant current source. So, the power consumption in the MCML topology is constant with respect to the operating frequency and is static in nature [10]. The power consumed in MCML topology can be reduced further by using another advanced circuit designing technique called BD-MCML topology. Similar to MCML topology, the nMOS transistors are used in logic implementation and the circuits will have constant current source. In BD-MCML topology, the transistors operate in the sub-threshold region by using the body terminal of the MOSFET. This property of BD-MCML helps in designing the circuits with low supply voltage which in turn helps in designing low-power circuits [11].

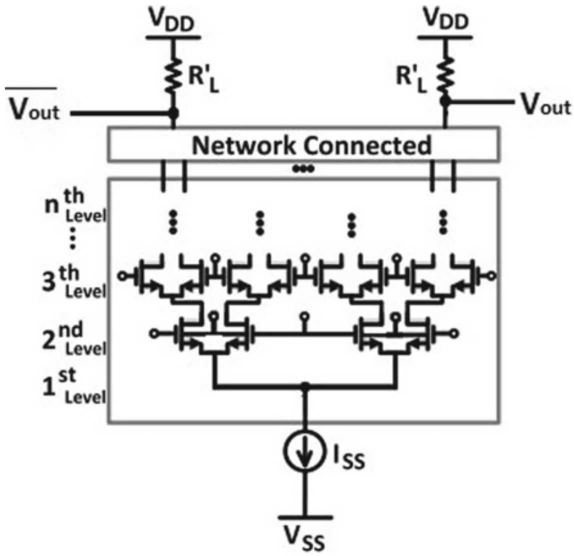
BD-MCML Topology

The combination of MCML topology and the bulk-driven technology forms BD-MCML topology. The transistors in BD-MCML topology work in the sub-threshold region, and this helps in designing circuits with low current and low supply voltages. The BD-MCML topology uses nMOS transistors to design logic function, and pMOS transistors help in current steering. The MCML topology and BD-MCML topology use a current mirror circuit which works as constant current source. The number of MOSFETs used in MCML topology is more compared to CMOS topology which can be resolved using BD-MCML topology by using the bulk terminal as one of the inputs to the MOSFET [12]. The power consumed in BD-MCML topology is always constant. The BD-MCML circuit is given in Fig. 1.

In Fig. 1, the nMOS transistors in different levels are used in implementing the logic function. The resistors in Fig. 1 are obtained by operating pMOS transistors in triode region of operation.

The MOSFETs bulk connections are classified into two categories: static and float bulk connections. In a static bulk connection, static voltages are connected to the bulk terminal. Generally, in static bulk connection, the bulk terminal of the MOSFETs is connected to the source terminal [13]. The MOSFET's threshold voltage depends on differential voltage between the Drain-Source terminal and Bulk-Source terminals for short-channel MOSFET as given in Eq. (1).

Fig. 1 BD-MCML circuit



$$V_{Th,p} = V_{T0} - \sigma V_{SD} + \gamma(\sqrt{|2\phi|} - \sqrt{|2\phi| + V_{BS}}) \quad (1)$$

where V_{SD} represents Drain-Source voltage, V_{BS} represents Bulk-Source voltage, V_{T0} denotes the threshold voltage of the MOSFET when the bulk terminal is shorted with source terminal. The symbols ϕ, γ, and σ represent the surface potential, coefficient of the body, and the coefficient of drain-induced barrier lowering effects, respectively.

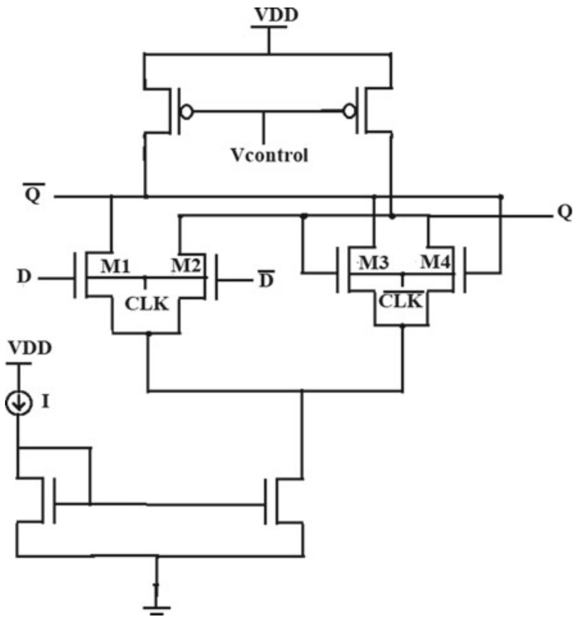
The bulk terminal in float bulk connection has different voltage levels. Using different voltage levels of the bulk terminal, the current through the MOSFET and the threshold voltage of the MOSFET can be altered. In BD-MCML circuits, this type of float bulk connection is used [14].

Existing BD-MCML D-Latch

The existing D-Latch circuit consists of eight MOSFETs; the logic function is implemented using four nMOS transistors, two nMOS transistors are used for the current mirror circuit, and they are operated in saturation region to make them work as a constant current source. Out of eight transistors, remaining two transistors are pMOS transistors and they act as load resistors [15]. The BD-MCML D-latch circuit is given in Fig. 2.

For this circuit, V_{DD} is the constant voltage and I_{SS} is the constant current source. The Q and Q' are differential outputs. If the clock and D inputs are high, then M1

Fig. 2 BD-MCML circuit



transistor works and the current flowing through this M1 transistor is more compared to the other transistors. So, the voltages at Q' and Q are low and high, respectively.

If D' is high (D is low) and clock is high, then the current flowing through the M2 transistor is high which makes the drop at load resistor connected to Q high so the voltage at Q is low and voltage at Q' is high.

If clock is low, then transistors M3 and M4 work. If M3 transistor is working, then more current flows into M3 so the output at Q' is low and less current passes through M4 that makes the output voltage at Q is high. If M4 transistor is working, then Q will be low. A small current will flow through M3 transistor so the voltage at Q' will be high.

Proposed BD-MCML D-Latch Circuits

The description and working of the three proposed bulk-driven MCML D-latch circuits are provided in this section. The truth table for D latch is shown in Table 1. Since one of the input is the bulk terminal, the number of transistors in the BD-MCML topology is less compared to MCML topology. With the help of the body terminal, the MOSFETs operate in sub-threshold region, which makes the circuit consume less power.

Table 1 Truth table of D-Latch

Clk	D	Q	$\overline{Q}$
0	0	Q	$\overline{Q}$
0	1	Q	$\overline{Q}$
1	0	0	1
1	1	1	0

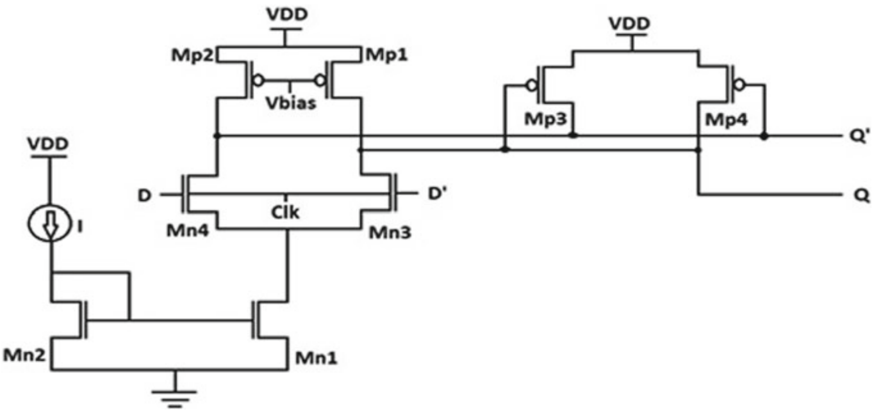


Fig. 3 Proposed 8-transistor BD-MCML D-Latch

Proposed 8-Transistor BD-MCML D-Latch

The proposed 8-Transistor circuit consists of 4 nMOS transistors and 4 pMOS transistors. A current mirror circuit is formed by using nMOS transistors Mn1 and Mn2, and this helps in providing the constant current to the rest of the circuit. The constant current source is important in the MCML topology; this is used to obtain the differential outputs. The nMOS transistors Mn3 and Mn4 are creating the pull-down network. This network operates when the clock input Clk = 1. The pMOS transistors Mp1 and Mp2 are load resistors. The other pMOS transistors, Mp3 and Mp4 work when clock input Clk = 0 and retain the previous outputs. The proposed 8-Transistor D-latch is given in Fig. 3.

If Clk = 1, the threshold voltage of the nMOS transistors Mn3 and Mn4 decreases. So the MOSFETs work, even if the input voltage is less than the nominal threshold. If Clk = 0, the nMOS transistors Mn3 and Mn4 threshold voltage are same as the nominal threshold voltage without body bias effect, so the MOSFETs do not work. In this case, the pMOS transistors Mp3 and Mp4 continue the previous output as present output.

The proposed circuit has static power less than the existing BD-MCML circuit. When Clk = 0, there is no complete path between supply and ground. So, the static power is consumed only during the positive half cycle of the clock signal. The

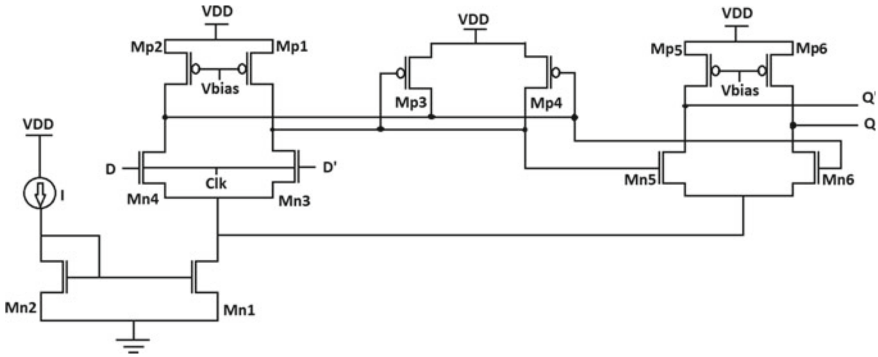


Fig. 4 Proposed 12-transistor BD-MCML D-Latch

proposed topology uses the voltage in the range [0 - V_{DD}], whereas in the existing model the voltage swing is in the range [$-V_{SS} - V_{DD}$].

Proposed 12-Transistor BD-MCML D-Latch

The proposed 12-Transistor D-latch circuit uses 6 nMOS transistors and 6 pMOS transistors. A current mirror circuit is formed by using nMOS transistors Mn1 and Mn2. The nMOS transistors Mn3 and Mn4 form the pull-down network. This network operates when the clock input Clk = 1. In this circuit, also pMOS transistors Mp1 and Mp2 act as load resistors. Mp3 and Mp4 pMOS transistors work when clock input Clk = 0 and retain the previous outputs. The remaining 2 nMOS transistors Mn5 and Mn6 with 2 pMOS transistors Mp5 and Mp6 work as MCML buffer. The proposed 12-Transistor D-Latch is given in Fig. 4.

The proposed 12-Transistor D-latch circuit works same as the proposed 8-Transistor circuit. The difference is in 12-Transistor circuit, and the output from the pMOS circuit is provided to the MCML buffer. This buffer circuit helps in reducing the rise time, fall time, and the distortion at the output.

Proposed 10-Transistor BD-MCML D-Latch

The proposed 10-Transistor D-latch uses 6 nMOS transistors and 4 pMOS transistors. Here also A current mirror circuit is formed by using nMOS transistors Mn1 and Mn2, and this helps in providing the constant current to the rest of the circuit to obtain the differential outputs. The nMOS transistors Mn3 and Mn4 are forming the pull-down network. This network operates when the clock input Clk = 1. The pMOS transistors Mp3 and Mp4 work when clock input Clk = 0 and retain the previous

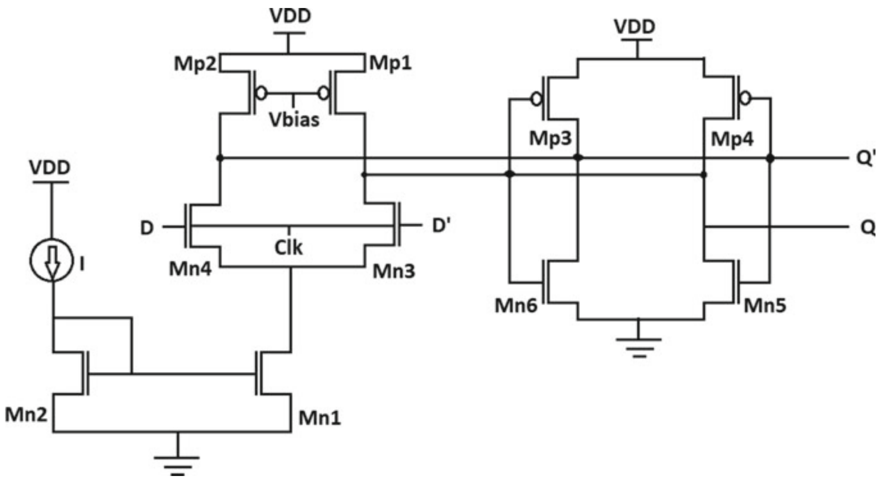


Fig. 5 Proposed 10-transistor BD-MCML D-Latch

outputs, whereas Mp1 and Mp2 act as load resistors. The proposed 10-Transistor D-Latch is given in Fig. 5.

The pMOS transistors Mp3 and Mp4 are used to restore the previous cycle high output, and the nMOS transistors Mn5 and Mn6 are used to restore the logic low state from the previous output when clock goes low.

Performance Analysis

The performance analysis of the proposed and existing D-latch circuits are done in 45 nm technology using Cadence Virtuoso. To compare the performance of the proposed and existing [13] BD-MCML D-Latch, we considered various parameters like delay, power, fall time, and rise time. The values of delay, power, fall time, and rise time for all the existing and proposed circuits are given in Table 2. All the circuits are simulated under same simulation environment with a supply voltage of 0.6 V, current source of 5 μA. The width of the transistor used is 120 nm and the length of the transistor is 45 nm. The value of parameters used in performance analysis is provided in Table 2.

The performance analysis and performance comparison of the proposed circuits with the exiting circuit are given in Tables 3 and 4, respectively.

The comparison of consumed power in existing and the proposed BD-MCML D-latches is provided as bar-graph in Fig. 6.

Figure 6 concludes that the power of the proposed 8-Transistor circuit is 55% less, the power of the proposed 12-Transistor circuit is 47% less, and the power of proposed 10-Transistor circuit is 35% less when compared with the existing circuit [13].

Table 2 Design specifications for existing and proposed D-Latch circuits

Design parameter	Existing BD-MCML D-Latch circuit [13]	Proposed BD-MCML D-Latch circuits
Supply voltage (V)	0.6	0.6
Constant current source value (μA)	5	5
Width of all the transistors (nm)	120	120
Length of all the transistors (nm)	45	45

Table 3 Performance analysis of existing and proposed BD-MCML D-Latch circuits

Parameter	Existing BD-MCML D-Latch circuit [13]	Proposed 8-transistor BD-MCML D-Latch circuit	Proposed 12-transistor BD-MCML D-Latch circuit	Proposed 10-transistor BD-MCML D-Latch circuit
Power (μW)	7.691	3.4	4.037	4.986
Rise time (ps)	381.4	348.6	75.17	106.3
Fall time (ps)	99.29	348.9	75.17	106.3
Delay (ps)	255.8	107.5	570.0	48.7

Table 4 Performance comparison between existing and proposed BD-MCML D-Latch circuits

% Change in proposed BD-MCML D-Latch w.r.t. Existing BD-MCML D-Latch [13]			
Parameter	8-transistor circuit	12-transistor circuit	10-transistor circuit
Power	55.4% Decreased	47.51% Decreased	35.15% Decreased
Rise Time	8.6% Decreased	80.29% Decreased	72.1% Decreased
Fall Time	71.4% Increased	24.29% Decreased	7.06% Increased
Delay	57.9% Decreased	55.15% Increased	80.96% Decreased

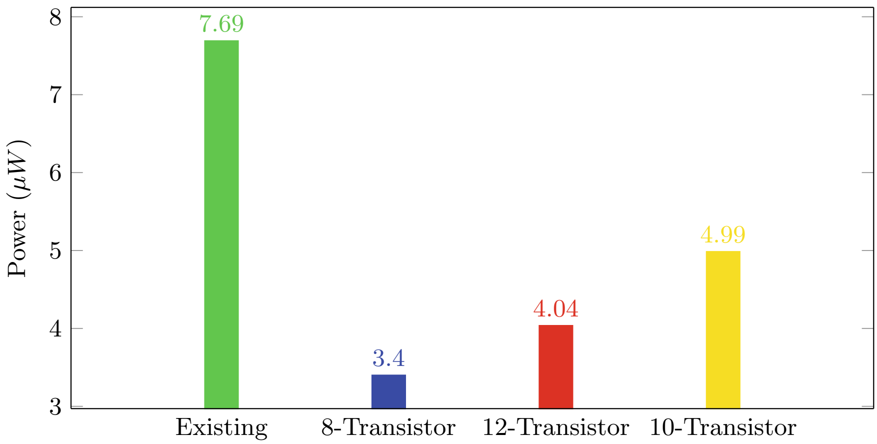


Fig. 6 Power consumed by D-Latch in existing and different proposed circuit

Figure 7 shows that the rise time of the proposed 8-Transistor, 12-Transistor, and 10-Transistor BD-MCML D-latch circuit is reduced by 8%, 80%, and 72%, respectively, compared to the existing BD-MCML D-latch. The rise time indicates that how much delay a circuits takes to change the output from 0 to 1. These results show that the 12-Transistor circuit is taking the minimum time to perform output transition from 0 to 1.

Figure 8 shows that the fall time of the proposed 8-Transistor BD-MCML circuit is 71% more than the existing BD-MCML circuit. The fall time of proposed 12-Transistor BD-MCML circuit is 24% less than the existing circuit and fall time of the proposed 10-Transistor circuit is 7% more than the existing BD-MCML circuit [13].

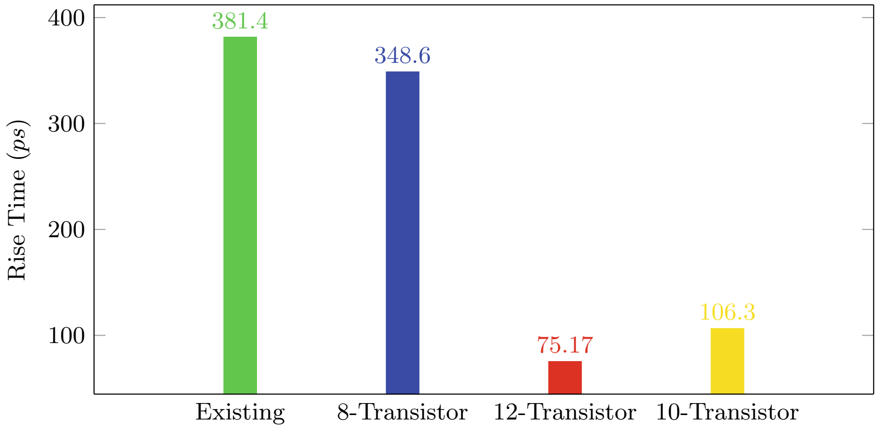


Fig. 7 Rise time of D-Latch in existing and different proposed circuits

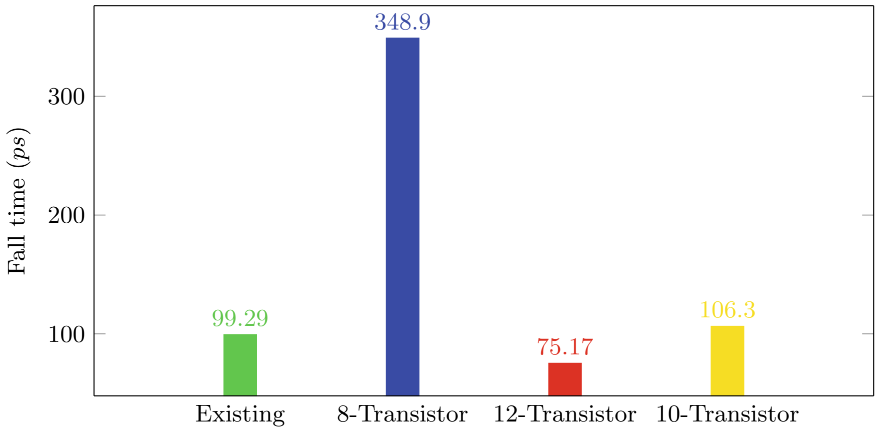


Fig. 8 Fall time of D-Latch in existing and different proposed circuits

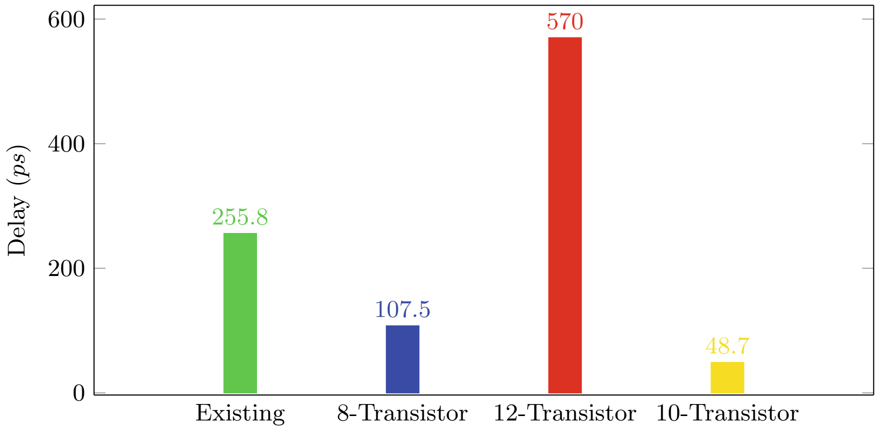


Fig. 9 Delay of D-Latch in existing and different proposed circuits.

The delay of the proposed 8-Transistor circuit is 57.9% less than the existing circuit. The delay of the proposed 12-Transistor D-latch circuit is 55% more than the existing circuit. The delay of proposed 10-Transistor circuit is 80% less than the existing D-latch circuit [13]. The delay comparison of the existing and proposed circuits is given in Fig. 9.

Conclusion

This chapter proposes three different BD-MCML D-latch circuits: 8-Transistor circuit, 12-Transistor circuit, and 10-Transistor circuit. The MOSFETs in BD-MCML topology operates in sub-threshold region by using body terminal as one of the input, and this also helps in reducing the area of the circuit. This also helps in designing the circuits with low supply voltage, which in turn helps in designing low-power circuits. As the current and voltage source values in the BD-MCML topology are constant, the power consumed by the circuit is also constant.

All the D-latch circuits are designed in 45 nm technology and analyzed using Cadence Virtuoso with supply voltage of 0.6 V and constant current source of 5 μA. The basic idea in proposed 8-Transistor circuit is that during the half cycle of the clock signal only the static power is consumed by the circuit. With this idea, the power consumed by proposed circuit has become less than half as compared to existing BD-MCML D-latch circuit [13]. In 12-Transistor circuit, an MCML buffer is connected at the output side to reduce rise and fall time but delay is increased. To reduce delay, 10-Transistor circuit is designed which consumes less power with less rise time and delay with a reasonable increase in fall time. This chapter concludes that the all the three proposed BD-MCML D-latch circuits are power efficient and independent of operating frequency. Finally, the proposed 10-Transistor circuit is the most efficient structure among all the proposed and existing D-latch structures for low-power and high-speed applications where the range of operating frequency is high. Further, to reduce rise and fall time in the BD-MCML D-latch, work can be done in the output capacitance part of the circuit.

References

1. Sarkar D, Bhattacharjee P, Majumder A (2018) Data-dependent clock gating approach for low power sequential system. In: 5th International conference on microelectronics, circuits & systems
2. Arya RK, Agrawal S (2019) Design of efficient 2–4 modified mixed logic design decoder. In: 2019 International conference on communication and electronics systems (ICCES), pp 29–34
3. Manikanth K, Ramesh SR (2018) Design of soft edge flip flops for the reduction of power delay product in linear pipeline circuits. In: IEEE International conference on communication and signal processing, pp 148–151, ICCSP 2018
4. Bansal M, Singh H, Sharma G (2021) A taxonomical review of multiplexer designs for electronic circuits & devices. J Electron 3(02):77–88
5. Kumar VS, Jaiswal SK, Verma K (2012) Design and analysis of CMOS inverter and D latch MCML inverter. Int J Adv Trends Comput Sci Eng
6. Yuxin B (2015) Architecting a MOS current mode logic (MCML) processor for fast, low noise and energy-efficient computing in the near-threshold regime. In: 33rd IEEE international conference on computer design (ICCD)
7. Mutukuri S, Pande KS (2020) Low power rail to rail D flip-flop using current mode logic structure. In: 4th International conference on electronics, materials engineering & nano-technology (IEMENTech), pp 1–6. 10.1109. https://doi.org/10.1109/IEMENTech51367.2020.9270082
8. Shapiro A, Friedman EG (2014) MOS current mode logic near threshold circuits. J Low Power Electron Appl 4(2):138–152
9. Cheng-Fang T, Lai J-L, Chen R-J (2006) Using bulk-driven technology operate in subthreshold region to design a low voltage and low current operational amplifier. In: IEEE international symposium on consumer electronics
10. Agarwal T, Sawhney A, Kureshi AK, Hasan M (2008) Performance comparison of static CMOS and MCML gates in sub-threshold region of operation for 32 nm CMOS Technology. In: IEEE international conference on computer and communication engineering ICCCE
11. Sharroush SM (2016) Performance optimization of MOS current-mode logic. In: International conference on electrical, electronics, and optimization techniques (ICEEOT)
12. Ismail Ayman H, Elmasry M (2003) A low power design approach for MOS current mode logic. In: IEEE International [Systems-on-Chip] SOC conference
13. Mahdi Y, Ghaznavi-Ghoushchi MB (2020) PDP and TPD flexible MCML and MTCML ultra low-power and high-speed structures for wireless and wireline applications. IEEE Trans Very Large Scale Integration (VLSI) Syst 28(8):1782–1795
14. Singh R, Pande KS (2018) 4-bit counter using high-speed low-voltage CML D-flipflops. In: 3rd International conference on communication and electronics systems (ICCES)
15. Giuseppe S, Trifiletti A, Palumbo G (2018) A novel 0.6 V MCML D-latch topology exploiting dynamic body bias threshold lowering. In: 25th IEEE international conference on electronics, circuits and systems (ICECS)

Prediction of Alzheimer's Disease Using Machine Learning

Krishna Kumar Varun, Shankar Hamritha, and Mavuthanahalli Channabasavegowda Vinay

Abstract The medical fraternity has for quite some time been an early adopter and furthermore profited immensely from the advances in technology. Nowadays, machine learning is seen to assume a vital part in numerous medical-related domains, including the improvement of new operations, forecast of diseases at the beginning, the treatment of patient records, and the therapy of persistent infections. An examination of the patterns and examples of patients will enable informed and educated decisions in the healthcare sector. The use of machine learning has a great deal of promise for diagnosing diseases early on. In the current work, dementia is studied. General term is describing a collection of symptoms related to memory loss or other reasoning impairment that might lessen the ability of an individual to deal with a variety of daily tasks. The patient's memory, thinking, and behavior are affected by Alzheimer's. The aim of this work is to utilize machine learning algorithms to process relevant data in order to detect symptoms of Alzheimer's disease at an early stage. The work targets to assemble a classifier model utilizing Scikit-learn instrument to detect Alzheimer's disease by employing logistic regression, support vector machine, random forest, extra trees, and gradient boosting algorithm. Experimental results for each model were evaluated, and the model with the highest accuracy was considered the most suitable model.

Keywords Advances in technology · Healthcare sector · Dementia · Machine learning algorithms

Introduction

In simple terms, the ability to ensure health is maintained or improved due to various diseases or physical damage to an individual through diagnosis, followed by treatment, which leads to the cure or recovery is known as healthcare. The people

K. K. Varun · S. Hamritha (✉) · M. C. Vinay
Department of Industrial Engineering and Management, Ramaiah Institute of Technology, Bengaluru 560054, India
e-mail: hamritha.shankar@msrit.edu

V. Bindhu et al. (eds.), *Proceedings of Third International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 844,
https://doi.org/10.1007/978-981-16-8862-1_50

responsible for ensuring optimum health is maintained in people are the healthcare professionals. Another term for healthcare is medical care [1].

The approach followed by various nations for medical care varies based on their communities, people, socioeconomic situations, and healthcare policies that are followed in their respective countries. An optimum healthcare institution is defined based on its ability to provide health-related services in the best possible manner, as well as in the best possible time [2]. There are many factors that may influence a healthcare institution to provide the required medical services. Some of them are barriers based on their location (the additional costs involved in transport and the availability of taking time off to avail of such services), limitations with respect to finance (e.g., coverage of insurance), and individual limitations due to the inability to communicate with the health-related service providers, poor health-related knowledge, and low revenue [3].

Benefits of Information Technology in Healthcare

The pros of data innovation in medical care are vast. To start with, IT empowers healthcare experts to store and recover information identifying with a patient's medical care records. It likewise upgrades the correspondence of patient data through a decipherable arrangement that anybody can utilize. Subsequently, it decreases the opportunity of medical mistakes. At long last, it makes it simpler to recover patient data through a dataset without new medical checks.

Attention to the health care system in terms of technological advancement is one of the key aspects as it should ensure patients safety. Especially, alerts or alarms can assist somebody with following explicit therapies and timetable for treatment. Likewise, an electronic account of information can prompt consistency of training across all medical professionals. At last, utilizing an electronic health record can work on giving care to normal conditions dependent on past proof [4].

Literature Review

Machine Learning in HealthCare

Machine learning, basically, is a sort of man-made reasoning wherein systems are modified to learn data without human involvement. In machine learning, the advancement of the basic calculations depends on computational insights. Computer systems are furnished with information and afterward the machine or computer system "learn" from that information. The information really "instructs" the machine by uncovering their unique designs and fundamental calculations. The bigger the example of information the "machine" is given, the more exact the machine's yield becomes [5].

Machine learning could potentially be a cost-effective solution for ensuring minimal infrastructure cost of the healthcare system as well as improve overall doctor–patient relationship [6].

Types of Machine Learning

Machine learning is classified into three general types, and they are as follows:

- Supervised learning: In supervised learning, the computer system is given model instances of data and their ideal results, provided by a "teacher," and the objective is to develop a mapping function to map the inputs to the outputs.
- Unsupervised learning: In unsupervised learning, no labels are given to the algorithm, which is left to look for structure in its input on its own. Finding hidden patterns in data can be a goal in itself or can serve as a means toward an end. (learning feature sets).
- Reinforcement learning: A reinforcement learning algorithm maximizes the notion of cumulative reward by taking the right action in the right environment in order to maximize the intelligence of the system [7].

Dementia and Alzheimer's Disease

Cognitive disorders (CDs), otherwise called neurocognitive disorders (NCDs), are a class of emotional wellness issues which fundamentally influence intellectual capacities including learning, memory, insight, and critical thinking. Neurocognitive issues incorporate incoherence and gentle and key neurocognitive issue (formerly recognized as dementia).

Dementia occurs as a bunch of related indications when the cerebrum is harmed by injury or illness. The side effects include reformist hindrances to memory, thinking, and conduct that influences the capacity to take care of oneself as a proportion of completing ordinary activities.

Alzheimer's disease amounts to 60–70% of instances of dementia worldwide [8]. The most widely recognized indications of Alzheimer's disease are momentary cognitive decline and word-discovering troubles. Normal initial symptoms of Alzheimer's incorporate repeating the same thing again and again, forgetting where they are, trouble following bills, issues with remembering new or intricate information, neglecting to take medicines, and difficulty in finding the right words to speak [9].

- Alzheimer's disease is by no means an ordinary aging process. In general, increasing age is the most recognized danger factor, and most Alzheimer's patients are 65 or older.

- Alzheimer's disease gets worse over time. Symptoms of Alzheimer's disease slowly worsen over time as the disease progresses. Cognitive decline is normally mild at the beginning of the disease, but with late-stage Alzheimer's, people lose the ability to carry on a conversation and react to the current situation. In general, an individual with Alzheimer's lives for four to eight years after diagnosis, but depending upon various factors, they may live up to 20 years.
- Though there is no existing cure for Alzheimer's, medicinal treatments are available and research is constantly progressing. Despite the fact that Alzheimer's medications cannot prevent the progression of disease, they can temporarily delay the effects of dementia and create a sense of satisfaction in individuals diagnosed with the disease and their family members [10].

Methodology

The machine learning approach followed in this work is supervised learning. Supervised learning is a form of machine learning that trains a model using labeled or training data, to detect patterns and insights between the input and output data to provide accurate results on data that has not yet been encountered by the model.

The following flowchart depicts the methodology followed to perform the comparative analysis in order to recommend the best algorithm for building a classification model in order to predict Alzheimer's disease (Fig. 1).

Procedure for Proposed Methodology

- **Step 1**: Dataset for Alzheimer's disease Prediction—The Alzheimer's dataset is obtained for preprocessing (Table 1).

  The dataset was obtained from PubMed Central submission: P50 AG05681, P01 AG03991, R01 AG021910, P20 MH071616, U24 RR0213.
  Step 2: Data Preparation and Visualization – Data visualization to detect relevant relationships between the attributes and insights into the data. This is followed by performing a data split of training data (70%) and testing data (30%). Hence, out of 373 data points, 261 random samples are considered as training data and 112 other samples are used as the testing data. Table 2 describes the attributes involved in the dataset for Alzheimer's prediction:

Box Plot

The following Fig. 2 depicts the box plot for the attributes in the dataset (Fig. 2).

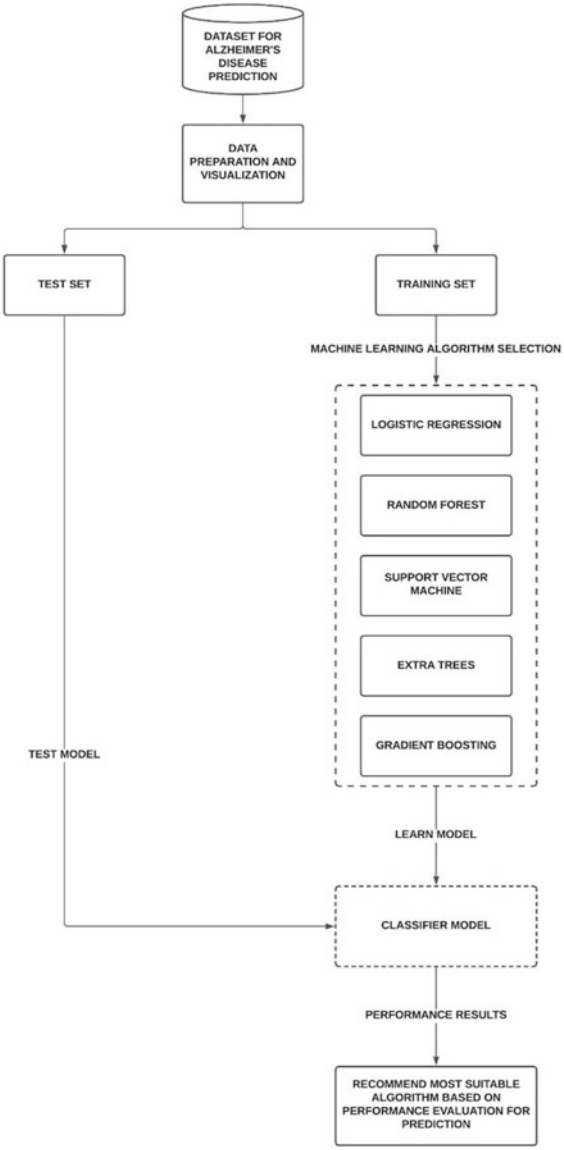


Fig. 1 Methodology flowchart

Correlation Matrix

See Fig. 3.

Figure 3 reflects the correlation matrix of the model parameters.

Table 1 Alzheimer patient dataset sample

Subject ID	MRI ID	Group	Visit	MR Delay	M/F	Hand	Age	EDUC	SES	MMSE	CDR	eTIV	mWBV	ASF
OAS2_001	OAS2_001_MR1	Nondemeneted	1	0	M	R	87	14	2	27	0	1987	0.696	0.883
OAS2_001	OAS2_001_MR2	Nondemeneted	2	457	M	R	88	14	2	30	0	2004	0.681	0.876
OAS2_002	OAS2_002_MR1	Demented	1	0	M	R	75	12	-	23	0.5	1678	0.736	1.046
OAS2_002	OAS2_002_MR2	Demented	2	560	M	R	76	12	-	28	0.5	1738	0.713	1.01
OAS2_002	OAS2_002_MR3	Demented	3	1895	M	R	80	12	-	22	0.5	1698	0.701	1.034
OAS2_004	OAS2_004_MR1	Nondemeneted	1	0	F	R	88	18	3	28	0	1215	0.71	1.444
OAS2_004	OAS2_004_MR2	Nondemeneted	2	538	F	R	90	18	3	27	0	1200	0.718	1.462
OAS2_005	OAS2_005_MR1	Nondemeneted	1	0	M	R	80	12	4	28	0	1689	0.712	1.039
OAS2_005	OAS2_005_MR2	Nondemeneted	2	1010	M	R	83	12	4	29	0.5	1701	0.711	1.032
OAS2_005	OAS2_005_MR3	Nondemeneted	3	1603	M	R	85	12	4	30	0	169	0.705	1.033
OAS2_007	OAS2_007_MR1	Demented	1	0	M	R	71	16	-	28	0.5	1357	0.748	1.293
OAS2_007	OAS2_007_MR3	Demented	3	518	M	R	73	16	-	27	1	1365	0.727	1.286
OAS2_007	OAS2_007_MR4	Demented	4	1281	M	R	75	16	-	27	1	1372	0.71	1.279
OAS2_008	OAS2_008_MR1	Nondemeneted	1	0	F	R	93	14	2	30	0	1272	0.698	1.38

Table 2 Model parameters

Parameters	Description
MR delay time	A delay time given before the actual image acquisition
Gender	Male/Female
Age	60–90 Years
EDUC	Years of Education
SES	Socioeconomic Status
MMSE	Mini—Mental State Examination
DCR	Clinical Dementia Rating
eTIV	Estimated Total Intracranial Volume
nWBV	Normalized Whole Brain Volume
ASF	Atlas Scaling Factor

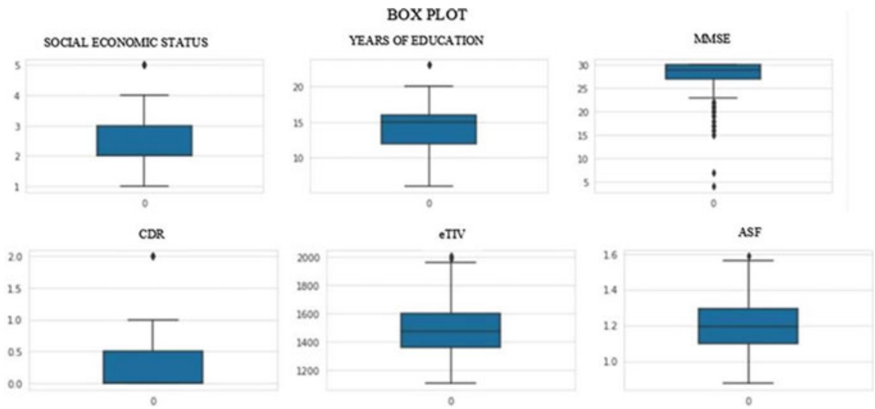


Fig. 2 Model parameters box plot graphs

Scikit-Learn Tool

- The Python Numerical and Scientific Libraries NumPy and SciPy are designed to interact seamlessly with scikit-learn tool and possess various classification, regression, and clustering algorithms.
- The Scikit-learn library for Python is an essential tool for machine learning. Including classification, regression, clustering, and dimensionality reduction, the sklearn library provides a plethora of machine learning and statistical modeling tools.

- **Step 3**: Machine Learning Algorithm Selection

  The models chosen for the data study are as follows:

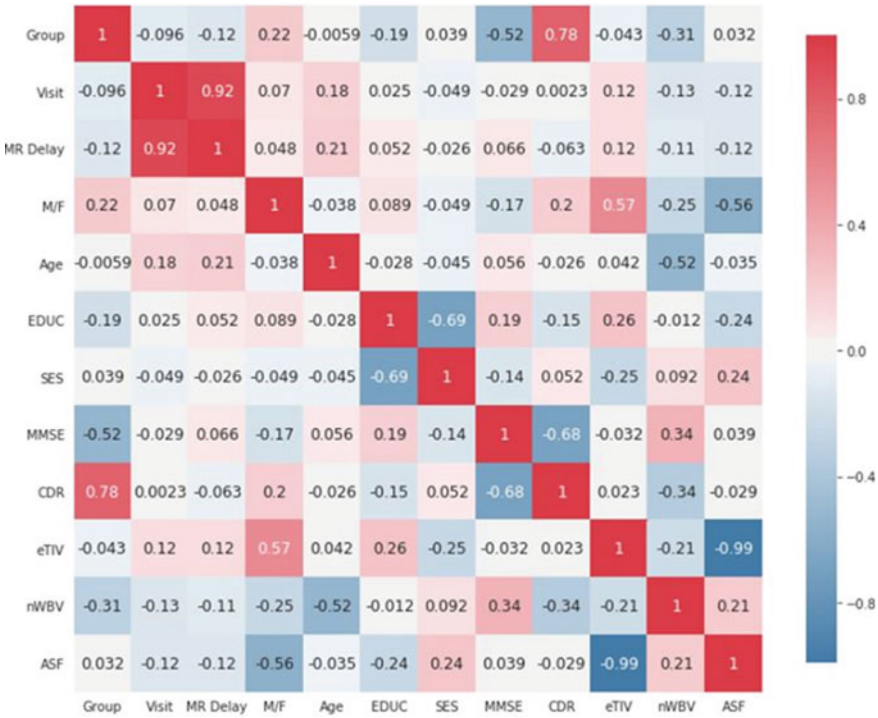


Fig. 3 Model parameters correlation matrix

- Logistic regression
- Random forest
- Support vector machine (SVM)
- Extra trees
- Gradient boosting.

Based on the literature review, the above classifiers were considered to have better accuracy over traditional classifiers due to the ability to obtain higher accuracy in prediction.

- **Step 4**: Building the Classifier Model—The model is built for each of the above-mentioned algorithms using the training data.
- **Step 5**: Testing the Classifier Model—The model is tested using the above algorithms using the testing data.
- **Step 6**: Conducting a Performance Evaluation of the Experimental Results of each Classifier—Evaluation of the results obtained from running each individual algorithm is conducted and compared with one another in order to determine which classifier is more suitable for the purpose of prediction.

- **Step 7**: Recommending Most Suitable Algorithm for Prediction—The most suitable model for prediction is concluded after the evaluation of experimental results.

The proposed classifier model has been built using scikit-learn tool, and based on successful execution of each step, we can evaluate in the experimental results.

Experimental Results

This section deals with the experimental results that are obtained after training logistic regression, support vector machine, random forest classifiers on the Alzheimer's disease dataset. The purpose of experimental results is for performance evaluation of all four classifier and to recommend the most suitable algorithm for prediction.

Confusion Matrix

A confusion matrix is an evaluation matrix composed of N-by-N elements used to assess the performance of a classification model with N being the number of target classes used. The actual or true values and the values obtained from the prediction of the model are compared with one another (Fig. 4).

Confusion Matrix: Understanding True Positives, True Negatives, False Positives, and False Negatives.

True Positive (TP)

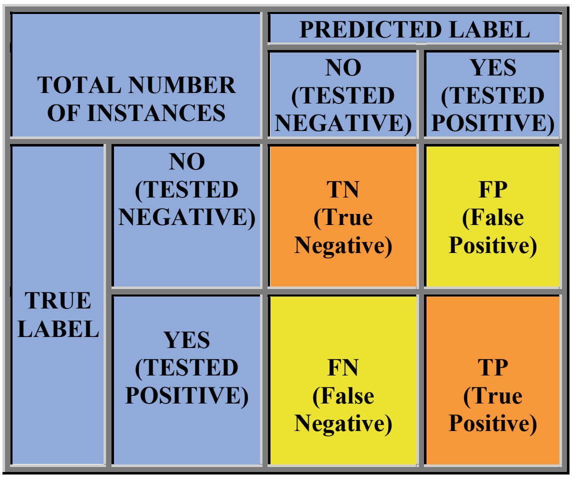


Fig. 4 Confusion matrix structure

Table 3 Confusion matrix values

Algorithm	True Positive (TP)	True Negative (TN)	False Positive (FP)	False Negative (FN)
Logistic regression	49	43	9	11
Random forest	51	47	5	9
Support vector machine	45	45	7	15
Extra trees	55	47	5	5
Gradient boosting	52	46	6	8

- Predicted and actual values are the same.
- Model predicted a positive value and the actual value was also positive.

True Negative (TN)

- Predicted and actual values are the same.
- Model predicted a negative value and the actual value was also negative.

False Positive (FP)—Type 1 error

- Predicted and actual values are different.
- Model predicted a positive value but the actual value was negative.

False Negative (FN)—Type 2 error

- Predicted and actual values are different.
- Model predicted a negative value but the actual value was positive.

As a result of using logistic regression, support vector machine, random forest, extra trees, and gradient boosting machine learning algorithms in Scikit to build classifiers, the following confusion matrix is obtained (Table 3).

Classification Report

Prediction quality from the classification algorithm is measured using a classification report. It is used to indicate the number of predictions that were correct and those that are not. A classification report's metrics are predicted by using true positives, false positives, true negatives, and false negatives.

Classification Report Measure Values

- **Precision**: Precision is defined as the ratio of the number of predicted positive instances to the total number of predicted positive instances.

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

- **Recall**: Recall is defined as the ratio between the predicted positive instances and the actual positive instances.

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

- ***F*-Measure**: *F*-Measure is defined as the weighted harmonic mean of the precision and recall and represents overall performance.

$$F\text{ - Measure} = \frac{2 * \text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}}$$

- **Support**: Support is measured by the number of samples that belong to each of the target value classes.

Table 4 represents classification report measure values for all five Classifiers, and they are as follows:

Table 5 depicts the classification accuracy of all five classifiers (Fig. 5).

Table 4 Classification report measure values

Algorithm	Precision	Recall	*F*-score	TP-rate	FP-rate
Logistic regression	0.84	0.82	0.83	0.81	0.17
Random forest	0.91	0.85	0.88	0.85	0.096
Support vector machine	0.87	0.75	0.80	0.75	0.13
Extra trees	0.90	0.92	0.90	0.92	0.096
Gradient boosting	0.90	0.87	0.88	0.86	0.115

Table 5 Classifier accuracy values

Algorithm	Accuracy
Logistic regression	0.82
Random forest	0.875
Support vector machine	0.8
Extra tress	0.91
Gradient boosting	0.875

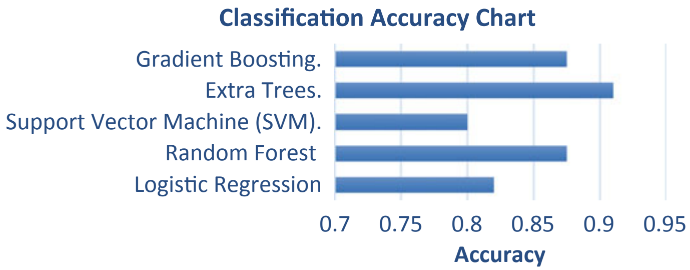


Fig. 5 Accuracy comparison between each algorithm

Receiver Operating Characteristic Curve (ROC)

A receiver operating characteristic curve (ROC curve) demonstrates how well a classification model performs at different thresholds. Two parameters are plotted on this curve as shown in Fig. 6.

- True Positive Rate
- False Positive Rate

True Positive Rate (**TPR**) is defined as follows:

$$\text{True Positive Rate} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}}$$

False Positive Rate (**FPR**) is defined as follows:

$$\text{False Positive Rate} = \frac{\text{False Positive}}{\text{False Positive} + \text{True Negative}}$$

Fig. 6 ROC curve

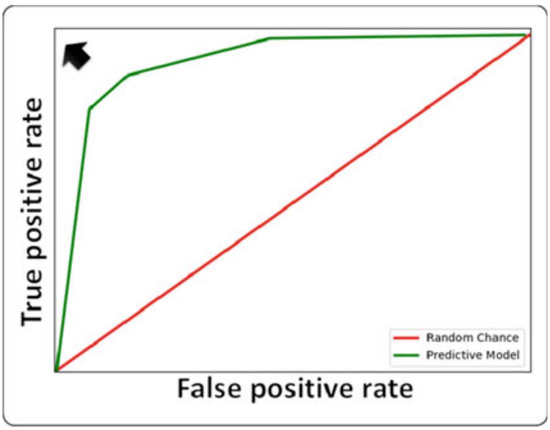


Table 6 AUC values after obtaining ROC curve

Algorithm	Area under the curve (AUC)
Logistic regression	0.82
Random forest	0.88
Support vector machine	0.81
Extra trees	0.91
Gradient boosting	0.87

Area under the Curve (AUC)

- AUC refers to area under the ROC Curve. AUC thus indicates the amount of two-dimensional area beneath the entire ROC curve (using integral calculus) between (0,0) and (1,1).
- AUC is a measure of performance based on classification thresholds across all possible categories. An AUC can be understood as the probability the model ranks a random positive instance higher than a random negative instance.

The ROC curve obtained after building the classifiers using logistic regression, support vector machine, random forest, extra trees, and gradient boosting machine learning algorithm in scikit tool are as follows (Table 6).

Model Versus Accuracy Comparison

A model versus accuracy comparison between logistic regression, random forest, support vector machine, extra trees, and gradient boosting classifier was conducted, and the results obtained are as follows (Fig. 7).

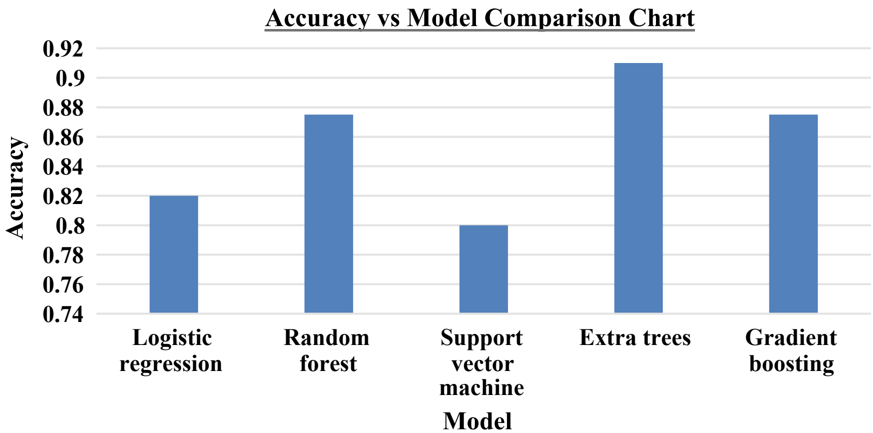


Fig. 7 Accuracy versus model comparison chart

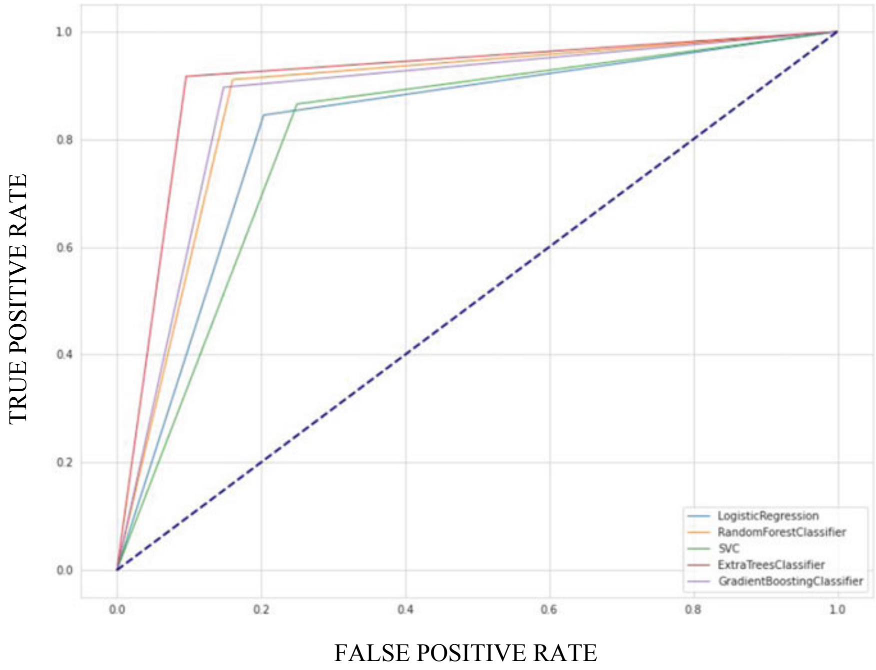


Fig. 8 ROC curve comparison graph

ROC Curve Comparison

An ROC curve comparison between logistic regression, random forest, support vector machine, extra trees, and gradient boosting classifier was conducted, and the results obtained are as follows.

From Fig. 8, it can be observed that extra trees classifier has the highest area under the curve (AUC) as compared to the remaining classifiers, and hence, it is considered the most suitable classifier for model prediction.

Conclusion

Today, people cannot negate that innovation in medical care is new. The selection of innovation in medical services over the course of the years has prompted better finding and treatment of patients. Analysis of patient treatment trends and patterns for determining specific infections in patients will lead to the establishment of educated and efficient choices to improve the general nature of the healthcare sector. Using machine learning to diagnose diseases is a highly promising and may assist in alleviating the delay in diagnosis and possibly allowing the decision-making process to proceed at an early stage. In this research work, five machine learning-based classifiers were selected for the experimentation using the scikit-learn toolkit to predict Alzheimer's disease. The five classifiers chosen were logistic regression, random forest, support vector machine, extra trees, and gradient boosting. The five classifiers have been compared by running them on the dataset, and the results of the classification report, confusion matrix, and ROC curve for each classifier were obtained. The overall performance of extra trees classifier was found to have the highest accuracy and hence considered the most effective classifier to predict Alzheimer's disease as compared to logistic regression, random forest, support vector machine, and gradient boosting.

- From Fig. 7, it can be inferred that the classifier having the highest accuracy is the extra trees classifier (92%).
- The random forest classifier and gradient boosting classifier have a joint high accuracy level of 87.5%.
- The logistic regression classifier has a lower accuracy score of 82% compared to support vector classifier having an accuracy score of 80% and both are lower than random forest classifier and gradient boosting classifier.
- Support vector classifier has the lowest accuracy compared to the other classifiers considered in this study.

Future Scope and Limitations

- The model generated in this study is used for predicting Alzheimer's disease at an early stage. However, Alzheimer's disease is a slow-progressing disease that has various stages.
- Hence, attempts can be made in identifying the stage of an Alzheimer's patient for more effective diagnosis. This will aid not only in detecting the stage of the Alzheimer's patient but also monitor the progress of the disease.
- This work can serve as a basis for future works that can potentially aid in obtaining a cure for Alzheimer's disease.
- The work carried out in this study is limited only to Alzheimer's disease, and hence, the same model cannot be used to determine the outcome for various other diseases.

References

1. Karunakaran P, Hamdan YB (2020) Early prediction of autism spectrum disorder by computational approaches to fMRI analysis with early learning technique. J Artif Intell 2(04):207–216
2. Palani U, Vasanthi D, Rabiya Begam S (2020) Enhancement of medical image fusion using image processing. J Innov Image Process (JIIP) 2(04):165–174

3. Breitner JC (2006) Dementia epidemiological considerations, nomenclature, and a tacit consensus definition. J Geriatr Psychiatry Neurol 19(3):129–136
4. Kiraly A, Szabo N, Toth E et al (2016) Male brain ages faster: the age and gender dependence of subcortical volumes. Brain Imaging Behav 10:901–910
5. Mesrob L, Magnin B, Colliot O et al (2008) Identification of atrophy patterns in Alzheimer's disease based on SVM feature selection and anatomical parcellation. Med Imaging Augmented Reality 5128:124–132
6. Glenn F, Jonathan S (2007) SVM feature selection for classification of SPECT images of Alzheimer's disease using spatial information. Knowl Inf Syst 11(2):243–258
7. Zhu X, Lei Z, Zi H (2014) A sparse embedding and least variance encoding approach to hashing. IEEE Trans Image Process 23(9):3737
8. Torabi M, Morad zadeh H, Vaziri R, Ardekani RD, Fatemizadeh E (2007) Multiple sclerosis diagnosis based on analysis of subbands of 2-D wavelet transform applied on MR-images. In: IEEE/ACS international conference on computer systems and applications. AICCSA '07, pp 717–721
9. Coppede F, Grossi E, Buscema M, Migliore L (2013) Application of artificial neural networks to investigate one-carbon metabolism in Alzheimer's disease and healthy matched individuals
10. Badnjevic A, Cifrek M, Koruga D, Osmankovic D (2015) Neurofuzzy classification of asthma and chronic obstructive pulmonary disease. BMC Med Inform Decis Mak 15(SuppI3)

Design, Simulation, and Development of DC-DC Converter for SPV-Powered DC Distribution System in Domestic Loads

Basanagouda F. Ronad

Abstract Solar photovoltaic (SPV) systems together with DC microgrids are promising option for remotely located domestic loads, and this combination has many advantages over conventional AC distribution system. This paper presents design, simulation, and development of single input multiple output (SIMO) DC-DC converter for SPV-powered DC distribution system employed for domestic loads. Simulation model is built for DC-DC converter with 24 V input and resulting 12, 24, 32, and 48 V DC output voltages. Regularly used domestic loads which basically operate on DC voltages are considered for the design, and thus, these loads are proposed to be operated in a DC distribution system. For the selected loads current and voltage assessment, during their full load performance is conducted. Based on these values, converter parameters for buck and boost operation are evaluated. Further, the circuit is successfully tested for its operation in Proteus software and in experimental setup with 300 Wp solar panels as input. Performance of the converter is tested with different loads and load combinations. Based on the performance results, it is concluded that successful implementation of proposed concept in large scale reduces the dependency on conventional energy sources, while satisfying the electricity requirement for remotely located domestic DC load.

Keywords DC-DC converter · DC distribution · Proteus software · SIMO converter · Solar PV

Introduction

Renewable energy sources together with DC microgrids have many advantages over conventional AC distribution system, in terms of cost, lesser conversion stages, efficiency, and the minimum complex operation. In DC distribution system, DC-DC converter acts as power conditioning units. Solar PV system and other renewable energy sources inherently generate DC voltages. This necessitates the development

B. F. Ronad (✉)
Department of Electrical and Electronics Engineering, Basaveshwar Engineering College (A), Bagalkot, India

V. Bindhu et al. (eds.), *Proceedings of Third International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 844,
https://doi.org/10.1007/978-981-16-8862-1_51

of DC energy system based on renewable electric systems. Choice of an appropriate DC-DC converter becomes significant as the overall system efficiency is strongly depending upon the converter's performance. Majority of domestic and commercial loads comprise of electronic loads that require DC power. Thus having DC system eliminates the rectification stages, thereby improving the system efficiency. In this regard, the conventional AC loads are switched to DC. Use of the DC loads to DC distribution system increase the efficiency and reliability of the entire system. DC-DC converter can also handle short circuits without the need for additional protection devices such as circuit breakers. In the current scenario, DC distribution systems are in the field of microgrids, telecommunication systems, DC buildings, and data centers due to its many advantages over conventional AC systems. On other hand, most parts of the India possess more than 300 clear sunny days in a year, which makes it to ideally suited for harnessing of solar energy for remotely located domestic loads.

Related Work in Literatures

Design, simulation, and implementation of different types of DC-DC converters are discussed in many literatures. S. Sheik Mohammed et al. presented simulation and analysis of microcontroller-based DC-DC converter using Proteus suite. DC-DC converter is simulated with constant DC input and variation of the voltage at the input side. Circuit is analyzed under open-loop condition for both cases. Pulse generator is directly used for constant DC input. Microcontroller is programmed to generate the firing pulse for variation of the voltage at the input side [1]. P. Arunthathi et al. presented isolated transformers connected between the solar PV array and SIMO DC-DC converter to avoid the short circuit between the two components. To improve the effectiveness of the converter, the PID controller is located. Proportional gain and the integral gain are varied such that the percentage error can be reduced. The relational operator and summer block are used along with PID controller, and thus, different voltage levels are obtained and supplied to multiple loads. Single input multiple output (SIMO) DC-DC converter block is driven by the driver circuit which is used to provide the various duty cycles to the respective switches. Variation in duty cycle results in variation in the output of the SIMO DC-DC converter, and thus, required output voltage can be accurately attained by duty cycle control [2] and also presented state-space averaged (SSA) modeling for non-isolated SIMO DC-DC converter. This hardware setup of Zeta–Buck–Boost (ZBB) converter is developed, and dynamic model is verified. Circuit dynamics are represented by a set of ordinary differential equations (ODEs) and differential algebraic equations (DAEs). Transfer functions for experimental setup are obtained using network analyzer [3]. Ilyass Abdillahi Aden et al. proposed non-isolated DC-DC buck converter for electric vehicles. The simulation of both dynamic model and static model is performed in MATLAB/SIMULINK. It has advantage of having three switches for two outputs. Further, control strategies and small signal analysis are verified [4]. Chinnu V. et al. presented non-isolated SIMO DC-DC converter which can provide step-up as well as step-down outputs from a single DC input simultaneously. Analysis of different modes of operation of converter is done in MATLAB/SIMULINK. Peripheral interface controller (PIC) is used for control the multiple outputs [5]. High efficiency DC-DC converter is obtained with the help of fuzzy logic control system. The behavior of the converter is analyzed using MATLAB/SIMULINK [6]. Further, single input multiple output converters are simulated [7], and tested for renewable energy applications [8, 9]. Investigation of energy conservation through DC distribution is demonstrated. In this study, the regular loads are deliberately modified to operate with DC voltages and operated at 48 V. It is proved that DC system reduces the power conversion stages and hence diminishes the associated power losses and standby losses [10]. Efficiency investigations and feasibility studies of SIMO converters are carried out [11]. In similar, many literatures presented the employment of various software's for analysis and simulation of Buck, Boost, and Buck–Boost converters, such as Pspice, Proteus design suite, and MATLAB/SIMULINK. Efficient SIMO DC-DC converters are designed for industrial applications, where multiple outputs are obtained from single input, and voltage gain is increased by coupled inductor [12–15].

Analysis and simulation of DC-DC converter using Proteus 8 professional are presented in the present work. SIMO DC-DC converter is suitable choice for PV applications, where the PV output is directly fed to the loads without the need of storage during the sufficient sunlight conditions. Concept of SIMO converter exclusively for domestic equipment is proposed in the work. Limitations associated with AC loads for DC distribution are identified, and an attempt to overcome issues is presented by using DC loads. In the proposed model, existing DC loads are listed and set the output voltages for SIMO DC-DC converter. It is designed for multiple outputs which can step-up as well as step-down from single DC input. Simulation is done for different DC loads, and results are verified.

Methodology

The functional block diagram of proposed model is shown in Fig. 1, and the interconnections of the system elements are shown in Fig. 2. Core objective in this

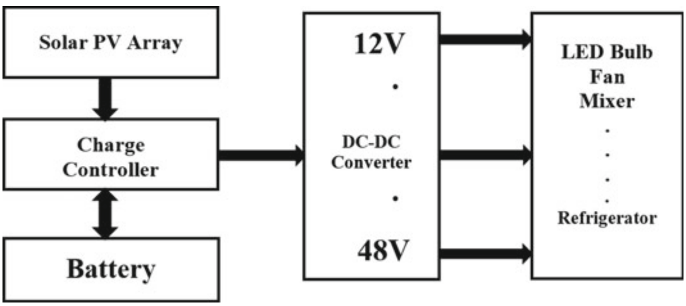


Fig. 1 Block diagram of the proposed SIMO converter

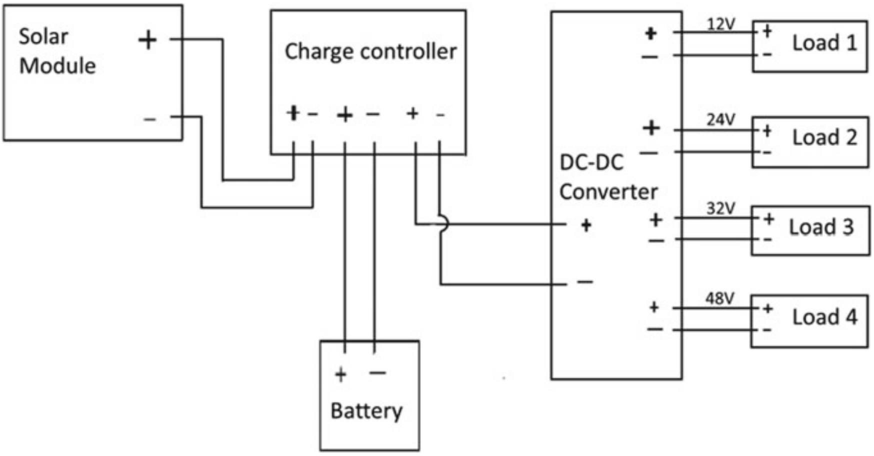


Fig. 2 Interconnections of the system components

methodology is to design single input multiple output DC-DC converter for domestic appliances employed with DC distribution system. System includes solar panel, charge controller, DC-DC converter, and DC loads. Solar photovoltaic array employed for the work includes two panels of polycrystalline type each rated 150 Wp connected in series to yield 24 V. Table 1 presents the specifications of solar panels employed in the present work. Small batteries of 7.2 Ah are employed to supply power to loads during non-sunshine hours, during the real-time operation.

The proposed methodology is taken up in three steps: selection of DC loads for domestic usage, design of proposed concept of DC-DC converter, and implementation/simulation of proposed converter setup.

Table 1 Specifications of solar PV panels

Power (P_{max})	150 Wp
Open circuit voltage (V_{oc})	21.4 V
Short circuit current (I_{sc})	8.95 A
Current at Max. power (I_{mp})	8.5 A
Voltage at Max. power (V_{mp})	17.7 V
Tolerance	±3%
Temperature coefficient—P_{max}	−0.46% /°C
Nominal voltage	12 V
Solar cells per module	36
No. of panels employed	2 (in Series)

Table 2 Loads considered for the design of converter

S. No.	Loads	Voltage (V)	Current (A)	Wattage (W)
1	Table fan	12	1.3–1.8	20
2	Mixer	24	13.4	250
3	LED bulb	32	0.15	5
4	LED bulb	48	0.188	9
5	DC ceiling fan	12	2.33	28
6	DC ceiling fan	48	0.5	24
7	DC refrigerator	12	2.08	25
8	LED bulb	48	0.25	12

Selection of DC Loads

There are many appliances in domestic and commercial applications, which basically use DC power. These loads can be directly operated through DC distribution system. In this regard, the loads are listed and respective DC ratings are assessed. Further, desired voltage levels for loads are selected. Table 2 presents loads to be considered for design of converter. The loads are fed directly from solar panels, and hence based on the wattage of loads listed for experimental setup, 300 W solar panels are employed for electricity generation.

Design of DC-DC Converter

Approach of Buck and Boost converters is employed in the present study for simulation to develop single input multiple output converter. Buck converter, Boost converter, and Buck–Boost converter approaches are presented in the following section:

Buck converter converts DC unregulated high voltage to regulated voltage at lower levels. It steps down voltage by stepping up the current. It is operated in continuous and discontinuous conduction mode. Operation of the Buck converter is done in two cases, switch ON and switch OFF. Figure 3 shows the circuit diagram of typical Buck converter [16]. During switch ON mode, current will flow through the inductor, and it will store the energy and remaining current will flow through the capacitor and load resistance with diode is in reverse biased mode. During switch OFF mode, inductor acts as source and energy stored in the inductor release through the capacitor and load resistance. This will circulate through forward-biased diode. Output voltage of Buck converter is the difference between input voltage and voltage across the inductor. For the given input voltage and output voltage, duty cycle can be obtained using Eq. (1):

Duty cycle,

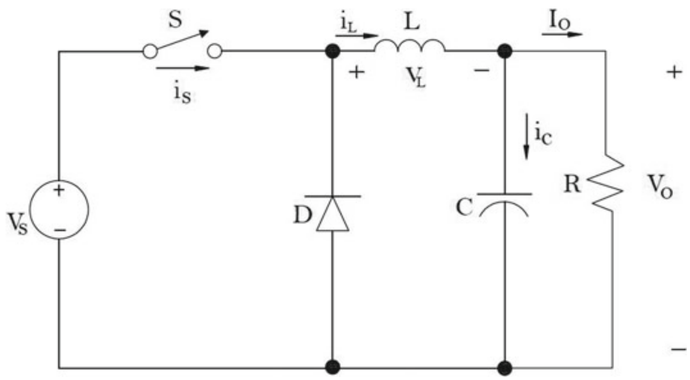


Fig. 3 Circuit diagram of buck converter [16]

$$D = \frac{V_O}{V_S} \tag{1}$$

where

D is duty cycle
V_o is output voltage in volts
V_s is input voltage in volts

Further, inductor and capacitor values are assessed using Eqs. (2) and (3), respectively:

$$L = \frac{D(1-D)}{\Delta I_L} \times V_S \times T \tag{2}$$

where

L is inductance in henry
ΔI_L is inductor ripple current
T is the switching time in sec

$$C = \frac{D(1-D)}{8 \times L \times \Delta V_C} \times V_S^2 \times T^2 \tag{3}$$

where

C is capacitor in Farad
ΔV_C change in output voltage.

Boost circuit converts unregulated low voltage DC into regulated high voltage DC. It steps up the voltage by stepping down the current. It is operated in continuous and discontinuous conduction mode. Operation of the Boost converter is in two cases,

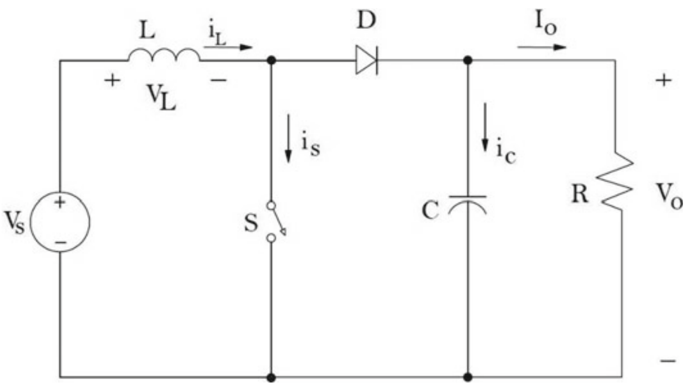


Fig. 4 Circuit diagram of Boost converter [16]

switch ON and switch OFF. Figure 4 shows the typical circuit diagram of Boost converter [16].

During Switch ON, current will flow through inductor, and it stores the energy. In switch OFF mode, diode is reversed biased and it blocks the reverse flow of current from load to source. In this mode, sum of the stored energy in inductor during on time and the supply current will flow through the load resistance and capacitor. Output voltage is zero during switch ON. Further, during switch OFF, output voltage is equal to sum of the voltage across the inductor and voltage across the source. For the given input supply voltage and output voltage, the duty cycle is assessed using Eq. (4):

Duty cycle,

$$D = 1 - \frac{V_O}{V_S} \tag{4}$$

Further, inductor and capacitor values are assessed using Eqs. (5) and (6), respectively:

$$L = \frac{V_S \times D}{f_s \times \Delta I_L} \tag{5}$$

$$C = \frac{I_o \times D}{f_S \times \Delta V_C} \tag{6}$$

Buck–Boost converter is used to get output voltage either less than or greater than the input voltage. Buck–Boost converter also operates in two modes, Switch ON and Switch OFF modes. Figure 5 shows the circuit diagram of Buck–Boost converter [16]. During switch ON, diode is in reversed biased condition, current flowing through the inductor increases and stores energy. During the switch OFF, diode is forward biased and inductor releases stored energy through capacitor and load resistance. The inductor current falls down until switch ON mode.

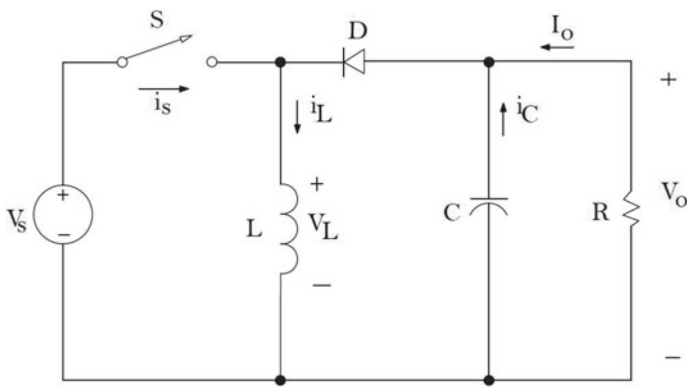
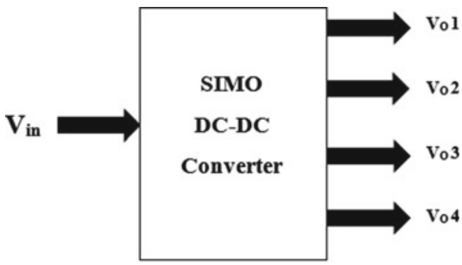


Fig. 5 Circuit diagram of Buck–Boost converter [16]

Fig. 6 Block diagram of SIMO DC-DC converter



The proposed SIMO DC-DC converter is the combination of one Buck and two Boost circuits. One output is connected directly from input source. They are connected in parallel to achieve four outputs from single input. 24 V input is provided from series-connected solar panels. DC-DC converter will result in different voltages of 12, 24, 32, and 48 V. Figure 6 shows the conceptual block diagram of SIMO DC-DC converter.

Inductor and capacitor values are calculated using Eqs. (2)–(6), for Buck converter and Boost converter, respectively, for selected DC loads. Calculated values are represented in Table 3. Further, for the real-time implementation, parameters associated with the largest wattage of the load are considered for implementation.

Simulation of SIMO DC-DC Converter

Simulation of single input multiple output DC-DC converter is conducted in Proteus 8 Professional Software. 24 V solar PV output is used as input source. One Buck and two Boost circuits are employed to get multiple outputs. Circuits are connected in parallel to source to attain different voltage levels. Parameters and components

Table 3 L&C value for selected DC loads

S. No.	DC loads	Capacitor (C) in μF	Inductor (L) in μH
1	12 V table fan	42.5	117.64
2	32 V LED bulb	0.39	1333.33
3	48 V LED bulb	0.65	2127.6
4	12 V DC ceiling fan	57.50	86.95
5	48 V DC ceiling fan	1.736	800
6	12 V DC refrigerator	52	96.15
7	12 V/24 V LED bulb	30	166
8	48 V LED bulb	0.8680	1600

of converter are selected as per mathematical calculation. MOSFETs employed in the converter operate at 100 kHz frequency through pulse generator property. Input voltage (15 V), duty cycle (%), and the operating frequency (kHz) are given as input to pulse generator. This designed circuit has a separate ground for each of the converters. Simulations are conducted for three circuits separately: Buck converter to get 12 V, Boost converter to get 32, and 48 V and the proposed SIMO converter with multiple output waveforms. The experimental setup with display of 12, 24, 32, and 48 V output voltages is shown in Fig. 7, and simulation circuit of proposed SIMO converter is shown in Fig. 8.

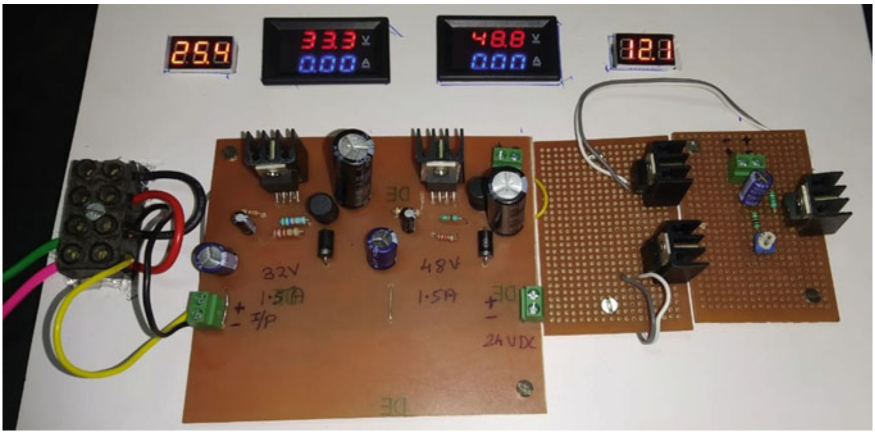


Fig. 7 Experimental setup with display of 12, 24, 32, and 48 V output

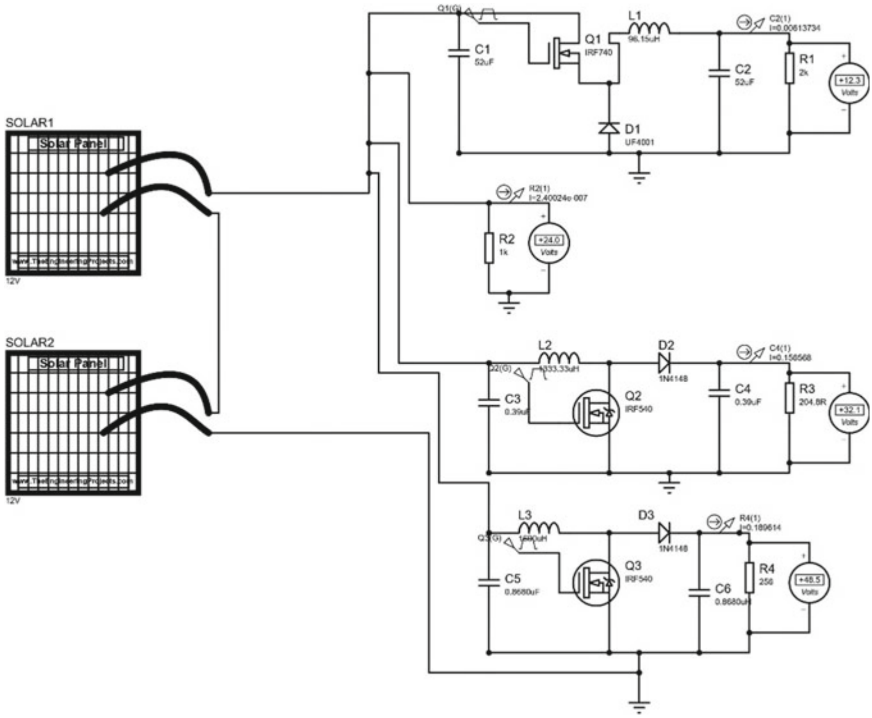


Fig. 8 Simulation diagram of proposed SIMO DC-DC converter

The proposed model is tested in the solar radiations ranging from 750 to 950 W/m^2. However, it is observed that, during lower solar radiations, power generated is not sufficient and hence requires batteries, leading to increase in system cost.

Results and Discussions

The input voltage of 24 V is received from solar PV panels and buck, boost voltages are attained at load terminals. Output voltage waveforms of 12 V, 32 V, and 48 V are presented in Figs. 9, 10, 11, respectively. It is observed from Fig. 10 that steady output voltage has reached in 8 ms after increased from 10.3 V to 12 V. Further, the current waveforms with different loading conditions are simulated and results are presented in Figs. 12, 13, 14, respectively. In similar, different combinations of load connected together are simulated to verify the successful operation of the propose converter scheme. Table 4 presents the current drawn by different loads.

Performance of the proposed system is tested with DC fans and DC LEB bulbs of different voltages. The desired voltages were attained at the load terminals. During

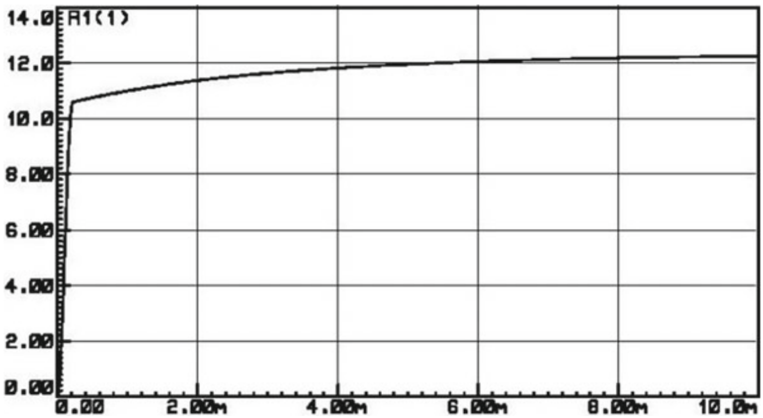


Fig. 9 24–12 V Output voltage

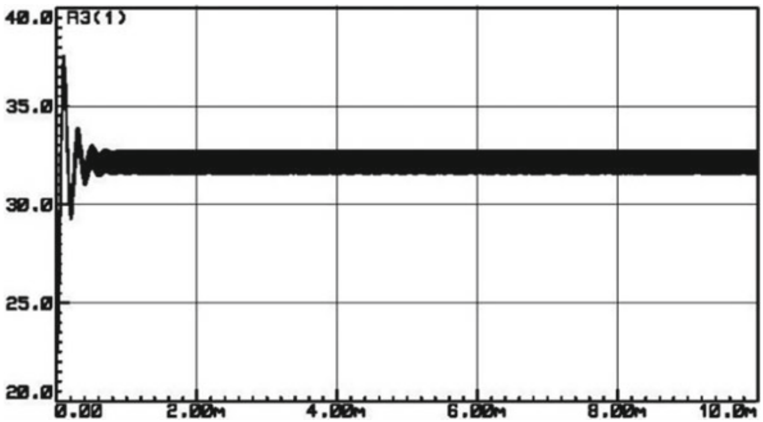


Fig. 10 24–32 V Output voltage

the off-load conditions, voltages were observed to be 13.2, 25.4, 33.3, and 49.2 V. These voltages were noted during the maximum solar radiation of 950 W/m^2.

Proposed DC-DC converter uses number of switches equal to number of outputs. This will further enhance the switching losses, thereby reducing the system efficiency. It is observed that efficiency of single input multiple output (SIMO) DC-DC converter reduce with number of outputs increased beyond three. Performance analysis of the SIMO DC-DC converter for different combination of loads is verified using simulation.

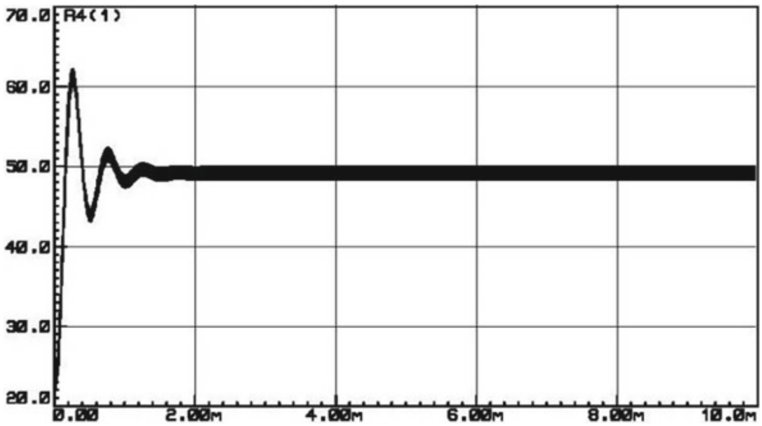


Fig. 11 24–48 V Output voltage

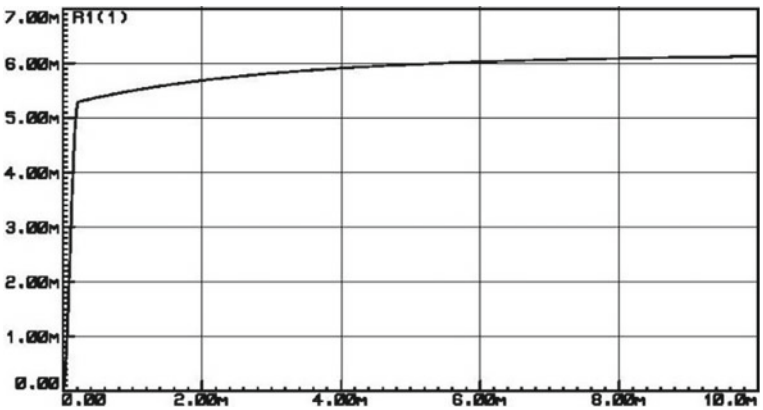


Fig. 12 Current waveform with 12 V DC Fan connected

Conclusions

Design and simulation of single input multiple output (SIMO) DC-DC converter for DC distribution system are presented in the paper. Performance testing of proposed converter is conducted for domestic loads. Simulation model and experimental setup are built for DC-DC converter with 24 V input and output of 12, 24, 32, and 48 V. Domestic loads such as DC fan, DC LED bulbs, DC refrigerator, DC TV are considered for the design and simulation of converter. For the selected loads and combinations of loads, current and voltage assessment during their full load performance are analyzed. Solar PV with 24 V is employed as input. It is concluded that successful implementation of proposed concept will be a boon for remotely located and grid isolated domestic appliances. However, it is observed that increase in number of

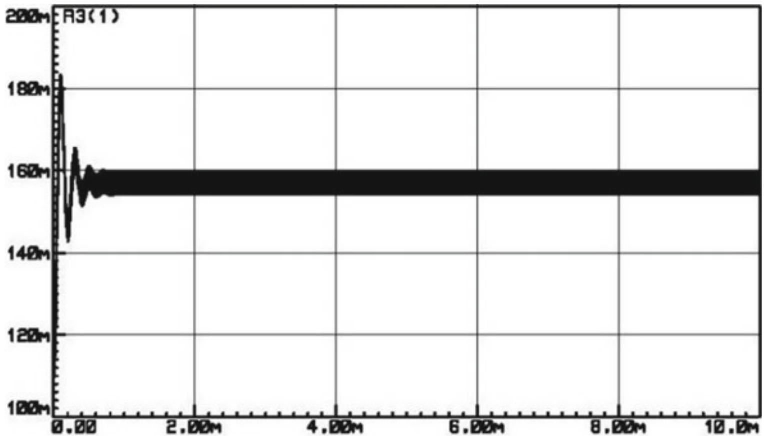


Fig. 13 Current waveform with 32 V LED Bulb connected

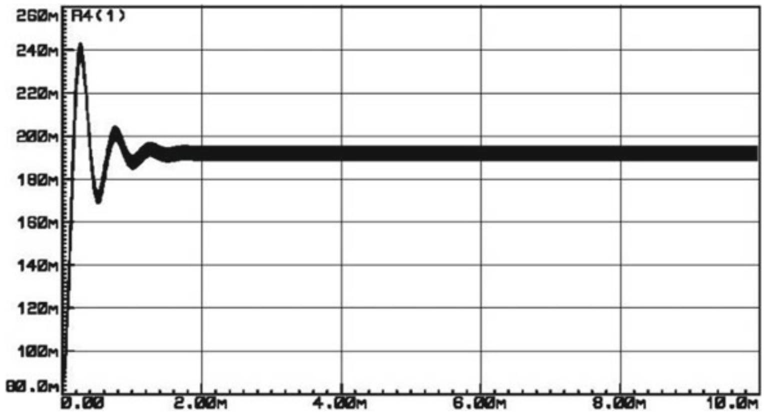


Fig. 14 Current waveform with 48 V LED Bulb connected

switches in converter circuit leads to further loss in switching losses, thereby reducing the system efficiency.

The major drawback of the system is electricity supply reliability. As solar radiation varying with time, clouds and other climatic conditions, loads are not fed satisfactorily. This necessitates the need of storage system, which in turn increase the system cost. Further, during off sunshine hours, system cannot be operated with full load conditions and with all load combinations. These issues can be addressed in future work with higher-rated loads and solar panels.

Table 4 Comparison of the actual current and current drawn by the different DC loads

S. No.	DC loads	Current rating (A)	Current drawn (A)
1	12 V table fan	1.3–1.8	0.61
2	32 V LED bulb	0.15	0.159
3	48 V LED bulb	0.188	0.19
4	12 V DC ceiling fan	2.33	0.6205
5	48 V DC ceiling fan	0.5	0.192
6	12 V DC refrigerator	2.08	0.6127
7	48 V LED bulb	0.25	0.189

References

1. Sheik Mohammed S, Devaraj D (2013) Design, simulation and analysis of microcontroller based DC-DC Boost converter using proteus design suite. In: International conference on advances in electrical and electronics, December 2013
2. Arunthathi P, Jothi Priya R, Muthu Chendu K, Pongeerthana M, Prince Winston D (2020) Development of SIMO DC-DC converter for PV applications. Int Res J Eng Technol (IRJET) 06(05). e-ISSN: 2395-0056
3. Markkassery S, Mahindrakar AD, Lakshminarasamma N, Pasumarthy R (2018) Modelling of non-isolated single input multi output DC-DC converter. In: IEEE ınternational conference on power electronics, drives and energy systems (PEDES). https://doi.org/10.1109/pedes.2018.8707836
4. Aden IA, Kahveci H, Sahin ME (2017) Single ınput multiple output DC-DC buck converter for electric vehicles. Turkish J Electromech Energy 30, ISSN 2547-975X
5. Chinnu VV, Abraham D, Joy N (2014) An ıntegrated single ınput multi-output DC-DC converter with simultaneous buck and Boost outputs. Int J Adv Res Electr Electron Instrum Eng 3(5)
6. Beulah Hematatha S, Sundararajan M (2017) High efficiency single ınput multiple output DC-DC converter. Int J Pure Appl Math 116(15):381–386. ISSN: 1311-8080
7. Kudchi F (2016) Single ınput multiple output DC-DC converter with ınverted output. Int J Innov Sci Eng Technol 3(4). ISSN: 2348-7968
8. Grace M, Nayana J (2020) Single ınput based multi-port DC-DC converter topologies for renewable energy applications. Int Res J Modern Eng Technol Sci 2(07). e-ISSN: 2582-5208
9. Ramteke ND, Burade PG (2016) Design of single ınput multiple output DC-DC converter. Int J Eng Res Technol (IJERT) ISSN: 2278-0181
10. Kamran M, Bilal M, Mudassar M (2017) DC home appliances for DC distribution system. Mehran Univ Res J Eng Technol. https://doi.org/10.22581/muer1982.1704.12
11. Kavitha HS (2016) High efficiency single ınput multiple output DC-DC converter. Int J Inf Futuristic Res (IJIFR) 3(5). ISSN: 2347-1697
12. Gomathy S, Senthilnathan N, Swathi S, Poorviga R, Dinakaran P (2020) Review on multi ınput multi output DC-DC converter. Int J Sci Technol Res 9(04). ISSN: 2277-8616
13. Lopa SA, Hossain S, Hasan MK, Chakraborty TK (2016) Design and simulation of DC-DC converters. Int Res J Eng Technol (IRJET) 03(01). e-ISSN: 2395-0056

14. Park H, Kim S (2020) Single ınductor multiple output auto-buck-boost DC-DC converter with error-driven randomized control. School of Integrated Technology
15. Sri Revathi B, Prabhakar M, Gonzalez-Longatt F (2017) High gain-high power (HGHP) DC-DC converter for DC microgrid applications: design and testing. Int Trans Electr Energy Syst 28(2):e2487. https://doi.org/10.1002/etep.2487
16. Bhattacharjeea S, Saharia BJ (2014) A comparative study on converter topologies for maximum power point tracking application in photovoltaic generation. J Renew Sustain Energy 6:053140

A System Design Perspective for Measurement of Parameters Using Different Scatterings Associated with Fibre Optic Sensors

Kadambari Sharma, J. M. Nair, and P. P. Vaidya

Abstract Fibre optic sensors offer enormous advantages over conventional sensors like immune to harsh and hazardous environments, transmission without losses, and easy integration in structures that make them a great substitute. Distributed fibre optic sensors (DFOS) allow measurement all along the fibre length and provide continuous measurement of physical parameters like temperature, pressure, vibration and strain. They can be easily implemented in coalmines, tunnels, pillars, etc., for structural health monitoring, for pipeline leakage detection and for scaling detection in distillation towers, boilers, furnaces, etc. This paper illustrates the working principle of the DFOS and the various scattering schemes utilised in DFOS for sensing the parameters. The paper provides clear understanding and important considerations needed for the conceptual design of the DFOS system using different scatterings. It also brings out the clear differences in the scattering schemes from which it can be concluded that the Brillouin scattering can be used for temperature and strain measurements, and it is implemented for long distances. Raman scattering provides better sensitivity to temperature but can be employed for shorter distances, and Rayleigh scattering can be used for dynamic measurements. Though DFOS are utilised in many commercial fields but still has to make its impact in the market due to its huge overall cost. With the advancement in the electronics and instrumentation technology, many new economical devices will help reduce this economical hindrance and make DFOS beneficial to the society.

Supported by Organization VESIT.

K. Sharma (✉) · J. M. Nair · P. P. Vaidya
VES Institute of Technology, Affiliated to University of Mumbai, Mumbai, India
e-mail: kadambari.sharma@ves.ac.in

J. M. Nair
e-mail: principal.vesit@ves.ac.in

P. P. Vaidya
e-mail: pp.vaidya@ves.ac.in

V. Bindhu et al. (eds.), *Proceedings of Third International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 844,
https://doi.org/10.1007/978-981-16-8862-1_52

Keywords Distributed fibre optic sensors · Backscattering · Optical time domain reflectometer · Rayleigh scattering · Raman scattering · Brillouin scattering · Stokes · Anti-Stokes

Introduction

The conceptual discovery of 'Total Internal Reflection' revolutionised the world. Inventions of lasers and emergence of techniques to reduce fibre losses led to the development of optical data communication at a great pace. With minimal losses and increased sensitivity of the fibre, variations in intensity, phase, and wavelength of light propagating through the fibre were observed caused by external perturbations on the fibre. This led to the conception of fibre optic sensors (FOS).

Non-optical parameters like temperature, pressure, vibrations, displacement, humidity, chemicals, etc., act on the optical fibre, changing the optical properties of light that are sensed and further processed. FOS have been widely accepted as it offers numerous merits over conventional senors. First, the fibre materials, i.e. silica and germanium, offer great immunity to radio frequency interference and high resistance to corroding substances that allow FOS to be implemented in harsh and dangerous environments. The optical signal caters to minimise transmission losses, which aids in remote sensing. The lightweight and small diameter tenders great advantage of being impregnated into composite structures in applications with limited space and weight requirements.

The FOS system (Fig. 1) comprises of a light source (namely LED, and LASER diode), an optical fibre and an optical detector (photodiodes, photo multiplier tubes etc.) followed by electronic processing to infer the parameter. The non-optical parameters are the physical parameters (e.g., temperature, strain, vibrations, pressure, etc.) that need to be measured. There are various types of FOS based on different operating principles. One such sensor that is gaining lot of research interest is the distributed fibre optic sensors (DFOS). In DFOS (Fig. 2) every point of the fibre acts as a sensor, and physical parameter variations that occur at any point on the fibre is sensed and detected. Thus, DFOS provides continuous measurements over the entire length. DFOS is useful for applications that require sensing lengths of many kilometers, e.g., tunnels, bridges, dams, pipelines, cables, and also in harsh environments like oil wells, mines, etc., making it important fibre optical technology developed for sensor applications.

This paper provides an insight into the principle employed in the DFOS, the different scattering schemes with their mathematical descriptions. It also imparts a detailed description about the designing aspects for the systems utilising DFOS. The comparisons of the different scattering schemes are also produced.

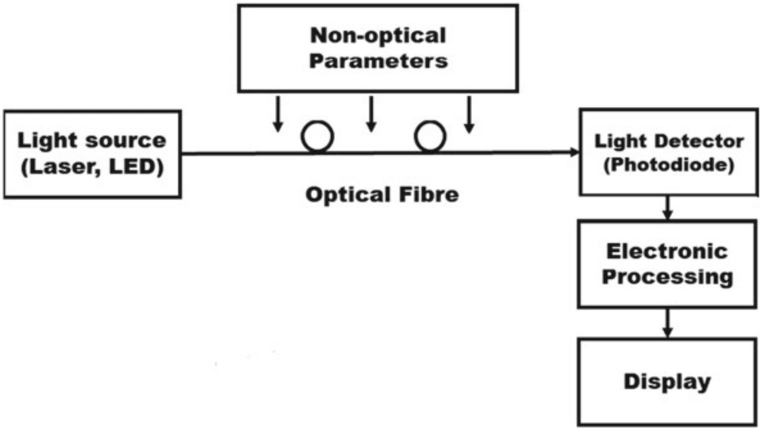


Fig. 1 Block diagram of fibre optic sensor (FOS) system

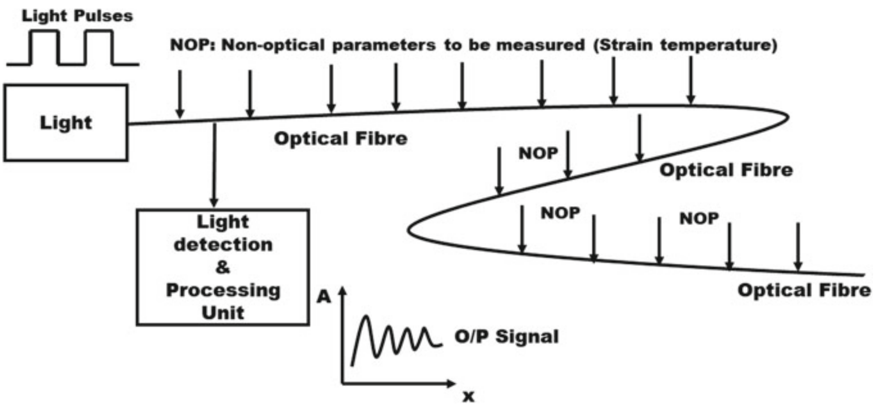


Fig. 2 Distributed fibre optic sensor system

Distributed Fibre Optic Sensors

Distributed fibre optic sensors (DFOS) are built on the phenomena of light scattering. The light propagating along the optical fibre undergoes scattering in all the directions, including backwards. The backward propagation is known as backscattering. This backscattered light is sensed and inspected for finding real-time parameter (temperature, pressure, etc.) variations in DFOS. In fact, distributed sensing employs the principle of optical time domain reflectometry (OTDR) along with light scattering event. In OTDR, the optical pulse is launched in to the fibre of a predetermined length, and the backscattered light is noticed and measured as a function of time [1–3]. The position of the backscatter point can then be located, which is characteristic of DFOS, to analyse any physical parameter variation. DFOS are mainly employed for temperature, strain, and vibration measurements as the backscattering occurrence is sensitive to these parameters.

The performance of the DFOS is governed by certain important parameters, namely [3]:

- Measurand Resolution: the ability of the DFOS to detect the smallest changes of the quantity being measured.
- Range: is the maximum length of the fibre over which the parameters can be sensed.
- Spatial resolution: is the smallest length of the fibre over which the parameter change can be detected.
- Speed: the time taken by the system to acquire the readings for all the points.

All these factors are linked together, and the interaction between these parameters depends on the particular measurement type and specification of the system design. Several kinds of research are being done in this field to improve the performance parameters of the DFOS.

Phenomenon of Scattering

When an incident light source is focused through an optical fibre, the light photons interact with the fibre molecules resulting in periodic agitation of fibre molecules with a frequency equal to that of the light source. These electronic perturbations lead to periodic disintegration of charge within the molecules emitted as electromagnetic radiation called scattered light [4] (Fig. 3). In short, scattering occurs when the medium through which the light propagates is in-homogeneous. As the light propagates in the forward direction, the power of the light decreases due to random scattering. Light scattering can be spontaneous or stimulated [2, 3, 5]. This paper focuses on spontaneous scattering. The DFOS concentrates mainly on three types of spontaneous scattering:

- Rayleigh scattering
- Raman scattering
- Brillouin scattering.

When the wavelength of backscattered light is identical to that of the incident light, it is termed as an elastic process and the resultant scattering is named Rayleigh scattering. In the Fig. 4, the Rayleigh is at the centre, coinciding with the incident light frequency [2, 3].

If the frequency of the backscattered light is different from the incident light, it leads to an inelastic process; the examples of this process are Raman and Brillouin scattering. In the figure (Fig. 4), the scattering on either side of the incident light implies a frequency shift that leads to two different frequency components called Stokes and anti-Stokes. The Brillouin frequency shift is less than that of Raman shift [2, 3] (Fig. 4).

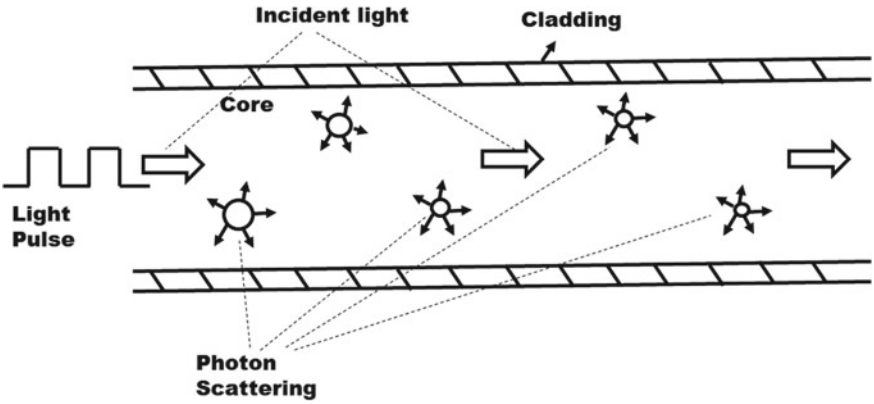


Fig. 3 Scattering phenomenon

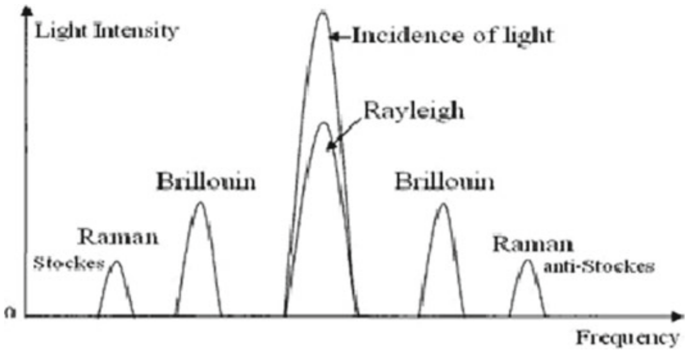


Fig. 4 Spontaneous scattering phenomenon

Optical Time Domain Reflectometer

DFOS employs the principle of optical time domain reflectometer (OTDR) [1, 3, 4] (Fig. 5), in which the small width and strong amplitude (power) optical pulses are launched into the fibre of known length. The propagating light loses its power due to the various attenuation factors like absorption, bending, splices, connectors and scattering. The small portion of the scattered light that falls within the angle of acceptance of the fibre in the reverse direction is recaptured and guided back towards the launching end, called the backscattered light. The backscattered light is a weak signal as it loses power through attenuation while returning. At the launching end, the backscattered light is focused on the receiver, which is further electronically processed, digitised and displayed which provides the detailed local loss information throughout the fibre length [5–7]. The acquisition and digitisation process needs to be synchronised with the pulsed light so that the travelling time of the backscattered signal and the location of the backscattering event happens accurately. The DFOS

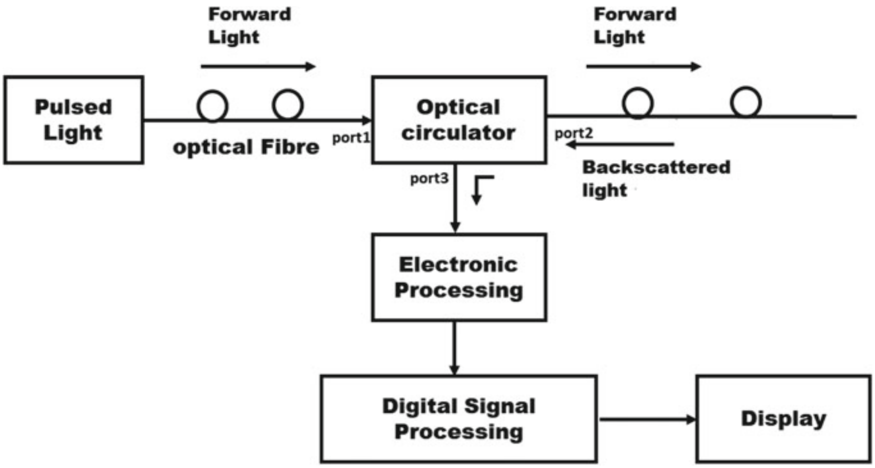


Fig. 5 Simple OTDR arrangement

utilises the concept of OTDR for monitoring the backscattered light from segments of fibre along its entire length.

The backscattering in OTDR is mainly due to Rayleigh scattering, which is a linear process [2]. The magnitude of the backscattered optical power is linearly proportional to the optical power at that location [8]. The measurement of the time-dependent backscattered power at the fibre input end provides the information about the loss distribution along the fibre [2, 3]. The fibre attenuation coefficient expressed as db/km is a function of the location along the fibre. For an input pulse of duration W_0, the backscattered power P_R is described as a function of time by the following equation [2, 3, 8, 9]:

$$P_R(T) = \frac{1}{2}\, P_i\ S\ \gamma_R\ W_0\ v_g\ exp(-\gamma\ v_g\ t) \tag{1}$$

where S = fraction of backscattered power captured, given by the equation

$$S = \frac{NA^2}{4n^2} \tag{2}$$

Parameters utilised in Eqs. 1 and 2 with the values considered for Silica fibre at a wavelength of 1.55 µm are present in Table 1.

v_g is the group velocity given by the equation:

$$v_g = \frac{c}{n} \tag{3}$$

Table 1 Parameter utilised in Eqs. 1 and 2

Description	Symbol	Value
Launched pump power (W)	Pi	As per the application
Numerical aperture	NA	0.12
Fibre refractive index	n	1.46
Rayleigh scattering coefficient (m^{-1})	γ_R	4.6×10^{-5}
Total attenuation coefficient (m^{-1})	γ	To be determined

where c is the velocity of light in m/s in vacuum. The spatial (location) information is obtained by mapping the time of flight of the light pulse travelling to and from the sensing location [10].

If x is the point on the fibre through which the backscattering occurred then as the light has to cover the distance to and fro in a given time t the length x is given by:

$$x = \frac{c\ t}{2\ n} = 0.5\ v_g\ t \tag{4}$$

Equation (1) as a function of distance can be expressed as:

$$P_R(x) = \frac{1}{2}\ P_i\ S\ \gamma_R\ W_0\ v_g\ \exp(-\gamma\ 2\ x) \tag{5}$$

Thus, the backscattered power can be obtained in terms of distance and in terms of time.

These are the basic OTDR equations utilised in DFOS for sensing the amplitude and location of fault variations. In order to improve the spatial resolution and measurement range of the distributed measurement system, numerous alternative techniques like phase- sensitive OTDR, optical frequency domain reflectometry (OFDR), optical low coherence reflectometry (OLCR), etc., have been developed.

Design Considerations for OTDR

The amplitude of the backscattered signal at the receiver end determines the signal-to-noise ratio (SNR) and the sensitivity of the OTDR systems. This amplitude is very weak and depends on the power of the launched pulse. The pulse peak power cannot be increased beyond a specific limit as it generates nonlinearities in the fibre resulting in erroneous measurement. Thus, it restricts improvising the signal-to-noise ratio.

The intensity of the backscattered signal measured at the receiver end also relies on the type of fibre utilised [11]. There are two broad categories of optical fibre, namely single-mode fibre and multimode fibre. The numerical aperture (NA) and the attenuation factor are the two parameters that account for fibre selection. Multimode fibres have high numerical aperture and attenuation. High NA implies high light-capturing ability, due to which it would capture more backscattered light, resulting in good intensity. However, an increase in attenuation factor causes the backscatter signal to undergo more losses while propagating through long distances. Thus, multimode fibres are restricted for short distances only up to 20 km. Single-mode fibres, though provide less intensity of the backscattered light but having less attenuation coefficient is utilised for long fibre lengths example hundreds of kilometres. The spatial resolution of an OTDR system is dependent on the input pulse width in the optical domain, the detection bandwidth and the sampling rate in the electronics domain [10]. The shorter the pulse width, greater is the spatial resolution, but the detection bandwidth should be high enough to fulfil the short pulse criteria.

The spatial resolution in terms of pulse width is expressed as:

$$\Delta x_{\text{pulse}} = \frac{c}{2n}\,\Delta T \tag{6}$$

where ΔT = pulse width in nanoseconds.

The spatial resolution in terms of detector bandwidth is expressed as:

$$\Delta x_{\text{detc}} = \frac{c}{2n}\,\frac{1}{2\,\Delta \upsilon} \tag{7}$$

where $\Delta \upsilon$ is the detector bandwidth in MeghaHertz.

For optimum signal-to-noise ratio, the

$$\Delta x_{\text{pulse}} = \Delta x_{\text{detc}} \tag{8}$$

therefore

$$\Delta T = \frac{1}{2\,\Delta \upsilon} \tag{9}$$

The pulsewidth ranges from nanoseconds to microseconds as per the application. For example, a spatial resolution of 1m would require a light having pulse width of 10 ns (nanoseconds). For such a short pulse width, the detector should have a bandwidth of at least 50 MHz (as per Eqs. 6, 7, 8, and 9). The repetition rate of these pulses should be such that the backscatter pulse returning from the other end to the launching end does not overlap with the forward pulses [12]. The repetition rate is dependent on the fibre length and is represented by Eq. 4. For example, a fibre length of 10 km would require the light pulses to be launched in the interval of 100 microseconds that yield the repetition frequency of 10 KHz. Similarly, a fibre length of 100 km would require the pulses to be launched with a frequency of 1 KHz. Thus, for long fibre lengths, the repetition frequency is low as compared to short lengths, and the repetition rate should be selected appropriately.

The receiver bandwidth should be such that it converts the whole of the backscattered optical power pulse into electrical without any dispersion. Thus, the receiver bandwidth should be selected in an optimised way, to maximise the signal-to-noise ratio [13]. The output current of the photodiode is proportional to the detected optical power; therefore, a preamplifier of transimpedance type is utilised. The gain and bandwidth of this amplifier should be selected considering the noise effects in the system in order to have a good signal-to-noise ratio. The photodiodes utilised are usually Avalanche photodiodes (APD), but due to improvement in technology, fast PIN photodiodes can also be implemented based on the applications.

The electrical signal is then digitised using analog to digital converters (ADC). The ADC sampling rate is also one of the factors in fixing the spatial resolution of the system. As, this sampling rate determines the distance of adjacent data to be sampled. For example, an ADC with 100 MHz sampling rate gives a spatial resolution of 1 m, and that of 500 MHz gives a spatial resolution of 0.2 m. However, it is not practical to increase the sampling rate to obtain high spatial resolution as spatial resolution also depends on the pulse width and receiver bandwidth. Thus, for a given spatial resolution, the pulse width, bandwidth and sampling rate should be selected accordingly. The ADC sampling frequency should be higher than the receiver bandwidth for effective conversion [12]. As the backscattered signal is very weak, averaging technique is employed to improve the signal-to-noise ratio. The number of samples digitised and the sampling frequency of ADC decide the measuring time of the OTDR system. High resolution 12 bit, 16 bit ADCs with high sampling rates can be utilised for conversion, which aids in lowering the measuring time of the system. Thus, taking into account all the design criteria a high resolution and fast OTDR as well as DFOS can be designed, so that it can respond to variations.

Rayleigh Scattering

Rayleigh scattering is an elastic process, i.e. the frequency of the scattered light is identical to the frequency of the source. It is caused by the inhomogeneities in the refractive index of the fibre that occurs during fabrication on the scale which is much smaller than the wavelength of the propagating light [3, 10]. It is characterised by an attenuation coefficient proportional to λ^{-4}, and therefore, Rayleigh scattering is at short wavelengths. The attenuation coefficient is given by [3, 10, 14]

$$\gamma_R = \frac{8 \prod^3}{3\,\lambda^4}\, n^8\, p^2\, \beta_c\, K\, T_F \tag{10}$$

Table 2 describes the various parameters employed in Eq. 10 and their values related to silica fibre. Although Rayleigh backscatter signals are used to identify the losses and locate the point of damage in the fibre, researchers have demonstrated that with slight modifications in the OTDR principle, Rayleigh backscatter can be utilised to monitor the spatial distribution of vibrational disturbances on the optical fibre.

Table 2 Parameter utilised in Eq. 10

Description	Symbol	Value
Boltzmann's constant	K	1.381×10^{-23}
Average photoelastic coefficient	p	0.286
Fibre refractive index	n	1.46
Rayleigh scattering coefficient (m^{-1})	γ_R	4.56×10^{-5}
Wavelength of the propagating light	λ	1550 nm
Isothermal compressibility at TF	β_c	$7 \times 10^{-11} m^2 N^{-1}$
Fictive temperature	T_F	1950 K

Raman Scattering

Raman scattering is an inelastic nonlinear scattering phenomenon in which the interaction between the incident photon and the molecular vibrations of the fibre material causes a transition between two vibrational states that either produce or absorb a phonon [3, 10, 15]. These are high-frequency optical phonons. Creation of a phonon in the incident leads to reduced scattered photon frequency giving rise to the Stokes wave, whereas absorption of a phonon causes increase in the scattered photon frequency termed as anti-Stokes wave (AS) [3, 16, 17]. Large frequency shift in characteristic of Raman scattering (typical tera hertz range 10^{12} for an incident wavelength of 1550 nm [16, 17]) of the incident photon to Stokes and anti-Stokes is due to the generation and absorption of optical phonons. Raman scattering is a temperature-dependent process caused by thermally driven molecular vibrations [18]. The AS Raman wave being highly sensitive to temperature provides the basis of temperature measurement in Raman-based DFOS [4, 17, 18]. The wavelengths of the anti-Stokes (λ_{AS}) and Stokes(λ_S) are related to the incident light of wavelength (λ_0) [3]

$$\frac{1}{\lambda_{AS}} = \frac{1}{\lambda_0 + \upsilon_R \text{ cm}^{-1}} \tag{11}$$

$$\frac{1}{\lambda_{ST}} = \frac{1}{\lambda_0 - \upsilon_R \text{ cm}^{-1}} \tag{12}$$

where (υ_R) represents the wavenumber; the frequency shift is related to the wavenumber (υ_R) which is given by

$$\upsilon_R = \frac{\omega}{2 \prod c} \tag{13}$$

where ω is the shifted frequency in radians/sec and c speed of light in vaccum in cm/sec.

The wavenumber for SiO_2 is 440 cm^{-1} and the wavenumber for GeO_2 is 460 cm^{-1} [3].

Table 3 Parameters utilised in Eqs. 14, 15, and 16

Description	Symbol	Value
Planck's constant	h	6.63×10^{-34} J s
Speed of light in vacuum	c	3×10^{8} m/s
Boltzmann's constant	K	1.381×10^{-23} J K $^{-1}$
Wave number	υ_R	440 cm^{-1} (for Silica fibre)
Absolute temperature	T	°K
Stokes wavelength	λ_S	nanometers (nm)
Anti-Stokes wavelength	λ_{AS}	nanometers (nm)

The Raman backscattered light provides the information of temperature variation around the point of scatter. As mentioned, the AS Raman signal is highly temperature sensitive, but the AS signal is also sensitive to fibre attenuation and local losses in terms of splices, connectors, etc. Any variations in the AS signal only due to these local losses and attenuation coefficient would be misinterpreted as fluctuations in temperature. This requires the AS Raman intensity to be normalised by a temperature-independent Stokes intensity [18]. The ratio of these two intensities provides information about the absolute temperature and cancels out the effect of the intrinsic fibre losses [17–20].

$$R = \frac{I_{AS}\ (T)}{I_S\ (T)} = \frac{\lambda_S}{\lambda_{AS}}^4 \exp\left(\frac{-h\ c\ \upsilon_R}{K\ T}\right) \tag{14}$$

The intensity of the anti-Stokes and Stokes scatter as a function of temperature is given by [3]

$$I_{AS}\ (T) = \frac{1}{\lambda_{AS}^4} \frac{1}{\exp\left(\frac{h\ c\ \upsilon_R}{KT}\right) - 1} \tag{15}$$

$$I_S\ (T) = \frac{1}{\lambda_S^4}\left(\frac{1}{\exp\left(\frac{h\ c\ \upsilon_R}{K\ T}\right) - 1} + 1\right) \tag{16}$$

the parameters used in the above equations are listed in Table 3.

Raman OTDR System

The basic principle of Raman OTDR temperature measurement system consists in filtering the Stokes and anti-Stokes component from the back scattered light, determining the ratio of the intensities of these components as a function of distance. The correct measurement of temperature is obtained by normalising this ratio with the intensity ratio at a known temperature of the calibration zone of the fibre [3].

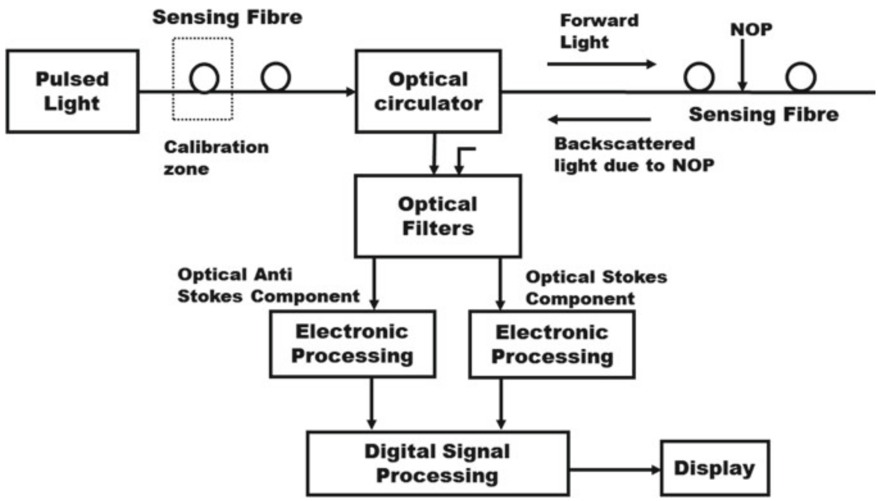


Fig. 6 Raman OTDR arrangement

The intense laser pulses are launched into the sensing fibre (of known length) through optical circulator (is a three-port device that ensures propagation of light is unidirectional, port1 is connected to port 2 and port2 to port3), used to separate the returning backscattered light from the forward travelling light. The backscattered light is directed to the detection system through the coupler (Fig. 6). In the detection system, the backscattered light is bifurcated into Stokes and anti-Stokes photons by the optical filters (optical filters are devices that transmit light of a particular wavelength through it, thus helps in filtering out unwanted wavelengths of light). The signals are then detected by fast photodetectors and further processed in signal processing unit. The signal processing unit acquires and stores the Raman anti-Stokes and Stokes component of every light pulse, averages the collected data and processes it, so that the ratio of anti-Stokes to Stokes signal is calculated which further helps in calculating the change in temperature. Finally, the change in temperature with respect to distance is displayed. In order to improve the spatial resolution, range of the system and response time of the system researchers have utilised various different techniques in the Raman OTDR. Also Raman OFDR is utilised that employs frequency domain techniques instead of time of flight measurement to establish relationship of the data with the fibre length.

Raman OTDR (ROTDR) Signal Processing

The basic steps of signal processing of ROTDR is the same as explained in Sect. 2.3. However, the important consideration in obtaining the backscattered signal is that the Raman signals are by far the weakest signals in the backscattered light. Therefore, an utmost care needs to be accounted during the optical filters selection to filter the

Stokes and anti-Stokes signals from the backscattered light [3]. The filters should have matched wavelengths concerning Stokes and anti-Stokes wave and should offer adequate gains in order to maintain the signal amplitude. A slight mismatch in the wavelength window leads to erroneous results despite having right electronics. For example, if a light source of wavelength 1550 nm is utilised then as per Eqs. (11, 12), the optical filters having centre wavelength of 1450 nm to obtain the anti-Stokes signal and an optical filter with a wavelength of 1650 nm for Stokes signal need to be utilised in the system.

Brillouin Scattering

Brillouin scattering, another type of inelastic scattering, was first observed in an amorphous, glass-like substance in 1950 by Krishnan [21].

Brillouin scattering occurs due to the interaction between an incident wave and thermally driven material density fluctuations that travel at the speed of sound [22, 23]. In the thermal agitation, acoustic phonons of all frequencies are created and reabsorbed. The acoustic phonons vary the refractive index of the glass and cause refractive index modulation [23]. Phonons whose acoustic wavelength matches the optical wavelength of the incident light interact with the probe by either adding or subtracting energy to the incident photon resulting in anti-Stokes or Stokes scattered wave, respectively [24]. The acoustic phonons have much lower energies that results in very small frequency shift, characteristics of Brillouin scattering (approximately 11 GHz for an incident light of 1530 nm) [3, 8]. On the basis of conservation of energy [10]:

$$\nu_p - \nu_A = \nu_S \tag{17}$$

$$\nu_p + \nu_A = \nu_{AS} \tag{18}$$

where $\nu_p, \nu_A, \nu_S, \nu_{AS}$ are the frequencies of the incident light, acoustic phonon, Stokes wave and anti-Stokes wave, respectively. The optical frequency of the Brillouin scattered light is closer to the incident light. The amount of the optical frequency shift is given by [25, 26]

$$\nu_B = \frac{2\, n\, \nu_A}{\lambda_0} \tag{19}$$

where n = refractive index of the fibre, λ_0 = wavelength of the incident light and v_A = velocity of the acoustic wave in the optical fibre.

The acoustic velocity is given by [26]

$$\nu_A = \sqrt{\frac{E\,(1-K)}{(1+K)(1-2K)\,\rho}} \tag{20}$$

where E is the Young's modulus, K is the Poisson's ratio and ρ is the fibre density.

Equation 19 depicts that for a given incident wavelength of light, the Brillouin frequency shift ν_B is dependent on the velocity of the acoustic wave and the refractive index of the fibre. Both these quantities are in turn dependent on fibre intrinsic characteristics (e.g., fibre composition and environmental factors such as temperature and strain that causes change in the fibre). Thus, the frequency shift of Brillouin scattering are caused by strain and temperature [21, 22, 24–27].

Brillouin Measurement for Temperature and Strain

Brillouin scattering may be used to measure temperature and/or strain distribution by virtue of the effect of Brillouin scattered power and frequency shift. If Brillouin power is utilised for measurement, then normalisation is necessary to compensate for intrinsic fibre losses and Rayleigh backscatter trace is utilised for normalisation [28]. If both the temperature and strain are to be measured simultaneously, then the normalised Brillouin power and frequency shift must be measured. The following equations reveal the method of extracting the temperature and strain parameters from the measurements. The Brillouin frequency is expressed by [29]:

$$\nu_B = \nu_0 + \frac{\partial \nu}{\partial \epsilon}\epsilon + \frac{\partial \nu}{\partial T}T \tag{21}$$

where v_0 is the frequency of the incident light, ϵ is strain expressed as micro strain and T is the temperature expressed in $^\circ$ Centigrade. The frequency shift of Brillouin scattering is [29, 30]

$$\delta \nu_B = C_{\nu\epsilon}\,\delta\epsilon + C_{\upsilon T}\,\delta T \tag{22}$$

The values of strain and temperature coefficient have been determined for an incident light of wavelength 1.5 micrometer given by [26, 29–34]

$$C_{\nu\epsilon} = 0.0483 \pm 0.0004 \left(\frac{\text{MHz}}{\mu\epsilon}\right) \tag{23}$$

$$C_{\nu T} = 1.10 \pm 0.02 \left(\frac{\text{MHz}}{K}\right) \tag{24}$$

The normalised Brillouin power is given by [29]

$$\Delta P_B = C_{P\epsilon}\,\delta\epsilon + C_{PT}\,\delta T \tag{25}$$

the values of C_{PT} and $C_{P\epsilon}$ (for 1.5 micrometer light) are [31, 32, 35]

$$C_{P\epsilon} = -9 \pm 1 \times 10^{-4} (\%\epsilon) \tag{26}$$

$$C_{PT} = 0.33 \pm 0.3(\%K) \tag{27}$$

In matrix form, Eqs. 22 and 25 give

$$\begin{pmatrix} \Delta P_B \\ \Delta \nu_B \end{pmatrix} = \begin{pmatrix} C_{PT} & C_{P\epsilon} \\ C_{\nu T} & C_{\nu\epsilon} \end{pmatrix} \begin{pmatrix} \Delta T \\ \Delta \epsilon \end{pmatrix} \tag{28}$$

The inverse of this equation exists, and therefore,

$$\begin{pmatrix} \Delta T \\ \Delta \epsilon \end{pmatrix} = \frac{1}{C_{PT}\, C_{\nu\epsilon} - C_{P\epsilon}\, C_{\nu T}} \begin{pmatrix} C_{\nu\epsilon} & -C_{P\epsilon} \\ -C_{\nu T} & C_{PT} \end{pmatrix} \begin{pmatrix} \Delta P_B \\ \Delta \nu \end{pmatrix} \tag{29}$$

Equation 29 can be utilised to calculate ΔT and $\Delta\epsilon$. Thus, by measuring, the Brillouin power and frequency shift at all points in the fibre, the temperature and strain distribution can be measured [28–32].

Brillouin OTDR System

Brillouin OTDR system is based on the spontaneous Brillouin backscatter power detection. In the basic or direct detection method, the Brillouin backscatter signal is segregated from the Rayleigh scatter using an appropriate and accurately stable optical filters as the Brillouin frequency shift is very small. After separation, the optical signals are converted to electronic signals, digitised and acquired for further processing and retrieving the temperature and/or strain variations along the length of the fibre (Fig. 7).

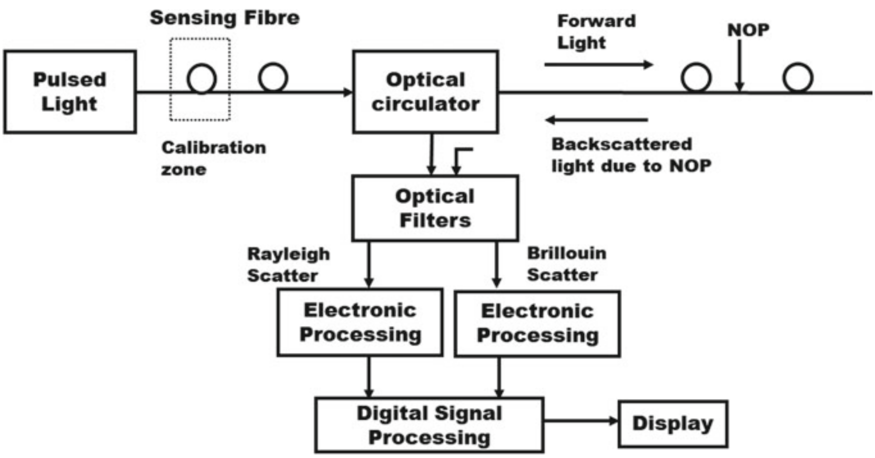


Fig. 7 Spontaneous Basic Brillouin OTDR arrangement

The Brillouin conversion efficiency is very low, so the direct detection method of BOTDR is suitable for low sensing range with low accuracy [36]. In order to increase the sensing range, accuracy and sensitivity, this basic technique has been modified which has lead to various other techniques like heterodyne based BOTDR [37, 38], pulsed coding technique-based BOTDR [39–41]. All these techniques employ spontaneous Brillouin scattering. Stimulated Brillouin scattering too has been employed in order to increase the spatial resolution and operating range of the system, Brillouin optical time domain analysis (BOTDA) [42, 43] and Brillouin optical correlation domain analysis (BOCDA) [44, 45] are examples of such systems.

Brillouin OTDR Signal Processing

Apart from following the basic design steps for electronic processing as described in Sect. 2.3, attention must be paid in the method of separating the Brillouin signal from the Rayleigh backscattered signal as the two signals have a small difference in wavelengths. High-resolution optical filters like Mach–Zender interferometers [3] need to be implemented for their separation, and also heterodyne techniques [46] have been implemented for the same. As the amplitude of Brillouin signal is better than the Raman signal, it provides an improved signal-to-noise ratio.

Comparison of Scattering Phenomenon

Table 4 gives a comparison of the three spontaneous scattering phenomenon. It shows that the Brillouin scattering principle for DFOS provides long-sensing range and an option of simultaneous strain and temperature measurement that are useful in various applications.

Table 5 lists the main features of the spontaneous anti-Stokes Raman scattering and Brillouin scattering at the wavelength of 1550 nm [47, 48]. Though the temperature sensitivity of Brillouin scattered light is less than the Raman scattering, the Brillouin scattering phenomena outweighs the Raman principle as a basis for DFOS in numerous ways. As mentioned, Brillouin scattering can be implemented to measure strain and temperature simultaneously, whereas the Raman scattering is insensitive to any type of strain variations. The magnitude of the Brillouin scatter is larger than the Raman scatter which provides a relatively good signal-to-noise ratio that results in increased spatial resolution and temperature range of the DFOS system. As the Raman signal has a broad bandwidth, the optical amplification techniques are not effective, thus limiting the range of the fibre, whereas in Brillouin scattering optical amplification is possible due to its narrow bandwidth that aids in increasing the overall sensing range of the DFOS system.

Table 4 Comparison of different Scattering phenomenon

Description	Rayleigh scattering	Raman scattering	Brillouin scattering
Type of scattering	Elastic	Inelastic	Inelastic
Cause of scattering	Interaction with fibre molecules	Interaction with optical phonons	Interaction with acoustic phonons
Wavelength of scattered light	Same as that of the incident light	Produces Stokes and anti-Stokes wavelength	Produces Stokes and anti-Stokes wavelength
Frequency shift	No shift in frequency	Large frequency shift (Tera Hertz range)	Small frequency shift (Gega Hertz range)
Band width	Same as that of the incident light	Broader	Narrower
Light intensity	strongest	Weak	strong
Sensitivity	Not much sensitive to temperature and strain, can be used for dynamic measurement	Sensitive to temperature only	Sensitive to temperature and strain
Sensing range	Metre, kilometre to tens of kilometre (depending on different schemes)	Metre, kilometre to tens of kilometre (depending on different schemes)	Metre, kilometre to hundreds of kilometre (depending on different schemes)

Table 5 Features of Raman and Brillouin scattering [10, 29, 47, 48]

Parameters	Raman scattering	Brillouin scattering
Frequency shift	13 THz	11 GHz
Bandwidth	~5THz	~20 to 100 MHz
Gain coefficient	7×10^{-14}m/W	5×10^{-11}m/W
Free-space wavelength shift	~104 nm	~0.09 nm
Scattering power ratio (wrt Rayleigh scattering)	~30 dB	~15 dB
Temperature sensitivity	~0.8% degree K	~0.3% degree K
Frequency shift temperature sensitivity	–	1.1 MHz/degree K
Intensity strain sensitivity	–	-9×10^{-4} %microstrain

Applications of DFOS

Temperature, pressure, vibrations and their variations with respect to location and time are key parameters for facility management systems and for the risk assessment in civil structures [49]. To identify and restrict the dangerous situations at an initial stage and to minimise possible damage to the society, temperature, vibration and strain inspecting system is required with which evaluations can be achieved simultaneously over huge distances with good range and space resolution. The sensing mechanism utilised should be free from causing any unwanted accidents. Looking into the above considerations, the DFOS offers huge solutions. The DFOS are extensively utilised in structural health monitoring of buildings, dams, bridges, tunnels and oil and gas leakage detection in pipelines [50].

Another field where DFOS are being utilised in is the electric power industry where temperature monitoring of the conductor, equipment is essential for safe reliable operation of the power plant, e.g. measurement of stator, rotor winding temperature, transformer temperature and generator circuit breakers [51].

DFOS is being employed in process industries to monitor the refractory lined performer vessels, blast furnaces and ceramic filters in pressurised, fluidised-bed combustion processes to detect the hot spots forming at the surface of the vessel [52].

The DFOS is being implemented in harsh fields like oilwells, coal mines, nuclear reactors, cement and concrete curing and many more such fields.

All IOT-based systems for temperature, pressure, acoustic measurements in process industries and power plants can be easily implemented using DFOS. Thus, DFOS have a wide range of applications in near future.

Limitations in the Field of DFOS

Although, DFOS offer wide applications in numerous fields but still have not emerged fully in the market due to few of its shortcomings. The major trouble with the fibre sensors in field applications is the process of installation as the durability of base fibre is very low and shielded fibres have low sensitivity to temperature and strain. The special cable arrangement and also the field installation cost is very high which restricts its adoption. With the availability of advanced technologies and economical instrumentation techniques, the implementation cost of DFOS will decline in near future. As fibre optic sensor technology continues to improve and more cost-effective and high-performance DFOS produced, the distributed sensing technology will have no boundaries.

Conclusion

This paper presents a systematic and comprehensive outline of the distributed fibre optic sensors including the different scattering mechanisms, namely Rayleigh, Raman and Brillouin. The design considerations for OTDR and DFOS, comparisons of the scattering, applications of DFOS and its limitations.

As mentioned, though the implementation cost of the DFOS at present is high but with the advancement of new cost-effective technologies, many distributed sensing system will be commercialised and utilised in various applications like aerospace, in civil structures, process industry, power industry, Internet of things, etc., making DFOS inevitable for human beings.

References

1. Taylor HF, Lee CE (1993) Apparatus and method for fiber optic intrusion sensing. U.S. Patent No. 5,194,847
2. Agarwal GP (1995) Non-linear fiber optics, 2nd edn. Academic Press, ISBN 0-12-045142-5
3. Hartog AH (2018) An introduction to distributed optical fiber sensors. CRC Press
4. Hahn DW (2007) Raman scattering theory. University of Florida, Department of Mechanical and Aerospace Engineering
5. Horiguchi T, Kurashima T, Koyamada Y (1993) Measurement of temperature and strain distribution by Brillouin frequency shift in silica optical fibers. In: Distributed and multiplexed fiber optic sensors II, vol 1797. International Society for Optics and Photonics
6. Bao X, Chen L (2012) Recent progress in distributed fiber optic sensors. sensors 12(7):8601–8639
7. Shimaponda M (2014) Numerical modelling of a Raman-Rayleigh distributed temperature fiber sensor implementing correlation techniques. ME thesis, University of Johannesburg
8. Hui R, Sullivan MO (2009) Fiber optic measurement techniques. Elsevier Inc.
9. Personick SD (1977) Photon probe-an optical-fiber time-domain reflectometer. Bell Syst Tech J 56(3):355–366
10. Maughan SM (2001) Distributed fibre sensing using microwave heterodyne detection of spontaneous Brillouin backscatter. PhD thesis, University of Southampton
11. da Silva LCB, Pontes MJ, Segatto MEV (2017) Analysis of parameters for a distributed temperature sensing based on Raman scattering. J Microwaves Optoelectron Electromagn Appl 16(1):259–272
12. Beller J (1997) Optical time domain reflectometer (OTDR) with improved dynamic range and linearity. U.S. Patent No. 5,621,518
13. Conduit AJ et al (1980) An optimized technique for backscatter attenuation measurements in optical fibres. Opt Quantum Electron 12(2):169–178
14. Senior JM (1992) Optical fiber communications, 2nd edn. Prentice Hall, ISBN 0-13-635426-22
15. Brinkmeyer E (1980) Analysis of the backscattering method for single-mode optical fibers. JOSA 70(8):1010–1012
16. Long DA (2002) The Raman effect: a unified treatment of the theory of Raman scattering by molecules, pp 541–564
17. Farahani MA, Gogolla T (1999) Spontaneous Raman scattering in optical fibers with modulated probe light for distributed temperature Raman remote sensing. J Lightwave Technol 17(8):1379
18. Soto MA et al (2011) Impact of loss variations on double-ended distributed temperature sensors based on Raman anti-Stokes signal only. J Lightwave Technol 30(8):1215–1222
19. Dakin JP et al (1985) Distributed optical fibre Raman temperature sensor using a semiconductor light source and detector. Electron Lett 21(13):569–570
20. Fernandez AF et al (2005) Radiation-tolerant Raman distributed temperature monitoring system for large nuclear infrastructures. IEEE Trans Nuclear Sci 52(6):2689–2694
21. Krishnan RS (1950) Fine structure of the Rayleigh line in amorphous substances. Nature 165(4206):933–934
22. Nikles M, Thevenaz L, Robert PA (1997) Brillouin gain spectrum characterization in single-mode optical fibers. J Lightwave Technol 15(10):1842–1851
23. Agarwal GP (2013) Non-linear fiber optics. Academic Press
24. Galindez-Jamioy CA, Lopez-Higuera JM (2012) Brillouin distributed fiber sensors: an overview and applications. J Sensors 2012
25. Culverhouse D et al (1989) Potential of stimulated Brillouin scattering as sensing mechanism for distributed temperature sensors. Electron Lett 25(14):913–915
26. Horiguchi T, Kurashima T, Tateda M (1989) Tensile strain dependence of Brillouin frequency shift in silica optical fibers. IEEE Photon Technol Lett 1(5):107–108
27. Lees GP et al (1998) Advances in optical fiber distributed temperature sensing using the Landau-Placzek ratio. IEEE Photon Technol Lett 10(1):126–128

28. Kee HH, Lees GP, Newson TP (2000) Technique for measuring distributed temperature with 35-cm spatial resolution utilizing the Landau-Placzek ratio. IEEE Photon Technol Lett 12(7):873–875
29. Zhang Z et al (2012) Recent progress in distributed optical fiber Raman photon sensors at China Jiliang University. Photon Sensors 2(2):127–147
30. Maughan Sally M, Kee Huai H, Newson Trevor P (2001) Simultaneous distributed fibre temperature and strain sensor using microwave coherent detection of spontaneous Brillouin backscatter. Measurement Sci Technol 12(7):834–842
31. Parker TR et al (1997) Temperature and strain dependence of the power level and frequency of spontaneous Brillouin scattering in optical fibers. Opt Lett 22(11):787–789
32. De Souza K, Wait PC, Newson TP (1997) Characterisation of strain dependence of the Landau-Placzek ratio for distributed sensing. Electron Lett 33(7):615–616
33. Izumita H et al (1996) Brillouin OTDR employing optical frequency shifter using side-band generation technique with high-speed LN phase-modulator. IEEE Photon Technol Lett 8(12):1674–1676
34. Kurashima T, Horiguchi T, Tateda M (1990) Thermal effects of Brillouin gain spectra in single-mode fibers. IEEE Photon Technol Lett 2(10):718–720
35. Wait PC, Newson TP (1996) Landau Placzek ratio applied to distributed fibre sensing. Opt Commun 122(4–6):141–146
36. Kee HH, Lees GP, Newson TP (2000) All-fiber system for simultaneous interrogation of distributed strain and temperature sensing by spontaneous Brillouin scattering. Opt Lett 25(10):695–697
37. Shimizu K et al (1994) Coherent self-heterodyne Brillouin OTDR for measurement of Brillouin frequency shift distribution in optical fibers. J Lightwave Technol 12(5):730–736
38. Kurashima T et al (1997) Performance improvement of a combined OTDR for distributed strain and loss measurement by randomizing the reference light polarization state. IEEE Photon Technol Lett 9(3):360–362
39. Sakairi Y et al (2008) Prototype double-pulse BOTDR for measuring distributed strain with 20-cm spatial resolution. In: SICE annual conference. IEEE, New York
40. Li Q et al (2016) High spatial resolution BOTDR based on differential Brillouin spectrum technique. IEEE Photon Technol Lett 28(14):1493–1496
41. Shibata R et al (2017) Improving performance of phase shift pulse BOTDR. IEICE Electron Express 14-20170267
42. Horiguchi T, Tateda M (1989) BOTDA-nondestructive measurement of single-mode optical fiber attenuation characteristics using Brillouin interaction: theory. J Lightwave Technol 7(8):1170–1176
43. Kishida K, Li C-H, Nishiguchi K (2005) Pulse pre-pump method for cm-order spatial resolution of BOTDA. In: 17th International conference on optical fibre sensors, vol 5855. International Society for Optics and Photonics
44. Fellay A et al (1997) Distributed sensing using stimulated Brillouin scattering: towards ultimate resolution. Optical Society of America, Optical Fiber Sensors
45. Hasegawa T, Hotate K (1999) Measurement of Brillouin gain spectrum distribution along an optical fiber by direct frequency modulation of a laser diode. In: Fiber optic sensor technology and applications, vol 3860. International Society for Optics and Photonics
46. Shimizu K et al (1993) Coherent self-heterodyne detection of spontaneously Brillouin-scattered light waves in a single-mode fiber. Opt Lett 18(3):185–187
47. Wait PC, De Souza K, Newson TP (1997) A theoretical comparison of spontaneous Raman and Brillouin based fibre optic distributed temperature sensors. Opt Commun 144(1–3):17–23
48. Horiguchi T et al (1995) Development of a distributed sensing technique using Brillouin scattering. J Lightwave Technol 13(7):1296–1302
49. Hurtig E, Großwig S (1998) Distributed fiber optics for temperature sensing in buildings and other structures. In: IECON'98. Proceedings of the 24th Annual conference of the IEEE Industrial Electronics Society (Cat. No. 98CH36200), vol 3. IEEE, New York

50. Bado MF, Casas JR (2021) A review of recent distributed optical fiber sensors applications for civil engineering structural health monitoring. Sensors 21(5):1818
51. Boiarski AA et al (1995) Temperature measurements in power plant equipment using distributed fiber optic sensing. IEEE Trans Power Del 10(4):1771–1778
52. Iida O, Hironaga T, Yanagisawa K (1994) Method and apparatus for monitoring temperature of blast furnace and temperature control system using temperature monitoring apparatus. U.S. Patent No. 5,356,220

Ensemble Model to Forecast the End of the Covid-19 Pandemic

S. Shwetha, Pramod Sunagar, S. Rajarajeswari, and Anita Kanavalli

Abstract The coronavirus disease 2019 (Covid-19) epidemic has caused a worldwide health catastrophe that has had a profound influence on how we see our planet and our daily lives. In this pandemic circumstance, machine learning (ML) based prediction models demonstrate their value in predicting perioperative outcomes to enhance decision-making on future course of action. Ensemble learning is used in the majority of ML based forecasting approaches. The ML models anticipate the number of patients who will be affected by Covid-19, and use this information to forecast the end of the pandemic is to be leveraged. Three types of predictions are made: the number of newly infected cases, the number of deaths, and the number of recoveries in the next '*x*' number of days. By combining one of the forecasting models with classifiers, we can predict the end of the pandemic. The proposed idea combines the SIRF model from epidemiology and a forecasting machine learning model named Prophet and a Naïve Bayes Classifier to predict the end of the pandemic. Using the theoretical equations of the SIRF model, we developed a formula for infectious growth rate. The classifier uses this infectious growth rate to check if the infection is fading. With confirmed, recovered and fatalities data, the infectious growth rate is calculated. Naïve Bayes classifier is used to check if the pandemic is about to end or not. If not then forecast the data for '*x*' number of days and do the calculations again. The process continues until we get a time frame where the pandemic may reach its end. The results are discussed for 2 countries India and Israel. The forecasts done for Israel were very accurate to the actual data, whilst for India it was less comparatively as India was hit by 2 waves of Covid-19 pandemic. By leveraging the forecasting and classification capabilities of machine learning models like FBProphet, Naïve Bayes

S. Shwetha · P. Sunagar (✉) · S. Rajarajeswari · A. Kanavalli
M. S. Ramaiah Institute of Technology (Affiliated to VTU), Bangalore, India
e-mail: pramods@msrit.edu

S. Rajarajeswari
e-mail: raji@msrit.edu

A. Kanavalli
e-mail: anithak@msrit.edu

V. Bindhu et al. (eds.), *Proceedings of Third International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 844,
https://doi.org/10.1007/978-981-16-8862-1_53

Classifier, and the mathematical equations of the SIRF model from epidemiology, the life span of the pandemic is determined.

Keywords Ensemble technique · Naïve Bayes · SIRF (Susceptibles · Infecteds · Recovered · Fatalities) model · FB Prophet (Facebook Prophet) model · Covid-19 pandemic

Introduction

At the beginning of December 2019, a virus started to spread rapidly across the mainland of China. In the following weeks, it spread across different countries and created a sense of global panic. The virus has been named "SARS-CoV-2," and the disease it causes has been named "coronavirus disease 2019," abbreviated "Covid-19." The coronavirus has infected millions of people worldwide. Infected people suffer from severe respiratory problems, may develop serious illnesses leading to a weak immune system. The World Health Organization (WHO), alarmed by this development, declared Covid-19 a pandemic. Many surveys are made on the prediction and detection of epidemic diseases outbreaks [1]. Countries declared partial or total lockdowns to contain the spread of disease, social distancing rules came into place, and wearing masks was made compulsory throughout the world. Totally, 221 countries and territories around the world have reported 219,979,049 confirmed cases of the Covid-19 and a death toll of 4,557,372 affecting roughly 10% of world's population. Over the last decade, Machine Learning (ML) has been established as a prominent area of research by solving very complex and sophisticated real-world issues. ML is applied in critical domains like healthcare, image and speech recognition, online fraud detection, stock market trading, etc. Forecasting techniques based on machine learning (ML) demonstrate their use in anticipating perioperative occurrences and improving decision-making on future course of action. The utilization of a model to predict future values based on previously observed values is known as time series forecasting. Most of the machine learning-based forecasting techniques will use ensemble learning. Mathematical models are useful for understanding how an infection behaves when it affects a population and determining whether it will be eradicated or persist under various situations [2, 3].

The proposed system operates on a novel coronavirus dataset. It fits the dataset to the SIRF model and derives an equation for infectious rate for a geographical area using theoretical concepts of this model. Then, the resultant data is passed to the Naïve Bayes classifier, which yields in 4 categories, one of which indicates the pandemic is fading for that geographical area. For example, considering today's active cases, recoveries, and fatalities of India, we can use this classifier model to check under which category it falls for the current day. Using one of the forecasting models, we can forecast to "x" number of days in the future and determine active cases, recoveries, and fatalities on the "xth day." The result of this forecasting model should be fit to the SIRF model. Then, it is passed on to the Naïve Bayes classifier to check if the forecasted result falls under the category of end of life. If not, the process is repeated further to determine the date at which the pandemic ends. As the proposed system uses multiple ML models, it falls under the category of ensemble learning. Given confirmed cases, active cases, fatalities, and recoveries for a period of time, transmission rate (β—the rate at which confirmed cases is increasing), recovery rate (γ—the rate at which infected people are recovering), and fatality rate (δ—the rate at which number of infected people deceases) are calculated, an equation for infectious growth rate is derived using these elements. The transmission rate, recovery rate, fatality rate, and infectious growth rate together are called transmission dynamics matrix. Naïve Bayes Classifier can use this matrix to check if the infection is fading. The end of life for pandemic can be obtained by considering transmission rate as 0 (when the infection is not spreading). The recovery rate is set to 1 (All infected patients are recovering), and the fatality rate is 0 or negligible. With this criterion, we can evaluate against the transmission dynamics matrix of forecasted data and conclude.

Related Works

Machine learning models have long been used to define and prioritize adverse factors for a threat in several application domains. Many research papers have demonstrated that machine learning algorithms can anticipate the number of future patients who will be impacted by Covid-19 [4]. These are also used to predict the number of deaths and recoveries for a certain number of days. The dataset utilized in the study was available from the GitHub repository provided by Johns Hopkins University's Centre for Systems Science and Engineering. Mainly, four models are considered in this research, namely: Linear regression (LR), Least Absolute Shrinkage and Selection Operator (LASSO), Support Vector Machine (SVM), and Exponential Smoothing (ES). The predictions of dangerous circumstances are made for the next 10 days by each model. According to the study's findings, using these strategies in the current Covid-19 pandemic scenario is a potential mechanism. In order to analyse the performance of each model, five metrics are considered: R-squared (R 2) score, Adjusted R-Square (R adjusted2), mean square error (MSE), mean absolute error (MAE), and root mean square error (RMSE). More research concentrates on using mathematical modelling techniques such as Rough Set Support Vector Machine (RS-SVM), Bayesian Ridge and Polynomial Regression, SIR model, and RNN [5].

A study conducted in [6] aims at using time series models like ARIMA and Prophet to predict the major parameters of Covid-19 data acquired from 10 world's most affected countries. The dataset is considered from January 22, 2020, to May 20, 2020. The data is sponsored by the ESRI living atlas team, the Applied Physics Lab (APL), and the Centre for Systems Science and Engineering (CSSE) at Johns Hopkins University in the United States. A study limited to the Iraq dataset for Covid-19 [7] aims to use the Box-Jenkins method to predict the number of people infected. The data from March 1 to July 31 were considered. The implementation of the model resulted in it matching with the ARIMA (2,1,5) model. Most attempts embrace variants of the research paradigm of the developed SIR that classify a population into fractions of "Susceptible", "Infectious," and "Recovered" and characterize their complex inter-relations with first-order difference equation [8]. A time-dependent Susceptible-Infected-Recovered (SIR) model is prepared for Covid-19 in [9]. This mathematical model is aimed at deriving transmission rate and recovery rate of infection at time t. As the Covid-19 started to spread in cities like Wuhan, Beijing, and Shanghai in China in early December 2019, it alarmed the immediate development of prediction models to cope with extreme circumstances. During this verge, a hybrid Artificial Intelligence model was proposed for predictions. An improved Infected-Susceptible (ISI) model was proposed, which utilizes the idea of the same infectious rate for all the people infected with the virus [10].

A study in [11] uses rule-based classifiers to predict the outbreak of dengue fever. The rule-based classifiers used are the Decision Tree, Rough Set Classifier, Naive Bayes, and Associative Classifier. Classifiers used in the modelling have shown up to 70% of accuracy. The results indicate that the performance of the multiple classifier classification and the ROC value was comparable to the single classifiers used. An enhanced version of the SIR model named as SIRD model is used in the study [12]. The Covid-19 dataset has all the features and should be analysed carefully for deriving meaningful insights useful for society. Hadoop offers solutions to handling huge amounts of data. Mahout is one amongst them and offers a Naïve Bayes Classifier to classify data [13]. Naive Bayes classifiers are a set of Bayes' Theorem-based classification algorithms [14]. The gap can be indentified where most of the research concentrates on just forecasting the critical parameters, the forecasters and classsifiers can be comined to propose a more realistc solution. Using the SIRF model, forecasting the infectious growth rate adds more value to the forecasted parameters. Deep learning models are used for predicting the Covid-19 infections based on the laboratory findings [15]. Totally, six deep learning models were implemented on different evaluation metrics and CNN-LSTM was found to be more accurate.

Proposed Methodology

Problem Statement

To design and develop an application that leverages ensemble learning techniques to forecast the critical parameters of Novel Coronavirus Dataset over a timeline and fits them to SIRF model from epidemiology in order to forecast the end of Covid-19 pandemic.

Covid-19 Dataset

The dataset considered is Novel Corona Virus 2019 Dataset from Kaggle. This dataset has daily level information on the number of confirmed cases, deaths, and recoveries for all the countries in the world. This is time series data, and it gets updated once in 15 days. The size of the dataset will be the sum of the total number of cases of Covid-19 all over the world. All the information in the files is dated and thus can be used for visualizing data and arriving at trends useful for society. Currently, the data is available from January 2020 to May 2021. For India, data is available till July 6, 2021.

Proposed Architecture

Figure 1 shows the high-level design flow of the proposed system. Initially, the dataset is cleaned by using pre-processing techniques and brought into a standard form. The data is then fit into The SIRF (Susceptible Infectious Recovered Fatality) model. Using this model, we can derive equations for transmission rate (β), recovery rate (γ), and fatality rate (δ). Then, we define an equation that must be true to declare the end of life of Covid19. Then, the Naïve Bayes Classifier is used to categorize the country level data into four categories described as follows:

A. Countries where the number of active patients exceeds the number of healed cases are aware that the recovery rate exceeds the mortality rate.
B. Countries where the number of active patients greater than the number of healed cases are aware that the recovery rate recedes the mortality rate.

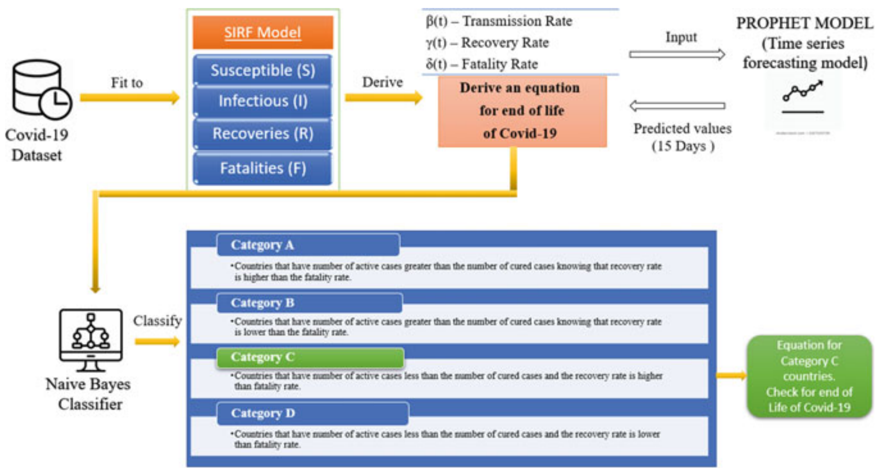


Fig. 1 Proposed architecture

C. Countries with a number of active patients lesser than the number of healed cases are aware that the recovery rate exceeds the mortality rate.
D. Countries where the number of active patients lesser than the number of healed cases are aware that the recovery rate recedes the mortality rate.

For the countries falling under category C, the pandemic appears to be fading. By forecasting the data of those countries that are not in category C repeatedly and repeatedly re-classifying, we can check if that country falls under category C. At this point, if the infectious growth rate value calculated by the transmission dynamics matrix reaches less than -1, then for that country, the pandemic has come to an end during that forecasted time frame.

Algorithm Used

A. **General SIRF Model**

In general, any pandemic has four important variables: Confirmed Cases, Active Cases, Recovered Cases, and Deaths. A classic epidemiological model is the SIRF model with standard incidence. It is divided into four subpopulations: Susceptibles *S*, Infectives *I*, Recovered people *R*, and Fatalities *F*. The susceptible has the potential to become infectious. The infectives can become recovered or fatalities. No other transitions are considered. The population *N* remains constant.

$$N = S + I + R + F \tag{1}$$

With the above terminologies, the general SIRF model in epidemiology is defined as follows

The rate at which susceptible is increasing over time

$$\frac{\mathrm{d}S}{\mathrm{d}t} = -\beta I \frac{S}{N}$$

The rate at which infection is spreading over time

$$\frac{\mathrm{d}I}{\mathrm{d}t} = \beta I \frac{S}{N} - (\gamma + \delta) I \tag{2}$$

Rate of recoveries over time

$$\frac{\mathrm{d}R}{\mathrm{d}t} = \gamma I \tag{3}$$

Rate of fatalities over time

$$\frac{\mathrm{d}F}{\mathrm{d}t} = \delta I \tag{4}$$

where β is the transmission rate, γ is the recovery rate, and δ is the fatalities rate.

B. **Practically Fit Data to SIRF Model**

In the context of Covid-19 in the world, we need to adjust the SIRF model to our data to approximate the behaviour of infection and label transmission rate and others. Suppose we consider (N) as the number of populations in a region (country or state) at time t.

$$population\_size = totalpositivecases + totalnegativecases$$
$$totalpositivecases = currentpositivecases + (recovered + death)$$

hence,

$$\begin{aligned}\text{population\_size} = {} & \text{totalnegativecases} \\ & + \text{currentpositivecases} + \text{recovered} + \text{death}\end{aligned}$$

From the above equation, the following identifications can be made.

- population_size is equivalent to total Population (N)
- totalnegativecases is equivalent to Susceptible (S)
- currentpositivecases is equivalent to an Infective (I)
- recovered + death is equivalent to a Recovered individual (R) + Fatalities (F).

And therefore, the general SIRF equations can be re-written as follows for Covid-19 context.

Equation (2) becomes the rate at which infection spread over time

$$\frac{\mathrm{d}I}{\mathrm{d}t} = (\beta - \gamma - \delta)I \tag{5}$$

Equation (3) becomes rate of recoveries over time

$$\frac{\mathrm{d}R}{\mathrm{d}t} = \gamma I$$

Equation (4) becomes the rate of fatalities over time

$$\frac{\mathrm{d}F}{\mathrm{d}t} = \delta I$$

where

$$\beta(t) = \frac{\text{the number of daily current confirmedcovid19patients at time}t}{\text{the number of accumulated confirmedcovid19patients at time}t}$$

$$\gamma(t) = \frac{\text{the number of daily recovered covid 19patients at time } t}{\text{the number of accumulated confirmed covid 19 patients at time } t}$$

$$\delta(t) = \frac{\text{the number of daily deaths covid 19 patients at time } t}{\text{the number of accumulated confirmed covid 19 patients at time } t}$$

C. **An Equation for Predicting the End of the Pandemic**

Let us re-consider Eq. (5)

$$\frac{\mathrm{d}I}{\mathrm{d}t} = (\beta - \gamma - \delta)I$$

The above equation can be re-written as

$$\frac{\mathrm{d}I}{I\mathrm{d}t} = (\beta - \gamma - \delta) \tag{6}$$

The LHS of the above equation is said to be the infectious growth rate. If this infectious growth rate is less than 0, then the transmission has come to an end, and the disease is no longer spreading; and hence, we consider

$$(\beta - \gamma - \delta) < 0$$

$$\beta < \gamma + \delta$$

$$\frac{\beta}{\gamma + \delta} < 1$$

Let us assume $R0 = \frac{\beta}{\gamma+\delta}$ and so $R0 < 1$.

Now sometime in the future, the pandemic will end. Considering this fact, we set time variable t to tend to ∞. Hence, the following equations are defined

$$\lim\ t \to +\infty\beta(t) = 0,$$

$$\lim\ t \to +\infty\gamma(t) = 1,$$

$$\lim\ t \to +\infty\delta(t) = \delta\,\text{threshold},$$

$$\lim\ t \to +\infty R0(t) = 0$$

Now the infectious growth rate from Eq. (5) becomes

$$\lim\ t \to +\infty \frac{\mathrm{d}I}{I\mathrm{d}t} = \lim\ t \to +\infty\,(\beta - \gamma - \delta) = \lim\ t \to +\infty\,\beta(t) - \lim\ t \to +\infty\,\gamma(t) - \lim\ t \to +\infty\,\delta(t)$$

Finally, we obtain

$$\lim t \to +\infty \frac{\mathrm{d}I}{I\mathrm{d}t} = -(1 + \delta\,\text{threshold}) \tag{7}$$

From the above equation, we see that sometime in the future, the infectious growth rate will be less than 0. This equation will be helpful to classify data recorded or predicted at an unknown time t will tend to end of a pandemic or not. Using the above methodology from the SIRF model, our algorithm repeatedly forecasts the data and classifies using Naïve Bayes Classifier and performs all the above calculations to check if the pandemic is fading. The algorithm is illustrated in Algorithm 1.

D. Pandemic Culmination Algorithm

Algorithm 1 Pandemic Culmination Algorithm

Input: dataset–Covid-19 world data
Parameter: n-Number of days in the future, country – selected country
Output: date_range–predicted date range where pandemic may end
n = 0
dataset (‘category’) = naïve_bayes (dataset)
country_data = dataset (“country” == country)
whilst country_data (‘category’) ! = ‘C’
    n = n + x
    forecasted_data = Forecast (data = country_data, days = n)
    Append forecasted_data to country_data
    β = country_data (data.confirmed/total_confirmed)
    γ = country_data (data.recovered/total_confirmed)
    δ = country_data (data.deaths/total_confirmed)
    infectious_growth_rate = β – γ – δ
    country_data (‘category’) = naïve_bayes (country_data)
    if infectious_growth_rate < −1
        return last row date in country_data
end whilst

The Algorithm 2 is used for prediction of critical parameters of Covid-19 like confirmed cases, deaths, and recoveries for “*x*” number of days in the future. The forecasting model used is FB Prophet Time Series Model. FB Prophet is a time series data forecasting method based on an additive model that fits non-linear trends with yearly, weekly, and daily seasonality, as well as holiday impacts. It performs effectively with time series with substantial seasonal influences and historical data from multiple years. Prophet is robust of missing data and pattern shifts, and it usually manages outliers well. The input to this algorithm is again the Covid-19 dataset and the number of days in the future for predictions. The algorithm pre-processes the data and converts it into required form and then splits the data into 80% training and 20% testing base. Once this split is done, the data is fit to the model and predictions are done. These predictions are then visualized using graphs.

E. Time Series Forecasting – FBProphet

Algorithm 2 FB Prophet forecast Algorithm

Input: dataset – Covid-19 world data, country, number of days in the future
Output: Prediction of critical parameters
1: Clean dataset for filling missing values, remove NAs
2: Reset the index and convert to time series data
3: Train, test = split (data, 80, 20)
4: FBProphet.fit(train)
5: Test_predicted = FBProphet.predict(test)
6: Data = Generate_data (number of days in the future)
7: Predicted_data = FBProphet.predict(data)
8: Return Predicted_Data

F. Naïve Bayes Classifier

Naïve Bayes is a classification methodology based on Bayes' Theorem and the premise of predictor independence.

Algorithm 3 Naive Bayes Algorithm

Input: dataset–Covid-19 world data, [forecasted data]
Output: Country list in 4 different categories A B C D
1: β (t) = dataset (active)/dataset (confirmed)
2: γ (t) = dataset (recoveries)/dataset (confirmed)
3: δ (t) = dataset (deaths)/dataset (confirmed)
4: transDynamics = new dataset (β(t), γ(t), δ(t))
5: transDynamics(category) = A (if $\beta(t) > \gamma(t) > \delta(t)$)
6: transDynamics(category) = B (if $\beta(t) > \gamma(t) < \delta(t)$)
7: transDynamics(category) = C (if $\beta(t) < \gamma(t) > \delta(t)$)
8: transDynamics(category) = D (if $\beta(t) < \gamma(t) < \delta(t)$)
9: Classifier = naïve_bayes.fit_predict(transDynamics)
10: Category = classifier.predict(dataset, transDynamics)
11: Return category

A Naïve Bayes classifier, in simplified way, implies that the existence of one feature in a class is independent of the presence of any other feature. The classifier classifies the data into 4 different categories. This classification is made based on the values obtained by calculations of transmission rate, recovery rate and fatality rate. Input to Algorithm 3 will be the Covid-19 dataset. The algorithm is used for actual data as well as forecasted data as shown in Fig. 4. Output will be the list of countries in their respective categories. Transmission dynamics of the pandemic is calculated using the spread of the disease (Figs. 2 and 3).

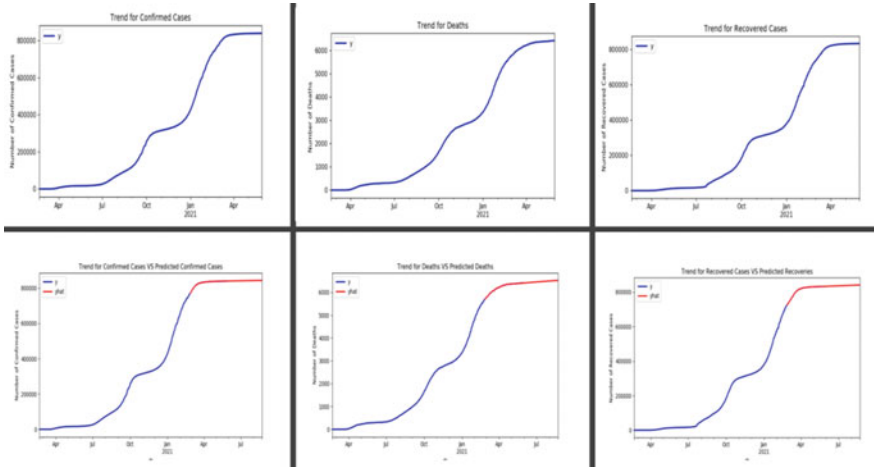


Fig. 2 Actual versus predicted trends of critical parameters for Israel

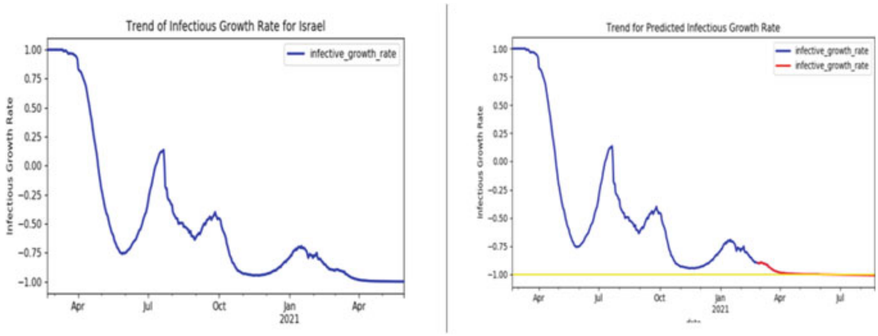


Fig. 3 Actual versus predicted trends of infectious growth rate for Israel

Results and Discussions

The Covid-19 dataset contains raw data for confirmed cases, deaths, and recoveries for every day, and in some countries, the data is available at the state/province level. The data is available from January 2020 to May 2021. We pre-process the data at the country level (i.e. if we have data for each state/providence for a date, we group the data by date and capture the sum of all states/provinces). The results are discussed for two countries: Israel (hit by only one wave of a Covid-19 pandemic) and India (hit by two waves of a Covid-19 pandemic). A comparison with the actual data is also discussed.

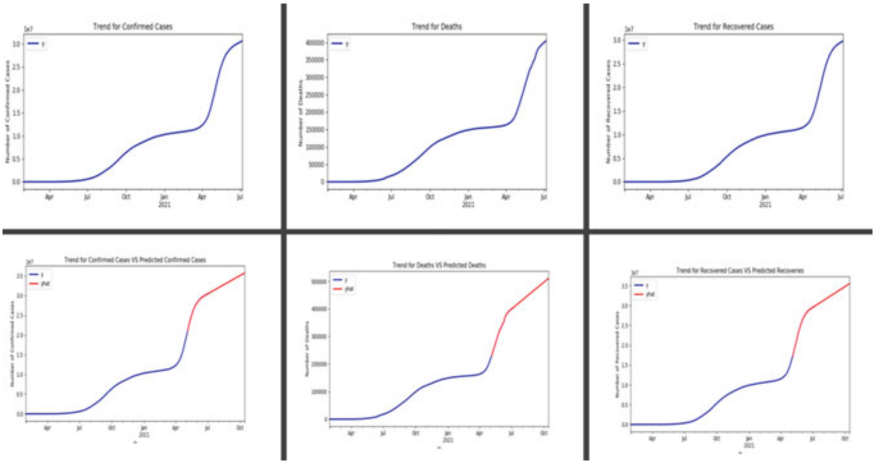


Fig. 4 Actual versus predicted trends of critical parameters for India

Country: Israel

The series of graphs in Fig. 2 shows time series trends for confirmed cases, deaths, and recoveries for Israel for one year. The graphs in the top row show the actual trend of critical parameters. By considering 100 days in the future and applying time series forecasting using FB Prophet, we get the trends for predicted confirmedcases and predicted deaths. They predicted recoveries, respectively, in the second row. The red line in the graphrepresents predicted results, whereas the blue line represents the actual data. By calculating the transmission dynamics matrix and calculating the infectious growth rate for the actual data, we get the trend as shown in Fig. 3. By calculating the same infectious growth rate on repeatedly predicted data as in Algorithm 1, we get a trend as shown in the same figure. This figure shows the trend for predicted data highlighted with red, which falls below −1 from June 2021 to July 2021. As per Eq. (7), this downfall of the trend indicates that the pandemic is fading. Now drawing a comparison with real-time data. Currently, Israel is reporting less than 500 cases per day which indicates that the infectious growth rate is minimal, and Israel has even announced to go mask-free.

Table 1 shows a comparative study of predicted cases by the proposed system and real-time data for the dates June 15, 2021, and June 28, 2021 are considered.

Table 1 Comparison with real-time data for Israel

Country—Israel	June 15th 2021		June 28th 2021	
	Predicted	Actual	Predicted	Actual
Confirmed case	8,40,681	8,39,878	8,41,348	8,42,218
Deaths	6427	6428	6446	6429
Recoveries	8,34,595	8,34,239	8,35,874	8,35,556

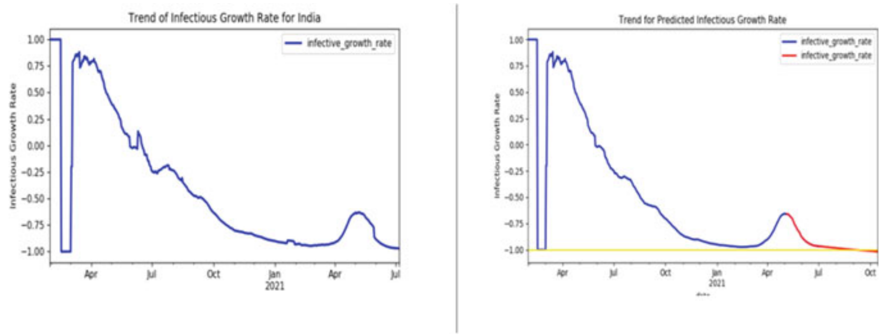


Fig. 5 Actual versus predicted trends of ınfectious growth rate for Israel

The difference between the numbers is almost very less, indicating that the system predicts with good accuracy. Figure 4 shows the trends for the Infectious growth rate in India. Due to the virus's mutations, the infectious growth increased drastically in April 2021 and May 2021. The predicted trend goes below −1 around September 2021 and October 2021, indicating that the pandemic would fade in that date range. Hence, the proposed system is less accurate in data concerning countries hit with more than one wave of Covid-19.

Country: India

The series of graphs in Figs. 4 and 5 shows the actual and predicted trends for the critical parameters as well as infectious growth rate of Covid-19 in India. The dataset for India is available till July 6, 2021. We can see the exponential rise of cases in April 2021 as India was hit by a second wave of the pandemic, clearly due to virus mutations. Because of this, the predicted trend slightly differs from the actual current day data.

Conclusion

The Covid-19 predictive analysis has become a challenging problem to help healthcare practitioners and authorities brace for and control the spread of infectious diseases. Modelling and predicting the contagious behaviour of virus will help healthcare systems to prepare for the expected influx of patients. Inaccurate disease predictions are a worry because they may impact government policies, containment guidelines, the health care system, and social life. The proposed system is utilized in contributing to the betterment of the current crisis going on. By leveraging the forecasting and classification capabilities of machine learning models like FBProphet, Naïve Bayes Classifier, and the mathematical equations of the SIRF model from epidemiology, the life span of the pandemic is determined. This ensemble model can be helpful for the respective countries to decide on preventive actions and measures. For example: resuming international travel, lockdowns, manufacturing testing kits, preparing hospital beds. The proposed idea need not be specific to the Covid-19 pandemic; it can be applied to any pandemic or epidemic disease. The covid protocols are a must in order to contain the pandemic. The predictions are only good when the public maintains all the covid protocols.

Acknowledgements This study is financially supported by Karnataka State Council for Science and Technology with Project Proposal Reference No. 44S_MTECH_040. The authors are thankful to M. S. Ramaiah Institute of Technology, Bangalore-560054, and Visvesvaraya Technological University, Jnana Sangama, Belagavi-590018, for their support..

References

1. Suggala RK, Anurag (2018) A survey on prediction and detection of epidemic diseases outbreaks. Int J Manage Technol Eng, DECEMBER/2018. ISSN NO: 2249-7455
2. Zeb A, Alzahrani E, Erturk VS, Zaman G (2020) Mathematical model for coronavirus disease 2019 (COVID-19) containing ısolation class. In: Research article biomed research ınternational, vol 2020, Article ID 3452402
3. Andreas A, Mavromoustakis CX, Mastorakis G (2020) Modified machine learning technique for curve fitting on regression models for COVID-19 projections. IEEE transactions © IEEE 2020
4. Rustam F, Rishi AA, Mehmood A, Ullah S, On B-W, Aslam W, Choi GS (2020) COVID-19 future forecasting using supervised machine learning models. IEEE. https://doi.org/10.1109/ACCESS.2020.2997311
5. Arun SS, NeelakantaIyer G (2020) On the analysis of COVID19—novel corona viral disease pandemic spread data using machine learning techniques. In: International conference on ıntelligent computing and control systems (ICICCS 2020) IEEE Xplore ISBN: 978-1-7281-4876-2
6. Kumar N, Susan S (2020) COVID-19 pandemic prediction using time series forecasting models © IEEE 2020, IEEE—49239
7. Mustafa HI, Fareed NY (2020) COVID-19 CASES in Iraq; forecasting ıncidents using box—Jenkins ARIMA model. In: 2nd Al-Noor ınternational conference for science and technology (2NICST2020) Baghdad , Iraq © IEEE 2020
8. Marmarelis VZ (2020) Predictive modeling of Covid-19 data in the US: adaptive phase-space approach. IEEE Open J Eng Med Biol. https://doi.org/10.1109/OJEMB.2020.3008313
9. Chen Y-C, Lu PE, Chang CS, Liu T-H (2020) A time-dependent SIR model for COVID-19 with undetectable ınfected persons. arXiv:2003.00122v6 [q-bio.PE] Apr 28th, 2020
10. Zheng N, Du S, Wang J (2020) Predicting COVID-19 in China using hybrid AI model. IEEE Trans Cybern 50:2168–2267 c 2020 IEEE
11. Bakar AA, Keifli Z, Abdullah S, Sahani M (2019) Predictive models for dengue outbreak using multiple rule-tbase classifiers. In: 2019 ınternational conference on electrical engineering and ınformatics 17–19 July 2011, Bandung, Indonesia. 978-1-4577-0752-0/11/$26.00 ©2019 IEEE
12. Singh S, Raj P, Kumar R, Chaujar R (2020) Prediction and forecast for COVID-19 outbreak in India based on enhanced epidemiological models. In: Second ınternational conference on ınventive research in computing applications (ICIRCA-2020) IEEE Xplore ISBN: 978-1-7281-5374-2

13. Almansouri HT, Masmoudi Y (2019) Hadoop distributed file system for big data analysis. 978-1-7281-1232-9/19/$31.00 ©2019 IEEE
14. Seref B, Bostanci E (2019) Performance comparison of Naïve Bayes and complement Naïve Bayes algorithms. In: 2019 6th ınternational conference on electrical and electronics engineering (ICEEE), 978-1-7281-3910-4/19/$31.00 ©2019 IEEE. https://doi.org/10.1109/ICEEE2019.2019.00033
15. Alakus TB, Turkoglu I (2020) Comparison of deep learning approaches to predict COVID-19 infection. Chaos Solitons Fractals 140:110120

Enhancement of IoT Security by Integration of Convolutional Neural Network and Image Processing

Yudhvir Singh and Rohit Dalal

Abstract This research is focusing on enhancement of Internet of Things security by integration of convolutional neural network mechanism to image processing. This research paper has considered various image processing mechanisms such as edge detection and used convolutional neural network for pattern classification. In the IoT environment, simulation has been performed by capturing images by camera, and face mask pattern has been detected by convolutional neural network after applying edge detection mechanism in order to increase the accuracy as well as performance. It has been concluded that the proposed simulation performed to detect the face mask is taking less space as compared to previous research. Moreover, the time consumption in case of proposed work is less than previous research. The edge detection mechanism used in research has increased the performance along with resolving storage issues. After elimination of unnecessary content, it has become convenient for convolutional neural network classifiers to train and test the image dataset rapidly without wasting storage space in order to get more accurate results.

Keywords Internet of Things · Convolutional neural network · Image processing · Edge detection · Accuracy · Space consumption · Performance

Introduction

The goal of this study is to improve IoT security by incorporating a convolutional neural network technique into image processing. This study looked at different image processing mechanisms, such as edge detection, and utilized convolutional neural network to classify patterns. In an IoT context, simulation was carried out by collecting an image with a camera and detecting a face mask pattern using convolutional neural network after using an edge detection method to improve accuracy and performance. It has been determined that the suggested simulation for detecting the face mask takes up less space than prior studies. Furthermore, the time required for

Y. Singh (✉) · R. Dalal
Maharshi Dayanand University, Rohtak, India

V. Bindhu et al. (eds.), *Proceedings of Third International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 844,
https://doi.org/10.1007/978-981-16-8862-1_54

the suggested study is shorter than that required for prior studies. The edge detection method employed in the study improved performance while also addressing a storage problem. After the irrelevant elements were removed, it became easier for the convolutional neural network classifier to train and test the picture dataset quickly without wasting storage space, resulting in more accurate results.

Image Processing Technology

The process of making changes in graphical content by rotation, scaling, cropping, edge detection, and compression is considered image processing. Several images processing mechanism are stated below:

1. **Scaling**: This is process of enlarging or reducing the image horizontally/vertically.
2. **Rotation**: This is process of changing the angle of image.
3. **Cropping**: Eliminating some portion of image during selection during preprocessing is termed as cropping.
4. **Edge detection**: This is a process of getting the edges or boundaries from image to present an outline of image.
5. **Compression**: During this process, the size of image consumed on storage space is reduced by applying lossless or lossy image compression mechanism.

Sensor for Image Processing Technology

A lens and an image sensor are part of a digital imaging system. On the image sensor plane, the lens creates an image. The image sensor has an array of light-sensitive pixels that generate a digital value that represents the number of light photons that have collected on the pixel during the exposure period.

The charge-coupled device (CCD) and the active-pixel sensor are the two major kinds of electronic image sensors (CMOS sensor). Metal–oxide–semiconductor (MOS) technology is used in both CCD and CMOS sensors, with CCDs using MOS capacitors and CMOS sensors using MOS field-effect transistor (MOSFET) amplifiers.

1. **CCD**: A charge-coupled device is integrated circuit. It is containing array of connected capacitors. Each capacitor is capable to transfer its electric charge to a nearest capacitor in control of external circuit. CCD sensors have been frequently used in digital imaging.
2. **CMOS sensor**: It is electronic chip which is converting photons to electrons in order to perform digital processing. Complementary metal–oxide–semiconductor (CMOS) **sensors** have been used to develop images. These images are developed in digital cameras as well as digital video cameras along with digital CCTV cameras.
3. **MOSFET**: Metal–oxide–semiconductor field effect transistor is acting as semiconductor device. It is frequently utilized to switch and amplify electronic

signals. These signals are amplified in the electronic devices. **MOSFET** is device with three terminals. These three terminals are source, gate, and drain.

How and What Type of Image Can Be Used in Prediction

The image captured by image sensors is stored in system. The images are raster images with joint picture expert graphic or portable network file format. Different algorithms are applied on images. The edge detection would be applied on these images to eliminate useless portion, and the convolutional neural network mechanism would be applied in order to train the network in order to perform prediction on the basis of trained dataset.

These images could be of following types:

1. **JPEG**—Joint photographic experts group
2. **PNG**—Portable network graphic
3. **GIF**—Graphical interchange format
4. **TIFF**—Tag image file format
5. **BMP**—Bitmap image.

Various Algorithms for Image Processing Technology

Scaling, resizing, comparing, and altering graphical elements are all common uses of image processing methods. The convolutional neural network (CNN) model may be used to identify face mask patterns from picture sets on a regular basis. However, there are a few problems with convolutional neural network-based categorization. During COVID-19, there was only a little amount of work done to enhance the performance of face mask detection.

Despite the fact that there are many studies in the field of image processing, it is clear that the time required for prediction has to be decreased. Furthermore, there is the problem of graphical material using too much space. The proposed study aims to reduce forecast time and space usage. The examination of current image processing studies and methods, as well as the elimination of their limitations, has been the subject of research. Using an edge-based convolutional neural network technique, the researchers suggest a method for detecting face masks. Time was saved by removing unnecessary information from the graphical picture before using the convolutional neural network. Furthermore, the graphical dataset's storage requirements have been decreased. Every comparison creates a significant disparity in size and comparison time as the number of datasets grows. The proposed effort will use MATLAB to accomplish the suggested approach. During simulation, the suggested approach and algorithm are compared to the conventional algorithm. The algorithms used during research are as follows:

1. **CNN algorithm** has been discussed in Sect. "Convolutional Neural Network (CNN)".
2. **Canny edge detection** has been presented in Sect. "Canny Edge Detection".

Application and Algorithms Used for Image Processing Technology

Research is making use of convolutional neural network and edge detection mechanism in order to improve the performance during image preceding.

Convolutional Neural Network (CNN)

Convolutional neural networks are well-known in the field of deep learning as a powerful machine learning mechanism [1, 2]. Convolutional neural networks are often taught using large sets of graphical pictures in a variety of formats. Convolutional neural networks might acquire rich characteristic representations for a range of graphical characteristics from such large datasets. Handcrafted features like histogram of oriented gradients (HOG), local binary patterns (LBP), or SURF frequently outperform such representations (speeded-up robust features). Using a single convolutional neural network as a characteristic's capturer is the easiest method to harness the capabilities of a convolutional neural network without investing a lot of time or effort on teaching.

The examination of current face mask detection methods and overcoming its limitations has been the subject of research. The study offers a technique for detecting face masks based on an edge-based convolutional neural network algorithm. The time required to apply convolutional neural network was decreased by removing unnecessary information from the graphical picture. Furthermore, the graphical dataset's storage requirements have been decreased. Every comparison creates a significant disparity in size and comparison time as the number of datasets grows. During simulation, the suggested approach and algorithm are compared to the conventional algorithm. When compared to conventional pattern detection methods, the suggested work is shown to be more efficient.

1. **Algorithm**: Parallel convolutional neural network
2. **Input**: d: dataset, l: dataset true labels, W: word2vec matrix
3. **Output**: score of parallel convolutional neural network trained model on dataset
4. **Let** f be the feature set 3d matrix
5. **For j in i do**
6. **Let** f_i be feature set matrix of sample i
7. **For** j **in** i **do**
8. $V_j \leftarrow$ vectorize $(_{j,w})$
9. **Append** v_j to f_i

10. **Append** f_i to f
11. f_{train}, l_{test}, $l_{train,}$ l_{test} ← splite feature set and labels into train subset and test subset
12. M ← Parallel Convolutional Neural Network (l_{test}, l_{train})
13. Score ← evaluation (I, $l_{test, M)}$
14. **return** score.

Canny Edge Detection

Edge detection of images is extremely important in today's society, because the style of life is quite contemporary. It is mostly used in the area of medicine, but it may also be used in defensive applications. As a result, it is thought that studying edge detection algorithms is extremely important. Edge detection plays a significant part in the idea of image processing. When the image's intensity value or pixel value changes sharply, edge detection can quickly identify the change in pixel value. Various kinds of edge detection methods are now available on the market. Edge detection plays a significant part in image processing techniques. It has a wide range of applications, including image distortion, sample verification, image segmentation and removal, and so forth. When the edge detector is applied to a picture, it produces a lot of odd edges in the extreme situation. Canny is concentrating on identifying a more effective edge detecting method. In such cases, an appropriate edge detector is one that performs well in detection. The programme marks the real edges that are present in the picture. When the indicated margins of the actual picture are extremely near to the edge, it is regarded to be in a favorable position. Because picture noise does not generate false edges, it responds in a single try. As a result, each image's edge is only shown once and only when feasible. The mathematical issue of constructing an optimum smoothing filter given requirements of detection, localization, and reducing numerous responses to a single edge was addressed by John Canny. Given these assumptions, he demonstrated that the best filter is a sum of four exponential terms.

Literature Review

Existing researches in field of image processing, convolutional neural network, and IoT have been presented in Table 1.

Table 1 Literature survey

S. No	Authors	Title	Objective	Mechanism	Limitation
1	Gonzalez [3]	Improved neural classifier for microscrew shape recognition	To improve the neural classifier recognizing microscrew shape	Convolutional neural network	The system is consuming more time and space during simulation
2	Manoharan [4]	Early diagnosis of lung cancer with probability of malignancy calculation and automatic segmentation of lung CT scan images	To provide the lung cancer using malignancy calculation	Automatic segmentation	Research has ignored the intelligent approach and ignored performance
3	Chen [5]	Analysis of the impact of mechanical deformation on strawberries harvested from the farm	To analysis the impact of mechanical deformation on strawberries	Mechanical deformation	Need to do more work in field of image processing and pattern detection
4	Meenpal [6]	Facial mask detection using semantic segmentation	To perform face mask detection	Semantic segmentation	This system needs to improve the accuracy and reliability of simulation
5	Gulve [7]	Implementation of IoT-based smart video surveillance system. In computational intelligence in data mining	To implement smart video surveillance system using IoT	Surveillance system	System has ignored the use of neural network for smart decision making
6	Valente [8]	Privacy and security in Internet-connected cameras	To provide security by making use of web based camera	Surveillance system	The performance of work is very slow
7	Ibrahim [9]	A comprehensive review on intelligent surveillance systems	Reviewing the smart surveillance systems	Surveillance system	There is need to reduce time consumption and increase accuracy

(continued)

Table 1 (continued)

S. No	Authors	Title	Objective	Mechanism	Limitation
8	Yang [10]	Study on a recurrent convolutional neural network-based FDTD method	To perform study of RNN for accurate prediction	Convolutional neural network	Results show that the processing speed of work is slow
9	Xiao [11]	Improving bug localization with character-level convolutional neural network and recurrent neural network	To improve the bug localization using neural network	Convolutional neural network	There is need of qualified technical to manage operation of such system
10	Zhang [12]	Investigation of image edge detection techniques-based flood monitoring in real time	To perform the edge detection in real-life application	Edge detection	Need to integrate edge detection to intelligent approach to build scalable solution
11	Lee [13]	Implementation of the high-speed feature extraction mechanism	Performing feature extraction with high performance	Feature extraction	This research has neither opted to reduce size of image nor opted to improve the performance

Problem Statement

Existing research works have considered several parameters to conclude the efficiency of models. Lot of them considered the time consumption during operations. However, some of them considered the performance and accuracy parameters. The size, type of image file has been also considered. The time consumption during image capturing from closed-circuit television (CCTV) camera along with dimension of graphical contents is also considered. The limitations of existing research found are the time consumption as well as file size of graphical content. There is need to improve the performance of tradition system. The image captured from CCTV camera took lot of space. It has been observed that there is need to preprocess such images by making use of image resizing and edge detection mechanism. Moreover, the issue of accuracy is also required to be managed in those researches.

Pattern detection has been tackled using methods such as support vector machine (SVM), convolutional neural network, and random forest. The primary goal of traditional research has been carried out in the current study. It was to double-check the accuracy of the data categorization. The accuracy of data has been assessed in terms of each algorithm's efficiency and effectiveness.

According to a literature study conducted by different researchers, SVM works best on textual data, while convolutional neural network does better with graphical assessments and categorization of graphical data. As a result, given the advantages of convolutional neural network, additional work on pattern detection models is required. However, the current convolutional neural network approach has limits in terms of space consumption and comparison time. When comparing graphical material, it takes a long time. As a result, the convolutional neural network model's performance must be enhanced.

1. While traditional research has shown that SVM is the best for textual data, convolutional neural networks (CNN) are effective for picture analysis and classification. As a result, given the advantages of CNN, additional work on pattern detection models is required.
2. If there is any overlap in the dataset, SVM is not the best option. In some situations, random forest may provide better results than SVM. As a result, a performance comparison of random forest and SVM is required.
3. The computational method is called as PSO. It was utilized to improve a challenge. It has been used to enhance a candidate solution according to the provided quality measure. PSO is also known as a met heuristic. The reason for this is because it makes no assumptions about potential problems. It has the ability to explore bigger areas of potential solutions. As a result, PSO may be used to optimize the resolution of problems that arise during its usage. Such difficulties may be irregular in nature, change over time, be loud, and so on.

Proposed Work and Simulation

Present research has process image dataset by making use of edge detection mechanism in order to speed up the training and testing time. Moreover, the accuracy of research would also increase by making use of such approach. This section is presenting the objective, algorithm, and flowchart of proposed work. Moreover, the parameters considered during simulation are also considered. The major objective of research is to consider the research in field of pattern detection mechanism in order to build more accurate and high-performance system for face mask detection and proposing more suitable approach. The objectives considered in research are considering working and limitation of existing research made to detect face mask using convolutional neural network-based training and testing mechanism and investigating the suitable edge detection mechanism in order to reduce size and training and testing time. Research is integrating canny edge detection in image processing during training and testing to improve performance and accuracy and performing comparative analysis to compare time consumption and size consumption in case of normal image and edge detected image.

Role of CNN in Proposed Work

The proposed study has employed CNNs that, due to their great accuracy, are used in picture classification and recognition. The CNN follows a hierarchical architecture that acts like a funnel to construct a network and eventually produces a fully connected layer in which all the neurons are linked, and output processed. Convolutionary layers are the filters used in the original picture and in other deep CNN characteristics. Most of the parameters used are located in the network. The number of kernels and the kernel size are the most significant characteristics. The many levels of the CNN are present. For a convolutionary neural network, four types of layers are available: (1) Convolution layer, (2) Pooling layer, (3) Reliability layer, and (4) Fully connected layer.

1. **Convolutional layer**: A convolutionary network (CNN) is a deep neural network class which is commonly employed for detecting visual patterns, but it is also utilized for spatial data analysis, computer vision, natural language processing, signal handling, etc.
2. **Pooling layer**: Pooling layers are used to minimize characteristic map size. It lowers the number of learning parameters and the calculation carried out within the network. The pooling layer summarizes all the characteristics of a convolution-generated layer in a feature map region.
3. **ReLU correction layer**: ReLU refers to the real nonlinear function defined by $\text{ReLU}(x) = \max(0, x)$. The ReLU correction layer replaces any negative values received as inputs with zeros visually. Visually it appears like this. It works as a function for activation.
4. **Fully connected layer**: It is simple to connect fully layer, feed neural networks. The last few levels in the network are fully connected layers. The input into the fully connected layer is the output of the final swimming pool, which is flattened and then fed into the completely linked layer.

CNN Working

The CNN is fitted with four overall layers, three max pool layers, two fully linked layers, and a softmax output layer. The input includes three 48×48 parts from the centerpiece of the voxel target from axial, sagittal, and coronal images.

Between the input layer size and the output level size, the number of hidden neurons should exist. The hidden neurons number should be 2/3 the input layer and the output layer size. There should be less than twice the input layer size of the hidden neurons. The results from the convolutionary layers show high-level characteristics in the data. While this output may be flattened and linked to the output layer, it is (usually) an inexpensive method to learn nonlinear combinations.

CNNs are taught to find the best pictures for the task in hand and to extract them. This is the strength of them. Due to the strength of their classification, the final layers of a CNN are fully linked. Though you may believe that these two architectures contain FC levels, they do not compete.

CNN Algorithm

A CNN is a deep learning system, which can capture the input picture, attribute importance in various aspects/objects in the image (learnable weights and bias), and can discriminate between them. A CNN is a multi-layered neural network with a unique design that detects complicated data characteristics. In image reconnaissance, power-to-view robots, and self-driving cars, CNNs have been employed. Once a CNN is created, the contents of various pictures may be classified (Fig. 1).

CNN is an effective algorithm for recognition of patterns and image processing that is utilized frequently. It has several benefits, including easy construction, little training, and adjustability parameters. In speech analysis and picture recognition, it has become an important issue.

CNN is a neural network with one or more convolutionary layers, which is mostly used for image processing, classification, segmentation, and other auto-correlated data. In essence, a convolution slides the filter over the input.

A deep neural network model has to be defined for the CNN classification techniques. This model was defined as a basic SVM model. Although the accuracy of the CNN is 94, 01%, the visual interpretation goes against the preciosity of SVM classifiers. As predicted for the prepared dataset, CNN exceeds SVM. CNN boosts

Fig. 1 Pseudocode of CNN

Input: negative set NG containing n negative instances, positive set PS with m_f positive instances
Output: N mini-batches $MB = \{mb_i | \ i = 1, 2, \ldots N\}$ for CNNs

1. Divide n negative instances into N subsets $X = \{x_i | \ i = 1, 2, \ldots N\}$, each with S_n negative instances
2. Set $Temp = PS$ //save a copy of all positive instances
3. for $i = 1{:}N$
4. $\quad pos = \emptyset$; //pos: positive subset in the mini-batch
5. $\quad$ if (length($Temp$)< S_m)
6. $\qquad Temp$ = random(PS);
7. $\quad$ for $1{:}S_m$
8. $\qquad$ randomly pick an instance t from $Temp$;
9. $\qquad pos = pos \cup t$;
10. $\qquad Temp = Temp - t$;
11. $\quad$ end for
12. $\quad mb_i = x_i \cup pos$;
13. end for;
14. return MB to train CNNs;

overall performance of categorization by 7.7%. Moreover, each class has a performance of more than 94%. This finding shows that CNN can be utilized to fulfill high precision criteria for the defense system.

It is largely because the functionality of the RNN is reduced, and it may accept various output/input lengths, therefore affecting the overall time and efficiency of your computer. The CNN has a fixed input, however, which provides a constant output that allows the results to be calculated more quickly. The architecture of a CNN is different from the RNN. CNNs are 'neural feed networks,' using filters and pooling layers, whereas RNN feeds return to the network (more on this point below). The input size and the results are determined in CNNs.

Process Flow of Proposed Methodology

When compared to conventional face mask detection methods, the suggested work is shown to be more efficient. The use of the suggested work in COVID-19 is expected to enhance the decision-making capacity of convolutional neural networks. In comparison with the conventional method, the proposed work is said to be more accurate. In order to enhance the performance of the face mask identification mechanism, the suggested study will combine the convolutional neural network method with edge detection algorithms (Fig. 2).

Results and Analysis

This section is considering the integration of image processing in order to reduce the space and time consumption. Image processing is a method for enhancing raw pictures gathered for various reasons from spacecraft, space probes, and aircraft cameras/sensors, as well as images recorded in daily life. Various image processing methods have been developed during the past four to five decades. The majority of methods have been developed to improve pictures acquired from unmanned spacecraft, space probes, and military reconnaissance missions. Image processing systems are growing more common as powerful personal computers, huge memory devices, graphics tools, and other tools become more widely available.

Integration of Image Processing Integration to Intelligent System

An image acquisition system may be a video camera that takes pictures. The picture acquired with the help of optical or analogue cameras may be utilized as the input.

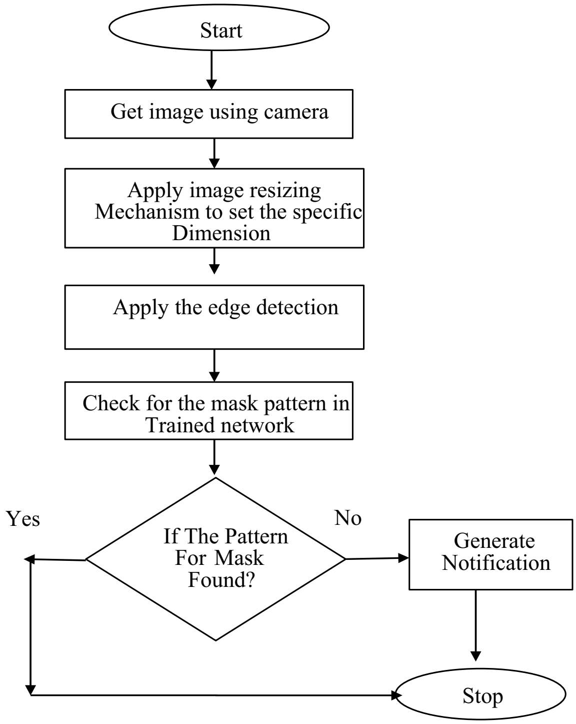


Fig. 2 Process flow for face mask detection

CCD or CMOS sensors are used in digital cameras that have a direct USB port connection to the PC.

MATLAB is presently using augmented reality to capture real-world live video feeds. It is directly connected to the PC (Image Processor), and MATLAB accesses or interacts with it via adapters, which are built-in tools in MATLAB. By eliminating part of the functional information from the gathered pictures, picture processing may be accomplished. If an item has to be defined, it is essential to take notice of many strong qualities of the thing, such as color, pattern, boundaries, strength, and shape. The Intelligent Protection System's function is based on image processing as well as an embedded system, which uses a microcontroller intended for control applications. In this technique, a camera is configured to focus on a single picture. The camera records the image on a regular and continuous basis at predetermined intervals. Following are the images that have been captured by camera for face mask detection (Fig. 3).

These collected pictures are sent as input to the MATLAB application. The software MATLAB compares the images collected to the original image produced on the device. The effect of the MATLAB program is the location of the unexpected element in the picture that is provided as an input to the microcontroller. The microprocessor spins the motors horizontally and vertically to target the unexpected item by concentrating a laser beam on its precise position. The intelligent protection system

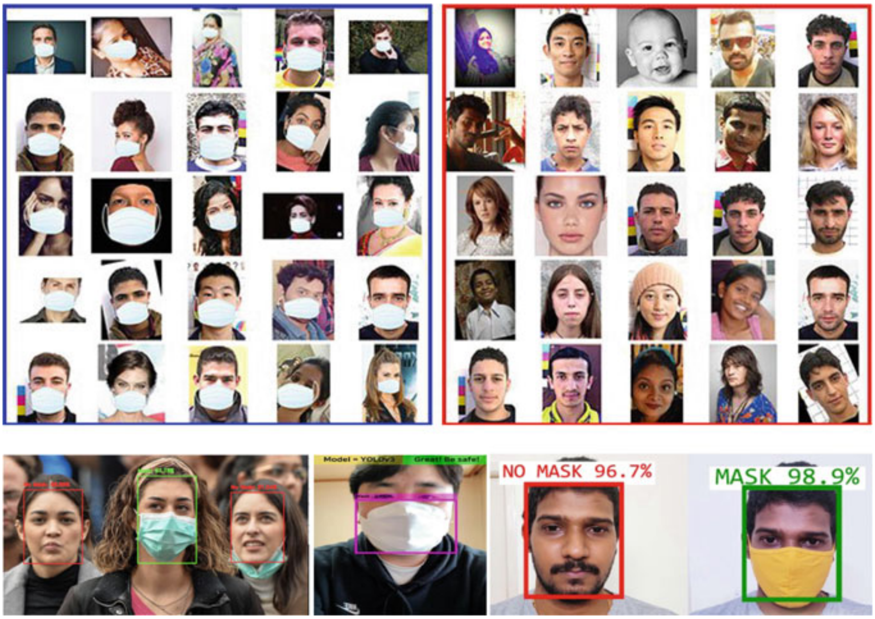


Fig. 3 Image base for face mask detection

ensures security by taking action on an unexpected item discovered in a picture during capture.

Simulation

During simulation actual image, I1 has been considered for face mask detection from dataset in Fig. 4 without applying any edge detection. Figure 5 is showing the space consumption of image I1 which is 171 kb. But after applying edge detection mechanism on image I1, the edge detected image I2 has been shown in Fig. 6. The size of edge detected image is shown in properties of image I2 in Fig. 7.

1. **Getting face mask image I1**

   In this phase, following image has been captured and termed as I1.
2. **Getting size of image s1**

   After getting image I1, the size of image is captured by getting the properties of image as shown in following windows:
3. **Applying canny edge detector on face mask image I2**

   Edge detection mechanism has been applied on the image I1, and I2 image is generated as shown below.

Fig. 4 Captured image I1

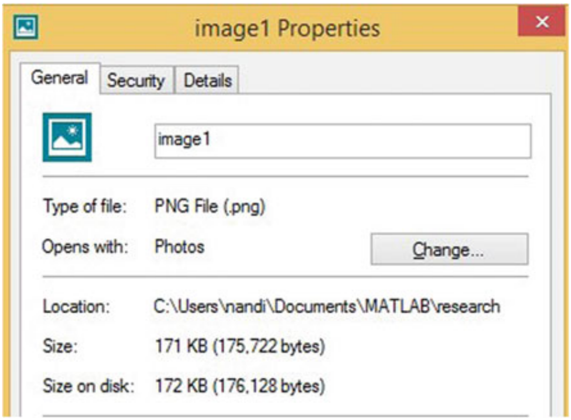


Fig. 5 Getting image size from properties

4. **Getting size of image s2**
   After getting image I2, the size of image is get from its property and saved in s2.
5. **Compare s1 and s2**
   Simulation of space consumption in case of normal and edge detected image is made in MATLAB environment. Figure 8 is showing the difference of size of file in kb when number of images gets increased. It is shown that the normal image is consuming more space as compared to edge detected image. The ratio of image size in case of normal and edge detected image is 171: 9.

Fig. 6 Edge-based image I2

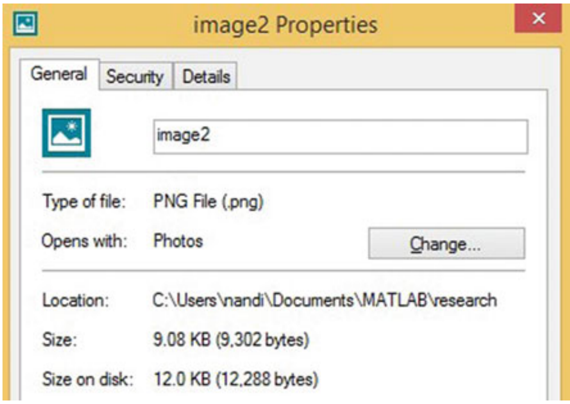


Fig. 7 Getting edge-based image size from properties

6. **Get comparison time in case of normal image t1**
   The time to get image matrix from I1, I2 is stored in t1 and t2, respectively. Simulation of time consumption in case of normal and edge detected image is made in MATLAB environment. Figure 9 concludes that the time consumption during comparison of normal image is more than that of edge detected image. The trained model is consuming ratio of 0.137766: 0.129526 in case of normal and edge detected image, respectively.

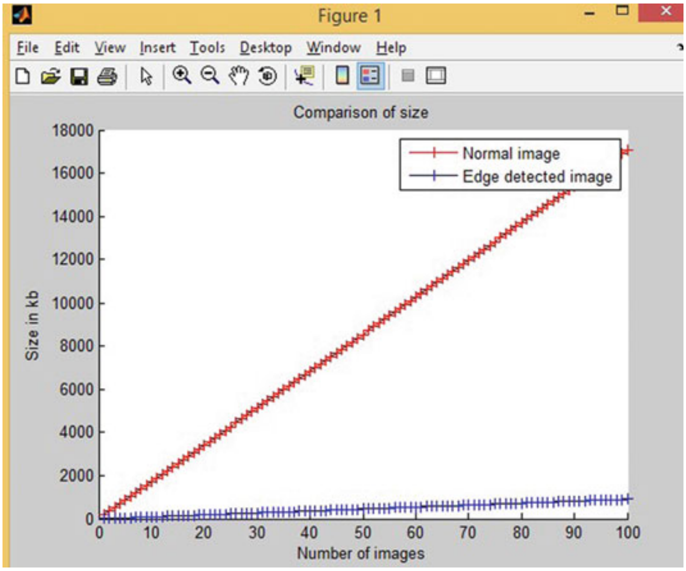


Fig. 8 Comparison of size

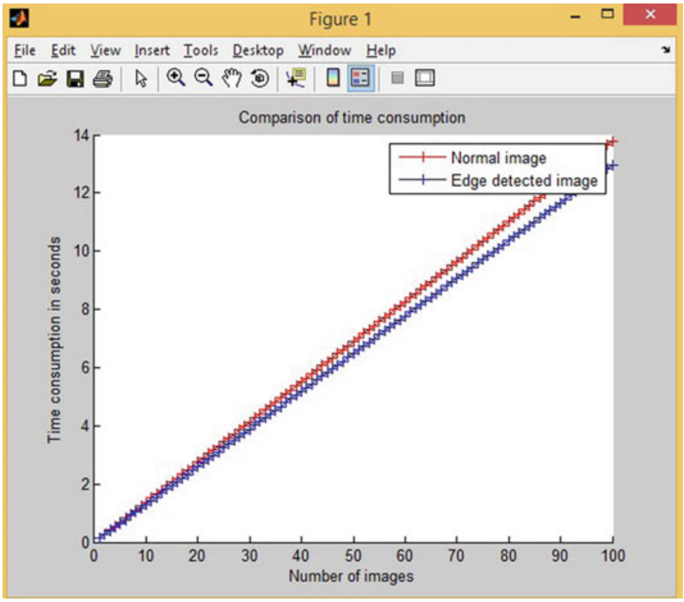


Fig. 9 Comparison of time consumption in normal and edge-based image

Table 2 Comparison chart for previous and proposed work

Feature	Previous research	Proposed work
Detection time	Comparatively high because real image pixel comparison is time-consuming process	Comparatively low due to edge detection because lesser pixels are compared
Space	More storage space is required because there is no elimination from image	Comparatively less space required because most of the useless portion of image is eliminated
Accuracy	Relatively less accuracy because useless portion is considered during training and comparison	Relatively high accuracy because only significant pixels have been considered during face mask detection operation
Edge detection mechanism	Edge detection mechanism is not applied to reduce the size of image and eliminate useless portion	Canny edge detection is applied to eliminate useless portion of image and improve the accuracy
Neural network	Previous model is applying traditional neural network	Proposed work is using convolution neural network that would provide for accurate solution
Flexibility	Lack of flexibility due to bulky graphics	High flexibility as it could be applied in another application
Scalability	Limited scalability due to heavy weighted images	Work could be implemented at huge scale due to lightweighted images
Performance	Relatively low because pixel comparison takes lot of time	Relatively high as pixel takes lesser time

Comparison of Proposed Work with Existing Technology

The comparison of proposed work has been made to previous research considering time, space, accuracy, flexibility, and scalability and performance factors (Table 2).

Conclusion

According to research, the applicability of edge detection mechanisms reduces time consumption. The suggested approach outperforms the previously used convolutional neural network model. On the other side, the amount of RAM used is decreased. Because the amount of graphical material is decreased and a large part of graphical content is removed, the edge detection method has enhanced convolutional neural network's capabilities. Simulation has a 14% higher accuracy than the current method. It has been discovered that the accuracy varies depending on the picture size.

Furthermore, changes in the picture dataset have an impact on the findings. Convolutional neural network has been studied for its benefits. During face mask identification, CNN employs layers that may be utilized to verify characteristics of a dataset including pictorial information. To enhance the efficiency of an existing CNN model, researchers used an edge detection method. During picture segmentation, an edge detection technique is used to remove unnecessary graphical information. Furthermore, only a small amount of study has been done in the field of graphical pattern recognition models. As a result, research has focused on edge-based convolutional neural networks for graphical content processing. As a consequence, the intelligent image processing protection system detects an odd item in a high-security picture and can take immediate action on the unexpected object in a fraction of a second.

Future Scope

Another IoT-based project may make advantage of the research's canny edges detection using convolutional neural network model. Such improvements are required for image processing tasks. Furthermore, the area used by a graphical sample may be readily controlled. After making these modifications, the categorization and prediction mechanisms would have no effect. The suggested study in medical research has the potential to improve convolutional neural network's functioning during decision making. The next work is expected to be more precise, and the accuracy may vary depending on the size of the picture. Furthermore, future modifications in graphical content are being considered through study.

References

1. Deshmukh et al (2013) Webcam based intelligent surveillance system. Res Inventy Int J Eng Sci 2(8):38–42
2. Gonzalez RC, Woods RE (2002) Digital image processing second edition; [15] Ciresan et al. (2011) Flexible, high performance convolutional neural networks for image classification. In: Proceedings of twenty-second international joint conference on artificial intelligence
3. Smys S, Basar A, Wang H (2020) Artificial neural network based power management for smart street lighting systems. J Artif Intell 2(01):42–52
4. Manoharan S (2020) Early diagnosis of lung cancer with probability of malignancy calculation and automatic segmentation of lung CT scan images. J Innov Image Process (JIIP) 2(04):175–186
5. Chen JIZ, Yeh L-T (2020) Analysis of the impact of mechanical deformation on strawberries harvested from the farm. J: J ISMAC 3:166–172
6. Meenpal et al (2019) Facial mask detection using semantic segmentation. In: Proceedings of 4th international conference on computing, communications and security, pp 1–5; [2] Ojala et al (2002) Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Trans Pattern Anal Mach Intell 24(7):971–987
7. Gulve et al (2017) Implementation of IoT-based smart video surveillance system. Comput Intell Data Min, 771–780

8. Valente et al (2019) Privacy and security in Internet-connected cameras. In: Proceedings of IEEE international congress on internet of things (IoT), pp 173–180
9. Ibrahim SW (2016) A comprehensive review on intelligent surveillance systems. Commun Sci Technol 1(1)
10. Guo et al (2019) Study on a recurrent convolutional neural network based FDTD method. In: Proceedings of international applied computational electromagnetics society symposium-China (ACES), vol 1, pp 1–2
11. Xiao Y, Keung J (2018) Improving bug localization with character-level convolutional neural network and recurrent neural network. In: Proceedings of 25th Asia-Pacific software engineering conference (APSEC), pp 703–704
12. Zhang et al (2019) Investigation of image edge detection techniques based flood monitoring in real-time. In: Proceedings of 16th international conference on electrical engineering/electronics, computer, telecommunications and information technology, pp 927–930
13. Lee et al (2020) Implementation of the high-speed feature extraction algorithm. Trans Electr Electron Mater 21(2):150–156

Conversational Image Captioning Using LSTM and YOLO for Visually Impaired

Prabhav Karve, Shalaka Thorat, Prasad Mistary, and Om Belote

Abstract According to the WHO, as of 2021, there are 2.2 billion people with distance or near vision impairment. The biggest challenge they face is to navigate around places. The main purpose is to build a guidance system for the visually impaired which will take imagery input from the camera, describe it, and give a speech-driven output to its users. This paper describes the process of image captioning based on artificial intelligence by showcasing some image preprocessing techniques, CNN and RNN. However, as improvisations on conventional methods, we have implemented CNN inspired from the YOLO algorithm and LSTM for captioning along with a text-to-speech system to dictate information about the surroundings to the user. This paper presents how the YOLO v1-LSTM model serves the purpose and delivers the faster output with minimum loss and less weight while maintaining comparable accuracy against conventional methods.

Keywords Image captioning · Encoding–decoding · YOLO-LSTM · Deep learning · Sequence-generation · Image summarizer · CNN-RNN · Visual assistance

Introduction

Mobility or movement in an environment which consists of various obstacles like potholes, bins, stairs, and various objects, as far as people with visual impairment are concerned, is a dangerous task. But if there is any system which would help such people to navigate around places, it would make their life simpler. The whole purpose is to make the mobility or movement of visually impaired people easy. The main concept of this paper is to present a system that would sense the surrounding environment and guide the user about the same. This can be done by image captioning. Image captioning is the process to describe the image scenario in human interpretable language. This can be done with convolutional neural networks (CNN) and recurrent

P. Karve (✉) · S. Thorat · P. Mistary · O. Belote
Department of Information Technology, International Institute of Information Technology, Pune, India

V. Bindhu et al. (eds.), *Proceedings of Third International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 844,
https://doi.org/10.1007/978-981-16-8862-1_55

neural networks (RNN). Our solution is based on generating image descriptions by encoding–decoding architecture using the YOLO encodings and LSTM, a type of RNN. The text-to-speech-driven output conveys the message to the user, according to which the user visualizes, based on the capability of the person.

Response time of the systems in the form of text-to-speech-driven outputs is equally important as the accurate caption for the images. The speech output should be fast, accurate, and meaningful for the user enough to give a good sense of the surrounding environment for a free and safer roadside walk. The result section of the paper explains how the meaningful captions are generated, how to interpret the results, and how our solution is better in speed.

The paper presents our study of conventional caption generation approaches and also the advantages of our solution over conventional methods of caption generation and will also give a glance of future scope, so that this guidance system would do even more to serve the purpose.

The rest of the paper is organized as follows:

Section "Related Work" presents the related survey study based on referred research papers followed by Sect. "Methodology", which presents the methodology with some tables and diagrams to help to understand the insights, and the working of the solution. The results and conclusion are presented in Sects. "Results" and "Conclusion", respectively, along with comparative study results and sample output of generated caption. And lastly, the future scope of the system is presented in Sect. "Future Scope" before the references list.

Related Work

This paper presents our method to use deep learning technologies/models namely long short-term memory (LSTM) and You Only Look Once (YOLO) to generate captions for images and use them to describe a sequence of images in real time. The text-to-speech system is used to verbally dictate the output. This section presents the study done on various approaches which helped in coming up with our findings and solution. Referring to various research papers, our findings are as follows:

Preprocessing

Starting with initial working on images which is also called preprocessing, following are our findings:

Some preprocessing techniques like mean normalization of the image is given by Pal and Sudeep [1]. Mean normalization refers to rescaling the pixel values, so that they lie within a defined range. This resolves the issue of propagating gradients.

Resizing and de-noising are used for applying fixed sizes and to remove noise given by Perumal and Velmurugan [2]. Resizing is done, because some images captured through the camera are not in proper dimensions and vary in size. So, to feed these images into the algorithm, the images need to be resized according to the input requirement of the algorithm. In de-noising, mentioned technique is Gaussian blurring to remove the noise in the image.

Thresholding (segmentation method), Gaussian noise elimination, grayscale conversion are preprocessing methods explained by Gandhi et al. [3] which are performed on the data running through the actual algorithm and are performed to reduce the complexity of the algorithm.

After preprocessing, we then move to study various research works related to captioning and various algorithms and methods related to it.

Captioning

An object detection algorithm called YOLO (YOU Only Look Once) by Redmon et al. [4] is one of the fastest object detection algorithms. It has 24 convolutional layers which are followed by two fully connected layers. It uses a 1×1 reduction layer which is followed by 3×3 convolutional layers. It is a high-speed object detection algorithm that uses a grid-based image to predict probabilities and bounding boxes.

As the YOLO algorithm is heavy to work with, improvisation is suggested by Huang et al. [5] which mentions a light version of YOLO. Using the Tiny YOLO v2 as a point of start, few layers were modified, and the model was trained on the Pascal VOC dataset of 2007. It has fewer fps (frames per second) and reduced accuracy concerning YOLO-v2.

A captioning methodology explained by Zeyer et al. [6] represents LSTM, RASR speech, recognition toolkit, 3:4-g language model, and NN-HMM hybrid acoustic model. Comparison and experimentation were done on RNN models for regularization, optimization, batching, layering, and all possible hyperparameter.

A comparative study of cloud-based captioning algorithms like Microsoft Cognitive Service, Clarifai, and Google Vision API is done by Ahmed et al. [7] which is also the base reference paper and talks about the generation of meaningful image captions with personalized feedback including voice command, haptics, and ringtone for visually impaired.

Speech

After captions are generated, they must be feed to suitable hardware for verbal dictation. For the same purpose, a walking stick is embedded with an ultrasonic sensor and a camera to the electronic unit is explained by Masal et al. [8]. It uses optical character recognition (OCR) for obtaining text and converting it to speech. A device to help blind people navigate in the environment without asking any person and is based on global positioning system (GPS) is suggested by Morad [9]. However,

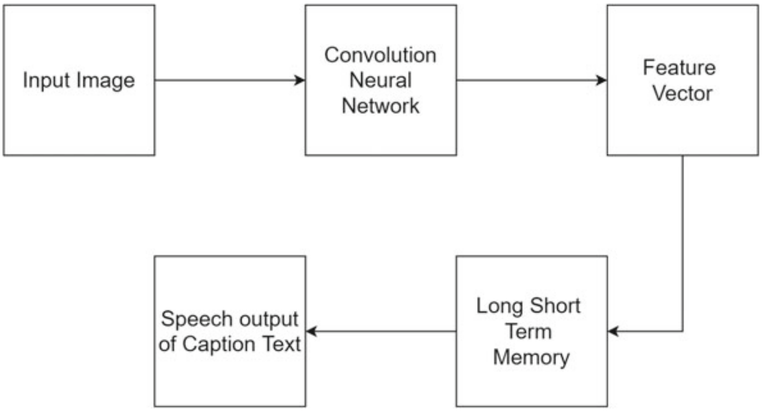


Fig. 1 YOLO v1-LSTM proposed system

its operational range is limited. A device to provide navigation aid to help visually impaired people to navigate easily, safely, and to detect any obstacles is given by Bousbia-Salah [10] which is specially designed with safety-first-approach.

By studying these papers, we have come up with few points which state issues with existing models. The existing models lack speed in the generation of output and also limits the usage where complex and heavy weighted models are not compatible. Moreover, the models which use simple RNN face problem of vanishing or exploding gradient problem which needs to be addressed, while the models which use conventional CNN algorithms can also be replaced by better versions of CNN. Considering all above factors, we have presented our solution which addresses these issues (Fig. 1).

Methodology

The whole process can be broadly divided into two parts: feature prediction and caption prediction. Feature prediction consists of preprocessing and feature extraction. Considering the Flickr-8k dataset, all the images and descriptions were loaded, and the following steps have been performed sequentially:

1. Feature extractor model (YOLO-v1 Tiny) was imported, which consists of 24 convolutional layers and has been trained on PASCAL VOC dataset. Every normalized image was passed for feature vector prediction, and all these vectors were stored (3.2).
2. Preprocessing of the captions/descriptions from Flickr-8k dataset was done which mainly included steps such as converting to lowercase, removing punctuations and hangings, and removing digits. All the descriptions were added with ‘startseq’ and ‘endseq’ to denote the starting and ending of caption for

Table 1 Dataset division (Flikr8k)

Training size	6000
Validation size	1000
Test size	1000

Table 2 Vocabulary and description details

Vocabulary size: 7579
Description length: 34

the model, respectively. Vocabulary count, i.e., unique words in the descriptions, was found and saved (7579 for Flickr-8k) also the maximum length of a description was found out (34 for Flickr-8k).

3. Training and test datasets were created for the image captioning model that took one input as the feature vector, and the other input as the sequence of words using LSTM (3.3). This sequence was padded to the maximum length (34), and each word was tokenized (given a specific representing number). The generated text of captions is dictated using text-to-speech.
4. Encoder–decoder image captioning model was built and trained on the dataset, and results were evaluated, and the BLEU scores were obtained (IV).

Dataset and Preprocessing

The dataset used is the Flickr-8K dataset with 8000 images along with the approximately five captions for each image from Kaggle. Flickr is a specially built dataset for image captioning and object localization. Further, the dataset was divided accordingly from training and testing requirements as shown in the dataset distribution in Table 1. The division contains 6000 training samples, 1000 validation, and testing samples each. We need validation. The validation set is for purpose of tuning the parameters of the model. The initial operation to be done on the dataset was image preprocessing [1–3]. As CNN-based YOLO is used, the preprocessing required is resizing and normalization. Resizing was used to format all the input in the same size and dimension because the graphics processing unit (GPU) applies the same instruction to a whole batch of images at the same time to process superfast (Tables 1 and 2).

YOLO-V1 Tiny

The problem of object detection consists of classification and localization. YOLO is a clever convolutional neural network for doing object detection in real-time as mentioned in III by Redmon et al. [4]. The CNN algorithm in YOLO applies a

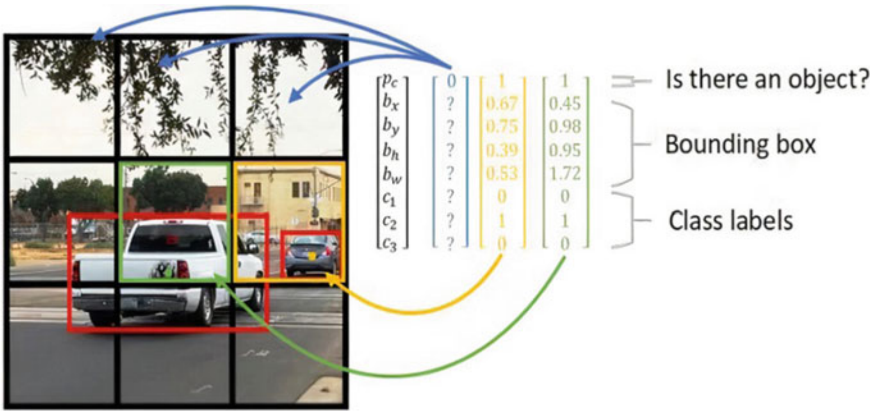


Fig. 2 YOLO output shape [11]

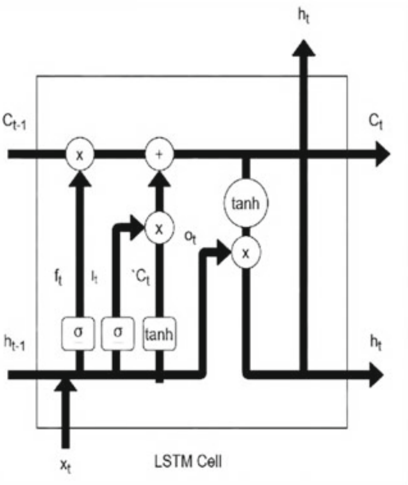
single neural network layer to the entire image and then divides the image into grids and predicts the boxes with a score of accuracy and probabilities for corresponding grids. The image gets divided into d X d dimension grids. Each grid is responsible for generating n number of anchor boxes. Each anchor box contains a probability score, bounding box dimensions, and class probability as shown in Fig. 2. In the model, the CNN block performs the encoding part as shown in Fig. 2. The images are encoded in a numeric form which is further processed by YOLO-based CNN. For the same purpose, we had to customize the YOLO model according to the output required. As we are interested in obtaining the feature vectors, we eliminate the last layer from our YOLO model, so we get the feature vector of the image in the second last layer.

Figure 3 (left) describes in detail all the layers of YOLO v1-tiny [5]. The first layer is of shape—448 × 448 × 3, which then after 33 layers gives an output of shape—7 × 30. The feature vector extracted from YOLO v1-tiny will be fed to LSTM.

LSTM

RNN is a type of neural network in which output of previous state is passed as an input to the current state along with the current input. These models perform well for sequential tasks, where there is a certain pattern and dependency between the inputs. RNN has previously processed output attached to the current input state/hidden unit. RNNs have two major problems: Vanishing gradient problem and exploding gradient problem. This focuses on the need of LSTMs, which are a type of RNNs capable of handling long-term dependencies. LSTM [12], the advanced RNN, is capable of handling the vanishing gradient problem with some added enhancements in its architecture. LSTM is explained by Gers et al. [13]. LSTM model consists of three

input_1 (Input Layer)	[(None, 448, 448, 3)]
convolutional_0 (Conv2D)	(None, 448, 448, 16)
bnconvolutional_0 (BatchNorm	(None, 448, 448, 16)
leaky_re_lu (LeakyReLU)	(None, 448, 448, 16)
max_pooling2d (MaxPooling2D)	(None, 224, 224, 16)
convolutional_1 (Conv2D)	(None, 224, 224, 32)
bnconvolutional_1 (BatchNorm	(None, 224, 224, 32)
leaky_re_lu_1 (LeakyReLU)	(None, 224, 224, 32)
max_pooling2d_1 (MaxPooling2	(None, 112, 112, 32)
convolutional_2 (Conv2D)	(None, 112, 112, 64)
bnconvolutional_2 (BatchNorm	(None, 112, 112, 64)
leaky_re_lu_2 (LeakyReLU)	(None, 112, 112, 64)
max_pooling2d_2 (MaxPooling2	(None, 56, 56, 64)
convolutional_3 (Conv2D)	(None, 56, 56, 128)
bnconvolutional_3 (BatchNorm	(None, 56, 56, 128)
leaky_re_lu_3 (LeakyReLU)	(None, 56, 56, 128)
max_pooling2d_3 (MaxPooling2	(None, 28, 28, 128)
convolutional_4 (Conv2D)	(None, 28, 28, 256)
bnconvolutional_4 (BatchNorm	(None, 28, 28, 256)
leaky_re_lu_4 (LeakyReLU)	(None, 28, 28, 256)
max_pooling2d_4 (MaxPooling2	(None, 14, 14, 256)
convolutional_5 (Conv2D)	(None, 14, 14, 512)
bnconvolutional_5 (BatchNorm	(None, 14, 14, 512)
leaky_re_lu_5 (LeakyReLU)	(None, 14, 14, 512)
max_pooling2d_5 (MaxPooling2	(None, 7, 7, 512)
convolutional_6 (Conv2D)	(None, 7, 7, 1024)
bnconvolutional_6 (BatchNorm	(None, 7, 7, 1024)
leaky_re_lu_6 (LeakyReLU)	(None, 7, 7, 1024)
convolutional_7 (Conv2D)	(None, 7, 7, 256)
bnconvolutional_7 (BatchNorm	(None, 7, 7, 256)
leaky_re_lu_7 (LeakyReLU)	(None, 7, 7, 256)
flatten (Flatten)	(None, 12544)
connected_0 (Dense)	(None, 1470)
yolo_reshape (Yolo_Reshape)	(None, 7, 7, 30)



LSTM (Input 1: YOLO)

input_2 (InputLayer)	[(None, 34)]
input_1 (InputLayer)	[(None, 1470)]
embedding (Embedding)	(None, 34, 256)
dropout (Dropout)	(None, 1470)
dropout_1 (Dropout)	(None, 34, 256)
dense (Dense)	(None, 256)
lstm (LSTM)	(None, 256)
add (Add)	(None, 256)
dense_1 (Dense)	(None, 256)
dense_2 (Dense)	(None, 7579)

Fig. 3 YOLO V1-tiny model summary (left), LSTM cell (top right), LSTM summary (bottom right)

gates: an input gate, an output gate, and a forget gate. It works much more like RNN with three parts: First part whether information from previous hidden unit needs to be remembered or can be discarded. Second part tries to learn and process new information from the input. And the last part updates and passes the information to next unit. These three parts are known as gates, first part corresponds to forget gate, second is the input gate, and third one is known as output gate [12]. Figure 3 (top right) is a block diagram of the LSTM cell. The LSTM block accepts features from the CNN block (Fig. 1) as input along with initial captions to generate human-readable text format resulting in a good and grammatically efficient string of sentences.

So basically, LSTM predictions are made by past experiences of the user's input which means the contextual information extracted from previous sequences is used for predicting the next output. In the decoder block (Fig. 1), features from the CNN block are inputs to the LSTM model along with initial captions. A dropout layer along with a dense layer with rectified linear unit (ReLU) as an activation layer is added. This LSTM model generates a sequence of words using word embeddings. Finally, after compiling the whole model, a sequence of words (image captions) is returned (Fig. 4) from the model and is made conversational using text-to-speech API.

The information/features are extracted and encoded by YOLO, and these feature vectors are passed to LSTM blocks for information processing, to find patterns and dependency, and thus ultimately predict the relevant word part of the caption. Figure 3 (top right and bottom right) shows the LSTM cell with various gates and describes in detail the summary of LSTM part of the model. Thus, by processing contextual data, it generates the caption strings which are dictated by text-to-speech system to the visually impaired user. Below given are the LSTM equations presented in comparing neural network-based decoders for the surface code [12].

$$i_t = \sigma\left(x_t U^i + h_{t-1} W^i\right) \tag{1}$$

Fig. 4 Output sample

$$f_t = \sigma\left(x_t U^f + h_{t-1} W^f\right) \tag{2}$$

$$o_t = \sigma\left(x_t U^o + h_{t-1} W^o\right) \tag{3}$$

$$\tilde{C}_t = \tan h\left(x_t U^g + h_{t-1} W^g\right) \tag{4}$$

$$C_t = \sigma\left(f_t * C_{t-1} + i_t * \tilde{C}_t\right) \tag{5}$$

$$h_t = \tan h(C_t) * o_t \tag{6}$$

f_t	Forget gate
i_t	Input gate
o_t	Output gate
x_t	Current input received
C_t	Cell state of current timestamp
C_{t-1}	Cell state of previous timestamp
h_{t-1}	Hidden unit for previous timestamp
h	Hidden unit of current timestamp
W	Weight matrix for current timestamp
U	Weight matrix for previous timestamp.

Results

A comparative study is done where VGG-16, YOLO v1 tiny, and YOLO v2 are compared. Based on the results given in Table 4 by BLEU score evaluation [14] and we have found that YOLO v1 tiny suffers a minimum loss, it is lightweight, and its runtime is minimum, and hence, is fastest among the other two.

As the comparative study table shows (Table 4), our YOLO v1 tiny-LSTM models is flexible to the input shape, the number of learning parameters is 27 million compared to the conventional VGG16-LSTM model which has 138 million learning parameters, the runtime of out YOLO v1 tiny-LSTM model is 3.4 s on the machine compared to conventional VGG16-LSTM model which is 5.3 s, it also suffers minimum loss which were detected in first and fifth epochs (also shown in Fig. 5 and Table 4).

As far as BLEU score is concerned, our results (see Table 3) of BLEU are within the below given benchmark ranges of BLEU scores set as the benchmark for skillful models [15, 16]:

"BLEU-1: 0.401–0.578,
BLEU-2: 0.176–0.390,
BLEU-3: 0.099–0.260,

Table 3 BLEU score comparisons with benchmarks

BLEU	Benchmark range	YOLO v1 tiny-LSTM	YOLO v2 tiny-LSTM	VGG-16-LSTM
BLEU-1	0.401–0.578	0.518	0.501	0.543
BLEU-2	0.176–0.390	0.264	0.240	0.292
BLEU-3	0.099–0.260	0.181	0.145	0.196
BLEU-4	0.059–0.170	0.086	0.066	0.090

Table 4 Comparative study results

Parameters	YOLO v1 tiny-LSTM	YOLO v2 tiny-LSTM	VGG-16-LSTM
Number of layers	24 Conv-layers, 3 fully connected layers	19 Conv-layers	13 Conv-layers, 3-fully connected layers
Number of learning parameters	≈27 million	≈16 million	≈138 million
Input feature vector length	1470	21,125	4096
Flexibility of image size	Flexible	Flexible	Fixed size (224,224)
Runtime	3.42364	3.96124	5.329513
BLEU-1	0.518	0.501	0.543
BLEU-2	0.264	0.240	0.292
BLEU-3	0.181	0.145	0.196
BLEU-4	0.086	0.066	0.090
Loss at 1st epoch	4.566	4.830	4.688
Loss at 5th epoch	3.210	3.632	3.395

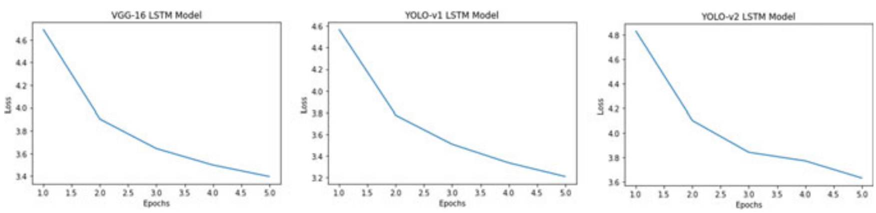


Fig. 5 Losses plots

BLEU-4: 0.059–0.170″ [16, p. 9].

The above ranges are referred to evaluate if the BLEU scores fall in the palatable range as given in Table 3. The result samples are shown in Fig. 4 which shows a sample caption of the image. Considering these parameters our model is lightweight, faster and suffers minimum loss, however, keeping the accuracy comparable against conventional counterparts as given in Table 3.

Conclusion

Considering the flexibility in the image input size, amount of loss, model weight with respect to the number of learning parameters, runtime speed, YOLO v1 tiny is fast, flexible, lightweight, and outperforms others while having the BLEU scores comparable with the VGG-16 model (see Table 3).

Hence, YOLO v1 tiny-LSTM model outperforms others and hence is preferable wherever a lightweight and fast model is required. In this way, our YOLO v1 tiny-LSTM model would generate and verbally dictates the captions using text-to-speech system to help the visually impaired navigate. The dictated captions will describe the scenario surrounding the person, so that person would have a better mental map of the surrounding conditions to make safe movements within the area.

Future Scope

This paper has discussed very little as to what can be done with this technology because of resource constraints. In the future, as improvisation to the current system, many more versatile adaptations can be made possible.

As far as implementation is concerned, single anchor boxes can be used in YOLO instead of multiple to increase to improve speed. Grid-based captioning can be implemented to generate long and more detailed captions. Dataset size can be increased with more versatility in image samples and their captions which may include specifications about scenarios.

As far as the application is concerned, the system can be modified to give specifications of the detected objects and scenarios like dimensions of objects, the number of steps detected the number of steps traveled, and many more. The system can be integrated with a map facility to give advice and suggestions based on scenarios, for example, danger zones with respect to the crime rate of an area, safe areas, areas with less traffic, short routes, nearby SOS services like hospitals and police stations, and many more. The system can be improvised to scan documents like newspapers, magazines, or monitor screens like TV or computer screens.

References

1. Pal KK, Sudeep KS (2016) Preprocessing for image classification by convolutional neural networks. In: 2016 IEEE international conference on recent trends in electronics, information &

communication technology (RTEICT). IEEE, pp 1778–1781
2. Perumal S, Velmurugan T (2018) Preprocessing by contrast enhancement techniques for medical images. Int J Pure Appl Math 118(18):3681–3688
3. Gandhi M, Kamdar J, Shah M (2020) Preprocessing of non-symmetrical images for edge detection. Augmented Hum Res 5(1):1–10
4. Redmon J, Divvala S, Girshick R, Farhadi A (2016) You only look once: unified, real-time object detection. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 779–788
5. Huang R, Pedoeem J, Chen C (2018) YOLO-LITE: a real-time object detection algorithm optimized for non-GPU computers
6. Zeyer A, Doetsch P, Voigtlaender P, Schlüter R, Ney H (2017) A comprehensive study of deep bidirectional LSTM RNNs for acoustic modeling in speech recognition. In: 2017 IEEE international conference on acoustics, speech and signal processing (ICASSP). IEEE, pp 2462–2466
7. Ahmed F, Mahmud MS, Al-Fahad R, Alam S, Yeasin M (2018) Image captioning for ambient awareness on a sidewalk. In: 2018 1st international conference on data intelligence and security (ICDIS). IEEE, pp 85–91
8. Masal U, Rajput A, Tate G, Bhong VS (2019) Virtual eye for visually blind people
9. Morad AH (2010) GPS talking for blind people. J Emerg Technol Web Intell 2(3):239–243
10. Bousbia-Salah M, Fezari M (2007) A navigation tool for blind people. In Innovations and advanced techniques in computer and information sciences and engineering. Springer, Dordrecht, pp 333–337
11. https://i.stack.imgur.com/aUcNf.jpg
12. Varsamopoulos S, Bertels K, Almudever CG (2019) Comparing neural network based decoders for the surface code. IEEE Trans Comput 69(2):300–311
13. Gers FA, Schmidhuber J, Cummins F (2000) Learning to forget: continual prediction with LSTM. Neural Comput 12(10):2451–2471. https://doi.org/10.1162/089976600300015015
14. Papineni K, Roukos S, Ward T, Zhu WJ (2002) Bleu: a method for automatic evaluation of machine translation. In: Proceedings of the 40th annual meeting of the association for computational linguistics, pp 311–318
15. Tanti M, Gatt A, Camilleri KP (2018) Where to put the image in an image caption generator. Nat Lang Eng 24(3):467–489
16. Noever D, Kalin J, Ciolino M, Hambrick D, Dozier G (2021) Local translation services for neglected languages. arXiv preprint arXiv:2101.01628

Performance Analysis of Docker Containerization and Virtualization

Pawanpreet Kaur, Jagroop Kaur Josan, and Nirvair Neeru

Abstract Data centers are the beating heart of Internet and cloud. Virtualization is one of the major technologies in data centers. Cloud computing is slower as compared with the on-premises or native environments. But its advantages weigh much more. Virtualization technologies have seen constant rise in recent years with cloud is booming the IT industry like never before. Applications are deployed using containers and virtual machines with faster and secure deployment environments. In the current industry, we are seeing containers and virtual machines as two major technologies to run application while reducing the wastage of computing resources. Performance and security of applications running in these environments are compared in our experimental work with parameters like CPU performance, file I/O, and memory utilization are compared among both the virtualization technologies. Apart from above parameters, we have also used HTTP load from different continents that is used to showcase how the application running on container and virtual machine perform with traffic around the world.

Keywords Docker · KVM · Amazon Web services · HTTP load · Microservices · Apache Sysbench

Introduction

DevOps is creating new revolution in IT industry and is the emerging driving force in cloud [1] and software industry along with technologies like Internet of things (IoT) [2], software defined networking (SDN), machine learning, and many more. The DevOps tooling also has entered into the server, network, and security industry and is rapidly changing the way deployments are done. Containerization is one of the most popular and used tool in DevOps compute tooling. Containers, since their inception, have been continuously making shifts in the data center and enterprise computing environments with its lightning fast deployment and portability. Virtual machines are

P. Kaur (✉) · J. K. Josan · N. Neeru
Department of Computer Science and Engineering, Punjabi University, Patiala, India

V. Bindhu et al. (eds.), *Proceedings of Third International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 844,
https://doi.org/10.1007/978-981-16-8862-1_56

ruling compute services for over a decade because of the benefits they provide over the OS deployments on the physical servers. With virtualization, organizations started utilizing their CPU, memory, disk, etc. computing services in a much efficient manner than before. Docker was the first popular containerization technology started in 2013, and since its inception, Docker along with different containerization platforms like Kubernetes, OpenShift, etc., were welcomed by enterprise and data center industry with open arms. During last 8 years, containerization has seen an exponential growth and is still a young technology with plenty of enhancements are carrying out in continuous manner. Containers and virtual machines are used to deploy applications and microservices. So, the question which arises when designing some application or microservice is what to use for compute services or which technology we should opt for better performance: containers or virtual machines.

Background

Containers are a virtualization at the OS-level and have become extremely popular and used technology in the recent years because of its distributed architecture without any VM instance. Using containerization technology, it has become possible to have multiple isolated instances run on a single host accessing single kernel increasing efficiency.

In the cloud, containers act on platform as a service (PaaS) and are mainly installed on the VMs running at infrastructure as a service (IaaS). VMs and containers both are virtualization techniques but with some differences. Containers are used to provide computing infrastructure resources, inter-operable, and distributed applications that need lightweight resources in the cloud. Different containers can be easily interconnected over large number of servers. Containers do not need any hypervisor and VM running over it as shown in Fig. 1.

Some of the most popular containerization platforms are Docker, Kubernetes, OpenShift, LXC, etc. Android uses LXC to run apps. Cloud vendors that offer containerized services include Amazon AWS, Micrsoft Azure, Google Cloud Platform, IBM Cloud.

Some of the popular container platforms are explained below:

A. **Docker**—Docker is one of the most popular platforms of containerization and is an open-source technology and its deployment is easier as compared with other container platforms. Docker containers can be created and pushed to the Docker Hub platform which is the repository-based system, where Docker images are stored. We can pull the Docker images from the Docker Hub and enable them to use as containers. Containers can be clustered by using Docker Swarm for better performance and reliability.
B. **KVM**—KVM is the Linux-based virtualization hypervisor available in almost all the Linux distributions for free.

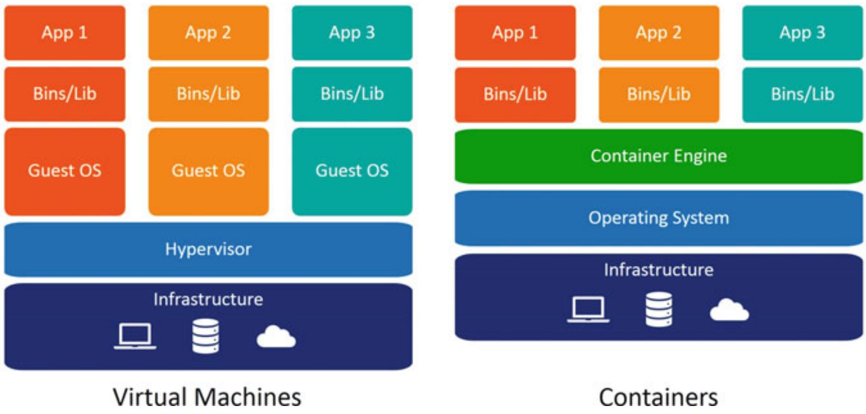


Fig. 1 VMs and containers

Related Work

Che et al. [3] compared OpenVZ, Xen, and KVM. Their study includes different benchmarks to measure the processor, RAM, disk, and network utilizations. According to them, processor virtualization is not the bottleneck for performance and major bottleneck is caused due to I/O, interrupt requests, and page table issues.

Xavier et al. [4] compared different container platforms like OpenVZ, LXC, VServer, and Xen and found that container-based platforms have native performance of processor, RAM, disk, and network utilization and the major difference found to be in the implementation of resource management and security.

Babu et al. [5] compared different virtualization techniques, i.e., Xen for paravirtualization, OpenVZ for containers and Xen for full virtualization and found that full virtualization has better performance in file copying, pipe-based context switching, shell scripting, while containerization has better performance than paravirtualization in some test cases.

Felter et al. [6] used different benchmarks as workload to get the performance of KVM and Docker and found that Docker container provides better performance than KVM hypervisor in almost all the cases. But, Docker's network address translation (NAT) adds to the overhead to the workload.

Xavier et al. [7] does the analysis of interference which is suffered during performance with large disk intensive workloads and it was found that LXC does not bring complete isolation as like KVM.

Jaikar et al. [8] assesses the performance of OpenVZ and OpenStack. It was found that containers provide better performance as compared with VMs. Intensive workload is created on processor, memory, and I/O.

Ruan et al. [9] differentiate between system and application-based containers. Extra layer on VMs with hypervisor is also analyzed for the overhead that it produce and it is compared with the containers among Amazon elastic container service (ECS) and Google container engine (GKE). Experimentation includes processor, RAM, bandwidth, and network delay between Docker, LXC, ECS, and GKE. As per the results, containers which are running inside the VMs had 42.7% degradation and 233% delay in the network and containers mainly counts on system container service in case of high I/O workloads.

Herbein et al. [10] explained different solutions to improve resource management for HPC apps like CPU, network/I/O bandwidth. Mechanism relates with context switching, and processor management, and I/O balancing in the Docker and network-based bandwidth. LINPACK is used as the benchmark to reevaluate the work.

Mardan and Kono [11] compared the performance of MYSQL I/O running on containers and VMs. It was found that KVM performed around 86% better as compared with LXC, which was mainly due to file system journaling that has a major impact on performance of the containers and it has become an I/O bottleneck for the containers.

Barik et al. [12] used different benchmarks in order to assess the usage of PaaS containers. After various tests, it was found that containers have slight edge over the VMs on the basis of the performance, but there are limitations related to security and isolation in the containerization as compared with VMs.

Recent research work [3–12] mainly revolves around single container environment, and no focus was no clustered container environments. Web applications and APIs run on either VMs or containers, and HTTP load is one parameter that showcases how the application is performing on some specific infrastructure.

Methodology

Experimentation done in our work consists of Docker containers and KVM VMs running on AWS. Apache2, along with PHP, is deployed on both the container cluster and VM and a PHP Web page is created calling phpinfo(); function that responds with a PHP manual. Performance is measured using parameters like CPU, memory, disk utilization along with Web page load time. AWS CloudWatch, Wireshark, and PRTG are the tools used for monitoring and analysis. Linux distribution used for both Docker and KVM VMs is Ubuntu Server 20.04. AWS instance type used is t2.micro for every instance that has 1vCPU, 1 GB RAM by default and we have extended the default 8 GB SSD to 60 GB for storage per Ubuntu VM. Hardware requirements are feasible to run any information Web page with 1–50 user traffic. Comparison is done using analysis graphs gathered from PRTG, Wireshark, and AWS CloudWatch.

Results

This section discusses the performance analysis of both containerization and virtual machines. Results are mainly categorized using four different parameters, i.e., CPU utilization, Apache load, disk utilization, memory utilization with comparisons been made using AWS CloudWatch, Apache Benchmark, and PRTG monitoring tools. Subsections are created to describe different parameter-related comparison.

CPU Performance

Different methods are used to measure the computing performance with major focus on completion of number of operations in a given amount of time or event per second. Number of virtual CPU cores used can impact the performance. Apache Benchmark is used to test the CPU performance, and AWS CloudWatch is used to test the CPU utilization.

Apache Sysbench Maximum Prime Number

Apache Sysbench tool is used to conduct the test that showcases the time needed to perform the maximum prime number. The maximum prime number is to be selected from the first 70,000 numbers. So, when this calculation is running with CPU workload, Apache Sysbench verifies the prime numbers by performing division operation between two and square root of number and if the number comes in output as 0, then the next number is calculated. This calculation will add the stress of CPU. Below is the output using a single thread of CPU to run the operation. Experimentation is performed on AWS EC2 t2.micro instance having 1 GB RAM and 1vCPU. Figures 2 and 3 shows the Apache Sysbench CPU testing with maximum 70,000 prime numbers problem on docker and KVM respectively.

Another case study has been performed using 100,000 prime numbers limit for both Docker container and KVM-based VM as shown in Figs. 4 and 5.

File I/O Performance Testing

When file input/output is in use, we need to create a test file with size larger than available memory to make sure caching of file does not make impact on server workload.

Apache Sysbench Software is used to test file I/O operations from making a file to run the test case. Two tests were conducted with 10 and 50 GB file size, and disk operations were calculated using Sysbench with outputs shown in Figs. 6, 7, 8, and 9.

Table 1 is also created to compare the file I/O operations on Docker container and KVM machine which show that KVM performs better than containers.

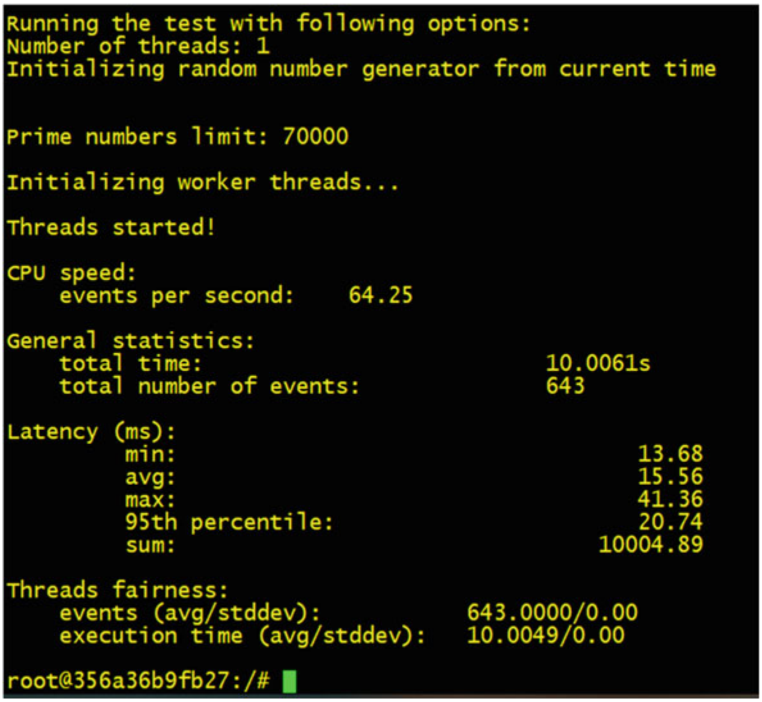


Fig. 2 Apache Sysbench CPU testing with maximum 70,000 prime numbers problem on Docker container

HTTP Load Test

HTTP load test is used to check what amount of time (msec) is needed to load the Website running on Docker container and KVM. PRTG is used as the monitoring tool to check the HTTP load on the container and virtual machine, and the results are depicted in the graph as shown in Figs. 10 and 11.

HTTP Load Test from 4 Continents

This HTTP load test is conducted to check how RTT is getting affected when traffic from different continents is hitting our Web server running on Docker container and KVM virtual machine on AWS Canada Central Ubuntu Server 20.04. Figures below show the RTT on Website running on Docker container and KVM machine (Figs. 12, 13 and Table 2).

After comparing both results of HTTP load test from 4 different locations, it is clear that KVM has a slight edge over Docker containers in terms of RTT.

```
Running the test with following options:
Number of threads: 1
Initializing random number generator from current time


Prime numbers limit: 70000

Initializing worker threads...

Threads started!

CPU speed:
    events per second:    63.27

General statistics:
    total time:                          10.0031s
    total number of events:              633

Latency (ms):
         min:                                    13.87
         avg:                                    15.80
         max:                                    41.26
         95th percentile:                        24.83
         sum:                                 10002.03

Threads fairness:
    events (avg/stddev):           633.0000/0.00
    execution time (avg/stddev):   10.0020/0.00
```

Fig. 3 Apache Sysbench CPU testing with maximum 70,000 prime numbers problem on KVM

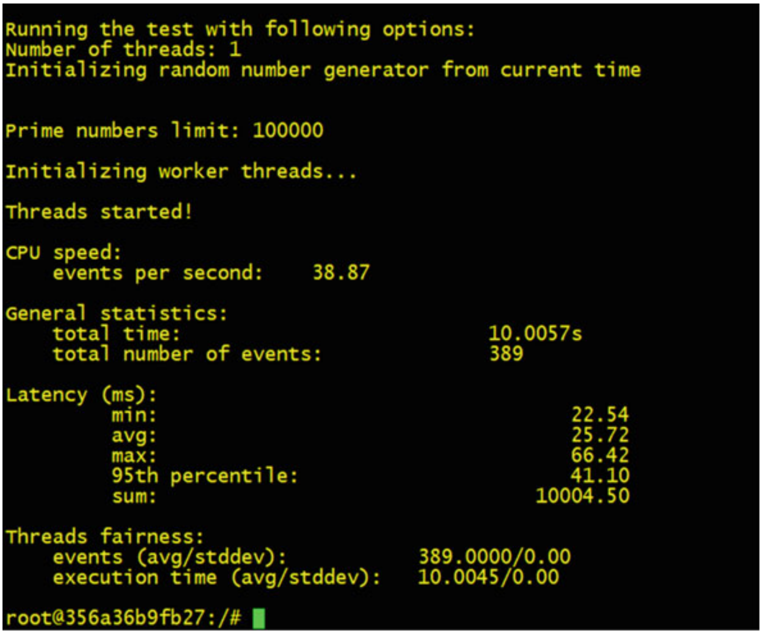
Memory Utilization Test

Memory utilization test is performed to check how efficiently Docker and KVM use memory resources and which technology uses more memory while running an Apache Web server with PHP application running over it. Figures given below highlight the usage of memory in the form of graph depicted in PRTG monitoring tool via SNMP protocol (Figs. 14 and 15).

So, when we check the above graphs, it is clear that memory utilization is almost similar while running PHP application on Apache Web server on both KVM and Docker container.

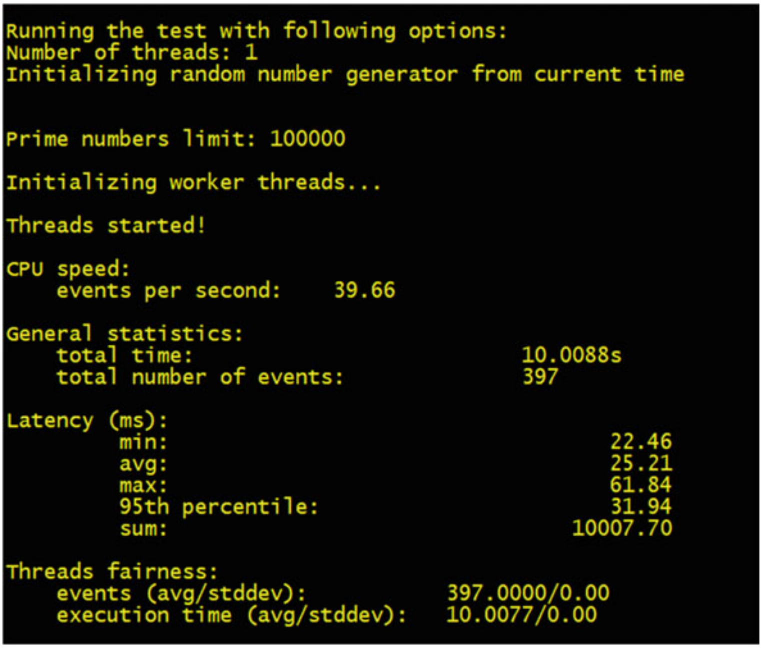
Conclusion and Future Scope

Virtualization and containerization are running the data centers worldwide and help in reducing resource utilization along with many different benefits. Containers have seen exponential rise in usage in last 5 years because of microservices and DevOps. When compared on the basis of different parameters related to performance like CPU, memory, file I/O, HTTP load from different locations, KVM results better



```
Running the test with following options:
Number of threads: 1
Initializing random number generator from current time


Prime numbers limit: 100000

Initializing worker threads...

Threads started!

CPU speed:
    events per second:    38.87

General statistics:
    total time:                          10.0057s
    total number of events:              389

Latency (ms):
         min:                                  22.54
         avg:                                  25.72
         max:                                  66.42
         95th percentile:                      41.10
         sum:                               10004.50

Threads fairness:
    events (avg/stddev):           389.0000/0.00
    execution time (avg/stddev):   10.0045/0.00

root@356a36b9fb27:/#
```

Fig. 4 Apache Sysbench CPU testing with maximum 100,000 prime numbers problem on Docker container



```
Running the test with following options:
Number of threads: 1
Initializing random number generator from current time


Prime numbers limit: 100000

Initializing worker threads...

Threads started!

CPU speed:
    events per second:    39.66

General statistics:
    total time:                          10.0088s
    total number of events:              397

Latency (ms):
         min:                                  22.46
         avg:                                  25.21
         max:                                  61.84
         95th percentile:                      31.94
         sum:                               10007.70

Threads fairness:
    events (avg/stddev):           397.0000/0.00
    execution time (avg/stddev):   10.0077/0.00
```

Fig. 5 Apache Sysbench CPU testing with maximum 100,000 prime numbers problem on KVM

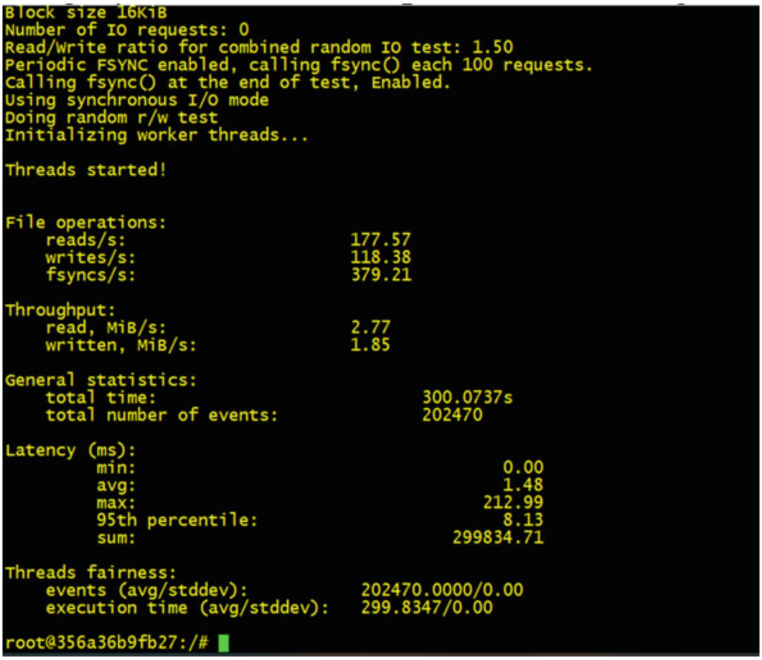


Fig. 6 Apache Sysbench file I/O performance testing with 10G file on Docker container

than Docker containers using Benchmark tooling like Apache Sysbench for CPU performance, file I/O. PRTG is used for memory utilization and HTTP load tests. Results also show that KVM uses less memory as compared to Docker containers. Also, HTTP load test shows that PHP-based Web application running over Apache Web server KVM has slight edge over Web server in Docker as KVM-based Web server has lower RTT. Overall, KVM has edge over Docker and provides better performance than Docker containers. In future, researchers should analyze performance comparison on microservices running over containerization and virtualization technologies as more and more applications are using microservices on containerized and virtualized environments. Apart from performance analysis, security is another aspect which is needed to be analyzed and lots of improvements are needed with rapid Internet growth.

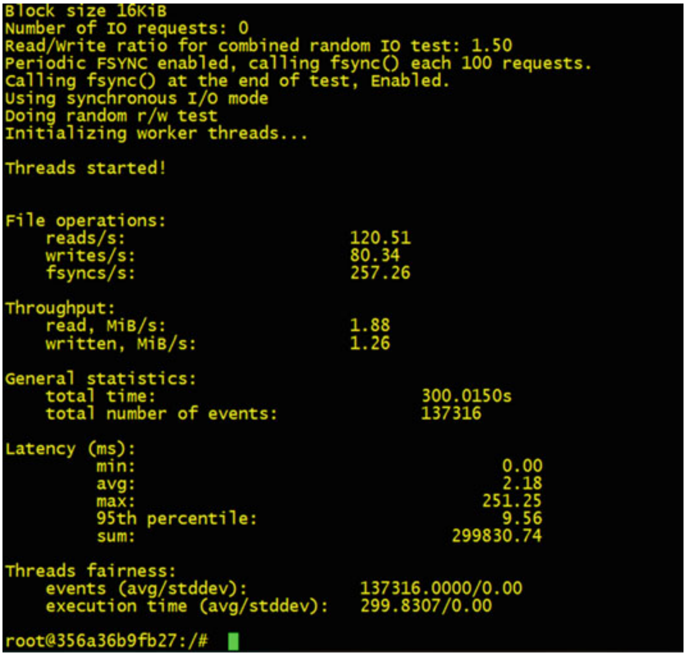


Fig. 7 Apache Sysbench file I/O performance testing with 10G file on Docker container

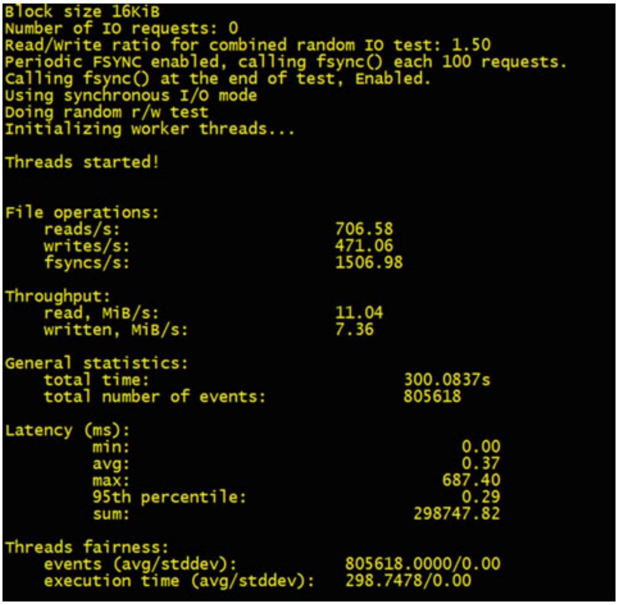


Fig. 8 Apache Sysbench file I/O performance testing with 10G file on Docker container

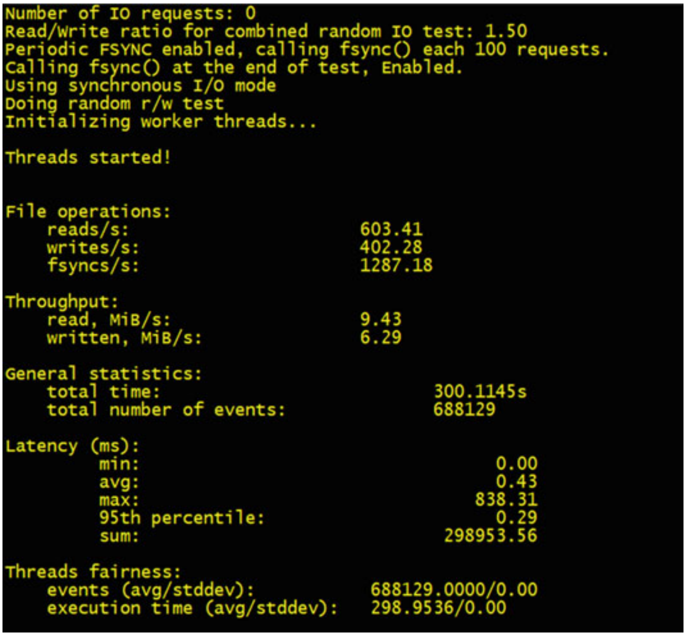


Fig. 9 Apache Sysbench file I/O performance testing with 10G file on Docker container

Table 1 Container versus KVM on the basis of file I/O

Operations	Docker (10 GB file)	KVM (10 GB file)	Docker (50 GB file)	KVM (50 GB file)
Read operations per second	177.57	706.58	120.51	603.41
Write operations per second	118.38	471.06	80.34	402.28
Other operations per second	379.21	1506.98	257.26	1287.18
Read throughput in MB per second	2.77	11.04	1.88	9.43
Written throughput in MB per second	1.85	7.36	1.26	6.29
Average latency in ms	1.48	0.37	2.18	0.43
Max latency in ms	212.99	687.40	251.25	838.31

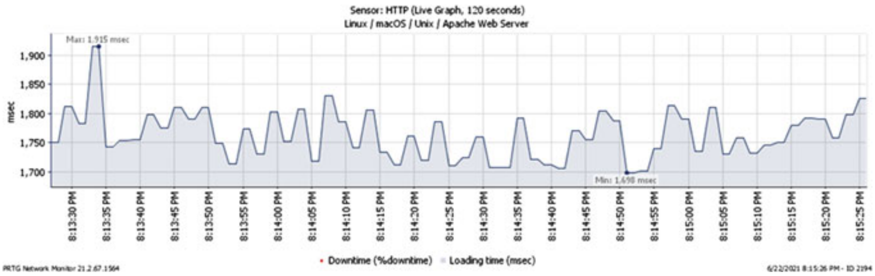


Fig. 10 HTTP load test on KVM

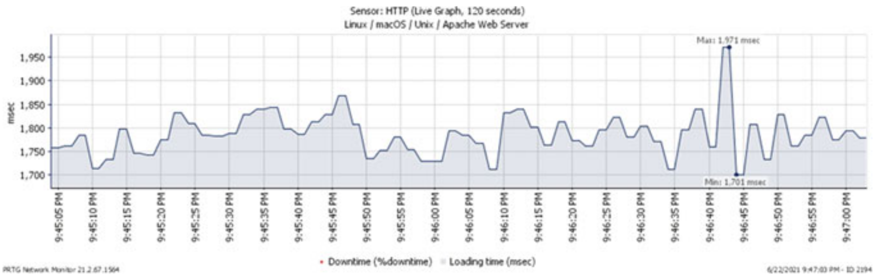


Fig. 11 HTTP load test on Docker

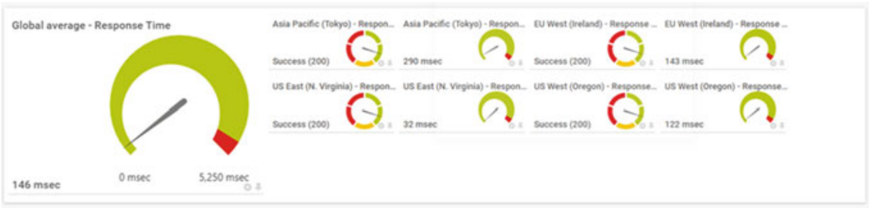


Fig. 12 HTTP load test from 4 continents on KVM

Fig. 13 HTTP load test from 4 continents on Docker

Table 2 HTTP load test from 4 different continents

Source location	Docker container apache server RTT (msec)	KVM apache Web server RTT (msec)
Global response time	148	146
Asia Pacific (Tokyo)	292	290
EU West (Ireland)	146	143
US East (North Virginia)	32	32
US West (Oregano)	123	122

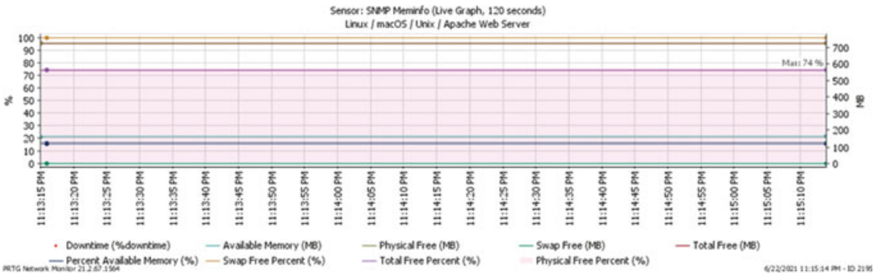


Fig. 14 Memory utilization test on Docker

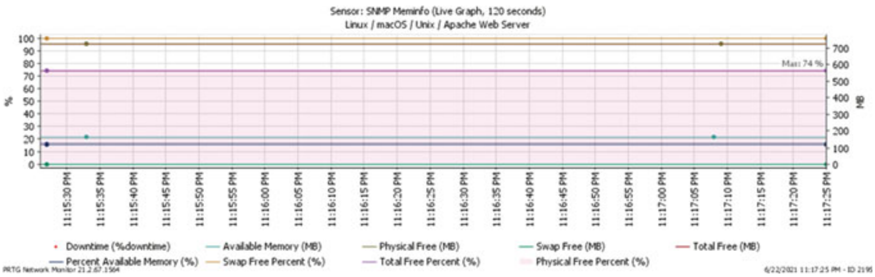


Fig. 15 Memory utilization test on KVM

References

1. Smys S (2020) A novel multi-tier architecture based mobile cloud computing for enhanced energy utilization. J ISMAC 2(01):62–72
2. Lai KL, Chen JIZ (2021) Development of smart cities with fog computing and internet of things. J Ubiquit Comput Commun Technol (UCCT) 3(01):52–60
3. Che J, Shi C, Yu Y, Lin W (2010) A synthetical performance evaluation of openvz, xen and kvm. In: 2010 IEEE Asia-Pacific services computing conference, Dec 2010, pp 587–594
4. Xavier MG, Neves MV, Rossi FD, Ferreto TC, Lange T, Rose CAFD (2013) Performance evaluation of container-based virtualization for high performance computing environments. In: 2013 21st Euromicro international conference on parallel, distributed, and network-based processing, Feb 2013, pp 233–240

5. Babu A, Hareesh MJ, Martin JP, Cherian S, Sastri Y (2014) System performance evaluation of para virtualization, container virtualization, and full virtualization using Xen, Openvz, and Xenserver. In: 2014 fourth international conference on advances in computing and communications, Aug 2014, pp 247–250
6. Felter W, Ferreira A, Rajamony R, Rubio J (2015) An updated performance comparison of virtual machines and linux containers. In: 2015 IEEE international symposium on performance analysis of systems and software (ISPASS), Mar 2015, pp 171–172
7. Xavier MG, Oliveira ICD, Rossi FD, Passos RDD, Matteussi KJ, Rose CAFD (2015) A performance isolation analysis of diskintensive workloads on container-based clouds. In: 2015 23rd Euromicro international conference on parallel, distributed, and network-based processing, Mar 2015, pp 253–260
8. Jaikar A, Shah S, Bae S, Noh S (2016) Performance evaluation of scientific workflow on openstack and openvz. Soc Inf Telecommun Eng, LNICST 167:126–135
9. Ruan B, Huang H, Wu D, Jin H (2016) A performance study of containers in cloud environment. In: Wang G, Han Y, Martínez-Pérez G (eds) Advances in services computing. APSCC 2016, vol 10065 , pp 343–356
10. Herbein S, Dusia A, Landwehr A, McDaniel S, Monsalve J, Yang Y, Seelam S, Taufer M (2016) Resource management for running hpc applications in container clouds. Lect Notes Comput Sci 9697:261–278
11. Mardan AAA, Kono K (2016) Containers or hypervisors: which is better for database consolidation?. In: 2016 IEEE international conference on cloud computing technology and science (CloudCom), Dec 2016, pp 564–571
12. Barik RK, Lenka RK, Rao KR, Ghose D (2016) Performance analysis of virtual machines and containers in cloud computing. In: 2016 International conference on computing, communication and automation (ICCCA), Apr 2016, pp 1204–1210
13. Docker (2019) https://docs.docker.com/2019. Online; Accessed 24 Mar 2019
14. Kengond S, Narayan DG, Mulla MM (2018) Hadoop as a service in openStack. In: Emerging research in electronics, computer science and technology , pp 223–233
15. Kominos CG, Seyvet N, Vandikas K (2017) Bare-metal, virtual machines and containers in OpenStack. In: 20th conference on innovations in clouds, internet and networks (ICIN), Paris, pp 36–43
16. Higgins J, Holmes V, Venters C (2015) Orchestrating Docker containers in the HPC environment. In: Kunkel J, Ludwig T (eds) High performance computing. Lecture notes in computer science, vol 9137, pp 506–513
17. Plauth M, Feinbube L, Polze A (2017) A performance evaluation of lightweight approaches to virtualization. In: CLOUD COMPUTING: the eighth international conference on cloud computing, GRIDs, and virtualization
18. Seo K-T, Hwang H-S, Moon I-Y, Kwon O-Y, Kim B-J (2014) Performance comparison analysis of linux container and virtual machine for building cloud. Adv Sci Technol Lett 66:105–111
19. Morabito R, Kjällman J, Komu M (2015) Hypervisors vs. lightweight virtualization: a performance comparison. In: IEEE international conference on cloud engineering, Tempe, AZ, pp 386–393
20. Rad BB, Bhatti HJ, Ahmadi M (2017) An Introduction to Docker and analysis of its performance. IJCSNS Int J Comput Sci Netw Secur 17(3):228–229
21. Kozhirbayev Z, Sinnott RO (2017) A performance comparison of container-based technologies for the cloud. Future Gener Comput Syst, 175–182
22. Felter W et al (2015) An updated performance comparison of virtual machines and linux containers. IEEE Int Symp Perform Anal Syst Softw (ISPASS), 171–172.
23. Sysbench (2019) https://wiki.gentoo.org/wiki/Sysbench2019. Online; Accessed on 24 Mar 2019

24. Phoronix Benchmark tool (2019) https://www.phoronix-test-suite.com/2019. Online; Accessed on 24 Mar 2019
25. Apache Benchmark tool (2019) https://www.tutorialspoint.com/apache_bench/2019. Online; Accessed on 24 Mar 2019
26. Ghatrehsamani D, Denninnart C, Bacik J, Amini Salehi M (2020) The art of cpu-pinning: Evaluating and improving the performance of virtualization and containerization platforms. In: 49th international conference on parallel processing-ICPP, pp 1–11

A Study on Recent Advances in Artificial Intelligence and Future Prospects of Attaining Superintelligence

Anudeex Shetty and Nivesh Raj

Abstract By the year 2029, computers will be as intelligent as humans, and in 2045, computers will attain superintelligence and lead to the singularity, as per the futurist Ray Kurzweil. The same was echoed by Elon Musk, who predicted artificial intelligence to overtake the human race by 2025. Time travel between different points in space-time is through a time machine. Time machine bends space-time to invert timelines and form loops. Likewise, the machines can learn similarly to a child's brain, i.e., by interacting with other objects in the environment and leveraging mental time travel and past experiences. This is feasible with the required hardware and appropriate programming. The computer acting as a time machine coupled with artificial intelligence (AI) might lead to superintelligence, and machines might take over humans. In this paper, we go through current and future AI systems coupled with time travel. We look at both recent and the path forward for software and hardware infrastructure progress. We also estimate the timeline when machines potentially could overtake humans (and performing almost all the work) by correlating AI market size and percentage of work done by machines along with experts' opinion analysis. To conclude, we also provide few policies to avoid the potential AI apocalypse.

Keywords Artificial intelligence · Time travel · Time machine · Superintelligence · Moore's law · Machine learning · Singularity · AI ethics · Policy · Special relativity

A. Shetty (✉)
Computer Science and Engineering, Birla Institute of Technology, Mesra, Ranchi, India
e-mail: be10708.15@bitmesra.ac.in

A. Shetty · N. Raj
Office of Sustainability, The Leadership 30, Mumbai, India

V. Bindhu et al. (eds.), *Proceedings of Third International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 844,
https://doi.org/10.1007/978-981-16-8862-1_57

Introduction

Superintelligence: Artificial Intelligence Apocalypse

Superintelligence means an intellect much smarter than the best human brains in just about all the areas, such as wisdom, scientific inventiveness, and social competencies. This leaves the question, how superintelligence is implemented openly. It can be anything, a network of computers, a digital computer, cultured cortical tissue, and many more [1–4].

Both hardware and software are needed for superintelligence. According to Moore's law [5], in 2025, considering the rate of change, a $1000 computer would have a processing speed of 10^{16} ops/sec, equivalent to that of a human brain. This addresses the superintelligence hardware requirement. Numerous approaches are requiring a varying amount of top-down direction [4]. The procedure is already underway with sophisticated artificial intelligence (AI) systems like DeepMind [6], IBM Watson [7], and Vicarious [8]. These systems continuously improve themselves, thereby becoming more intelligent. For example, the advancements in the AI field stagnated in the 1970s and 1980s. In retrospect, the AI projects of that period could not have succeeded. The hardware available was not at the level needed by AI.

The development of AI will increase multifold once they get to human-level intelligence because of the feedback loop in them. This would be like a chain reaction, wherein AI systems keep on building superior AI systems, also known as "intelligence explosion" [9] or "seed AI" [10]. We would have an AI apocalypse once we have a human-level AI system [4].

Artificial Intelligence Systems

AI systems are better with more data, with expanding Internet of everything (IoE), by 2025, we would have 100 billion devices with each having more than a dozen sensors connected. With these sensors collecting data from wearables, cameras, satellites, cars, and many more, we would have a perfect knowledge dataset. Furthermore, global high-speed Internet plans by SpaceX (Starlink [11]), Facebook and Qualcomm (Internet.org [12]), and Virgin (OneWeb [13]) to have orbital Internet would further facilitate it.

Companies and other groups will continue to invest heavily in augmented and virtual reality that will lead to a new generation of user interfaces and display with scope revolutionizing travel, entertainment, education, and science. Similarly, AI assistant bots like Siri, Alexa, and Google Assistant have a high IQ as of now, and they would continue evolving and we delegating more to them [14].

Time Travel

Time travel is a notion for movement from one point to another in time, analogous to an entity's movement in space, generally done via a device known as time machine. A time machine is a device that produces closed time-like curves, therefore enabling time travel [15].

Time travel, other than the accustomed awareness of time, is possible as per Einstein's special relativity and general relativity. As for time travel in the backward direction, as per general relativity, there is plausibility. Given the limitation in theoretical physics, traveling to a random point in space-time is usually linked with wormholes or quantum mechanics [2].

Related Work

There have been several tracks of research on using time travel capabilities in a computer to tackle a wide range of problems like diagnosis of issues, debuggability, ML feature generation, ML model feedback loops, VR games, virtual time travels, and many more. They are covered in Sects. "Computer as a Time Machine", AI and Time Travel", and "Virtual Time Travel".

Many researchers in the past have predicted when human intelligence and superintelligence will be achieved [1, 4, 9, 16]. But they are pretty old and they have not corelated to the recent progress made in AI. Plus, no one has thought time travel capabilities could also play a pivotal role in getting AI to that level which we will cover in the following section.

Proposed Work

Computer as a Time Machine

Time travel is a potent hypercomputation criterion. A computer can travel to the future and send itself back in time to the present, creating solutions to otherwise computationally intractable problems [17].

Time machine computing is a technique in which information is arranged based on time in computers, making it possible for the user to look up different (both past and future) circumstances of the computers [18]. The computer as a time machine is helpful for various purposes like diagnosing and debugging issues, virtual reality, human–computer interaction (HCI), and many more.

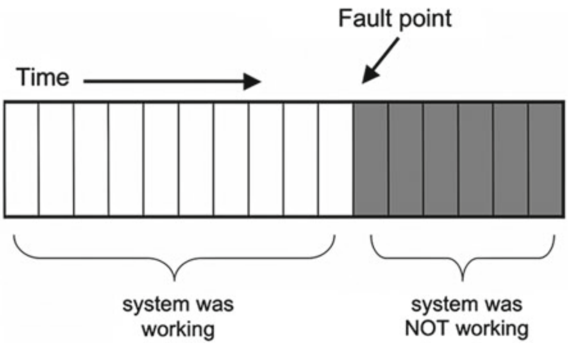


Fig. 1 Chronus using time travel for diagnosis [21]

Diagnose and debug computer issues

Time travel virtual machine (TTVM) leverages time travel in two steps. First, it restores to a checkpoint before the potential fault point. Then, it re-runs from the checkpoint to the fault point. This provides an ability for programmers to navigate backward and forward in the execution history and replay random segments of the past [19]. This is more convenient and efficient than classic debugging because it does not require the entire run to be repeated. Other uses of TTVM are improving availability, forensics, system administration, and boot speedup [20].

TTVM has been utilized for overcoming many difficulties (non-deterministic errors, direct hardware impact, corrupted state, and longer runs) associated with debugging operating systems. In Fig. 1, Chronus, a TTVM, maintains the logs of all system state changes such that it could later simulate the states for the random points in the past. It then uses binary search to find out the fault point in the system [21].

Time travel debugger (TDD) empowers the developer to precisely track and understand the order of commands and corresponding values causing the error quickly. In addition, TTDs also provide transparent application migration, deterministic replay, idempotent program checking, and automatic fault recovery [22].

For instance, SPYDER uses execution backtracking techniques along with dynamic slicing as depicted in Fig. 2 to automate debugging [23]. Some other TTDs are ODB [24], TOD [25], Whyline [26], ReTrace [27], EXPOSITOR [28], IGOR [29], Mugshot [30], Timelapse [31], Undangle [32].

AI and Time Travel

AI is time travel to past and future information and using it to demonstrate intelligence. Although they are just predictions, AI predictions could fetch us insights from the future or the past. A superintelligent model would go highly accurate in predicting the events to be expected of its future. It would be easier for the same model to calculate the possible story of the past based on the present situation of that specimen.

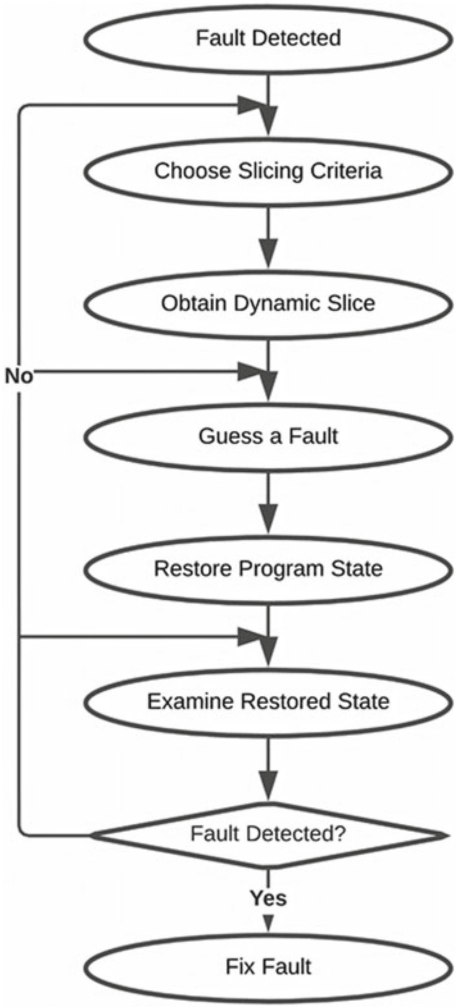


Fig. 2 SPYDER's debugging paradigm [23]

Netflix's time machine for feature generation

Netflix aims to have a platform where its users can easily find the content they want and satisfy their needs. They do so by taking an algorithm-based approach backed by data. They leverage machine learning using time travel-based feature generation on historical data and effortlessly try out new ideas.

The time machine creates snapshots and uses them to reconstruct the inputs and compute any new feature values for a time in the past. It travels back to fetch all the required data points from the snapshots based on the experiment plan. Then, it produces a dataset of labels and features for that time in the past to train the ML models. Most importantly, paradoxes are not possible (e.g., the label cannot be in the features) [33, 34] (Fig. 3).

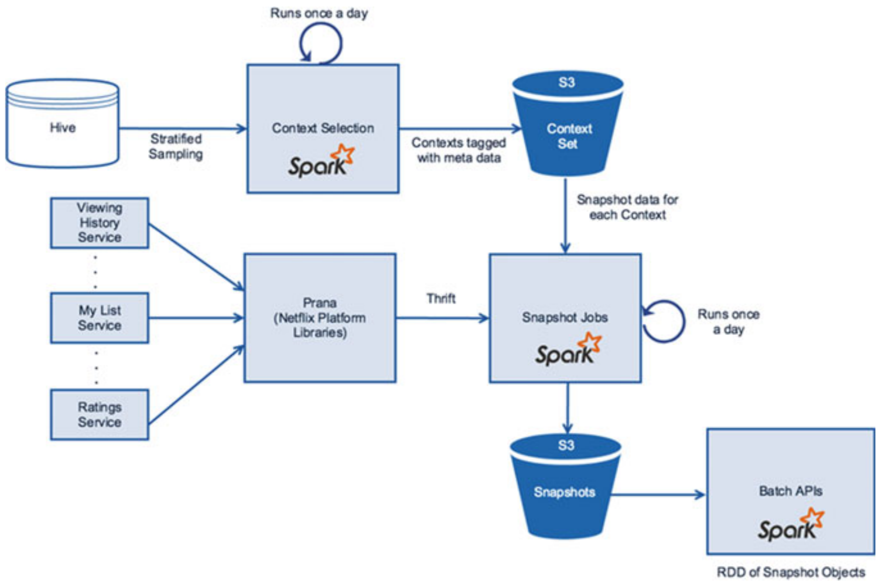


Fig. 3 Netflix's feature generation time machine architecture

Virtual Time Travel

Virtual time travel can make our imagination of being able to progress back to the days of our antecedent to experience how people were living at that time a reality. There are various applications of virtual time travel in the field of entertainment, museums [35], gaming, archeology, heritage, education, science, history, and tourism.

For example, experiential archeology using virtual time travel is becoming possible. Prehistory sites that are no longer in existence or inaccessible due to time and space could now be accessed by anyone anywhere [36]. Kartta Labs, Google's open-source initiative, aims to rebuild cities virtually for different periods and making them available for numerous use cases [37].

Similarly, in gaming, Time Warp [38]—a mobile outdoor pervasive augmented reality game for exploring the history of a city, which enables them to see the elves and visit different time periods—roman, medieval, new age, and the future. The game aims to amalgamate virtual and real aspects and make users feel they are part of different time periods.

The European time machine project

The European time machine project is a joint effort to build a virtual time machine between major industrial companies such as Ubisoft and 225 European research institutions from 32 countries. It would be a large database for storing, interpreting, and connecting historical information and rebuilding plausible past views [39]. A

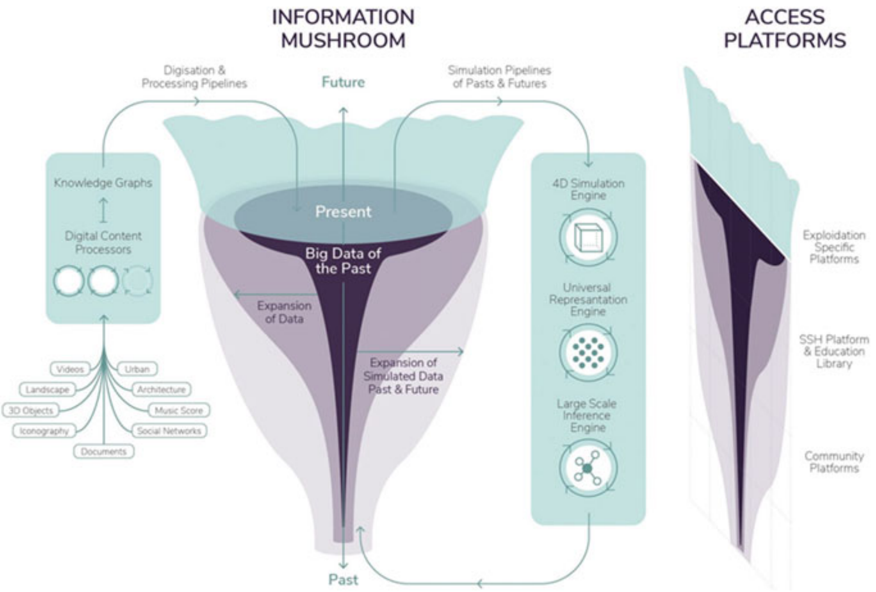


Fig. 4 European time machine architecture

proof of concept was done using digitized archival records to show Venice city's evolution—Venice time machine [40].

As per Fig. 4, AI will be pivotal at each step, from digitization planning to document interpretation and fact-checking. Large digitization frameworks and high performance computing (HPC) will be combined with ML techniques to build a multi-resolution simulation of the history of several thousand years. Tackling issues of quality of digital intellects drawn, extracting knowledge and large-scale digitization and data processing [41].

Beyond Moore's Law

For the past half-century, Moore's law has been the chief motive for the technology sector. It was given by Gordon Moore, the co-founder of Intel and Fairchild Semiconductor. It mentions that the integrated circuits crammed on a microchip double every 18–24 months [5]. As evident from Fig. 5, the process has slowed down in recent years, and we are nudging toward a limit. Gordon Moore and others expect this to last till 2025 [42]. Reasons being thermodynamic, thermal, and quantum tunneling limits [43].

As a result, we are going beyond it with other ways to speed up overall performance, like neuromorphic computing, 3D integrated circuit (IC) stacking, quantum computing, and AI. And strategies like silicon photonics and carbon nanotubes are

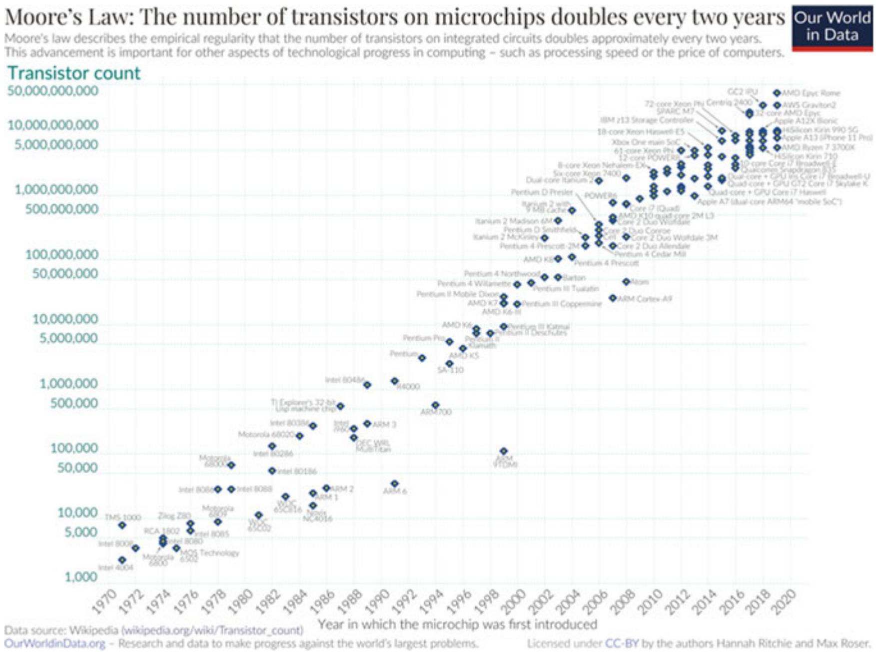


Fig. 5 Moore's law over the years with different chips scatter plotted [44]

also pursued [45]. The innovation will carry on with the development of new IC packaging and architectures, system architectures, and parallelism techniques [46]. So, we are on track to have the hardware needed for a superintelligent system, even with the end of Moore's law.

AI Technical Performance

AI Index annual reports by Stanford Institute for Human-Centered Artificial Intelligence have discovered AI development has surpassed Moore's law in 2012. As shown in Fig. 7, it resembled Moore's law growth till 2012 doubling every 24 months now it doubles every 3.4 months (Fig. 6).

The research analyzed the improvements in image identification algorithms, and most of them are supervised ML algorithms. It benchmarked on model training time and cost to evaluate advances in AI infrastructure and quantify it. As per that in 2017, the ResNet model needed 2 weeks of training to get 93% accuracy costing \$2323. The same model was trained to cost a mere \$12 in 2018 on the GCP cloud. Similarly, the training time for huge image classification models has improved from 3 h in October 2017 to 88 s in July 2019 in a year and a half period [48].

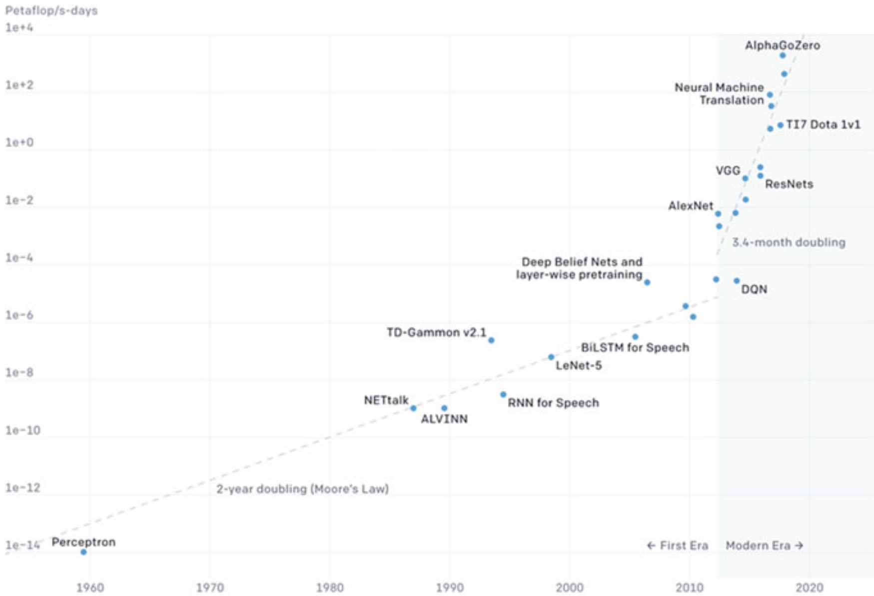


Fig. 6 Two periods of compute (pfs-day[1]) needed for training AI systems [47]

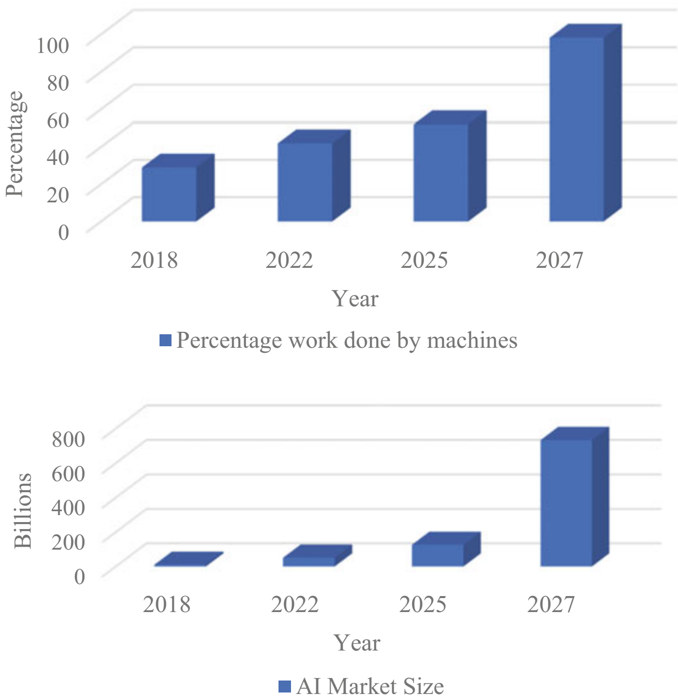


Fig. 7 AI market size and percentage of work done by machines plotted for recent years

Table 1 Data points for past years. 2018, 2022, 2025 referred from Trantica, 2027 referred from grand view research [52, 53]

Year	Market size of AI (in billion)	Percentage work done by machines
2018	10.1	29
2022	51.27	42
2025	126	52
2027	733.7	? (x)

Similarly, autonomous vehicles (AVs) have shown multifold growth in recent years. In between 2015 and 2018, the numbers of companies getting involved and the number of test miles have increased seven times. For instance, more than 2 million miles were driven by 500 AVs of 50 companies in California.

Now, AI generative models can generate a high-quality image, text, and audio which even humans struggle to identify the difference in the outputs. Some famous generative models are AlphaFold [49] by DeepMind for protein molecular structure and GPT-3 [50], a human text generator by OpenAI [51]. From this, we can be confident that software will be at the level needed for superintelligence.

Correlation between Market Size of AI and Percentage Work Done by Machines

See Table 1.

Calculating estimated percentage work done by machines in 2027.

Between year 2022 and 2025,

Percentage increase in work done by machines = (52 − 42) % = 10%

Increase in market size = 126 − 51.27 = 74.73 billion

Therefore, a 10% increase in machine work translates to a 74.73 billion market size.

Similarly, a 1% market share translates 7.47 billion AI market size.

For the year 2027.

Estimated AI market size = 733.7 billion

Considering market size per percentage work calculated above, 7.47 billion/year[2]

Estimated percentage work done by machines in 2027 (x) = 733.7/7.47 = 98.2% (Fig. 7).

[1] A petaflop/s-day (pfs-day) equivalent to 10^{20} executions, or 10^{15} neural net executions/second for one day.

[2] *Due to lack of data points, giving priority to the recent time* window.

Forecasts and Surveys

AI scientists forecast machines to reach the intelligence of the human level. Müller and Bostrom surveyed prominent researchers in 2012–13 asking them about human-level intelligence and superintelligence. According to it, human-level intelligence will be attained in 2040–50 time period with 50% chances. They also predict super-intelligence will be achieved within 30 years of post-human-level intelligence. Moreover, one-third of them felt this would turn out bad for humans, probably leading to AI Apocalypse [54]. A similar poll was conducted by AI@50 in 2006, according to it, 18% of them felt we will get to human intelligence by 2050 [55]. The confidence has increased significantly since.

Similarly, Loup Ventures, a research company from New York, published a report testing the IQ of famous digital personal assistants—Google Assistant, Siri, and Alexa asking all of them the same 800 questions. Google Assistant was best with 93% correct answers and understood all questions correctly. Next, Siri misunderstood two questions and had 83% correct answers. Then Alexa misunderstood one question only but had 80% correct answers. The results were remarkable. Notably, in the last few years, there has been an improvement in the accuracy along with the skillset of the assistants due to continuous advances made here [61].

Policy Recommendation and AI Ethics

- UN should make sure through policymaking that AI systems must contribute to humanitarian goals. The government, corporations, foundations, and private sector should collaborate to develop better AI systems to achieve SDGs. At least 25% of CSR funds should be spent on research and developing AI systems in every country.
- Coding and artificial intelligence must not be restricted to only developed countries. The policy funding structure of every country needs to develop a home talent pool from the youth in the AI sector if they want to be independent of countries like the USA and China, which hold the current dominant position across the globe in AI expertise.
- Global dialogs on the best AI policies should be established between countries to protect and maintain human rights.

Conclusion

The research by World Economic Forum (WEF) prognosticates that by the year 2025, machines will carry out longer current work tasks than humans, compared to 71% being performed by humans today [53]. And as per our estimations done above (Sect. 3.6), 98.2% of work would be done by machines in 2027. This is on the same lines as Ray Kurzweil, John von Neumann, Nick Bostrom, Vernor Vinge, Stephen Hawking, Elon Musk, and Bill Gates predicted and echoed [1, 3, 4, 9, 16].

We also suggested few policies considering AI ethics, which is currently in the discussion phase. But this should be prioritized and would become crucial in the near future. We should learn from our mistakes and be proactive to avoid repeating human complacency concerning climate change.

We mentioned few wonders of AI in this paper, and they are not an exhaustive list. And this is not the end of the road for this era. DeepMind, for example, is pushing the AI boundaries to further dimensions. It may now be possible to have a machine that can solve any knotty problem without training how or one that can correct or renew itself. It would be one of the prime and paramount scientific advances ever invented when this comes to achievement. The AI Avenue is still wide open for future improvement [56]. And, machine takeover will not be long enough, leading to AI apocalypse.

References

1. Kurzweil R (2005) The singularity is near: when humans transcend biology. Penguin
2. Time travel (2021) Wikipedia
3. Elon Musk claims AI will overtake humans "in less than five years" (2020) In: The independent. https://www.independent.co.uk/life-style/gadgets-and-tech/news/elon-musk-artificial-intelligence-ai-singularity-a9640196.html. Accessed 19 Aug 2021
4. Bostrom N (1998) How long before superintelligence? Int J Futur Stud 2
5. Moore GE (1998) Cramming more components onto integrated circuits. IEEE J Mag|IEEE Xplore. In: Proceedings of the IEEE vol 86, No. 1, pp 82–85
6. DeepMind—What if solving one problem could unlock solutions to thousands more? In: Deepmind. https://deepmind.com/. Accessed 19 Aug 2021
7. IBM Watson (2021) https://www.ibm.com/in-en/watson. Accessed 19 Aug 2021
8. Robotics Integrator Powered by AI—Vicarious. https://www.vicarious.com/. Accessed 19 Aug 2021
9. Vinge V (1993) Vinge V. Technological singularity. In: VISION-21 symposium sponsored by NASA Lewis research center and the Ohio Aerospace Institute, pp 30–31
10. Yampolskiy RV (2015) Analysis of types of self-improving software. In: Bieger J, Goertzel B, Potapov A (eds) Artificial general intelligence. Springer International Publishing, Cham, pp 384–393
11. Starlink. In: Starlink. https://www.starlink.com. Accessed 19 Aug 2021
12. Facebook Connectivity. https://www.facebook.com/connectivity/. Accessed 19 Aug 2021
13. Homepage. In: OneWeb. https://oneweb.net/. Accessed 19 Aug 2021
14. Diam PH, IS, MD (2015) The World in 2025: 8 predictions for the next 10 years. In: Singul. Hub. https://singularityhub.com/2015/05/11/the-world-in-2025-8-predictions-for-the-next-10-years/. Accessed 19 Aug 2021
15. Earman J, Wüthrich C, Manchak JB (2004) Time machines
16. Shanahan M (2015) The technological singularity. In: The technological singularity. MIT Press, p 233
17. Akl SG (2009) Time travel: a new hypercomputational paradigm
18. Rekimoto J (1999) Time-machine computing: a time-centric approach for the information environment. In: Proceedings of the 12th annual ACM symposium on User interface software and technology. Association for Computing Machinery, New York, NY, USA, pp 45–54

19. King ST, Dunlap GW, Chen PM (2005) Debugging operating systems with time-traveling virtual machines. In: Proceedings of the annual conference on USENIX annual technical conference. USENIX Association, USA, p 1
20. Ta-Shma P, Laden G, Ben-Yehuda M, Factor M (2008) Virtual machine time travel using continuous data protection and checkpointing. ACM SIGOPS Oper Syst Rev 42:127–134.https://doi.org/10.1145/1341312.1341341
21. Whitaker A, Cox RS, Gribble SD (2004) Using time travel to diagnose computer problems. In: Proceedings of the 11th workshop on ACM SIGOPS European workshop. Association for Computing Machinery, New York, NY, USA, pp 16-es
22. Barr ET, Marron M (2014) Tardis: affordable time-travel debugging in managed runtimes. ACM SIGPLAN Not 49:67–82. https://doi.org/10.1145/2714064.2660209
23. Agrawal H, Demillo RA, Spafford EH (1993) Debugging with dynamic slicing and backtracking. Softw Pract Exp 23:589–616. https://doi.org/10.1002/spe.4380230603
24. Lewis B (2003) Debugging backwards in time. arXiv:cs/0310016
25. Pothier G, Tanter É, Piquer J (2007) Scalable omniscient debugging. ACM SIGPLAN Not 42:535–552. https://doi.org/10.1145/1297105.1297067
26. Ko AJ, Myers BA (2008) Debugging reinvented: asking and answering why and why not questions about program behavior. In: Proceedings of the 30th international conference on Software engineering. Association for Computing Machinery, New York, NY, USA, pp 301–310
27. Xu M, Malyugin V, Sheldon J et al (2007) Retrace: collecting execution trace with virtual machine deterministic replay. In: Proceedings of the 3rd annual workshop on modeling, benchmarking and simulation, MoBS
28. Khoo YP, Foster JS, Hicks M (2013) Expositor: scriptable time-travel debugging with first-class traces. In: Proceedings of the 2013 international conference on software engineering. IEEE Press, San Francisco, CA, USA, pp 352–361
29. Feldman SI, Brown CB (1988) IGOR: a system for program debugging via reversible execution. ACM SIGPLAN Not 24:112–123. https://doi.org/10.1145/69215.69226
30. Mickens J, Elson J, Howell J (2010) Mugshot: deterministic capture and replay for Javascript applications. In: Proceedings of the 7th USENIX conference on networked systems design and implementation. USENIX Association, USA, p 11
31. Burg B, Bailey R, Ko AJ, Ernst MD (2013) Interactive record/replay for web application debugging. In: Proceedings of the 26th annual ACM symposium on User interface software and technology. Association for Computing Machinery, New York, NY, USA, pp 473–484
32. Caballero J, Grieco G, Marron M, Nappa A (2012) Undangle: early detection of dangling pointers in use-after-free and double-free vulnerabilities. In: Proceedings of the 2012 international symposium on software testing and analysis. Association for Computing Machinery, New York, NY, USA, pp 133–143
33. Sadekar K, Jiang H (2018) Time travel based feature generation
34. Blog NT (2017) Distributed time travel for feature generation. In: Medium. https://netflixtechblog.com/distributed-time-travel-for-feature-generation-389cccdd3907. Accessed 19 Aug 2021
35. Fabola A, Kennedy S, Miller A et al (2017) A virtual museum installation for time travel. In: Beck D, Allison C, Morgado L et al (eds) Immersive learning research network. Springer International Publishing, Cham, pp 255–270
36. Ch'ng E (2009) Experiential archaeology: Is virtual time travel possible? J Cult Herit 10:458–470.https://doi.org/10.1016/j.culher.2009.02.001
37. Tavakkol S, Chiang Y-Y, Waters T et al (2019) Kartta Labs: unrendering historical maps. In: Proceedings of the 3rd ACM SIGSPATIAL international workshop on AI for geographic knowledge discovery. Association for Computing Machinery, New York, NY, USA, pp 48–51
38. Herbst I, Braun A-K, McCall R, Broll W (2008) TimeWarp: interactive time travel with a mobile mixed reality game. In: Proceedings of the 10th international conference on Human computer interaction with mobile devices and services. Association for Computing Machinery, New York, NY, USA, pp 235–244

39. Maier A (2020) Will machine learning and virtual reality enable time travel? In: Medium. https://towardsdatascience.com/will-machine-learning-and-virtual-reality-enable-time-travel-b66eb1c848ca. Accessed 19 Aug 2021
40. Kaplan F (2015) The Venice time machine. In: Proceedings of the 2015 ACM symposium on document engineering. association for computing machinery, New York, NY, USA, p 73
41. Time Machine Europe. https://www.timemachine.eu/. Accessed 19 Aug 2021
42. Moore's law (2021) Wikipedia
43. Kumar S (2015) Fundamental limits to Moore's Law. ArXiv151105956 Cond-Mat
44. Roser M, Ritchie H (2013) Technological progress. Our World Data
45. Satell G 3 Reasons To Believe The Singularity Is Near. In: Forbes. https://www.forbes.com/sites/gregsatell/2016/06/03/3-reasons-to-believe-the-singularity-is-near/. Accessed 8 Sep 2021
46. (2018) What the GlobalFoundries' Retreat Really Means. In: IEEE Spectr. https://spectrum.ieee.org/what-globalfoundries-retreat-really-means. Accessed 8 Sep 2021
47. AI and Compute (2018) In: OpenAI. https://openai.com/blog/ai-and-compute/. Accessed 8 Sep 2021
48. AI Index 2019. In: Stanf. HAI. https://hai.stanford.edu/ai-index-2019. Accessed 8 Sep 2021
49. Jumper J, Evans R, Pritzel A et al (2021) Highly accurate protein structure prediction with AlphaFold. Nature 596:583–589. https://doi.org/10.1038/s41586-021-03819-2
50. Brown TB, Mann B, Ryder N et al (2020) Language models are few-shot learners. ArXiv200514165 Cs
51. Zhang D, Mishra S, Brynjolfsson E et al (2021) The AI index 2021 annual report. ArXiv210306312 Cs
52. Artificial Intelligence market size/revenue comparisons 2021. In: Statista. https://www.statista.com/statistics/941835/artificial-intelligence-market-size-revenue-comparisons/. Accessed 19 Aug 2021
53. Machines Will Do More Tasks Than Humans by 2025 but Robot Revolution Will Still Create 58 Million Net New Jobs in Next Five Years. In: World Econ. Forum. https://www.weforum.org/press/2018/09/machines-will-do-more-tasks-than-humans-by-2025-but-robot-revolution-will-still-create-58-million-net-new-jobs-in-next-five-years/. Accessed 19 Aug 2021
54. Müller VC, Bostrom N (2016) Future progress in artificial intelligence: a survey of expert opinion. In: Müller VC (ed) Fundamental issues of artificial intelligence. Springer International Publishing, Cham, pp 555–572
55. Moor J (2006) The Dartmouth College artificial intelligence conference: the next fifty years. AI Mag 27:87–87. https://doi.org/10.1609/aimag.v27i4.1911
56. Would machines overtake humans? (2020) In: Jordan Times. http://www.jordantimes.com/opinion/talal-abu-ghazaleh/would-machines-overtake-humans. Accessed 19 Aug 2021

Secured Home Automation with Encrypted Fast Fourier Transformation

Venkatarao Dadi, Naresh Pathakamuri, Satya Venu Abhishek Kolluri, K. Suryadeep, Satya Tarun Gurugubelli, and Prabhu Teja Mandala

Abstract Home automation is one of the significant developing businesses in controlling home appliances as it simplifies the way of individual's life. Automation is particularly helpful to the people who are unable to move like old or crippled one. Wireless operating modes of automation include HC-05 Bluetooth module for the transmission of voice signals and remotely operated infrared (IR) module for the transmission of non-voice signals. The key feature of home automation system for controlling the home appliances wirelessly is to recognize the secured voice signals decoded in microcontroller. Encrypted voice commands (EVCs) are implemented with Fast Fourier Transform (FFT) and used in control unit for authentication of user. Finally, encrypted real part of the voice string is compared with the existing EVC string to control the corresponding loads in the home. Graphical and virtual simulation of controlling of loads for voice signals is analyzed using the Proteus 8 Professional software. Status of wireless controlled appliances is stored in Excel using parallax data acquisition (PLXDAQ) tool. This work is also validated with the real-time design.

Keywords Home automation · Voice signal encryption · Graphical simulation · Fast Fourier Transform · PLXDAQ

V. Dadi · N. Pathakamuri (✉) · S. V. A. Kolluri · S. T. Gurugubelli · P. T. Mandala
Department of Instrument Technology, A.U.C.E(A), Andhra University, Visakhapatnam 530003, India

K. Suryadeep
Department of Electronics and Communication Engineering, Malla Reddy Engineering College and Management Sciences, Medchal 501401, India

N. Pathakamuri
Department of Electrical and Electronics Engineering, Raghu Engineering College(A), Visakhapatnam 531162, India

V. Bindhu et al. (eds.), *Proceedings of Third International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 844,
https://doi.org/10.1007/978-981-16-8862-1_58

Introduction

Wireless voice and non-voice activated home automation are a single portable unit which allows controlling appliances or equipment, etc. Appliances are switched to ON or OFF by a programmable control unit. Appliances or equipment are interfaced to control unit and communicated with each other over a local wired or wireless network. Home automation plays a major part in creation of smart homes for efficient control of appliances from anywhere. Waleed [1] designed an automation system that controls home appliances remotely by using an Android smart phone and monitored the status of appliances using Wireless Fidelity (Wi-Fi) to reduce the energy consumption.

Yue [2] controlled the home appliances by interconnecting with wireless devices such as Wi-Fi, Bluetooth, IPv6 low-power wireless personal area network (6LoWPAN), radio frequency identification (RFID), and ZigBee. Basanta [3] controls the devices in a smart home with voice and gestures. Khan [4] controlled the home apparatus using voice recognizer (in Android cell phone), Bluetooth module and Arduino UNO micro controller for the handling of client orders and control the exchanging information among the gadgets. Tejesh [5] implemented an Internet of things (IoT)-based smart system with sensors and control unit for monitored and controlling the appliances remotely using Wi-Fi. In the transmitter section, the information of appliances is gathered at the central server of Raspberry Pi which provides graphical interface to user.

Asadullah [6] designed home automation system using Arduino microcontroller control unit and smart phone. The system interconnected with Bluetooth, sensors like moisture sensor and ultrasonic sensor. An ultrasonic sensor is used for water level detection, and soil moisture sensor is used for automatic plant irrigation system. The system allows the users to control up to 18 devices or appliances. Palaniappan [7] designed an interactive Python program used in the Arduino prototype controller communicating via Bluetooth module HC-05 and Bluetooth in smart phone. Vinay Sagar [8] implemented a multi-touch mobile device for smart home automation system with cloud networking to control the home appliances. With mobile application, the consumer is interfaced to personal computer and handheld the automation system remotely for eldercare. Ramlee [9] implemented a smart home system accessed by using Bluetooth wireless connectivity.

Sen [10] controlled home appliances using an Arduino microcontroller, an Android smart phone, Bluetooth module, and relays. Android voice recognizer converts the speech to text service (STT), and text command is transmitted to the serial ports of Arduino UNO via Bluetooth connectivity. Based on the text commands received by control unit, it will switch ON the relay of particular appliance. Gayatri [11] designed a home automation system by storing 80 voice samples with no group of instruction. Also, discussed that if any problem with voice recognition occurs in the system, then the appliances are controlled by using the keypad. Humaid [12] designed a voice recognition-based home automation system and uses low-power RF ZigBee modules for wireless data transmission. Spectrum of finger print data is computed using FFT in automation system. Nombulelo [13] presented an IoT-based speech-controlled home automation system using Google Assistant. Home appliances are controlled using speech recognition and observed the nature of speech recognition by operating the appliances in different situations of quiet, empty, noisy, and fully furnished room environments of a particular room size and distance from transmitter to receiver section. Pandya et al. [14] proposed a system that contains microphone, speech processing computer and microcontroller. Voice signal is extracted using Mel Frequency Cepstrum Coefficients and vector quantization then it is matched with the stored speech in computer to control the particular appliance.

Liu [15] discussed uncovered security issues in the case of home automation in the Belkin WeMo ecosystems. Implemented a logic on WeMo devices in reverse engineering process and showed that eavesdropping on user traffic when appliances are controlled by mobile application. Also explains the misleading process of a user to disclose his private information with cross-site scripting using a fake device to a WeMo user. Seo et al. [16] implemented ZigBee wireless network for home automation with improved safety exchanged messages and appropriate security level. Messages are exchanged based on attribute-based proxy re-encryption (ABPRE) technology. Robert [17] discussed a system to control home appliances with authentication and data encryption based on JSON Web token system.

From the literature of voice-based home automation, authors did not discuss graphical simulation analysis to control the appliances using voice signals. In this paper, wireless voice-controlled home automation is implemented in real time along with graphical and virtual simulation. Two level security is provided with customized Bluetooth (password protected) and encrypted FFT algorithm for voice authentication is implemented in controller. The Arduino ATmega 2560P controller with four channel relay driver module, customized Bluetooth, IR TSOP sensor, Manual one-way switches and a smart phone application and remote are used in this work.

Methodology

Encryption of voice signals using Fast Fourier Transformation (FFT):

Microcontroller receives the voice signal processed with FFT encryption in frequency domain complex form. Voice signal in time domain is converted into voice frequency domain using FFT. Performance of FFT [18] is fast with numerical precision, better in the low-quality audio at low bit rates, narrow bandwidth and reduces the quantization errors of voice signal. Initially, EVC string for various voice commands of an authenticated user is implemented with FFT and that data are stored in control unit.

Received voice signal string is converted into encrypted voice string by the controller using FFT calculation. Flow of virtual simulation is shown in Fig. 1. The controller processes the conversion process of voice signal into frequency domain at sampling frequency (f_s) of 5000 Hz using FFT calculations. Encrypted voice string

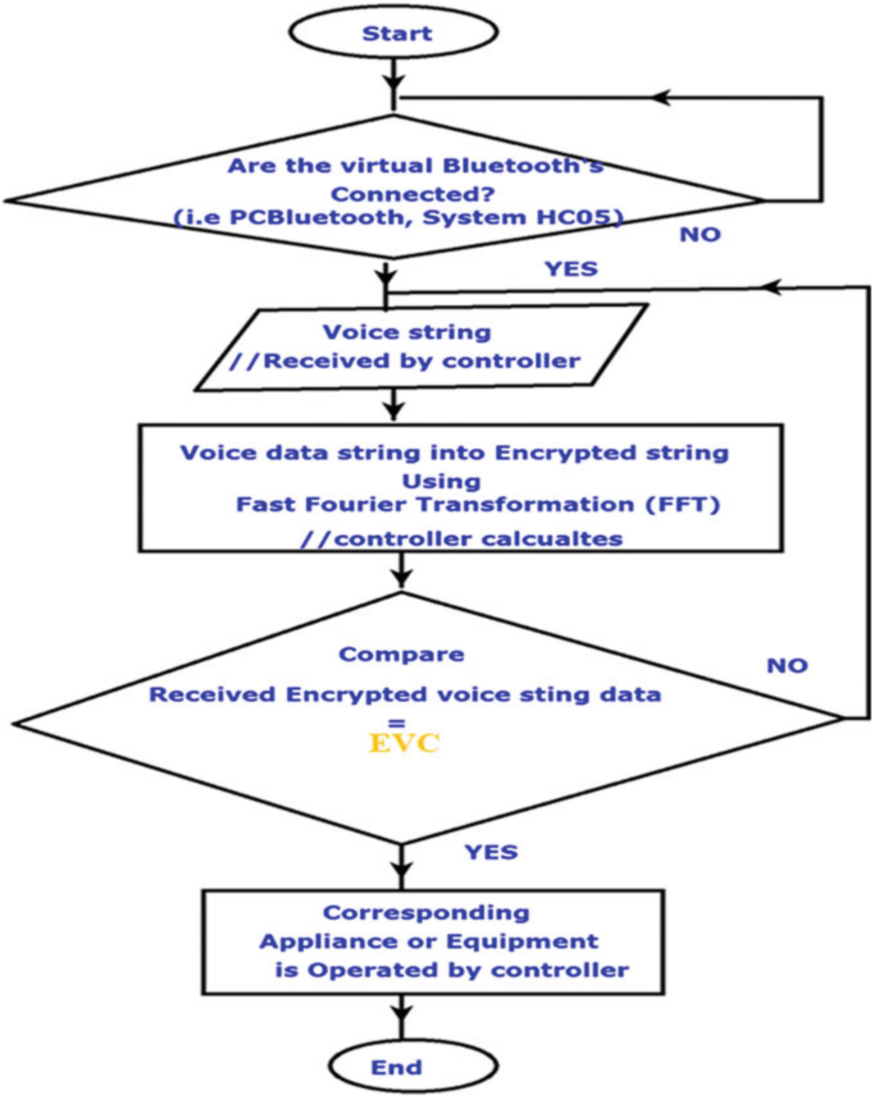


Fig. 1 Flowchart for controlling the appliances using FFT

data are compared with the stored EVC string, if it is matches than the corresponding appliance is controlled by the controller.

Experimental Setup

An Arduino Mega 2560, a Bluetooth HC-05 module, 4-channel relay driver, and bulbs L1, L2, F1, F2 are (are exemplary names taken in the place of home appliances like light1, light2, fan1, fan2 etc.) interconnected as shown in Fig. 6. A 12 V 1A DC power adapter is used as power supply to controller. The mobile Bluetooth is paired with the customized HC-05 Bluetooth for wireless connectivity. Customized HC-05 Bluetooth means user can set the password himself. First level of user authentication is provided by customized password of system HC-05 Bluetooth (Figs. 2, 3, 4 and 5).

In the second level of security, controller converts the received voice signal into encrypted voice string using FFT and compared with the stored EVC strings, if they are matched then controller controls the corresponding appliance. TSOP IR sensor with remote is used for non-voice controlling of loads, where IR signals are decoded by controller. Manual controlling of appliances during failure of system or voice connectivity then a non-voice connectivity (Remotely or manual) is added in the real-time circuit design. Voice commands are designed using the MIT application and transferred from Bluetooth to HC-05 for controlling of appliances. Application is designed using various building blocks known as list pickers as shown in Fig. 2.

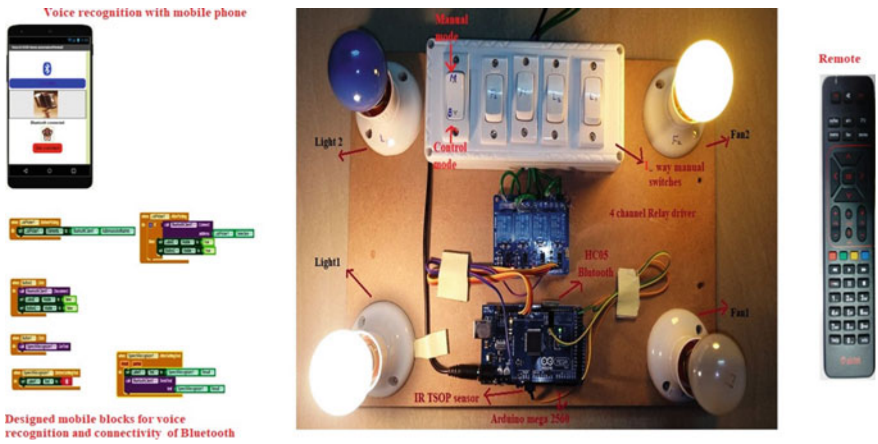


Fig. 2 Real-time implemented design

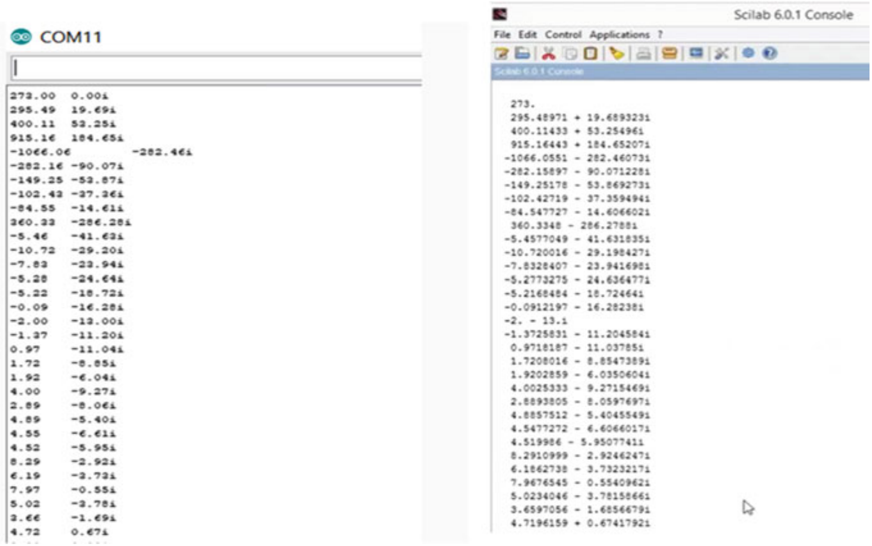


Fig. 3 Verification of designed program of FFT in Arduino IDE with calculate FFT in Scilab

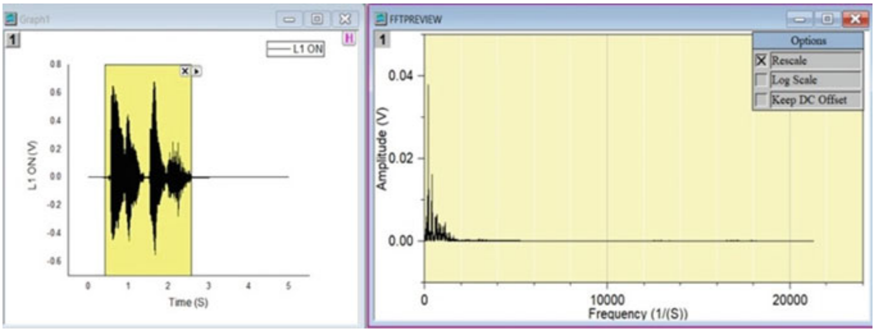
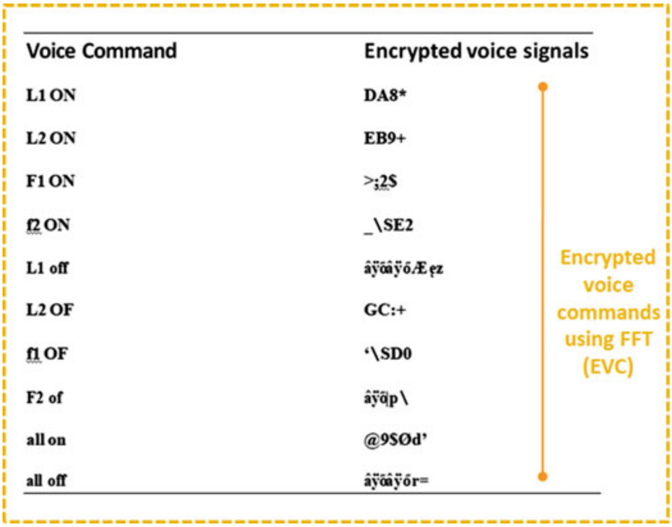


Fig. 4 “L1 ON” voice signal processing using FFT in Origin pro

Results and Discussion

Testing of FFT Encryption Designed in Arduino IDE with Calculated FFT in SCI Lab

Designed programming code for FFT calculated up to two floating points in Arduino IDE which is monitored on serial monitor and is similar to numerically calculated FFT in Scilab for the same input sequence as shown Fig. 3.



Voice Command	Encrypted voice signals
L1 ON	DA8*
L2 ON	EB9+
F1 ON	>;2S
f2 ON	_\SE2
L1 off	âÿöâÿõÆęz
L2 OF	GC:+
f1 OF	‘\SD0
F2 of	âÿöp\
all on	@9SØd’
all off	âÿöâÿõr=

Fig. 5 Voice commands and corresponding EVC strings

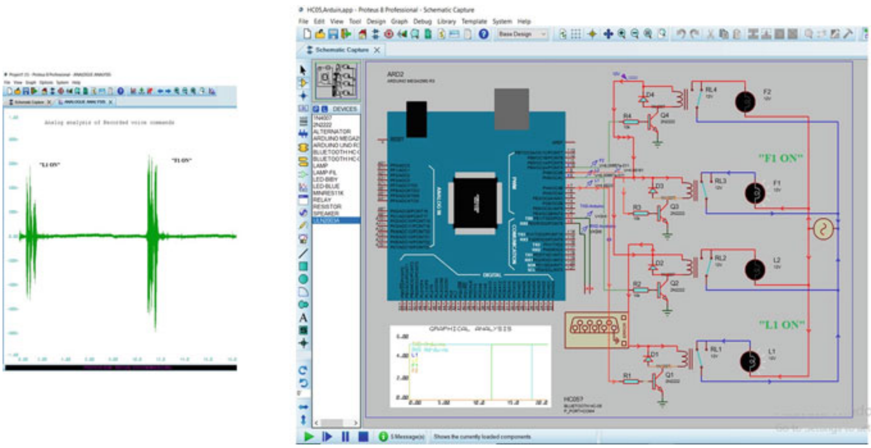


Fig. 6 Virtual simulation for voice signals “L1 ON,” “F1 ON”

FFT Analysis for Voice Signal

A pre-recorded voice signal “L1 ON” with respect to time interval is taken for FFT signal processing analysis. “L1 ON” is a voice signal to turn on light L1 which is given as input signal to mobile voice recognizer in mobile application. This string data are converted into FFT. Voice signal simulation is analyzed using FFT signal processing technique in Origin Pro software as shown in Fig. 3. From the figure sampling frequency (f_s) of 5000 Hz is taken for FFT calculation in real-time Arduino code.

EVC strings implemented using FFT for various voice signals of a user as shown in Fig. 5. It is implemented in Arduino IDE using FFT with f_s 5000 Hz which is observed on Arduino serial monitor at baud rate of 115,200. Array size of four is taken from real part of encrypted FFT of each voice signal as predefined voice signal to control corresponding load for a voice signal of specified person. Observed that each voice string instruction is not same even they contain some part of common letters. So that specified person voice signals are only decoded, extracted, and compared with predefined stored encrypted FFT string of that particular voice levels. This predefined encrypted voice string is the main parameter to control the appliances, and voice signal of unauthorized persons is not recognized to control the appliances.

Virtual Simulation of Wireless Voice-Controlled Appliance

Virtual devices like an Arduino UNO R3, a HC 05 Bluetooth module, four channel relay driver module and virtual lamps L1, L2, F1, and F2 are connected in Proteus 8.6 SP2 Professional application. For analog graphical simulation, voltage probes are connected to digital output ports of controller unit.

Voice instruction "L1 ON" is given as input, and virtual simulation is observed as shown in Fig. 4. Voice signal "L1 ON" has wave amplitude exists in the time interval of 600 ms to 2 s. This specific voice instruction is given through mobile application and mobile voice recognizer recognized the voice and converted the voice into string, and data are transmitted through its synchronized mobile Bluetooth to the system Bluetooth HC-05. Transmitting data (TXD) port and receiving data (RXD) ports of Arduino get enabled and acknowledged to HC-05 module when it receives the signal. FFT algorithm is implemented on voice string "L1 ON" and compared with the predefined stored voice encrypted data of "DA8*." When the encrypted strings are equal, then the corresponding digital input relay driver is switched to high. The relay is now switched from "off state" to "on state" and its corresponding light L1 is switched to ON.

The schematic diagrams are given for the instructions "L1 ON" and "F1 ON," respectively, as shown in Fig. 6. Similarly, we can follow the same procedure for the appliances L2 and F2. As the encrypted voice signal strings are differed from "L1 ON" to "F1 ON," so observed virtual simulations are also differed when compared to each other. For voice signal "F1 ON" encrypted predefined voice string is ">;2$." This received encrypted voice instruction is compared with predefined stored encrypted voice by controller. When the encrypted strings are equal then the corresponding light (i.e., F1) is activated to ON. "F1 ON" instruction is observed in the time interval of 580 ms to 2 s. Generally, voice signal of specified person depends on amplitude (i.e., pitch) and space between discrete levels of voice signal.

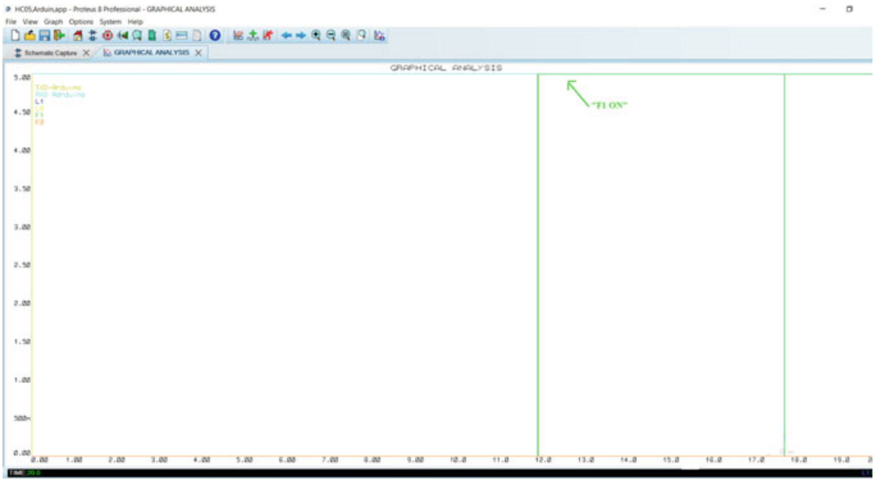


Fig. 7 Graphical simulation for "F1 ON" voice command

Graphical Simulation of Wireless Voice-Controlled Appliance

The analog graphical simulation analysis of voltage values at digital output ports of Arduino with respect to corresponding voice instruction as shown in Fig. 7. Analog analysis is observed by connecting the voltage probes at digital out ports, TXD, and RXD port of Arduino micro controller.

The analog graphical simulation is analyzed for two different voice instructions like "L1 ON" and "F1 ON" which are encrypted by FFT in Arduino controller. Encrypted voice string is compared with predefined stored voice string. TXD, RXD, L1, and F1 voltage probes are denoted with red, white, blue, and pink lines, respectively. When the voice commands are given, the graph for "F1 ON" is observed at 12 ms with enabled TXD, RXD lines.

Appliances Status Monitoring

Wirelessly controlled appliances (L1, L2, F1, and F2) data are monitored and stored in Excel sheet using parallax data acquisition (PLXDAQ) as shown in Fig. 8. Status update of wirelessly controlled appliance is directly transferred via synchronized system Bluetooth and HC Bluetooth module and stored into Excel within the time interval of 2 ms (Table 1).

Fig. 8 Data stored and monitored in EXCEL using PLXDAQ tool

Table 1 Comparison of present work with previous voice-based automation systems

Refs.	Controlling	Voice command	Speech recognization algorithm/security	Graphical simulation
[1]	Wireless voice	Mobile application		No
[2]	Wireless voice	Mobile application	Artificial intelligence (Alexa)/one level security	No
[11]	Wireless voice, physical key pad module	Mobile application		No
[14]	Wireless voice	Mobile application	Mel Frequency Cepstrum coefficients/one level security	No
[17]	Wireless virtual keys	Virtual mobile application keys	JSON Web token (JWT) and advanced encryption standard (AES)/one level security	No
Present	Wireless, wired	Mobile application	Encrypted voice commands (EVC) using FFT/two level security	Yes

Conclusion

Wireless secured voice and non-voice controlling of home automation are one of the emerging technologies. It reduces the human effort to a great extent. The rapid growth of present technology motivates us to use smartphone to control the home appliances wirelessly. Secured control of home appliances with the voice authentication plays the key role. Encryption of voice string with real part of FFT in Arduino programming helps to control the appliances with the specified person voice only. This system not only reduces human effort, but also provides secured automation. AI-based voice recognition with FFT is implemented in the futuristic approach.

References

1. Waleed J, Abduldaim AM et al (2018) Smart home as a new trend, a simplicity led to revolution. In: 3rd international scientific conference of engineering science (ISCES), IEEE. https://doi.org/10.1109/ISCES.2018.8340523
2. Yue CZ (2017) Voice activated smart home design and implementation. In: 2nd international conference on frontiers of sensors technologies (ICFST), IEEE, pp 489–492. https://doi.org/10.1109/ICFST.2017.8210563
3. Basanta H et al (2017) Assistive design for elderly living ambient using voice and gesture recognition system. In: IEEE international conference on systems, man, and cybernetics (SMC), IEEE, pp 840–845. https://doi.org/10.1109/SMC.2017.8122714
4. Khan IU et al (2017) Voice controlled home automation system. Int J Res Comput Commun Technol T 6(5). https://www.researchgate.net/publication/317386416_Voice_Controlled_Home_Automation_System
5. Tejesh BS, Neeraja S (2018) A smart home automation system using IoT and open source hardware. Int J Eng Technol 7(2.7):428–432. https://doi.org/10.14419/ijet.v7i2.7.10856
6. Asadullah M et al (2017) Smart home automation system using Bluetooth technology. In: International conference on innovations in electrical engineering and computational technologies (ICIEECT), IEEE, pp 1–6. https://doi.org/10.1109/ICIEECT.2017.7916544
7. Suseelan AD et al (2015) Home automation systems—a study. Int J Comput Appl 116(11):11–18. https://doi.org/10.5120/20379-2601
8. Vinay Sagar KN, Kusuma SM et al (2015) Home automation using internet of things. Int Res J Eng Technol (IRJET).https://www.academia.edu/14381897/IRJET_Home_Automation_Using_Internet_of_Things
9. Ramlee RA et al (2013) Smart home system using android application. In: International conference of information and communication technology (ICoICT), IEEE, , pp 277–280. https://doi.org/10.1109/ICoICT.2013.6574587
10. Sen S et al (2015) Design of an intelligent voice controlled home automation system. Int J Comput Appl 121(15):39–42. https://doi.org/10.5120/21619-4904
11. Shinde GR et al (2015) Voice and non-voice control based wireless home automation system. Int J Innov Res Sci Eng Technol 04(07):5063–5067. https://doi.org/10.15680/ijirset.2015.0407005
12. AlShu'eili H et al (2011) Voice recognition based wireless home automation system. IEEE, pp 1–6. https://doi.org/10.1109/ICOM.2011.5937116
13. Noruwana N et al (2020) Interactive IoT-based speech-controlled home automation system. In: 2nd international multidisciplinary information technology and engineering conference (IMITEC), IEEE, 2020, pp 1–8. https://doi.org/10.1109/IMITEC50163.2020.9334081
14. Pandya S et al, Interactive home automation system using speech recognition. Int J Eng Res Technol (IJERT) 3(04):2295–2299. https://www.ijert.org/research/interactive-home-automation-system-using-speech-recognition-IJERTV3IS042264.pdf
15. Liu H et al (2019)Uncovering security vulnerabilities in the Belkin WeMo home automation ecosystem. In: IEEE international conference on pervasive computing and communications workshops (PerCom Workshops), IEEE, pp 894–899.https://doi.org/10.1109/PERCOMW.2019.8730685
16. Seo H et al (2011) Zigbee security for visitors in home automation using attribute-based proxy re-encryption. In: 15th international symposium on consumer electronics (ISCE), IEEE, pp 304–307.https://doi.org/10.1109/ISCE.2011.5973837

17. Sowah RA, Boahene DE et al (2020) Design of a secure wireless home automation system with an open home automation bus (OpenHAB 2) framework. J Sens, 1–22. https://doi.org/10.1155/2020/8868602
18. Makwana V et al (2013) Analysis of performance of fast Fourier transformation of an audio file. Int J Appl Innov Eng Manage (IJAIEM) 2(11)

Retrospective Review of Activation Functions in Artificial Neural Networks

Manjunatha Badiger and Jose Alex Mathew

Abstract Deep learning is a subfield of machine learning and artificial intelligence technique. It employs neural network tasks like image processing, computer vision, voice recognition, machine translation, medical information processing, self-driving vehicles, predictive forecasting, robotics and control, cybersecurity, natural language processing, bioinformatics, and countless others. The performance of a neural network is determined by a variety of factors, and activation functions are an essential element in the design of a neural network. The hidden layer's activation feature defines the extent to which the network model learns the training data set. The type of prognostications made by the model is regulated by the activation functions present at the output layer. This paper presents a comprehensive review of research studies on different activation functions aimed toward deep learning applications.

Keywords Deep learning · Neural network · Machine learning · Artificial intelligence · Activation function

Introduction

Deep learning is a form of hierarchical learning. It comprises algorithms and topologies to solve a wide variety of issues. Deep learning is a feature-learning method with many levels of representation. Thus, it will be easy to understand nonlinear representations one layer at a time [1]. The lower-level features are minor details that are used to transform the representation at one level to build high-level features on top of it [2]. Thus, the complex functions can be learned using such transformations. Over the last few decades, deep learning has become a very popular and most powerful tool as it can handle a huge quantity of data. Deep learning architectures, which are outstandingly, expanded the number and types of problems that neural networks can

M. Badiger (✉)
Sahyadri College of Engineering & Management, Adyar, India

J. A. Mathew
Srinivas Institute of Technology, Mangalore, India

V. Bindhu et al. (eds.), *Proceedings of Third International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 844,
https://doi.org/10.1007/978-981-16-8862-1_59

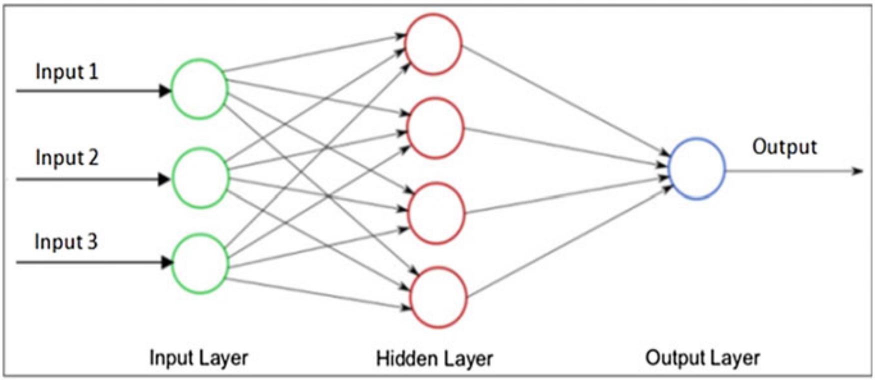


Fig. 1 Simple neural network

handle. Deep learning has shown several advancements by researchers and academicians in the last two decades [3]. Neural networks are a novel architecture imitating biological neural networks [4]. The rudimentary building block of artificial neural networks is a neuron. A neuron is a mathematical function that simulates the functioning of a biological neuron [5]. Neural networks are composite computer codes written with many basic, highly interconnected computing components that mimic human biological brain structure operations to simulate data models of human brain functioning and processing. Figure 1 shows the simple neural network. The neural network comprises three layers, namely activation function, learning technique, and weights. All these layers include neurons that are interlinked to form a network [6]. Its elements are validated depending on whether the neuron is used for input, output, or in one of the hidden layers.

The two key hyperparameters which control the architecture or topology of neural networks are the total number of layers and the number of total nodes in each hidden layer as shown in Fig. 2.

Input layer: It receives input either through an outside source or through other neighboring nodes. Every node is attached to another node of the succeeding layer. Each connection has a specific weight. Depending on its degree of importance, weights are assigned to a neuron in relation to other inputs. Once the entire node values from the input side are multiplied by their corresponding weights and totted up, the value for hidden layers is generated [6]. The output of the input layer can be given by the equation:

$$y_i = \sum(\text{weights} * \text{input} + \text{bias}) \tag{1}$$

It can range from −infinity to +infinity. So it is necessary to bound the output to get the desired prediction or generalized results.

Hidden layers: Hidden layers are always found between the input and output layers. It is always shrouded from the outside world. Hidden layers may vary from

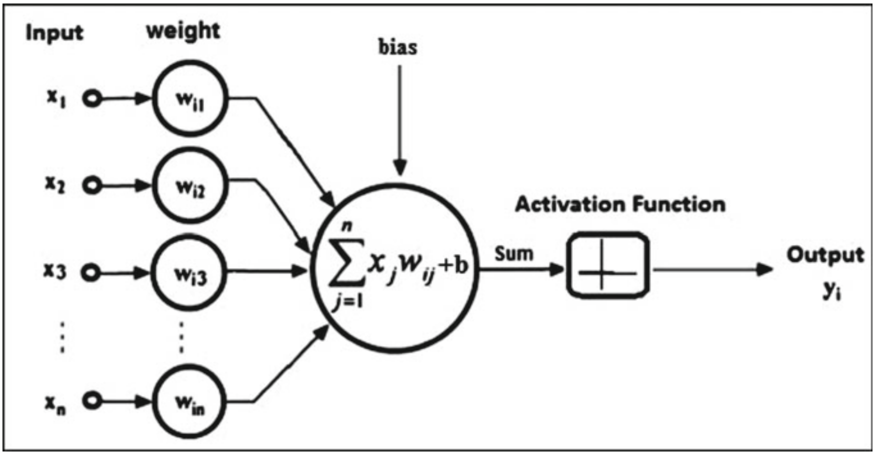


Fig. 2 Complex neural network

network to network that we selected. The number of hidden layers in a neural network is determined by the problem's nature and size. Once the hidden layer gets information from the input side, it performs all the computational tasks and provides the result [6, 7]. This result is then forwarded to the output layer. Hidden layers refine the input weightings until the marginal error of the neural network is small.

Output layer: The output nodes are known collectively as the output layer, and they are responsible for bringing out the final result. This output layer is designed differently to contour and improve the final results of the iterative task [8]. The output layer acquires the input from the hidden layers and uses its neurons to complete the computations, after which the output is generated.

Activation Function

In neural networks, activation functions are mathematical functions used to represent each neuron present in the network. The activation function of a neuron decides whether it should be turned on or turned off depending on the input or set of input values [9]. Activation functions facilitate normalizing the output of all the neurons and map them into a range within 1 and 0 or −1 and 1.

The activation functions can be classified mainly into the following types:

1. Identity or linear activation function
2. Nonlinear activation functions.

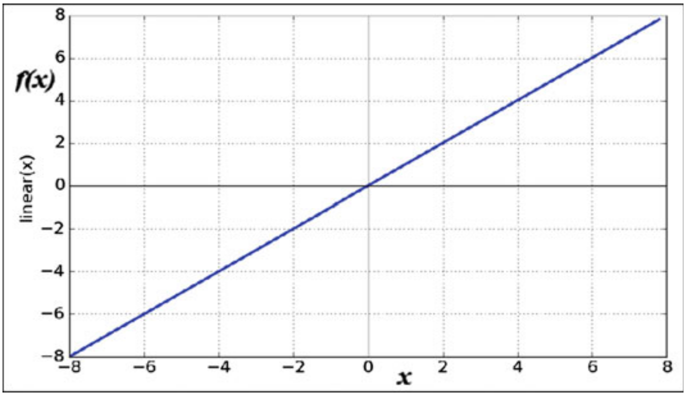


Fig. 3 Linear function

Identity or Linear Activation Function

The input values multiplied by their corresponding weights from each neuron are given as input to an activation function. The activation function produces an output corresponding to the input values [10]. A neural network is nothing but a linear regression model without a linear activation function. Linear activation function has confined power and also has limited capacity to solve the composite input data. The equation of a linear function is alike to that of a straight line, i.e., $f(x) = a * x$ and it ranges from $-\infty$ to ∞. The graphical representation is shown in Fig. 3.

When a linear activation function is employed in the multi-layered neural network, irrespective of how many layers are present in the network all the layers will be linear. Therefore, the final layer is purely a linear transformation of the final layer [11]. The linear function $f(x)$ has an invariable derivative and also it does not rely on the input value x. Thus, the linear function is unable to perform backpropagation every time to train the model. The gradient is pretty much the same hence it is not possible to improve the error.

Nonlinear Activation Function

Nonlinear activation functions can be used to represent any imaginable process as a computational function in the neural network [12]. They enable the model between the network inputs and outputs, to develop a composite mapping. These are essential to learn and model the nonlinear type of composite data such as images, video, audio, and data sets. Thus, to resolve the problems of a linear type of activation function, often nonlinear types of activation functions are used [13]. They have an input-related derivative function that permits backpropagation. Nonlinear functions often allow a deep neural network to be built up by stacking multiple hidden layers of neurons.

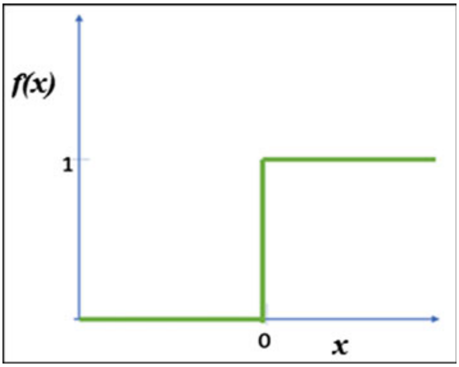


Fig. 4 Binary step activation function

Thus, it is possible to understand the composite data sets with high levels of precision using deep neural networks.

Binary Step Activation Function

The step function is among the most basic activation function available, which provides binary output [9]. That is why it is also called a binary step function. Here, we consider a threshold value, when the input passes the threshold limit the function produces a value 1(true) and then the neuron is activated. If the input does not pass the threshold value, the function produces a value 0(false) and then the neuron is deactivated. That is why they are very useful for binary classification studies. A graphical representation of the binary step activation function is shown in Fig. 4.

Mathematically, binary step activation function can be described as

$$f(x) = \left\{ \begin{matrix} 0, & \text{for } x < 0 \\ 1, & \text{for } x \geq 0 \end{matrix} \right\} \tag{2}$$

Sigmoid or Logistic Activation Function

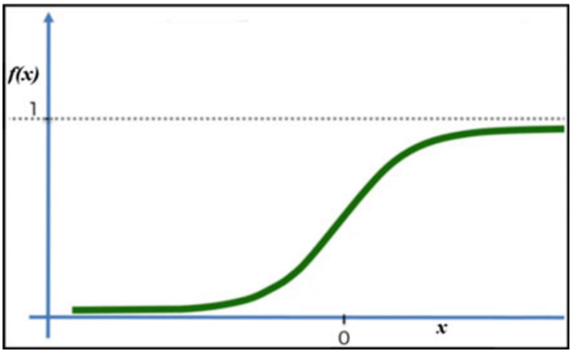
The sigmoid function resembles an "*S*" shaped curve. It can be used to represent the anticipated values to probabilities. The sigmoid function distributes the input values of any size to output values in the interval between 0 and 1, normalizing the output of each neuron [14]. Graphical representation sigmoid activation function is shown in the below fig. Here, the output is not zero-centered as shown in Fig. 5.

Mathematically, sigmoid activation function can be described as

$$f(x) = \frac{1}{(1 + \mathrm{e}^{-x})} \tag{3}$$

The above function is exclusively monotonic in its entire region and it is easily differentiable. However, its derivative is not monotonic. There is almost no ambiguity in the estimation of very low or very high variation in the values of x, which creates

Fig. 5 Logistic/sigmoid function

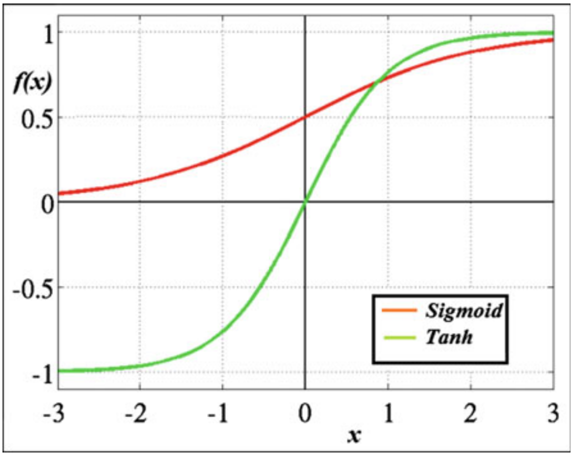


a problem of vanishing gradient. That leads to a situation where the network refuses to learn further or being too sluggish to succeed in an accurate prediction. During the testing period, the logistic sigmoid function may end up causing the neural network to become stuck.

Tan*h* Function

Tan*h* activation function works almost always better than the sigmoid function. It is a mathematically modified version of the sigmoid function that transforms input to output with values ranging from -1 to 1 [15]. The gradient is stronger for tan*h* than the sigmoid activation function. The major advantage of the tan*h* activation function is that its negative inputs are always represented as strongly negative; zero inputs are represented near to zero which is not the same for sigmoid function as the range for it is between 0 and 1 as shown in Fig. 6.

Fig. 6 Tan*h* function



The mean of the activations coming out of the hidden layer is closer to having a zero mean. Therefore, the data are more centered, making learning easier, and faster for the next layer. The function and its derivative are both monotonous.

Mathematically, tan*h* activation function can be described as

$$\tan h(x) = \frac{1 - \mathrm{e}^{-2x}}{1 + \mathrm{e}^{-2x}} \tag{4}$$

The main disadvantage of the tan*h* activation function is that its gradient will experience a very small value and even it may accomplish a near-zero value. This can thwart the gradient descent. The most difficult aspect of implementing these functions is that it demands the exponential term, which results in nonlinear behavior.

Arctan Function

The arctan function is similar to the sigmoid and tan*h* function and is obtained by the inverse of the tangent function. This activation function maps input to accelerating and decelerating output values ranging between $(-\pi/2, \pi/2)$. The arctan function graph is a slightly flattered *S*-shape compared to the tan*h* function, which provides better classification power [11]. Arctan function can be mathematically described as

$$f(x) = \tan^{-1}(x) \tag{5}$$

Figure 7 shows the graphical representation of arctan function.

For the larger values of the input, its derivative converges to zero. Contrarily, the derivative of the sigmoid activation function converges exponentially to zero.

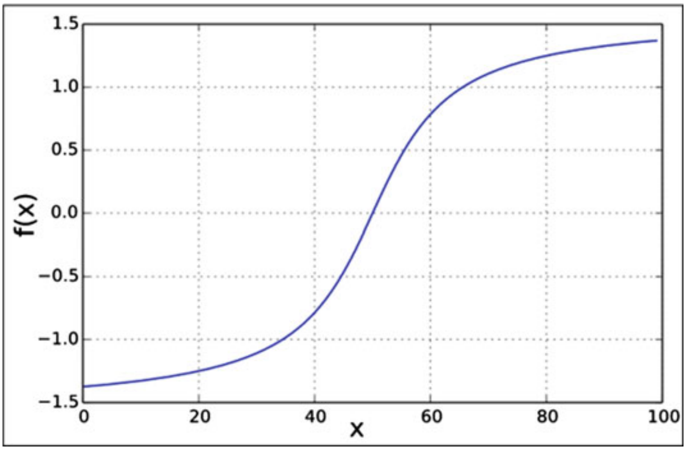


Fig. 7 Arctan function

Softmax Function

The Softmax function is another type of mathematical function that always produces an output ranging from 0 to 1 irrespective of whether the input values are positive, negative, zero, or greater than one. The cumulative of all the probabilities is equals to 1.Thus, the Softmax function is used to compute the normalized output probability distribution comprised of K probabilities from the input vector consisting of K real numbers. The formula for the Softmax functions can be specified as follows [16]:

$$\sigma(z_i) = \frac{e^{z_i}}{\sum_{j=1}^{K} e^{z_j}} \tag{6}$$

The above expression calculates the input exponential value as well as the sum of all input exponential values. The Softmax function's output is proportional to the exponential input value as well as the sum of the exponential values.

The graph shown in Fig. 8 shows the variation of output probabilities corresponding to the variation in the input values. To overcome the issues of multi-class classification, the Softmax function can be used as the activation function in the output layer of neural network models.

Softsign Function

Softsign mathematical function is again a different type of activation function employed in neural networks. It almost resembles the hyperbolic tangent activation function but the main difference between them is that, unlike the tan$h$ function which converges exponentially the softsign function converges in a polynomial form [17]. The value of softsign function is zero-centered and it ranges between -1 and $+1$, so the network learns effectively.

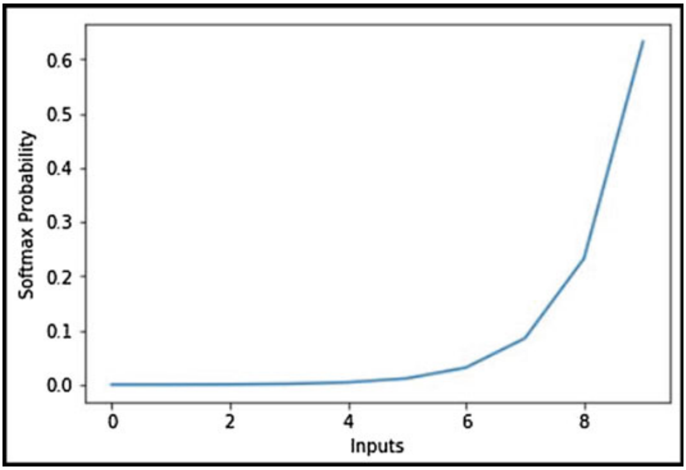


Fig. 8 Softmax function

Fig. 9 Softsign function

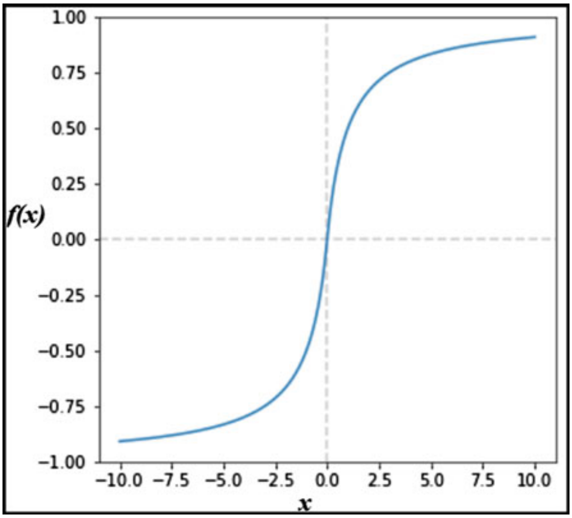


Figure 9 shows the graphical interpretation of the softsign function. The formula for softsign functions can be specified as follows:

$$f(x) = \frac{1}{(1 + |x|)} \tag{7}$$

Softsign activation function is characterized by a high degree of nonlinearization and good error tolerance. Thus, it can be used in neural networks to transform the input into nonlinear distribution. The main advantage of the softsign function is that its output is centered on zero and its asymptote lines are smoother [11]. Thus, the output saturation reaches steadily to 0 on both sides. This alleviates the problem of gradient vanishing to some degree.

Rectified Linear Unit (ReLU) Function

In DL models, the rectified linear unit (ReLU) is one of the most famous and oftenly utilized activation functions. This function conserves the characteristics of a linear function. It also prevents the vanishing gradient problem seen in earlier forms of activation functions by rectifying the values of the inputs less than zero to zero otherwise; it will direct the input to output [18]. ReLU maps output ranging between 0 and 1 and it can be represented as

$$f(x) = \begin{Bmatrix} x, \text{ if } x \geq 0 \\ 0, \text{ if } x < 0 \end{Bmatrix} = \max(x, 0) \tag{8}$$

The below graph shown in Fig. 10 represents the equation of the ReLU activation function.

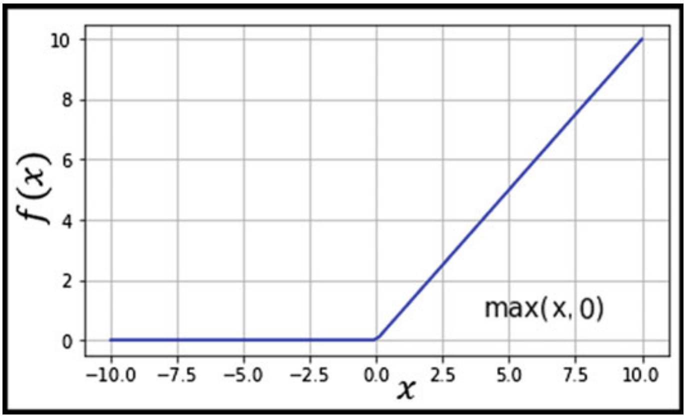


Fig. 10 ReLU function

ReLU incorporates faster AF learning that outperforms other AFs including the sigmoid and tan*h* functions in terms of efficiency and generalization [11]. It will be easier to train the varieties of neural network models using ReLU as a default activation function.

Exponential Linear Units (ELUs) Function

ELU also known as exponential linear unit is another form of activation function which is similar to the ReLU with certain variations. The ELU activation function can be specified mathematically as follows:

$$f(x) = \left\{ \begin{array}{ll} \alpha(e^x - 1), & x \le 0 \\ x, & x > 0 \end{array} \right\} \tag{9}$$

The equation strictly outputs x-value for positive values of input x, which is the same as ReLU. In the case of negative input, the output will be α times $(e^x - 1)$. Where α is a hyperparameter that controls the value of negative inputs for which ELU saturates. This is an excellent way of handling the negative inputs [18]. The graphical contrast between the ReLU and ELU activation functions is shown in Fig. 11.

ELUs have negative values that try to bring the mean of the activations closer to 0. This enables quicker learning when the gradient is closer to the natural gradient. It does not experience the issue of dying neurons because the gradient of ELU is non-zero for all negative values. ELU is a steady and differentiable activation function at all points.

Swish Function

The swish activation function is an innovative activation function that can be effectively used in deep learning models across a variety of complicated data sets. Mathematically, the swish function can be defined as follows:

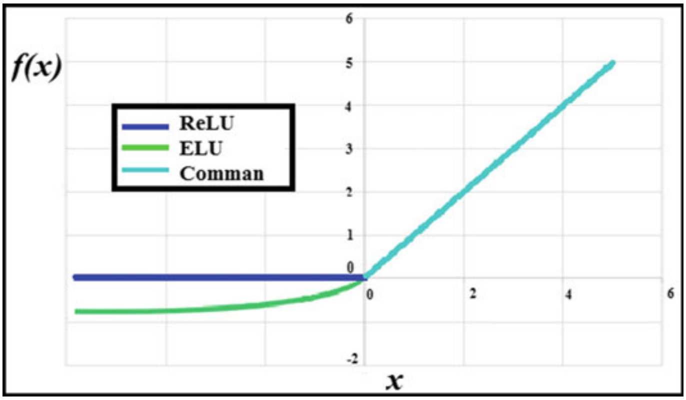


Fig. 11 Comparison between ReLU and ELU activation functions

$$f(x) = \frac{x}{(1 + e^{-x})} \tag{10}$$

The function is just the multiplication of the input x with the sigmoid function and its graphical representation is shown in Fig. 12.

Swish is unrestricted in the upper portion of the graph, therefore, the output would not be saturated to the maximum for large values of input [19]. However, it is restricted in the lower portion of the graph; therefore, for negative inputs, it does not return a zero as is the case for ReLU. Swish is smooth, non-monotonic, and continuous at all points; this differentiates it from most of the common activation functions.

Flatten-T Swish (FTS) Function

FTS, or flatten-T Swish (FTS), was introduced by Chieng as a novel activation function. Flatten-T swish incorporates activation features of both swish and rectified

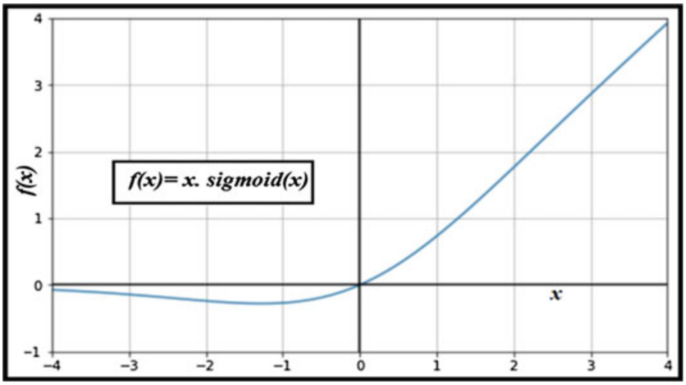


Fig. 12 Swish function

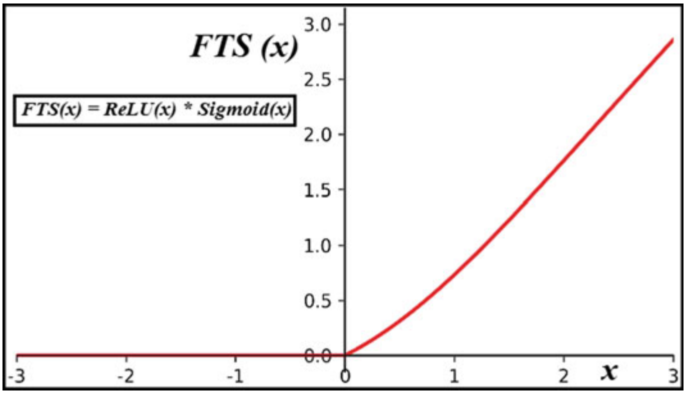


Fig. 13 Flatten-T swish function

linear units (ReLUs) activations functions together into an innovative one [20]. It is used to cope with the negative cancelation property in ReLU. Mathematically, FTS is formulated as follows:

$$\mathrm{FTS}(x) = \left\{ \begin{array}{ll} \frac{x}{1+\mathrm{e}^{-x}}, & x \geq 0 \\ 0, & x < 0 \end{array} \right\} \tag{11}$$

When the value of $x \geq 0$, the FTS function has properties identical to that of the swish activation function. If the value of x is less than zero, then the function acts as ReLU. The graphical representation of the FTS function is shown in Fig. 13.

FTS has network limitations like dynamicity, pliability, and nonlinear representation capacity. Table 1 gives a summary of activation functions and their corresponding equations, derivatives, and applications.

Conclusion

Deep learning approaches employ neural networks consisting of several hidden layers to perform complex tasks. In the design of neural networks, the activation function plays a vital role. The hidden layer's activation function dictates how well the network model learns the training data set. The kind of predictions the model will offer is determined by the activation function employed in the output layer. Vanishing gradient is an unstable behavior that inhibits the training of deep neural networks with saturated activation functions. As the network's layers become deeper, the training efficiency and precision encounter numerous challenges which stimulate the development of different kinds of activation functions. Thus, activation functions are an important component of networks and selecting proper activation functions and analyzing their impact on the network will assist in optimizing the efficiency of the DL model.

Table 1 Activation functions and their corresponding equations

Function	Computation equation	Derivative	Application
Binary step activation function	$f(x) = \begin{Bmatrix} 0; & \text{for}\, x < 0 \\ 1; & \text{for}\, x \geq 0 \end{Bmatrix}$	$f(x)' = \begin{Bmatrix} 0; & \text{for}\, x \neq 0 \\ ?; & \text{for}\, x = 0 \end{Bmatrix}$	Perceptron linear classifier
Sigmoid activation function	$f(x) = \frac{1}{(1+e^{-x})}$	$f(x)' = f(x)(1 - f(x))$	Logistic regression classification
Tan*h* function	$\tan h(x) = \frac{1-e^{-2x}}{1+e^{-2x}}$	$f(x)' = 1 - f(x)^2$	Classification between two classes
Arctan function	$f(x) = \tan^{-1}(x)$	$f(x)' = \frac{1}{1+x^2}$	Learn complex patterns in the data
Softmax function	$\sigma(z_i) = \frac{e^{z_i}}{\sum_{j=1}^{K} e^{z_j}}$	$\sigma(z_j)' = \sigma(z_j)(1 - \sigma(z_j))$	Normalize the output of a network
Softsign function	$f(x) = \frac{1}{(1+\lvert x\rvert)}$	$f' = \frac{1}{(1+\lvert x\rvert)^2}$	Predict the multinomial probability distribution in the output layer
Rectified linear unit function	$f(x) = \begin{Bmatrix} x; & \text{if}\, x \geq 0 \\ 0; & \text{if}\, x < 0 \end{Bmatrix} = \max(x, 0)$	$f(x)' = \begin{Bmatrix} 1; & \text{if}\, x \geq 0 \\ 0; & \text{if}\, x < 0 \end{Bmatrix}$	Prevent the exponential growth in the computation
Exponential linear units function	$f(x) = \begin{Bmatrix} \alpha(e^x - 1), & x \leq 0 \\ x, & x > 0 \end{Bmatrix}$	$f'(x) = \begin{Bmatrix} f(x) + \alpha; & \text{if}\, x < 0 \\ 1 & \text{if}\, x \geq 0 \end{Bmatrix}$	Introduces nonlinearity into the output of a neuron
Swish function	$f(x) = \frac{x}{(1+e^{-x})}$	$f'(x) = f(x) + \frac{1}{1+e^{-x}}(1 - f(x))$	Achieves higher test accuracy in very deep networks

(continued)

Table 1 (continued)

Function	Computation equation	Derivative	Application
Flatten-T swish function	$f(x) = \left\{ \begin{matrix} \frac{x}{1+e^{-x}}, & x \geq 0 \\ 0, & x < 0 \end{matrix} \right\}$	$f'(x) = \left\{ \begin{matrix} f(x) + \frac{1}{1+e^{-x}}(1 - f(x)); & x \geq 0 \\ 0; & x < 0 \end{matrix} \right\}$	Improved classification accuracy and converges twice as fast as ReLU

References

1. Min E, Guo X, Liu Q, Zhang G, Cui J, Long J (2018) A survey of clustering with deep learning: from the perspective of network architecture. IEEE Access 6:39501–39514. Author, F.: Article title. J 2(5):99–110 (2016)
2. Yan B, Han G (2018) Effective feature extraction via stacked sparse autoencoder to improve intrusion detection system. IEEE Access 6:41238–41248
3. Hatcher WG, Yu W (2018) A survey of deep learning: platforms, applications and emerging research trends. IEEE Access 6:24411–24432
4. Gogoi P, Roy S, Bujarbaruah SM (2019) Modeling a H-H neuron based spiking neural network incorporating multiple pre-synaptic inputs. In: 2nd international conference on innovations in electronics, signal processing and communication, Shillong, India
5. Paugam-Moisy H, Bohte S (2010) Computing with spiking neuron networks. In: Rozenberg G, Back T, Kok JN (eds) Handbook of natural computing, 1st edn. vol. 1. SpringerVerlag, Heidelberg, Germany, pp 1–47
6. Shrestha A, Mahmood A (2019) Review of deep learning algorithms and architectures. IEEE Access 7:53040–53065
7. Xin Y et al (2018) Machine learning and deep learning methods for cybersecurity. IEEE Access 6:35365–35381
8. Zhao Z, Zheng P, Xu S, Wu X (2019) Object detection with deep learning: a review. IEEE Trans Neural Netw Learn Syst 30(11):3212–3232
9. Yu Y, Adu K, Tashi N, Anokye P, Wang X, Ayidzoe MA (2020) RMAF: Relu-memristor-like activation function for deep learning. IEEE Access 8:72727–72741
10. You W, Shen C, Wang D, Chen L, Jiang X, Zhu Z (2020) An intelligent deep feature learning method with improved activation functions for machine fault diagnosis. IEEE Access 8:1975–1985
11. Nwankpa C, Ijomah W, Gachagan A, Marshall S (2018) Activation functions: comparison of trends in practice and research for deep learning. arXiv:1811.03378
12. Wang Y, Li Y, Song Y, Rong X (2020) The influence of the activation function in a convolution neural network model of facial expression recognition. Appl Sci
13. Dlugosz Z, Dlugosz R (2018) Nonlinear activation functions for artificial neural networks realized in hardware. In: 25th international conference mixed design of integrated circuits and system (MIXDES), Gdynia, Poland, pp 381–384
14. Tan TG, Teo J, Anthony P (2011) A comparison of two sigmoidal-type activation functions in video game controller evolution. In: IEEE conference on sustainable utilization and development in engineering and technology
15. Pasca B, Langhammer M (2018) Activation function architectures for FPGAs. In: 28th international conference on field programmable logic and applications, Dublin, Ireland, pp 43–437
16. Wang M, Lu S, Zhu D, Lin J, Wang Z (2018) A high- speed and low-complexity architecture for softmax function in deep learning. In: IEEE Asia Pacific conference on circuits and systems (APCCAS), Chengdu, China, pp 223–226

17. Le P, Zuidema W (2015) Compositional Distributional semantics with long short term memory. arXiv1503.02510 [cs]
18. Qiumei Z, Dan T, Fenghua W (2019) Improved convolutional neural network based on fast exponentially linear unit activation function. IEEE Access 7:151359–151367. https://doi.org/10.1109/Access.2019.2948112
19. Tripathi GC, Rawat M, Rawat K (2019) Swish activation based deep neural network predistorter for RF-PA"TENCON 2019. In: 2019 IEEE region 10 conference, pp 1239–1242. https://doi.org/10.1109/TENCON.2019.8929500
20. Chieng HH, Wahid N, Pauline O, Perla SRK (2018) Flatten-T Swish: a thresholded RElU-swish-like activation function for deep learning. Int J Adv Intell Inform 4(2):76–86

Early Detection of Android Locker Ransomware Through Foreground Activity Analysis

Shina Sheen and S. Gayathri

Abstract The usage of Android smartphones is rapidly increasing. The users are unaware of the malware activities which may target their mobile phones. This paper focuses on a particular kind of malware named ransomware that turned out to be a massive security threat to end-users, large organizations, and enterprises. In recent times, locker ransomware has been playing a major havoc in Android families. Locker ransomware blackmails victims for ransom by compulsorily locking the devices. We propose a model based on dynamic analysis to detect the locker ransomware variants using foreground analysis.

Keywords Ransomware detection · Android · Dynamic analysis-foreground analysis—Locker ransomware

Introduction

Android phones have become a popular and lucrative target for hackers as the smartphone's users are growing rapidly, and these users tend to store most important personal information on their devices. This attracts the hackers' community to specifically target the mobile devices with various malware attacks. Mobile malware has become a potent and growing threat to enterprise [1]. Malicious apps utilize multiple methods to evade the existing detection mechanisms provided by Android operating system or existing anti-virus software. Ransomware is a growing problem for users of mobile devices. The idea of scamming Inter net users and forcing them to pay money goes back to the late'90s [2], and the early appearances of ransomware on Android were cases in which the extortion functionality was added to fake anti-viruses in 2012. Lockscreen types and file-encrypting ransomware have been creating major problems in the Android platform. Once the ransomware attacks the device, it demands for ransom payment to be made by the users in order to regain access to their devices. The payments are usually made via Bitcoin. Recently due to the pandemic situation

S. Sheen (✉) · S. Gayathri
PSG College of Technology, Coimbatore, India
e-mail: ssh.amcs@psgtech.ac.in

V. Bindhu et al. (eds.), *Proceedings of Third International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 844,
https://doi.org/10.1007/978-981-16-8862-1_60

raised by corona virus, there is lot of apps in the play store to provide tracking and statistical information regarding COVID-19 updates. The hackers' community took up this situation and created the CovidLock [3], the malicious app which claims to be the corona virus tracker. Once the app is installed and acquires the required admin privileges, it locks the device and demands for ransom in the form of Bitcoin. It threatens with the message of leaking the victim's private data including photographs, videos, and more if failed to make the payment.

The number of ransomware attacks is increasing exponentially every year, and all the existing approaches terribly fail to safeguard the mobile devices due to the advanced evasion methods employed by the hackers. Ransomware comes in two different variants.

- Locker Ransomware—Denies the access to the device by locking the screen through hijacking activities or by disabling buttons. This form of ransomware commonly attacks smartphones or tablets.
- Crypto Ransomware—It takes control over the victim's device to encrypt the files and prevents the user from accessing their device. This form of ransomware commonly attacks PCs.

Although ransomware has been playing havoc into windows environment [4, 5] and Linux environment [6], it has turned out to be as threatening to mobile devices also. Mobile ransomware creeps into your phone using social engineering tactics that camouflage users into downloading malicious contents such as fake apps from third party app stores or through software updates or even by clicking on a spam link sent via SMS or mail. Once the ransomware gets into the device, it gains full access through acquiring required permissions from the user and displays a ransom note with fake message demanding for payment.

This paper primarily focuses on an analysis of Android ransomware families and the detection of locker variant of ransomware using foreground activity analysis. The main contributions of this paper are as follows:

- A detailed analysis of Android ransomware samples identifying the most common permissions and APIs used by Ransomware.
- Analysis of the foreground activity of malicious and benign files and identifying the difference in activities between the two types of files. Experiments show that the proposed system is capable of detecting locker ransomware like behaviors with high accuracy and efficiency.

Related Work

Alsoghyer et al. [7], proposed an application programming interface (API)-based ransomware detection system (API-RDS) to provide a static analysis paradigm for detecting Android ransomware applications. It focuses mainly on examining API calls. This approach used random forest classification algorithm for ransomware detection. Sharma et al. [8] designed RansomwareElite application to detect ransomware based on static approach. It checks for the permission requested by installed apps and threat messages in the app code. It works on text recognition as well as image processing to capture threat messages.

Andronio et al. [9], proposed HelDroid model to detect Android ransomware using static analysis and natural language processing. Two of the feature sets such as sequence of API calls to lock or encrypt the device and strings indicating threat messages are used. NLP detects threat messages displayed by the hackers by comparing with database of extracted threat words. A two-phase detection mechanism using static and dynamic approach is also given by Gharib et al. [10] which also make use of NLP in extracting strings present in ransomware files. Zheng et al. [11] proposed a GreatEatlon tool which is a novel encryption-detection approach that enhances the Heldroid model. He deployed the static analysis technique to discover device administration APIs abuse and to detect the most common patterns used by malware. Maiorca et al. [12], proposed a static approach and ML classifiers to distinguish ransomware apps from benign apps. Faris et al. [13] proposed a static approach for efficient detection of ransomware using machine learning. The ransomware behavior patterns are captured using sequences of application API calls and permissions to build the detection framework for ransomware. Yang et al. [14] and Chen et al. [15] proposed a dynamic approach to illustrate systematic characterization and real-time detection of Android ransomware. A novel system to detect encrypting ransomware in real-time, which can monitor a device's sensitive files and determine the user intention by three UI indicators. Song et al. [16] proposed a ransomware prevention technique on Android platform using dynamic analysis approach. This model is implemented on the kernel level which can detect ransomware apps faster than those at application level. It monitors for the process, memory, and storage I/O usages and utilizes statistical methods to prevent the intrusion of ransomware into the Android devices. Alzahrani et al. [17] introduces RanDroid which is a lightweight hybrid approach for detecting locker ransomware on Android platform by measuring the structural similarity between the information collected from legitimate apps and ransomware apps which utilizes the threatening messages.

Analysis of Android Ransomware Families

The Android ransomware families have evolved continuously over decades. In this section, few of the important variants of the Android ransomware families have been analyzed. Even though all the families exhibit the similar functionality of attacking the victim's device and demanding for ransom payment, they differ by certain factors. These ransomware families mainly differ by their type (locker/crypto), flow of events, APIs, permissions, threat messages, payment gateway and execution time. We have collected the dataset for seven different ransomware families—Charger, Fusob, Jisut, Koler, Simplocker, Svpeng, and Wannalocker with ten unique samples in each after eliminating the repeated samples from the collected dataset to acquire accurate and consistent data with most relevant features.

For further analysis of ransomware families in depth, we have considered permissions and API packages. To perform sensitive operations (e.g., tasks that can cost money or access private data), apps need specific permissions. Required permissions are defined in a manifest file (i.e.,AndroidManifest.xml), which is stored in the app's apk file. In order to obtain the data from those files, androguard packages (ApkAnalyzer) in Python was used to extract all the permissions and API methods being used in these ransomware families.

Permissions and API

The objective of ransomware applications acquiring Android permissions is to take control over the victim's device. The permissions used in the Android applications are mostly similar in all the ransomware families. All the permissions from the apk's have been collected and top ten permissions among those are considered for further analysis. Figure 1 shows the percentage of occurrences of the collected top ten permissions in each of the ransomware families. However, these permissions occur commonly in all the ransomware families. There exist few permissions specific to that particular kind of ransomware, and it is required to perform that specific task such as SET ALARM permission in charger family which was unavailable in other families.

A detailed analysis on characterizing malware families has been carried out by Mirzaei et al. [18], and the top ten permissions seen in benign samples are shown in Fig. 3. It is seen that `INTERNET`, `READ_PHONE_STATE`, and `ACCESS_NETWORK_STATE` permissions are widely requested in both

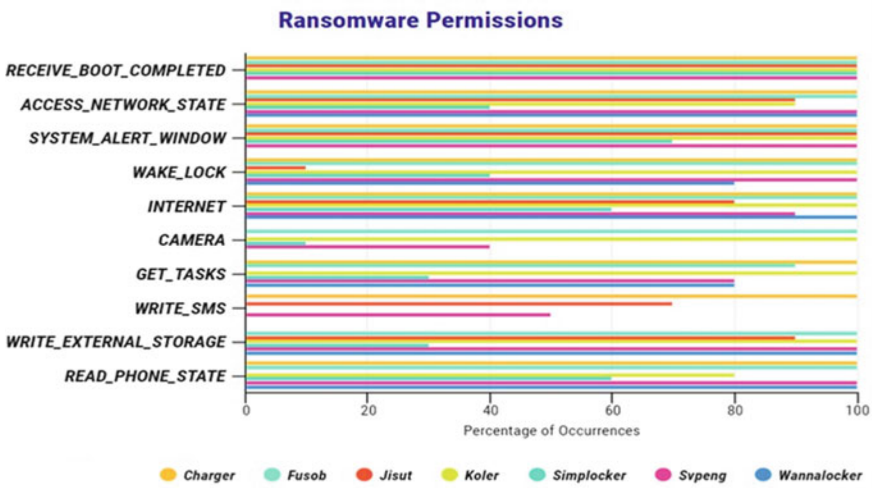


Fig. 1 Top ten permissions used by ransomware

ransomware and benign samples. The first two are typically needed to communicate with the remote server. In particular, we observe that ransomware tends to request more frequently on the device-related permissions, such as WAKE_LOCK, DISABLE_KEYGUARD, and SYSTEM_ALERT_WINDOW (Fig. 2).

The API packages are the collection of Java classes, packages, and interfaces used in the application. The top API packages used in the Android applications are mostly similar in all the ransomware families. The API packages are extracted from the DEX files, and it was found that common API packages in these ransomware families cover only up to 20 of the total API calls made in the application and other API packages are specific to that particular family. Figure 2 shows the percentage of occurrences of the collected top ten API packages in each of the ransomware families. Notably there exist few of the other API packages which are different in the each of the ransomware family with different functionalities. Example the API package com/caf consists of FM radio service which is available only in Simplocker

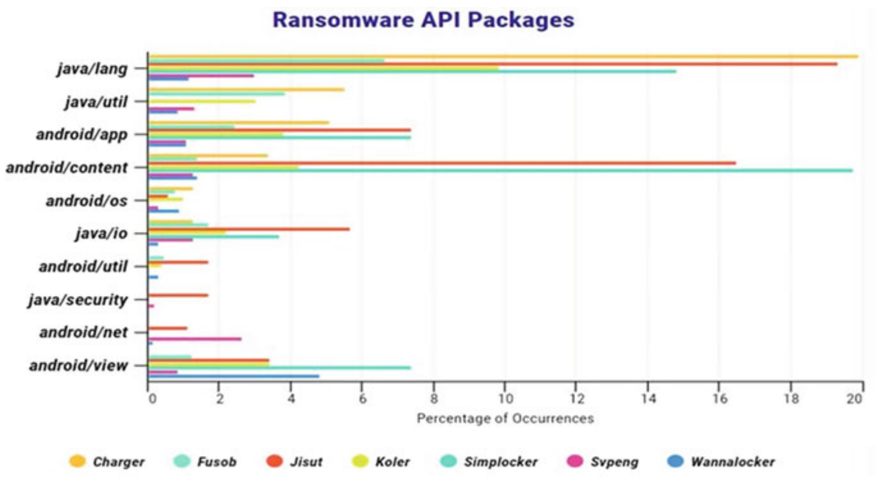


Fig. 2 Top 10 API packages used by ransomware

Fig. 3 Top 10 requested permissions from 1000 benign samples. The ranking corresponds to the findings in [19]

Permissions	Frequency
INTERNET	856
ACCESS_NETWORK_STATE	651
WRITE_EXTERNAL_STORAGE	471
READ_PHONE_STATE	388
VIBRATE	261
ACCESS_COARSE_LOCATION	245
WAKE_LOCK	234
ACCESS_FINE_LOCATION	221
RECEIVE_BOOT_COMPLETED	180
ACCESS_WIFI_STATE	176

and not in other families. Figure 2 shows the top API package used specifically in each of the family.

On analyzing the permissions and APIs used in ransomware and benign files, it is seen that most of the top permissions and APIs used by ransomware and benign files are almost the same as shown in Fig. 3. So discriminating between the two cannot be achieved using the permissions or APIs as features. Hence, a dynamic analysis approach based on the behavior of user activities is used in the proposed model.

Malware Versus Ransomware

Ransomware, a family of malware, differs from other kind of families with its unique characteristics and tendency of attacking victims' device. Other kinds of malware exhibits unusual behaviors like hiding itself, stealing the private data and most often works in the background. However, the behavior of ransomware is similar to that of usual benign apps such as popping-up screens, displaying notifications, and accessing the internal files, and it usually takes control of the device's top activity.

Rather than working in background like other malware families, the ransomware interacts with the user with the help of threat messages and by sending notifications in the form of screen overlays requesting for ransom payment to be made in order to regain access. The victim can easily ascertain once the ransomware starts infecting the device. On contrast, the other families of malware resides inside the device and works in the background without the knowledge of the victim and makes it difficult to discover the malicious activity.

The ransomware distinguishes itself from other malware family with their behavioral patterns. For instance using the foreground activity, it can be concluded. If the app remains in the foreground for an excessive time beyond a specified threshold value without the users' interaction, then the particular application is interpreted to be ransomware.

Another instance would be the number of processes requested by the ransomware application would be always on the higher side compared to other malwares.

Threat Model

According to the threat model, an attacker installs ransomware on victim's mobile device. Once the ransomware has gained access to the device and is executed, it does its malicious activity following it will then lock the device from use, using the new ransom notice on the lockscreen. Although locking ransomware looks not so harmful and could be defeated by ADB and root, it is actually not easy to be removed by ordinary users. Thus, in the history of ransomware, locker plays important roles in ransomware evolution. This is the reason why we have proposed the model on locking ransomware.

Table 1 List of factors to distinguish ransomware from malware

Factor	Malware	Ransomware
Behavior	Hides from users	Exposes to users
Activity	Background	Foreground
Discoverability	Harder	Easier
Device accessibility	Users can access	Users cannot access
Top permission	INTERNET	BOOT_COMPLETED
Admin privilege	Not mandatory	Mandatory
Data activity	Steals data	Locks/encrypts data
Payment	Not always	Demands ransom

Proposed Model for Locker Ransomware Detection

The proposed model utilizes the dynamic analysis of the app's behavior to detect ransomware. Unlike other malware, ransomware works in the foreground and let the users to identify it immediately once it invades the device rather than working in the background. Figure 4 shows the proposed model with the three modules,

1. Foreground analysis;
2. Monitoring user clicks;

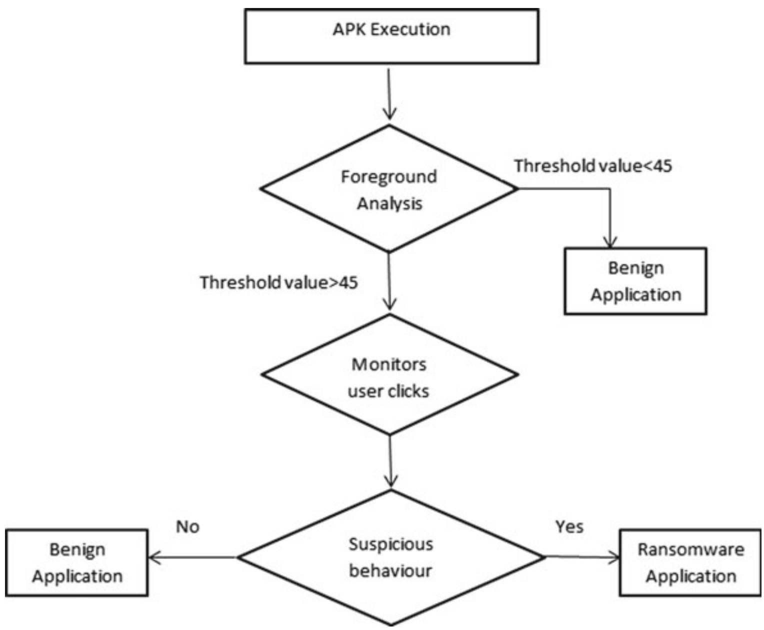


Fig. 4 Locker ransomware detection framework

```
ActivityManager am = (ActivityManager) this.getSystemService (ACTIVITY_SERVICE);
List<ActivityManager.RunningTaskInfo> taskInfo = am.getRunningTasks (1)          ;
Log.d ("topActivity", "CURRENT Activity:"+ taskInfo.get (0).topActivity.getClassName ());
ComponentName componentInfo = taskInfo.get (0).topActivity;
```

Fig. 5 Foreground activity capture

3. Suspicious behavior detection.

The samples from both ransomware and benign apps are executed in a sandbox in a controlled manner.

Foreground Analysis

In this module, the app is executed to capture the activities running mostly in the foreground, i.e., the top activity is captured as shown in Fig. 5, and the timer is being implemented such that it captures the duration of top screens' inactivity until the user clicks and the corresponding time durations are recorded.

Monitoring Clicks and Suspicious Behavior Detection

The next step in execution of the ransomware detection algorithm is monitoring and suspicious behavior detection. Basically, all the clicks/events/state transitions are recorded in the logs generated while executing the application. The state transitions from idle state of top activity are shown in Fig. 6.

For each of the default user's clicks such as home button, back and forward buttons the corresponding activities and events are stored in the repository for comparison. Once the user attempts to make a click, the corresponding event would be captured and compared with the repository. On analyzing the ransomware application, the user

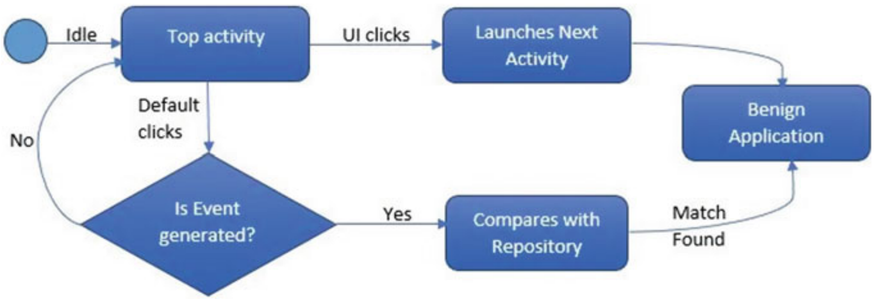


Fig. 6 State transition diagram of user click behavior

1. Executing APK sample in Virtual environment
2. Performing Foreground Analysis
   2.1. Recording Top activity screen overlay over home screen until the user interacts
   2.2. Comparing time duration with Threshold value
      2.2.1. if Threshold value falls less than 45 sec, then it's benign app
      2.2.2. else, it can be continued for further analysis
3. Monitor the logs to extract user clicks
   3.1. Record the user clicks and the corresponding events
   3.2. Monitor for the State Transitions when each click is detected
4. Detecting suspicious behavior from the collected data of previous modules
   4.1. Get output from the behavior analysis module
      4.1.1. If the output is found to be suspicious, alert the user
      4.1.2. If the output is found to be good ware, it can execute in the real device.

Fig. 7 Pseudocode for locker ransomware detection

clicks of default buttons does not yield any outcome and the same state appears. If there are any generated events, then it is compared with the repository which detects the benign samples. The detection mechanism is given in the pseudocode as shown in Fig. 7.

Experiments and Results

The experiments performed are mainly classified into two main modules: Foreground analysis and monitoring the user clicks through log traces and identifying the suspicious behavior from it. The datasets used for our experiment includes samples from the ransomware applications and from benign applications as well.

Finding 1

The benign samples of 50 applications with varied functionality are collected, and the foreground analysis is done, where the top activities screen's idle standby duration is recorded till the user start interacting with the device. The threshold value is calculated with the average mean of the time durations recorded by the executed samples. Based on observations, the lower threshold value is set up to be 5 s, and the upper threshold value is set up to be 45 s, where the benign app's top activities timeout does not go beyond the upper threshold value. When the ransomware apps are executed, the top activities timeout is always found to exceed the upper threshold value and no transitions occur with default button clicks say home button. Figure 8 shows that for benign samples, the recorded duration is below 45 s, while observing the locker ransomware variants it is found to be exceeding the upper threshold value of 45 s.

Finding 2

With the next module, the suspicious behavior can be detected by monitoring the events and state transitions of both the benign and ransomware applications. The

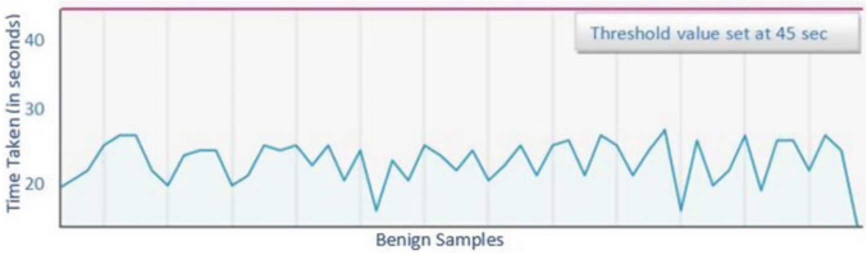


Fig. 8 Time taken for foreground analysis of benign samples

samples are executed, and the results have been analyzed for detecting the suspicious behavior of the application. In benign application for each of the user clicks, the corresponding activity or event is executed, and the state transition can be observed. For example, when the user clicks Home button, corresponding Home screen pops up in the device, whereas in the ransomware application for any of the user clicks, the corresponding event has miserably failed to execute and also the state transition has never happened with any of the default buttons execution. Rather it has remained idle for the entire execution cycle. This behavior is found to be suspicious and get be declared as ransomware application. The ransomware detection system is implemented, the results obtained while executing the collected samples are evaluated, and it is found to achieve lower false positive rate and yielded better accuracy of up to 98.2% for detecting locker ransomware variants.

Conclusion and Future Work

The paper does a detailed analysis of the various ransomware families. Based on the various factors like permissions requested and API calls analyzed, it was concluded that these sets of features were not suitable to discriminate between ransomware and benign files as many of the top permissions and API calls are also used by benign applications. So the paper proposes a method for detecting locker ransomware based on the foreground activity. Based on the samples analyzed, the proposed model is found to achieve lower false positive rate and yielded better accuracy of up to 98.2% for detecting locker ransomware variants.

Although the proposed work gives a good accuracy, it has certain limitations. It may lead to a false positive on apps that either stay in the foreground a lot without user interaction or where the user interaction does not trigger a new activity. A multilevel approach where the ransomware escapes the first level of foreground analysis can be done as a future work. Another direction is that the existing works in Android ransomware are designed to detect the ransomware only in English variants, whereas most often the hackers design the ransomware targeting only the users from specific countries in its language. Internationalization with all the languages utilizing NLP can be adopted as a future work.

Acknowledgements This research was funded by the Department of Science and Technology (DST), Govt. of India, under the Cyber security Research of Interdisciplinary Cyber Physical systems (ICPS).

References

1. McAfee Threat Report: https://www.mcafee.com/enterprise/en-us/lp/threats-reports/apr-2021.html
2. Young A, Yung M (1996) Cryptovirology: Extortion-based security threats and countermeasures. In: Proceedings 1996 IEEE symposium on security and privacy. IEEE
3. Covidlock: https://www.domaintools.com/resources/blog/covidlock-update-coronavirus-ransomware
4. Kharraz A et al (2015) Cutting the Gordian knot: a look under the hood of ransomware attacks. In: International conference on detection of intrusions and malware, and vulnerability assessment. Springer, Cham
5. Kharaz A et al (2016) UNVEIL: a large-scale, automated approach to detecting ransomware. In: 25th USENIX security symposium (USENIX Security 16)
6. G´omez-Hern´andez JA, Á lvarez-Gonz´alez L, Garc´ıa-Teodoro P (2018) R-Locker: Thwarting ransomware action through a honeyfile-based approach. Comput I Secur 73:389–398
7. Alsoghyer S, Almomani I (2019) Ransomware detection system for android applications. Electronics 8(8):868
8. Sharma G, Johri A, Goel A, Gupta A (2018) Enhancing RansomwareElite app for detection of ransomware in android applications. In: Proceedings of the 11th international conference on contemporary computing, pp 1–4
9. Andronio N, Zanero S, Maggi F (2015) Heldroid: dissecting and detecting mobile ransomware. In: International symposium on recent advances in intrusion detection. Springer, Cham
10. Gharib A, Ghorbani A (2017) Dna-droid: a real-time Android ransomware detection framework. In: International conference on network and system security. Springer, Cham
11. Zheng C, Dellarocca N, Andronio N, Zanero S, Maggi F (2016) Greateatlon: fast, static detection of mobile ransomware. In: Proceedings of the international conference on security and privacy in communication systems, Italy, pp 617–636
12. Maiorca D, Mercaldo F, Giacinto G, Visaggio CA, Martinelli F (2017) R-PackDroid: API package-based characterization and detection of mobile ransomware. In: Proceedings of the symposium on applied computing. Chicago, pp 1718–1723
13. Faris H et al (2020) Optimizing extreme learning machines using chains of salps for efficient Android ransomware detection. Appl Sci 10(11):3706
14. Yang T et al (2015) Automated detection and analysis for Android ransomware. In: 2015 IEEE 17th international conference on high performance computing and communications, 2015 IEEE 7th international symposium on cyberspace safety and security, and 2015 IEEE 12th International conference on embedded software and systems. IEEE
15. Chen J et al (2017) Uncovering the face of android ransomware: characterization and real-time detection. IEEE Trans Inf Forens Secur 13(5):1286–1300
16. Song S, Kim B, Lee S (2016) The effective ransomware prevention technique using process monitoring on Android platform. Mob Inf Syst 1:1–9
17. Alzahrani A, Alshehri A, Alshahrani H, Alharthi R, Fu H, Liu A, Zhu Y (2018) RanDroid: structural similarity approach for detecting ransomware applications in android platform. In Proceedings of the IEEE international conference on information technology. Michigan, USA, pp 1–536

18. Mirzaei O et al (2019) Andrensemble: leveraging API ensembles to characterize Android malware families. In: Proceedings of the 2019 ACM Asia conference on computer and communications security
19. Yerima SY, Sezer S, McWilliams G (2013) Analysis of Bayesian classification-based approaches for Android malware detection. IET Inf Secur 8(1):25–36

Self-Regulating Real-Time Server Log Monitoring Using Software-Defined Networking

A. Vishnupriya, Hirankumar Singh, and V. Surya Prakash Reddy

Abstract Monitoring the real-time log is the essential practice for network auditors to take proactive decisions. Collecting and analyzing raw log files are taking up a great deal of time. Software- defined networking is the prominent networking technology for providing the centralized network management solutions. The decoupled network architecture supports for the rapid enforcement of network policies in IT infrastructure. The programmable API- supported SDN controllers are collecting the flow statistics from the network nodes periodically. Here, we proposed the self-regulating real-time server log monitoring using SDN, and it retrieves the logs from controllers and switch components using ELK (Elastic search, Logstash, and Kibana) stack. Finally, we visualize the logs in graphical format using Kibana dashboard. The system can send the notification to network auditor whenever the suspicious activities invoked. Finally, we audit system by invoking the DDoS attack and test the performance of the model.

Keywords Server log · Software-defined networking · ELK stack

Introduction

The expeditious growth of high-speed mobile networks, generating the huge amount of network traffic in enterprise and data centre networking. Everyday, millions of requests are hitting the rack servers in the data centre. Managing these servers and providing the physical access control will be the unamenable duties for the cloud providers. Auditing the sever log is the one way for investigating the operations and

A. Vishnupriya (✉) · H. Singh · V. Surya Prakash Reddy
Department of Electronics & Communication Engineering, VelTech Rangarajan Dr. Sagunthala R&D Institute of Science & Technology, Chennai, India
e-mail: vishnu.priya@veltech.edu.in

H. Singh
e-mail: hksingh@veltech.edu.in

V. Bindhu et al. (eds.), *Proceedings of Third International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 844,
https://doi.org/10.1007/978-981-16-8862-1_61

performance of servers. This aid for the enterprise can take better decision about organization security practices [1].

Comprehensively, log contains the essential information about the server issues and the track of user access record. Server log contains the potential insights of request count, duration of access, and other relevant data over a period of time. In general, server generates the common log format (CLF) files, which is in incomprehensible format and difficult to understand the raw data [2]. The logs are periodically update the information, and the user can manually configure the updating interval. The log also persists the information of client IP address, frame size, time and date, and HTTP reply code. Since log files contain the wide collection of sensitive information, examine the log files that are permitted to the privilege users. Obtaining and analyzing the logs in non-invasive manner may reduce the performance and will take time for determination.

Centralized server log management is the wise choice and easy way to capture and inspect the logs. Even any server in the network goes down, centralized server can act as a redundant server and can replicate the server objects. Software-defined networking (SDN) is the intelligent network paradigm which can manage the orchestration on commercial off-shelf servers (COTS) and gears up the network programmability in central controller network. The three-plane SDN model offers the vendor neutral service in data centre, cloud, enterprise, and network operators. Scalable infrastructure, agile maintenance of virtual resources, and handling of big data are key advantages of software-defined networks.

The general server logs are classified as error log, access log, and agent logs that furnish the knowledge of the traps available in the webservers, activities of hacker's attempts, or presence of malicious broken links in the Web pages. Apart from this, there are no information about tracking of user activities, sessions, and geographical information. Hence, turn our focus toward the intelligent platform which can able to track the system performance and user activities in addition with general system logs. The proposed software-defined network-based log monitoring model runs in the application layer of SDN, and it can connect SDN controllers with ELK stack and collect the real-time data from controllers and stored in Elasticsearch database. The application program received logs can be analyzed and observe in Kibana dashboard. The great advantage of using ELK stack is that the auditor can extend the analysis with machine learning, so they can perform the predictive actions, before network goes down. It also extend feedback operations like sending notification or alarm to the administrator.

This paper is organized as follow: Section "Technology Stack" discusses the details of technology stack and Section "Proposed System Model" presents the procedure of conducting experiment and creation of SDN network followed by result and conclusion are discussed.

Related Works

Persisting the logs alone will not get rid of errors and network issues. We need to perform some additional actions to resolve the network issues. Siniarksi et al., proposed the real-time monitoring of SDN network using cloud-based logging tool Logentries. It is the real-time intelligent network monitoring and analyzing platform. This can make advance analysis from the system generated logs. The author discussed the performance of CPU usage for the taking decision. But the model not suggested the priority-based log handling procedures. Guye Liu et al. presented the application performance monitoring with SDN and NFV. In general, collecting data from individual components will be more expensive and make the performance degradation [2]. In this, the author developed the NetAlytics platform to monitor the application packet flows in the software-based network resources. The information has been collected using DPDK and NetVM in each device. Content popularity also analyzed using Apache storm real-time analytic engine data.

Anit Aeri et al., proposed the comparative study of network-based system log management tools [3]. There are plenty of log server tools are there. In that Snare, Kiwi, Spectorsoft , Manage Engine, and Splunk are some of the noteworthy log-monitoring tools. In that, Splunk is HTTP-based tool with the support of email notification and SQL database storage capabilities. Kiwi is a best centralized tool for inspecting system log files compared to others. Risto Vaarandi et al., use the Unix event collection tools for finding the anomalies in network event. They also described the different case studies for persisting the meaningful information from unorganized dump log file [4] by mining the text-based event pattern. The framework model can send email to administrator after identifying the attack. The approach can support the more number of servers and applications. Arjun Roy et al., reviewed the challenges of cloud data center virtual network monitoring using VNET Pingmesh. They examined the VNET addressing and flow handling using SDN. VNET pingmesh provides better fault detection and cross-layer monitoring framework [5].

Mohd Ameen Imran et al., analyzed the effective role of logs in data center and cloud networking. They created log as a security service (LAAS) by collecting the log files from various sources like application log, server log, network log, and VM logs [6]. The authors discussed the privacy-preserving procedures for log events. Ahamad Ali et al., described Blockchain-based log management approach for providing the log security, access control, and trust management. Distributed ledger technologies provide the better auditing of logs and make sure the privacy of the log files access [7]. Two stages of storing of log files in Elasticsearch and Blockchain increase the storage overhead. Gwi-sook Jang et al., presented the operator log monitoring system for nuclear power plant. The regular operation of power plant information has collected and stored in database [8]. The framework displays the operation events visually. This also performs the log preprocessing and analysis using data mining algorithms. Surendra Gaur et al., projected the centralized log mining framework for IT infrastructure. Authors suggested various log resources from all kind of network devices before taking the security actions [9]. The priority-based log handling has been included by including the severity details in header. Robert

Mayers recommended central log management (CLM) that reduced the manual network log maintenance and OPEX in enterprise networks. The framework suggested the filtering method when log volume increased [10].

The details of literature propounded the various network monitoring tools and frameworks. In the continuation, we author discussed the SDN- based centralized log-monitoring approach which not simply inspect logs and also automates the monitoring approach by close loop procedure of passing the control instructions to the central controller. It displays for collected data using Web user interface by three stack of tools Elasticsearch, Logstash, and Kibana.

Technology Stack

Ryu Controller

Ryu is an open source, API-supported Python-based SDN controller [11]. It is having rich libraries and interfaces to connect various modules in controller. It manages the networking devices like switches and routers by OPENFLOW and NETCONF interfaces. It consist of topology manager, event handler, application manager, switch manager, and Openflow Virtual Switch Data Base (OVSDB) modules. Ryu also compatible with Sflow and Netflow aggregator to collect flow statistics from connected devices. The event handler handles the asynchronous events of various applications [12]. Modules in Ryu can work collectively and create the centralized programmable networks. It can able to collect flow statistics per device basis and which can consider for further analysis by the application [13].

Elasticsearch

Elasticsearch is the Apache license, open source, no SQL database to store the big volume of data centrally. Advance analysis makes the application more reliable and easy deployment in Web and mobile platform. The rapid searching procedure puts together it in the prominent way of collecting the logs from various switch components. It persists the logging information in the database for further analysis [14].

Logstash

It is the next level of stack in ELK; it acts as an aggregate tool for collecting the data input from the agent source. The collected data is persisted in Elasticsearch database. It carry out the preprocessing of aggregated data for the effective real-time log analysis. Logstash performs three stages of operation (i) input, (ii) filter, and (iii) output. Input stage of the data is processed in machine-compatible format. Filter brings about the required analytical event like cleaning the data, parsing the date, and modifying the data format. The output stage is making the final decisions of the event.

Kibana

Kibana is a final stack of ELK to create the dashboard to observe the results of network events. The network administrator gets the quick intuition of the network metrics by referring the dashboards. Referring the dashboard components like diagrams, maps, tables, and graphs is much easier compared to understanding the complex log files. It also interact with the Elasticsearch database for advanced searching and monitoring the log metrics. Advanced elastic machine learning can also predict the event based on the historical statistics data [15].

Proposed System Model

The basic understanding of the proposed model is depicted in Fig. 1, the logs are collected periodically from Ryu controller by using log agents. Log agents are the software code that automates the network monitoring. The collected data drive directly to the ELK stack by TCP transport layer protocol [16]. In general, the logs consist of device interfaces, date and time, flow duration, frame count, CPU utilization, and flow statistics. The logs collected from Ryu controller using API are in simple JSON format, and it can be easy to interpret compared to complex server logs. By preconfiguring, the agent source and ELK stack can extract the necessary information for the analysis and display the details in Kibana dashboard. SDN can automate the monitoring platform by providing the feedback control as sending the notification to the administrator or triggering the flow rule update to controller. The historical flow statistics data can perform the proactive prediction activities using machine learning support in ELK stack. In the continuation, we tested the monitoring platform by experimenting DDoS attack in terms of flooding of ICMP packets in the server-connected network [17]. The monitoring ELK stack collects the flow and frame counts from controllerm and it identifies the anomalies in the traffic. We configured the Logstash to filter the frame and byte count in the flow for spotting the event abnormalities [18]. The DoS mitigation script running in application layer of the SDN interacts with the controller and modifies the flow entries after detecting the attack.

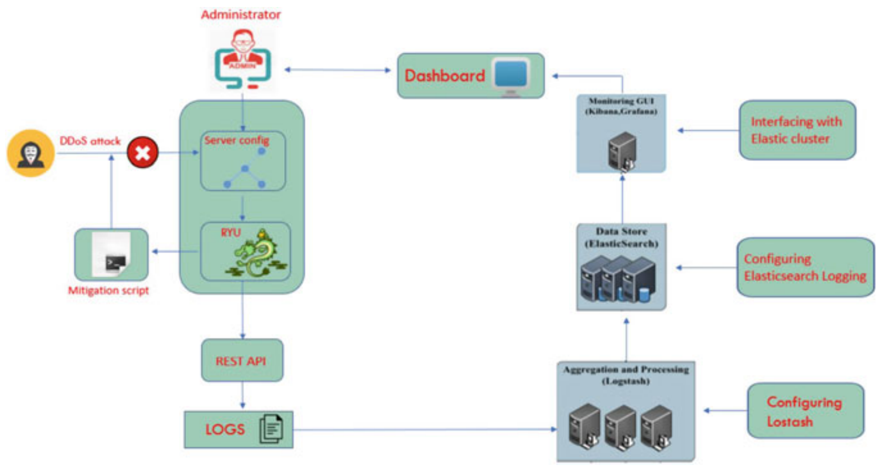


Fig. 1 Automated network monitoring model

Experimentation and Result

The emulation of server network is hosted in Mininet in virtual machine. The network monitoring model is tested in 16 GB RAM Intel core –C7-10750H CPU @2.60 GHz system. The virtual machine configured with Ubuntu operating system in 2 core with 4 GB RAM. As depicted in Fig. 2, a simple tree network topology is created in Mininet with 13 OVS switch and 13 host nodes in depth of 2. A simple HTTP server is emulated in h1 node. Client machine runs in the same network topology. In view of the fact that the objective of the proposed model is to analyze the efficiency of log based network monitoring, we create the client and server nodes in same network. Ryu controller connects all switches and aggregates the flow statistics from all connected nodes [19].

Logstash integrates with the ryu.app.ofctl_rest app module in SDN controller via REST API. It extends it for retrieving the switch statistics and obtains the flow details using HTTP_Poller plugin. It filters the unrelated items in the log, then pipelines the data to the Elasticsearch. A dashboard is created in Kibana to show to charts of flow count, byte count, and frame count values. The descriptive log also included as a part of dashboard for the detail representation.

Monitoring of DDoS Attack

The proposed SDN-connected ELK stack effectively monitoring the events in the local network. It can automate the monitoring platform, and it acts as a central aggregate to collect the events and makes the monitoring most effective in real time. The performance of the monitoring system has studied in the occurrence of DDoS

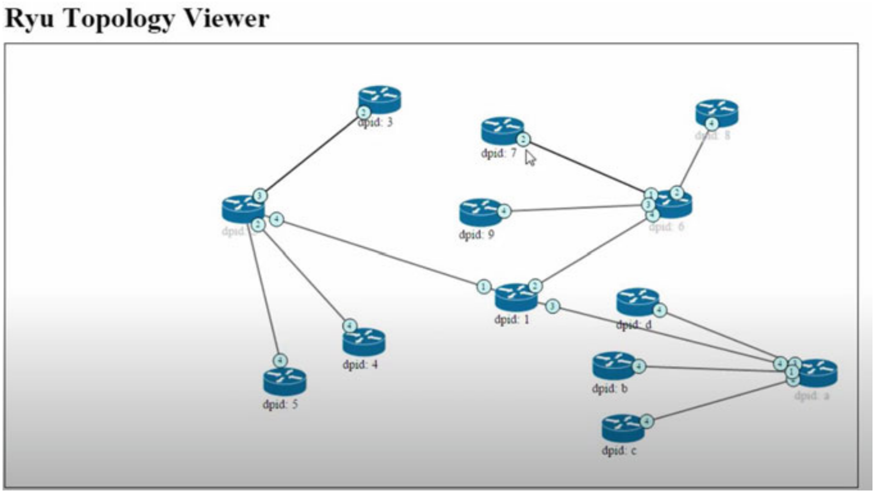


Fig. 2 Network topology

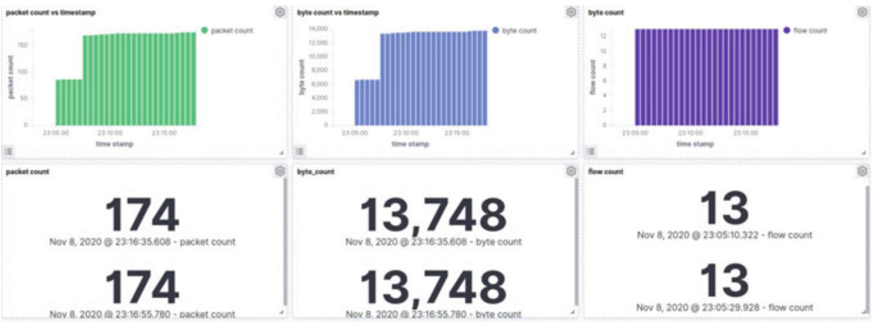


Fig. 3 Kibana dashboard

flooding attack in the network. ICMP flooding is initiated from SDN client nodes to the server. This makes the abrupt increase in number of packet count in the received flow statistics. The information retrieved by ELK stack displays that in dashboard as depicted in Fig. 3, at the same time, DDoS mitigation application retrieved the byte count values and compared it with threshold value. In this case, the threshold value is fixed as a static entity, and when the value exceeds the threshold value, it displays the notification message in dashboard. Once attack identified, the DDoS mitigation program instantiated flow rule modification instruction to controller for closing the flooded ports. After successful execution of mitigation program, Ryu controller drops all the packets from the client, and the controller output is mentioned in Fig. 4.

DDoS mitigation application requests to access the flow-table statistics of the switch, and it extracts the packet count data periodically. Using the packet count data obtains the traffic-rate. Based on the traffic-rate, DDoS mitigation is done by

```
------------------------------------------------------------
0000000000000001        1 00:00:00:00:00:03        3   422187 17731854
The packets are dropped from below IP to server
10.0.0.1
packet in 1 00:00:00:00:00:03 00:00:00:00:00:01 3
packet in 1 00:00:00:00:00:01 00:00:00:00:00:03 1
datapath              in-port eth-dst       out-port packets bytes
------------------------------------------------------------
0000000000000001        1 00:00:00:00:00:03        3  1067686 44842812
The packets are dropped from below IP to server
10.0.0.1
0000000000000001        3 00:00:00:00:00:01        1        3      210
```

Fig. 4 Network flow blocked in Ryu controller

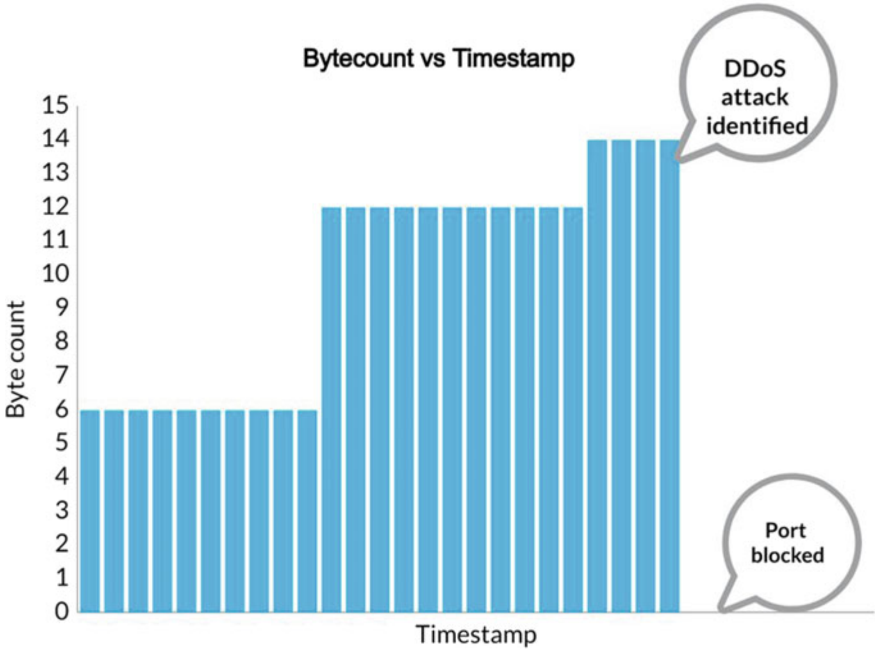


Fig. 5 Byte count after DDoS mitigation

dropping the packets from the scr_ip address. Figure 5 shows the dropping of flow can be visualized in dashboard.

Conclusion and Future Scope

In general, server log monitoring framework provides the event details and visualizes the event occurrence. The proposed work introduced the intelligent, autonomous server log monitoring using software-defined networks. ELK stack is the real-time advanced log- monitoring platform which can interact with the centralized SDN controller. It learns the flow details from switches periodically and displays the details in dashboard. The model has given the scope to identify the security violation in the network. It automates the mitigation procedure by sending the feedback control to the SDN controller for changing the flow rule. Here, we demonstrate the model with simple and straight forward approach of blocking the attack-initiated port by threshold value. In most cases, blocking the ports based on static threshold is not recommended. The future work covers to identify the attacks using statistical-based approach and reduce the allowable byte count in the infected ports.

References

1. Siniarski B, Olariu C, Perry P, Parsons T, Murphy J (2016) Real-time monitoring of SDN networks using non-invasive cloud-based logging platforms. In: 2016 IEEE 27th Annual international symposium on personal, indoor, and mobile radio communications (PIMRC). IEEE, New York, pp 1–6
2. Liu G, Wood T (2015) Cloud-scale application performance monitoring with SDN and NFV. In: 2015 IEEE international conference on cloud engineering. IEEE, New York, pp 440–445
3. Aeri A, Tukadiya S (2015) A comparative study of network based system log management tools. In: 2015 International conference on computer communication and informatics (ICCCI). IEEE, New York, pp 1–6
4. Vaarandi R, Kont M, Pihelgas M (2016) Event log analysis with the LogCluster tool. In: MILCOM 2016 IEEE military communications conference. IEEE, New York, pp 982–987
5. Roy A, Bansal D, Brumley D, Chandrappa HK, Sharma P, Tewari R, Arzani B, Snoeren AC (2018) Cloud datacenter SDN monitoring: experiences and challenges. In: Proceedings of the internet measurement conference, pp 464–470
6. Imran MA, ur Rahman MA (2020) Log as a secure service scheme (LASS) for cloud. J Sci Res 64(1)
7. Ali A, Khan A, Ahmed M, Jeon G (2021) BCALS: blockchain-based secure log management system for cloud computing. Trans Emerg Telecommun Technol e4272
8. Jang G, Son K, Seong S (2021) Operator action log based monitoring and control and its verification of nuclear power plants. Ann Nuclear Energy 151:107939
9. Gaur S, Rahman N (2020) A secured log mining approach to collection, monitoring, rotation, and analysis of frequent and heterogeneous logs. In: ICIDSSD, p 285
10. Meyers R (2020) Data highway and the digital transformation: arguments for secure, centralised log management. Network Security 10:17–19
11. Mishra S, AlShehri MAR (2017) Software defined networking: research issues, challenges and opportunities. Indian J Sci Technol 10(29):1–9
12. Duraipandian M (2020) Long term evolution-self organizing network for minimization of sudden call termination in mobile radio access networks. J Trends Comput Sci Smart technol (TCSST) 2(02):89–97
13. Ryu Documentation. https://ryu.readthedocs.io/en/latest/
14. Nishant R (2017) Visual logging framework using ELK stack. PhD diss
15. Bindhu V (2020) Constraints mitigation in cognitive radio networks using cloud computing. J Trends Comput Sci Smart technol (TCSST) 2(01):1–14
16. Bavaskar P, Pranita OK, Sinha AK (2019) A survey on: log analysis with Elk stack tool. IJRAR-Int J Res Anal Rev (IJRAR), E-ISSN: 2348-1269

17. Xu Y, Liu Y (2016) DDoS attack detection under SDN context. In: IEEE INFOCOM 2016-the 35th annual IEEE international conference on computer communications. IEEE, New York, pp 1–9
18. Oenflow concepts. https://overlaid.net/2017/02/15/openflow-basic-concepts-and-theory/
19. Singh MP, Bhandari A (2020) New-flow based DDoS attacks in SDN: taxonomy, rationales, and research challenges. Comput Commun 154:509–527

IoT-Enabled Home Automation System

Subhanullah Omarkhil, C. Shoba Bindu, and E. Sudheer Kumar

Abstract Home automation is a crucial component of the Internet of things, sometimes known as Internet of things (IoT). For decades, home automation has been utilized to control basic household goods such as lighting and small appliances. The globe will be connected with a touch of a fingerprint or simple voice instructions, according to current technology. The existing system was not secure and safe because it did not have fire detection and notification system. And it was high cost and high power consumption. The well-known IoT home automation system is typically straightforward to implement in a real home, allowing for real-time monitoring of home conditions and control of home equipment. Several sensors and actuators were connected to the NodeMCU controller, and the updated temperature, humidity, motion, and gas data could be seen on laptops and PCs using the MQTT Dash mobile application and the Adafruit IO website. For security and safety concerns, users receive warnings on their mobile phones when there is an odd scenario via the IFTTT server. Household appliances are regularly controlled in a simple and effective manner using the MQTT/Adafruit IO GUI or voice commands with Google Assistant. The proposed system aims to expand the Home Automation system with additional sensors, actuators and instead of utilizing batteries, solar panels are used to power the control box, making the system safe, secure, energy-efficient, and environmentally beneficial.

Keywords Smart home · IoT · Arduino Uno · GSM · Node MCU · Solar panels · Blynk

S. Omarkhil (✉) · C. Shoba Bindu · E. Sudheer Kumar
Department of CSE, JNTUA College of Engineering, Ananthapuramu, Andhra Pradesh 515002, India

C. Shoba Bindu
e-mail: shobabindhu.cse@jntua.ac.in

V. Bindhu et al. (eds.), *Proceedings of Third International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 844,
https://doi.org/10.1007/978-981-16-8862-1_62

Abbreviations

AC	Alternating current
COM Port	Communication port
DC	Direct current
GPRS	General packet radio service
GSM	Global system for mobile communication
GUI	Graphical user interface
HTTP	Hypertext transfer protocol
IDE	Integrated development environment
IoT	Internet of things
NodeMCU	Node microcontroller unit
ROM	Read-only memory
WiFi	Wireless fidelity

Introduction

Home automation systems have attracted considerable attention with the advancement of communications technology. A smart home (SH) is an Internet of things (IoT) application that utilizes the Internet to monitor and control appliances using a home automation system [1]. Automation involves introducing a degree of computerized or automatic control to certain electrical and electronic systems in a building. These include lighting, temperature control, etc. The past decade has seen significant advancement in the field of consumer electronics. Various intelligent appliances, such as cellular phone, air-conditioners, home security devices, home theaters, etc., are set to realize the concept of a smart home. They have given rise to a personal area network in home environment, where all these appliances can be interconnected and monitored using a single controller [2]. Smart home itself does not means smart when the home is built friendly to the environment, how space it uses, or using solar power and recycling wastewater, but what makes it smart is the interactive technologies that it contains [3].

In the present day, security systems play an important role in the protection of lives and investment. This is achieved by the incorporation of various subsystems into the security system with a single control unit such as surveillance, intruder control, access control, and fire detection. A smart home is one that is equipped with lighting, heating, and electronic devices that can be controlled remotely by smart phone or via the Internet [4]. Home automation refers to remotely monitoring the conditions of home and performing the required actuation. Through home automation, household devices such as TV, light bulb, and fan are assigned a unique address and are connected through a common home gateway. These can be remotely accessed and controlled from any PC, mobile, or laptop. This can drastically reduce energy wastage and improve the living conditions besides enhancing the indoor security [5].

In nowadays, development and changes of technologies are happening daily as well as continuous improvement of people's living standards is increasing. The mobile phones are the inspirable part of human lives today. The mobile phone is the most important part of human lives today. With the help of this smart gadgets, human can do many works with or without Internet like here we can make our home as well as organization smarter or more luxurious [6]. The concept of smart homes is becoming more and more popular. It is anticipated that radio frequency identification (RFID) technology will play a major role in such environments [7].

The Internet of things is the network of "things" which are connected to a common network path in order to communicate, exchange data, or control each other. The network path can be interconnected or interconnected with the "things" being either embedded software, hardware, or any sensor. It refers to the state where the things will have more and more data and information associated with them and have an ability to communicate, produce new information, and become the integral part of the free World Wide Web. It not only features Internet connectivity but also features cloud and data management, security management and all other fields concerned with the era of Internet [8].

In today's modern and smart trend people expect new devices and new technologies to simplify their day-to-day life. The innovators and researchers are always trying to find new things to satisfy the people but the process is still infinite. In the early 1990s, Internet connectivity began to proliferate in enterprise and consumer markets, but was still limited in its use because of the low performance of the network interconnects. In the 2000s, Internet connectivity became the norm for many applications and today is expected as part of many enterprise, industrial and consumer products to provide access to information [9]. We are experiencing a new era of Internet of things (IoT), where many electronic devices surrounding us are interconnected by a network. This paradigm enables copious amounts of data to be stored, processed, and conferred in a proficiently interpretable form without human invention. The emerging of IoT also sheds new light on the concept of a "smart home." IoT-enabled house equipment allows for a smart home to be more intelligent, remote controllable, and interconnected [10].

The theory of home automation has been around since 1970s. As much as IoT helping in automating tasks, the benefits of IoT can be extended for build up the current safety standards. In this modernized world use of home automation system has been increased due to numerous benefits in the terms of energy saving, safety, flexibility, and comfort, these significant new technologies offer [11]. Today we are living in twenty-first century where automation is playing an important role in human life. Home automation allows us to control household appliances like light, door, fan, and AC. It also provides home security and emergency system to be activated. Home automation not only refers to reducing human efforts but also energy efficiency and time-saving [12]. Figure 1 shows simple IoT devices for smart home environment.

The proposed system is a low-cost, home management, and monitoring system. Via NodeMCU board link to the Internet and several sensors controlled remotely using an android application on a smart phone. The NodeMCU microcontroller acts as a user to hardware interface all the atmospheric parameters like temperature,

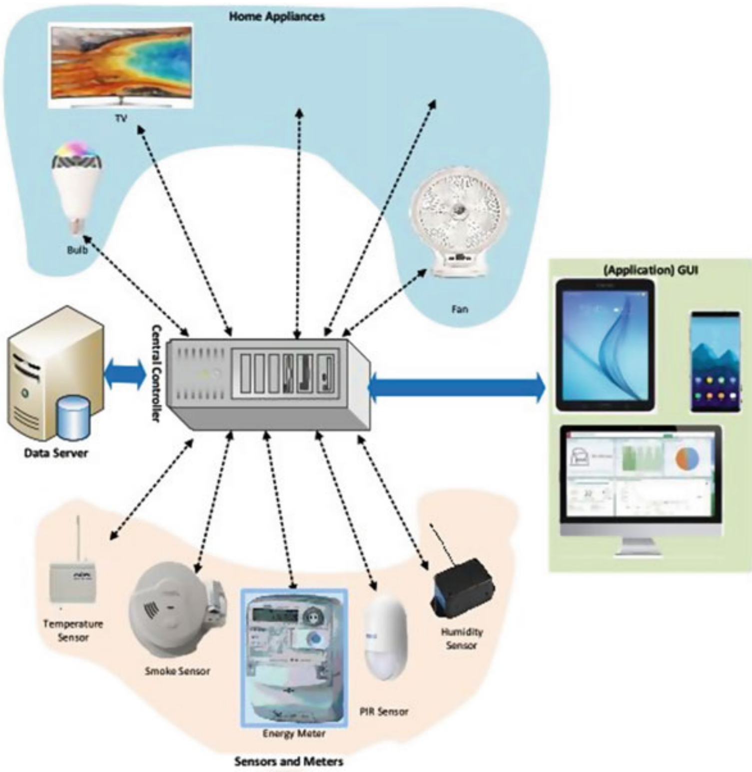


Fig. 1 Smart home environment

humidity, smoke/gas leakage sensor, and flame. We are calculating with the help of a sensor connected to Arduino Uno module. Then we are uploading the value on the server. NodeMCU is used as the wireless module by which we can provide the command with the help of an android application. Flame detection sensor and water pump sprinkle the water when the flame sensor detects the fire. Suppose any parameters exceeding the threshold value, then with the help of the GSM module, we can send a message to the concerned person. Solar panel is used to give power to the power supply instead of using batteries to reduce energy.

Several actuators are utilized to control household appliances such as light, fan, and water motor switching. In addition, a user-friendly user interface is being created to let consumers engage with the smart home. The proposed system intends to automate household appliances, minimize energy consumption using solar panels, improve safety and security, and improve overall quality of life and convenience.

Literature Survey

The theory of home automation has been around since 1970s. As much as IoT helping in automating tasks, the benefits of IoT can be extended for build up the current safety standards. In this modernized world use of home automation system has been increased due to numerous benefits in the terms of energy saving, safety, flexibility, and comfort, these significant new technologies offer.

I. Engineering. [2] Proposed the design and implementation of the smart home automation controller using Bluetooth for android mobile phone has been discussed. The purpose of this is to use mobile phone's inbuilt Bluetooth, Bluetooth serial module for automation of home appliances. The different hardware and software section of our system is described. The android application software has been designed using Eclipse and Android Studio software is used to write and burn the C program into microcontroller.

Davidovic and Labus [3] proposed a smart home system based on sensor technology. It is a new approach to utilize technology in a practical and meaningful manner within a smart home system that can be widely deployed into residential settings based on: group of sensors, Raspberry Pi device as a server system, and Bluetooth as a communication protocol.

David et al. [4] presented a low cost and flexible home control and environmental monitoring system. It employs an embedded micro web server in Arduino Mega 2560 microcontroller, with IP connectivity for accessing and controlling devices and appliances remotely. These devices can be controlled through a web application or via Bluetooth android-based smart phone application.

Kodali and Soratka [5] presented MQTT-based home automation system using ESP8266 which reports an overview of a lightweight Message Queuing Telemetry Transport (MQTT) protocol. In the prototype, we attempt to implement MQTT on ESP8266, a WiFi-based development board. Sensors and actuators are connected to ESP8266 and a mosquito-based MQTT broker is established for remote monitoring and control.

Kumar and Mehta [6] proposed home automation system enhances mobility and supports monitoring and control of devices from any remote location within WiFi range. Being a simple and user-friendly application, it serves as an application of great help to the old-aged or physically disabled people.

Ozeer et al. [13] presented the designing and implementing resilient IoT applications in the fog: A smart home use case is the design and implementation of a resilience approach on a realistic testbed in a Fog-IoT smart home environment. The resilience approach is designed taking into account the specificities of the environment.

Konindala et al. [7] proposed security framework for RFID-based applications in smart home environment that describes some of the RFID-based applications that are applicable to smart home environments. We then identify their related privacy and security threats and security requirements and also propose a secure approach, where

RFID-tagged consumer items, RFID-reader enabled appliances (e.g., refrigerators), and RFID-based applications would securely interact among one another.

Guptha and Chabara [8] presented an IoT-based smart home design using power and security management. It aims as multiple benefits of saving on electricity bills of the home as well as keep the users updated about their home security with an option of controlling the switching of the devices by using their voice or simple toggle touch on their smart phone, and last but most importantly, monitor the usage in order to conserve the precious natural resources by reducing electrical energy consumption.

Ganesh [9] proposed the implementation of IoT architecture for smart home using GSM technology that presents a web-based architecture for IoT using GSM for the implementation of smart home applications and GSM-based design of smart home controlling system.

Badabaji and Nagaraju [11] proposed an IoT-based smart home service system that uses IoT technology for the enhancement of safety standards. By using a WiFi module, the interfacing is done between transducers and the sensor network on a single chip solution wirelessly. They have used the IoT technology for gas leakage detector, fire detection and increase of temperature, which have alerting techniques for involving and sending text message to the particular mobile phone and taking safety measures.

Kaur et al. [12] presented home automation and security system. It is concluded that all the home automation system techniques use wireless technology. Arduino, GSM, and android-based home automation techniques have been implemented in order to provide ease to the people to control their home appliances (Table 1).

Proposed Methodology

The proposed system aims to expand the home automation system with additional sensors, actuators, and solar panels to power the control box instead of batteries to make the system safe, secure, energy-efficient, and environmentally friendly. The module consists of Arduino Uno as a primary controller, and NodeMCU is used to upload the data on the webpage and control the home appliances via Blynk application. The home appliances like fan, bulb, and water pump, we can control with the android application. Relay is acting as an electromagnetic switch, and with the help of this, we can maintain the electrical devices.

All the atmospheric parameters like temp, humidity, smoke/gas sensor, and flame. We are calculating with the help of a sensor connected to Arduino Uno module. Then we are uploading the value on the server. NodeMCU is used as the wireless module by which we can provide the command with the help of an android application. Flame detection sensor and water pump sprinkle the water when the flame sensor detects the fire. Suppose any parameters exceeding the threshold value, a notification message can be sent to the concerned person. Solar panel is used to give power to the power supply instead of using batteries to reduce energy. As shown in the block diagram in Figs. 2, 3.

Table 1 Summary of previous smart homes

SH system	Indoor Control	Outdoor Control	Security	Safety	Monitoring	Energy Management	Wireless Interface	Controller	Real Implementation	Smart phone	Web-based
Engineering [2]	Yes	Yes	No	No	No	No	Bluetooth/GSM	PIC	Yes	Yes	No
Davidovic and Labus [3]	Yes	No	No	No	Yes	Yes	Bluetooth/WiFi	Raspberry Pi	No	Yes	No
David et al. [4]	Yes	Yes	No	Yes	Yes	Yes	Bluetooth/WiFi	Arduino Mega	No	Yes	Yes
Kodali and Soratka [5]	Yes	No	No	No	No	No	WiFi	NodeMCU	No	Yes	yes
Kumar and Mehta. [6]	Yes	No	No	No	No	No	Ethernet	Arduino Mega	No	No	Yes
Ozeer et al. [13]	Yes	Yes	No	No	Yes	Yes	Fog-IoT	Raspberry Pi	Yes	No	Yes
Konindala et al. [7]	No	No	Yes	No	No	No	RFID	PC Server	No	Yes	No
Guptha and Chabara [8]	Yes	Yes	Yes	No	Yes	Yes	Ethernet	Galileo Board	No	Yes	Yes
Ganesh [9]	Yes	Yes	No	No	No	No	GSM	8051 μc	No	No	Yes
Yang, et al. [10]	No	No	Yes	No	No	Yes	WiFi	PC Server	No	No	No
Badabaji and Nagaraju [11]	No	No	No	Yes	Yes	No	GSM/WiFi	PC Server	No	No	Yes
Kaur et al. [12]	yes	No	Yes	No	No	Yes	GSM	Arduino	No	Yes	No

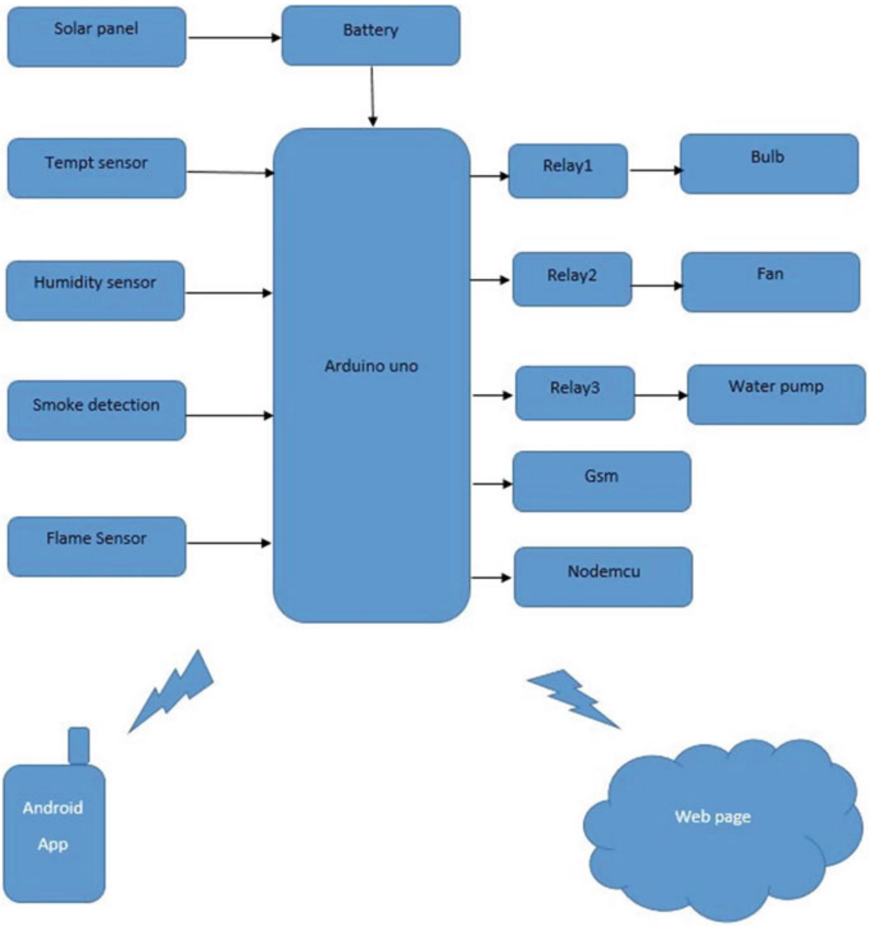


Fig. 2 IoT-enabled home automation block diagram

The solar panel generates energy by falling sunlight and gives charge to the battery for storage. After that the power supply will power the system via the relay and the NodeMCU ESP8266 modules, allowing all gadgets to work correctly. The sensor will be read by the NodeMCU ESP8266 microcontroller, sending the TCP/IP format data to the Blynk server. The Blynk server will also transmit TCP/IP instructions to the NodeMCU ESP8266 microcontroller, which will control the on/off of the house lights by relaying logic "HIGH" or "LOW" on specified pins. Cloud (Internet) by utilizing WiFi becomes the central connection between the Blynk application and NodeMCU. Blynk application and Arduino IDE preparation and this project is running by Blynk application.

Blynk application and Arduino IDE preparation and this project is running by Blynk application. Download the application to smart phone from the Google play

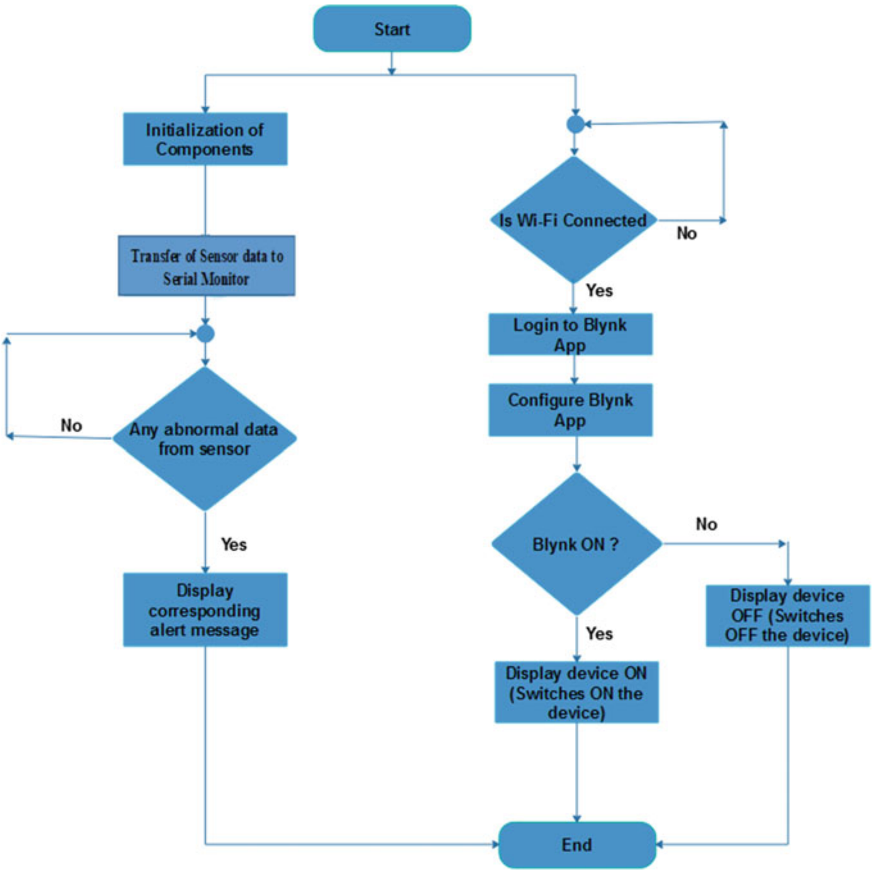


Fig. 3 Flowchart of system operations

store and then create a project on it with three switches. Set buttons to be switches on D1, D2, and D3. To control the appliances like bulb, fan, and water pump by Blynk application. The notification would be sent to the concerned person. The result of humidity and temperature sensor would be monitor through serial monitor and webpage. To program the NodeMCU using Arduino IDE, first add the NodeMCU to the Arduino IDE library by adding this address to the Arduino IDE settings. After adding this reference to the Arduino IDE, go to board management and choose NodeMCU 1.0. (ESP12EModule). Uploading code after adding NodeMCU to the Arduino IDE library and updating the hotspot name and password and the token code. The hotspot name and password are identical to those on the android because the relay input is included in [Blynk Run ();], the code does not need to identify it. When the Blynk program sends the auth (auth token) as an email, the SSID is the name of the smart phone hotspot.

Components Description

MQ-2 Sensor

We have used a MQ2 gas sensor in this system. MQ2 gas sensor is very popular sensor module to detect gas leakage. They are sensitive to a range of gasses and are used indoors at room temperature. The output is an analog signal or analog voltage which can be read with an analog input of the Arduino. If the predefined range of this analog value crosses, then Arduino will update an emergency status into the web server through the WiFi module [14, 15] (Fig. 4).

Flame Sensor

Fire sensor is the device used to detect the presence of flame or fire in the area where it is placed. Immediately after an area has been caught with fire, there will be significant increase in temperature of surrounding area and concentration of carbon dioxide and carbon monoxide increases in the atmosphere. User can then cause several actions to happen based on their requirement by interfacing it to several other components. A message can be sent to the concerned person and the water pump would be on through Blynk application by concerned person to sprinkle the water on fire [16, 17].

DHT 11 Sensor

See Fig. 5.

It is a digital sensor to monitor and gather information about the temperature and humidity that is prevailing in the smart home. It is operated in the voltages of the 3–5 V with the maximum current of 2.5 mA. Range for the temperature is fixed to zero degree to fifty degree, and the humidity percentage is set in between 20 and 80 percentages. The thermistor present inside engages a negative temperature coefficient and component to sense the humidity to identify the moisture content in the air [18] (Fig. 6).

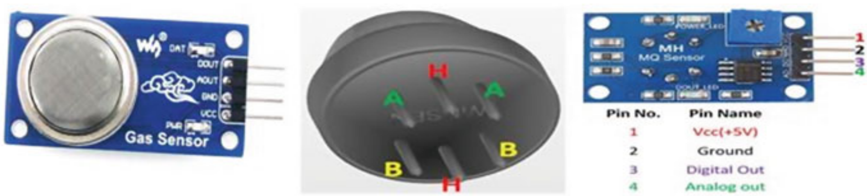


Fig. 4 MQ-2 gas and smoke sensor

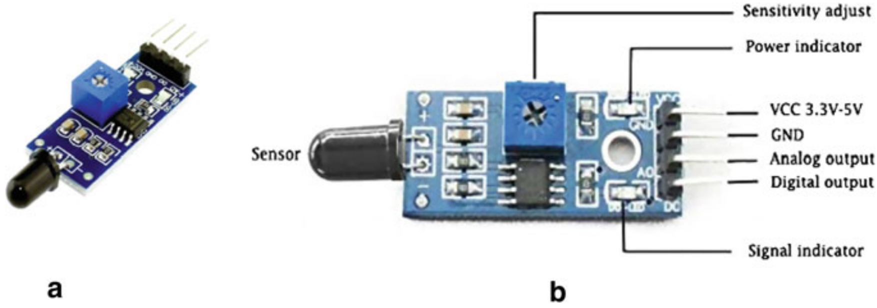


Fig. 5 Flame sensor

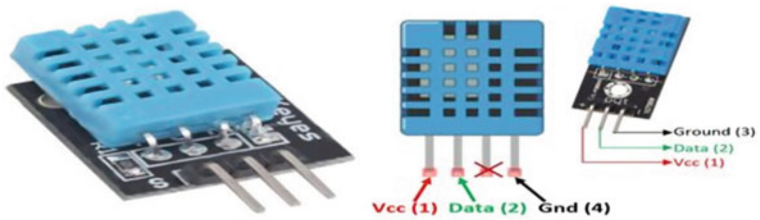


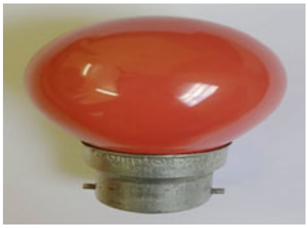
Fig. 6 DHT 11 sensor

Cooling Fan

A cooling fan that transfers energy to create the necessary pressure to maintain a constant flow of air. It is an axial kind, which means the air intake and outflow travel in a straight line along coaxial cylindrical surfaces. When the NPN transistor to which it is linked saturates, the brushless DC motor will be on (Fig. 7).

Fig. 7 Cooling fan

Fig. 8 Led light



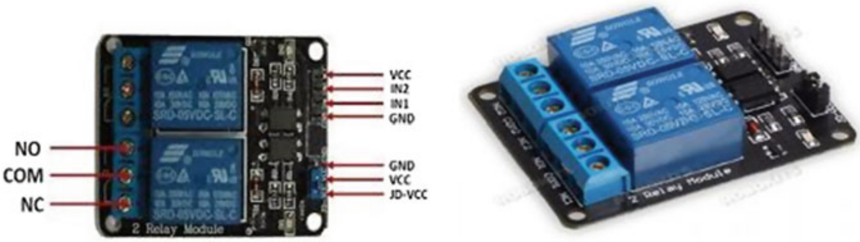


Fig. 9 Two-channel relay

Led Light

An LED lamp, often known as an LED light bulb, is an electric light that emits light from one or more light-emitting diodes (LEDs). The Blynk application is used to operate this LED, which is utilized for home automation. The smart phone is also used to control the LED automatically. We can, for example, turn on the lights before entering a room or home when traveling upstairs (Fig. 8).

Relay

A relay is utilized to carry out the exchanging activities for AC/DC devices. For example, in the recommended architecture, the cooling fan is controlled by a relay. As a result of the transfer, the cooling fan will switch on via Blynk application if the room temperature exceeds the breaking threshold (Fig. 9).

GSM SIM 800c

The SIM800c module is a 5–18 V power supply that can interact with a computer through a USBTTL module. It comes with a set of battery power supply interfaces for mobile phones, with a volume of 4.3 * 5.2 cm and two LED indications for simple troubleshooting. GSM supports text messages, telephones, GPRS data, HTTP protocol, DTMP decoding, MMS, recording function, TTS function, and more (Fig. 10).

Fig. 10 GSM SIM 800c

NodeMCU ESP8266

See Fig. 11.

The ESP8266 is a low-cost WiFi chip with full TCP/IP stack and microcontroller unit (MCU) capability produced by a Shanghai-based Chinese manufacturer, Espressif Systems. The low cost, compact size, and the presence of an inbuilt WiFi module were the reasons for selecting this microcontroller.

Arduino Uno

The Arduino Uno microcontroller board uses the ATmega328P CPU. There are 14 digital I/O pins on this board, as well as six analog inputs and a variety of additional ports and jacks. The Arduino may be powered directly from a computer with the included USB cable, or via an AC-to-DC adapter or battery. The Arduino IDE, which is based on the C programming language, may be used to program the Arduino. It is a small, user-friendly gadget that can handle a wide range of tasks [19, 20] (Fig. 12).

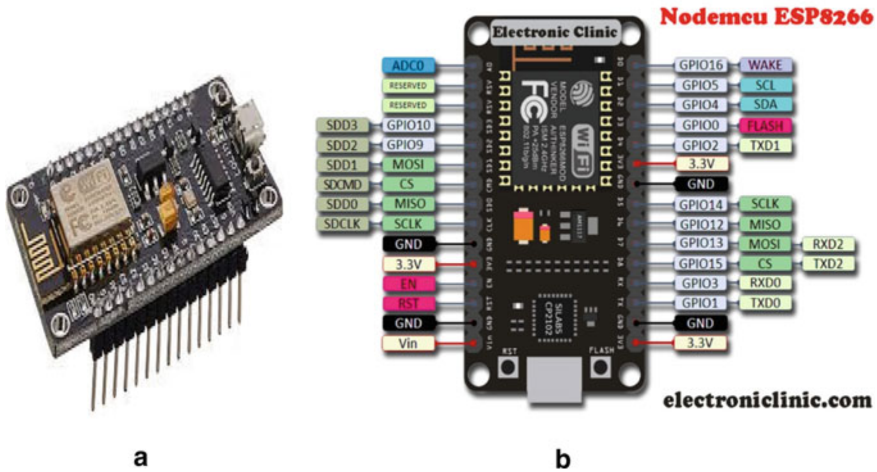


Fig. 11 NodeMCU ESP8266

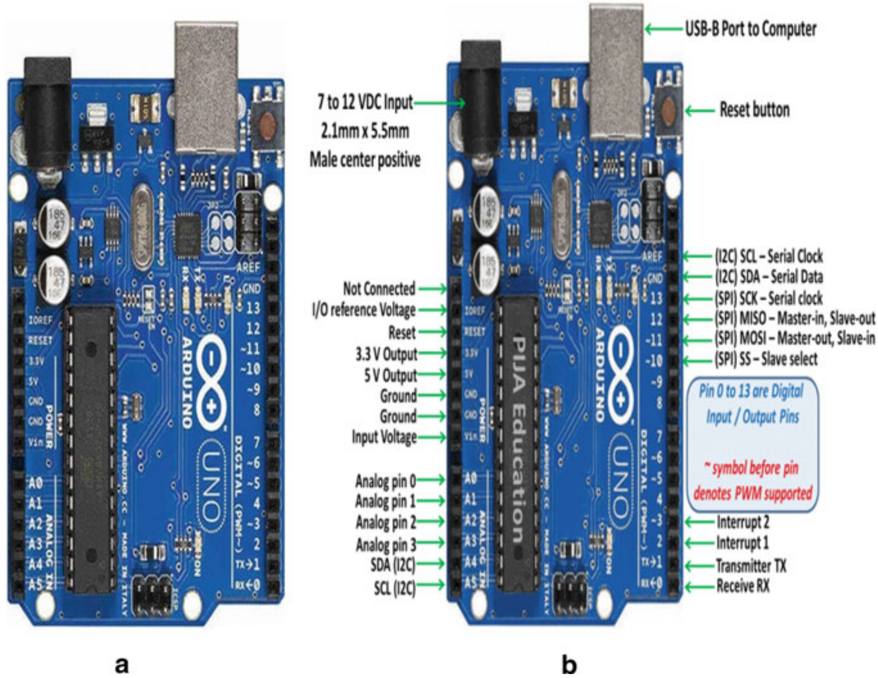


Fig. 12 Arduino Uno AT mega 328P

Software

Arduino IDE the ESP8266

It is programmed using an open-source Arduino IDE, which makes writing and uploading code to the ESP8266 a breeze. This software is design in an easy to understand language and runs on Windows, Linux, and Mac operating system. It comes with a Java-based processing environment and straightforward installation procedure [16].

Experimental Results

The design and implementation of the IoT-enabled home automation system have been discussed.

The different hardware and software section of our system are described. The purpose of this study is to develop a user-friendly, secure, energy-efficient, and low-cost automation system for smart homes based on IoT using an Arduino board with NodeMCU WiFi being remotely controlled by any android OS smart phone. As technology is advancing so houses are also getting smarter. Modern houses are gradually shifting from conventional switches to centralized control system, involving remote-controlled switches. Presently, conventional wall switches located in different parts of the house makes it difficult for the user to go near them to operate. Even more it becomes more difficult for the elderly or physically handicapped people to do so. Remote-controlled home automation system provides a most modern solution with smart phones. In order to achieve this, a NodeMCU WiFi module is interfaced to the Arduino board at the receiver end while on the transmitter end, several sensors and actuators were connected to the NodeMCU controller, which updated the data to the IoT server. The obtained data from the sensors (temperature, humidity, flame, gas, and smoke) can be monitored via serial monitor via laptops/PC. A GUI application on the smart phone sends ON/OFF commands to the receiver where loads are connected. For security and safety purposes, the user receives notifications on their mobile phones about any abnormal condition at home. The solar panel and rechargeable battery are used to give power to the power supply instead of using batteries to reduce energy and make the smart home more energy-efficient (Figs. 13, 14, 15, 16).

Conclusion and Future Work

This study presented the design and implementation of a user-friendly, secure, energy-efficient, and low-cost automation system for smart homes based on IoT. The developed IoT-enabled home automation system can be easily implemented

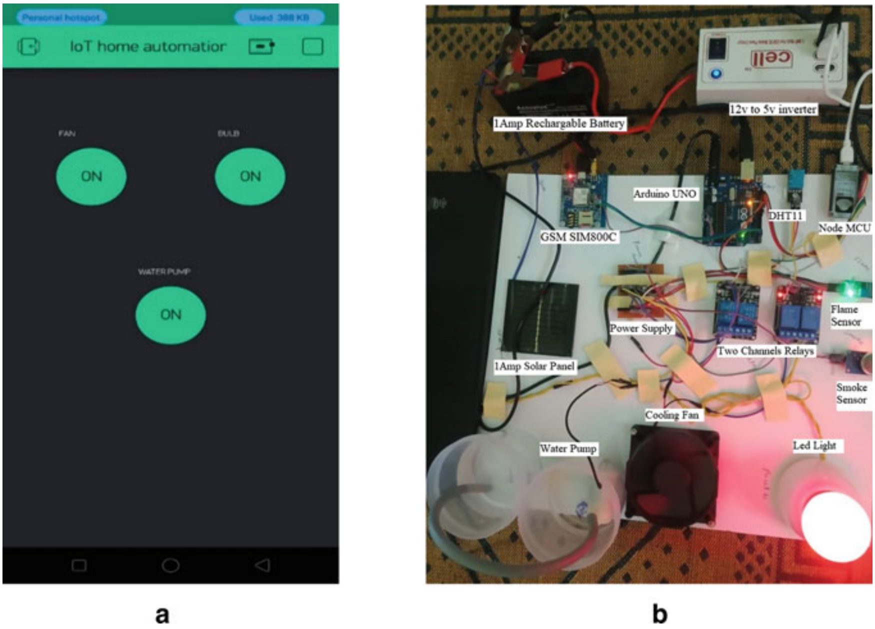


Fig. 13 Controlling developed module appliances via Blynk application

Fig. 14 COM port result for humidity and temperature notification

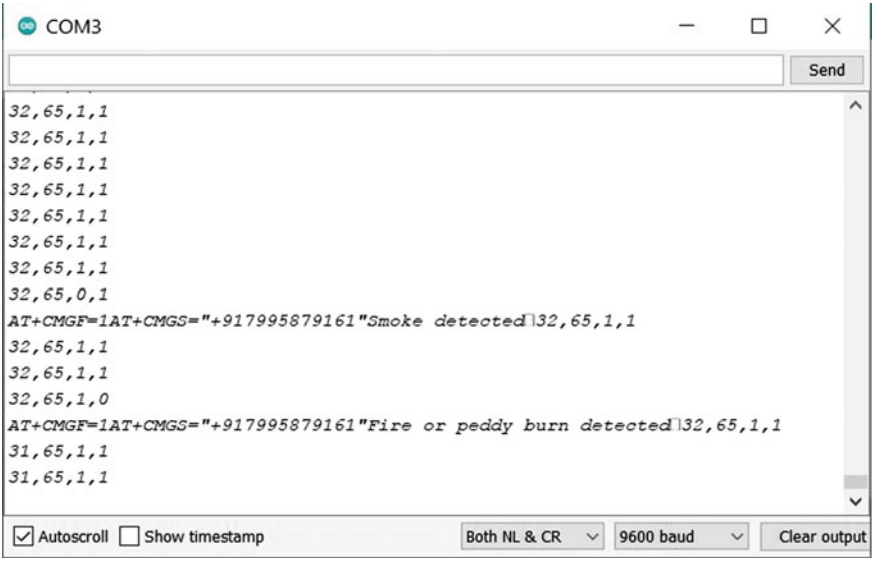


Fig. 15 COM port result for smoke and flame detected reading

in a real house to allow real-time monitoring of home conditions and control of home appliances. Several sensors and actuators were connected to the NodeMCU controller, which updated the data to the IoT server. The obtained data from the sensors (temperature, humidity, flame, gas, and smoke) can be monitored via serial monitor via laptops/PC. For security and safety purposes, the user receives notifications on their mobile phones about any abnormal condition at home. Control of home appliances can be easily and efficiently conducted by using android OS Smart Phones with the help of Blynk application. The results of this study are promising, and the developed system can increase the safety, security, intelligence, energy efficiency, and comfort of users. For the future work, the proposed system can be expanded with RFID Door Lock system and with additional sensors and actuators. The developed system can also be improved to make it suitable for future commercialization. Our next study will optimize all circuits using printed circuit boards to save space and minimize the risk of connection losses or short circuits.

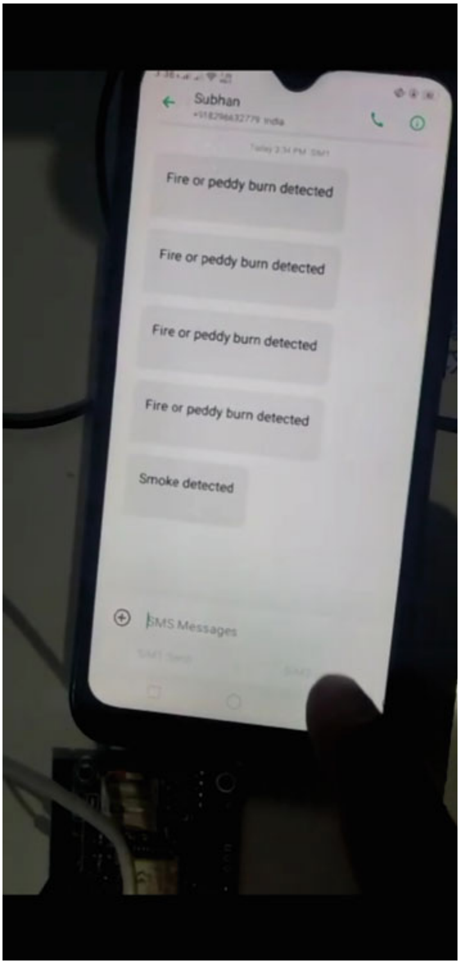


Fig. 16 Flame and smoke detected notification SMS via GSM

Acknowledgements The authors would like to express their gratitude to the contributors of the GITHUB repository's publicly accessible frameworks, libraries, and sketching datasets, which aided training, testing, and validation.

References

1. Jabbar WA, Member S, Kian TEEKOK, Zamrizaman NSM, Balfaqih M (2019) Design and fabrication of smart home with internet of things enabled automation system. IEEE Access 7(Figure 1):144059–144074. https://doi.org/10.1109/ACCESS.2019.2942846
2. I. Engineering (2017) Website : www.ijareeie.com Home Automation System Using Android via, pp 2298–2301. https://doi.org/10.15662/IJAREEIE.2017.0604016

3. Davidovic B, Labus A (2017) A smart home system based on sensor technology A smart home system based on sensor technology Feb. https://doi.org/10.2298/FUEE1603451D
4. David N, Chima A, Ugochukwu A, Obinna E (2015) Design-of-a-Home-Automation-System-Using-Arduino.doc 6(6):795–801
5. Kodali RK, Soratkal S MQTT based home automation system using
6. Kumar S, Mehta N (2017) Smart home system using IOT, pp 1108–1112
7. Konidala DM, Kim D, Yeun CY (2011) Security framework for RFID-based applications in smart security framework for RFID-based applications in smart home environment. https://doi.org/10.3745/JIPS.2011.7.1.111
8. Chhabra J (2016) IoT based smart home design using power and security management Iciccs:6–10
9. Ganesh EN (2017) Implementation of IOT architecture for SMART HOME using GSM technology abstract Mar
10. Yang H, Lee W, Lee H (2018) IoT smart home adoption : the importance of proper level automation 2018
11. Badabaji S (2018) An IoT based smart home service system 119(16):4659–4667
12. Kaur S, Singh R, Khairwal N, Jain P (2016) H a s s 3(3):17–23. https://doi.org/10.5121/acii.2016.3303
13. Ozeer U et al (2019) Designing and implementing resilient IoT applications in the fog : a smart home use case to cite this version : HAL Id : hal-01979686 Designing and implementing resilient IoT applications in the fog : a smart home use case
14. Mahamud S, Zishan SR, Ahmad SI, Rahman AR, Hasan M, Rahman L (2020) Domicile—an IoT based smart home automation system. In: 2019 international conference on robotics, electrical and signal processing techniques (ICREST) No February 2019, pp 493–497. https://doi.org/10.1109/ICREST.2019.8644349
15. Sarmah R (2019) SURE-H : a secure iot enabled smart home system. In: 2019 IEEE 5th world forum internet things, pp 59–63
16. Jeniton1 S, Thulasi K, Sivagami3 S, Aripreetha D (2020) IoT based smart home FGM detection system. Int Res J Eng Technol (IRJET) 07(05), e-ISSN: 2395–005, May 2020 www.irjet.net
17. Hamdan YB (2021) Smart home environment future challenges and issues—a survey. J Electron 3(01):239–246
18. Raj JS (2019) Automation using Iot in greenhouse environment 01(01):38–47
19. Singh H, Pallagani V, Khandelwal V, Venkanna U (2018) IoT based smart home automation system using sensor node. In: 2018 4th International Conference on Recent Advances in Information Technology (RAIT), pp 1–5
20. Hrestha S, Shakya S (2020) A comparative performance analysis of fog based smart surveillance system. J Trends Comput Sci Smart Technol (TCSST) 2(02):78–88

Generating Recommendations for Various Problems Using Data Mining and Machine Learning Algorithms

L. M. R. J. Lobo and Kale Manoj Birbal

Abstract The system available today which generates recommendations is supported by the user's information collected within the past. It does not include intention at a specific time. The utilization of data mining and machine learning algorithms offers real-time recommendations employing a fitness function and models that estimate the suitability of recommended lists. The utilization of a social network provides user-generated content during a much more convenient scenario. It also uses a user experience. Recommendation systems play an outsized role in providing quality to those. In this paper, we present how recommendations are often produced using data processing and machine learning algorithms. Three systems with varied applications are addressed during this paper. First, a recommendation is used for better crop cultivation using certain parameters of concern with the assistance of association rule mining and genetic algorithms. Second, a way to recommend videos for advertisements from those available with titles, descriptions, and hashtags as extracted features is presented. These use machine learning-associated multi-label classification algorithm. Finally, an anti-vice recommendation system that uses neural networks to get recommendations is brought forward. Altogether the three cases, it is observed that the accuracy and efficiency of recommendations are upgraded.

Keywords Genetic algorithm · Neural network · Facial image · Association rule mining · Recommender system · Multi-label

Introduction

Recommender systems have gained popularity since society to an excellent extent relies on recommendations. Now, an equivalent idea is employed within the past decade for E-commerce. Data collection within the explicit and implicit form helps in generating this recommendation. A system that generates recommendations should assign a high rating or prediction to an item counting on the characteristics of the item,

L. M. R. J. Lobo · K. M. Birbal (✉)
Department of Computer Science & Engineering, Walchand Institute of Technology, Solapur, India

V. Bindhu et al. (eds.), *Proceedings of Third International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 844,
https://doi.org/10.1007/978-981-16-8862-1_63

and therefore, the views expressed by people that have an interest within the item. Such recommender systems use context and collaboration in their processing. There are three types of recommendation systems: content-based recommender systems (CBRS), collaborative filtering recommender systems (CFRS), and hybrid recommender systems. These systems are used for various problems to generate appropriate recommendations.

Recommender Systems for Agriculture

Nowadays, the storage of digital computers has enabled us to manage a large amount of information with its use in almost all fields. Information technology is changing the face of the planet since it inherently depends on digitization. India being an agriculture-dependent country needs support for its development which depends on the result of agriculture. Though we have a large amount of data about agriculture, still there is a dirth in the amount of useful knowledge it is converted. A farmer is enabled with the output which the analysis is completed on the available information and knowledge to direct him to require proper timely recommendations to realize a profit on the produce grown. This will be through with the utilization of knowledge mining algorithms.

Recommender Systems for Advertisements

Machine learning provides systems the power to find out and improve from experience automatically without being explicitly programmed. Machine learning is aimed toward understanding the info structure and fitting that data into models that folks can understand and use. These algorithms need support to coach the system using predefined labels. As the availability of knowledge is gigantic and it is increasing rapidly, so there is an important aspect which is to be organized. Multi-label classification may be an expansion of multiclass classification; multi-class is a single-label classification problem that will be categorized into two or more classes. However, in multi-label classification, there is no restriction on the number of labels; one single instance is often assigned to multiple classes. Multi-label is emerged from the analysis of the text categorization problem, where each report could even be a neighborhood of several predefined categories; simultaneously an enormous problem is that the multi-label identification of text and image data. Definitions vary from pieces of media to messages. In contrasting multi-class versus multi-label, dissimilarity in multi-class and multi-label is often explained with the subsequent justification. The components belong to at least one and therefore the only group in multi-class, while one or more groups are often allocated to every element in multi-label.

Recommender Systems for Vice Detection

The accidents caused due to a driver being drunk have emerged to be a social problem. An equivalent is observed in other vices too. In these systems, a neural network with three layers is often used to identify a drunk person based on portions of the face. In the training of the neural network, a rear propagation algorithm is used. Face images are used as input. These images consist of both drunk and non-drunk individuals. Neural networks are popularly used for speech, image and pattern recognition involving synthesis and applications.

Related Work

The agriculture practice today is modernized and not the same as practiced by our ancestors. Global warming as produced effects on climate and climatic changes has become unpredictable [1]. The development of technology and data processing methodologies provides better results and services. Processing of data finds its application within the domain of agriculture within the near past [2]. The type of soil and associated factors make the farmer decide upon the position and the type of crop to be grown along with fertilizers and pesticides to be used. This causes a highly improve yield [3–5]. Data processing is equipped with tools and techniques to deal with very forms of data of which agricultural data is one with a huge contribution.

The different methods of knowledge extraction that are set over volumes of knowledge generate hidden information as the pattern from it [6]. A genetic algorithm is used as an optimization technique [7–9].

Nasira [10] worked on the prediction of vegetable prices and used the properties of artificial neural networks that self-adapt, self-learn, and are tolerable to errors to create the higher models of neural networks. Veenadhari [11] used data processing methods for finding crop yield within the field of agriculture. Sumitha and Kirubakaran [12] worked on an intelligent digital information management system. This technique was to be utilized by Indian farmers communicating weather information on their mobiles. ArvindJaiswal, GauravDubey [13] have used a method of creating association rules and optimizing them using a genetic algorithm. The proposed genetic algorithm-based method for locating frequent item sets. They transform individuals using the basic step of simple genetic algorithms and other machine learning algorithms [14]. The most advantage of using machine learning is that when an algorithm knows what to try to with the knowledge, it can do its job automatically. The multi-label model details, also because most areas of operation, provided us with the context needed to understand the works examined.

Jiang provided a multi-label object categorization approach using convolution and RNNs jointly [15]. For the classification of dataset enclosed images with one label, CNNs are extremely efficient. Kwangsoo suggested a system using multi-labels for identifying videos on YouTube data [16]. This was enhanced by using the VLAD and

FV models on the network. Piotr introduced scikit-multi learns: a multi-label Python environment supported a scikit [17]. This assortment is firm with the environment of the scikit/scipy, and sparse matrices are used for all internal operations.

Hermosilla [18] showed the potential to recognize the identity of a private person using only information from thermal images of the face to show whether the person is drunk. The proposed system has a dual-step procedure of face recognitions and classifications. In the facial recognition phase, the test images were recognized with robust facial recognition techniques: Weber's local descriptor (WLD) and native binary pattern (LBP). For classification, Fisher linear discrimination was used to scale back the dimensionality of the features, and people features were classified employing a classifier supported a mixture model that was of a Gaussian type sorting room for the person that expands on the cutting-edge concept of a "Drunk Room Sorter."

El-Harby [19] elaborated on the active research areas using face recognition. Zhao [20] discussed several successful applications of image analysis and image understanding; Facial recognition has received a lot of attention lately, especially in recent years. They provided an up-to-date critical overview of still image and video-based facial recognition research. Using the general public Faces of Meth database, Tekkam Gnanasekar [21] presented that facial attribute recognition is also used to identify selected attributes that indicate ongoing illegal drug abuse. A deep learning network was used for extracting features and attributes; this was then supported by a Bayesian network. Yadav's [22] highlighted the effect of illicit substance abuse on the countenance. Pandey [23] used compression as a comparatively recent technique that supported the representation of a picture by operations of special transform on space of the image. Liew [24] considered a strong four-layer CNN architecture for recognizing a face and working on the images. Pranav kilobyte and Manikandan [25] found that the arrival of high-speed processors and high-resolution cameras had spearheaded analysis on the look of face recognition systems for numerous applications. Manojkrishna [26] demonstrated that image classification may be a classic problem in image processing, computer vision, and machine learning. Xin and Wang [27] commented supported the backpropagation algorithm which tries to reduce the errors and trained the neural network.

Comparison Between Existing Work and Our Work

Table 1 gives the comparison between existing work in the literature review and our work.

Table 1 Comparison between existing work and our work

Problem	Existing work	Our work
Recommendation using association rule mining and genetic algorithm	GA is used as an optimization technique, Apriori algorithm is used to generate rules, use of ANN	Eclat algorithm-based method for locating frequent item sets and generating rules and GA for optimization
Advertisement recommendation related to the content of the video	Multi-label object categorization approach using convolution and RNNs jointly, use of the VLAD and FV models	Use of tongue processing clean technique and lemmatization. Multi-class and multi-label combination for classification
Anti-vice recommendation system based on facial images	Only considered as color balance change of personal face, in the existing system check, the chin, neck, ear, hand, etc., these features used neural network performed for only drinking	Along with chin, neck, ear, and hand, the check, eye, and nose are also used CNN & backpropagation algorithm. Performed for drinking, drug abuse, and tobacco chewing and smoking

Methodology

The methodology utilized in this paper is split into three parts first for generating recommendations using data processing algorithms like association rule mining blended with an optimization algorithm, namely genetic algorithm, second for generating recommendations for advertisement associated with the content of the video, and therefore, the third generating recommendation for anti-vice detection supported face images using machine learning.

Generating Recommendations Using Association Rule Mining and Genetic Algorithm

An optimized agricultural crop recommender system was developed employing a data processing technique, Éclat algorithm, for agriculture application. We generate recommendations on the use of association rule mining and the use of the genetic algorithm approach for agricultural historical databases to optimize the result. This system providing recommendations considers a feature map of the sector and provides a stimulating result that can be mapped to recommend a crop to the farmer. The amount of recommendations varies from farm to farm. It is a general approach that supports and confidence should be high to urge the foremost efficient rules. These two parameters identify items that are used regularly from datasets and generate required rules. The system flow is seen in Fig. 1.

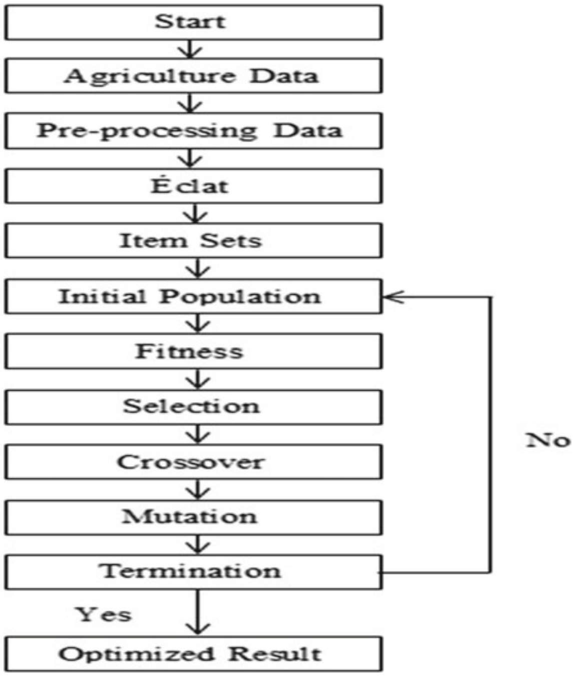


Fig. 1 System architecture for generating recommendations using association rule mining and genetic algorithm

Éclat algorithm is applied to the dataset of agriculture which contains tuples collected in the past. It is then used to generate a frequent item set. By the application of appropriate support to every rule, a final item set is generated. Data of Solapur Districts of Maharashtra state was taken for experimentation. It contains data of a specific period for crops cultivated there. This was used as the dataset in raw form. This was then converted to the required form using preprocessing technique of data mining. Six parameters within the database as Ph for soil Ph., E.C for soil E.C., O.C for soil organic contain, N for soil nitrogen, P for soil potassium, and K for soil phosphorus were used. Before giving the agricultural datasets to the association rule generator Éclat algorithm, we preprocess that data by converting the values of classification to the specified format. From Éclat algorithm, we get the item sets which further are going to be employed by the genetic algorithm. In preprocessing, we classify values that depend upon the range of the soil parameter in two categories, high and low. We apply the Éclat algorithm for obtaining the item sets. To urge association rules, on each item set, we apply support and confidence. At now, we apply a threshold value of support to urge item sets that produce good association rules. The six values are converted into string values as given in Table 2.

The necessary transaction from a defined transaction set FK for every item was identified. For every item {a}, the list containing {a} is identified for all I_jFk where $i < j$, $N = I_i \cap I_j$ we then compute the worth of support S. $S(X \rightarrow Y) = \sigma(X \cup Y)/N$ if N.sup $\geq$ min − sup then $Fk + 1 = ((Fk + 1) \cup (N))$; we repeat this step till we find good results. We apply the intersection operation to sets of generated elements.

Table 2 String values for parameters

Parameter	Range	Siring value
Ph(i)	3.5–6.51	Low
..	7.50–9.1	High
E.C(i)	0.41–0.81	Low
..	1.60–2.6	High
O.C(i)	0.21–0.40	Low
..	0.81–1.00	High
N(i)	141–281	Low
..	281–701	High
P(i)	7.0–14.0	Low
..	14.1–36.0	High
K(i)	101–1 0	Low
..	151–300	High

The genetic algorithm is used to optimize the results of the first stage that is within the initial population (Éclat generated item sets are used because of the initial population for genetic algorithm) from which association rules are generated.

Advertisement Recommendation Related to the Content of the Video

The proposed system for advertisement recommendation associated with the content of the video is going to be developed as seen in Fig. 2.

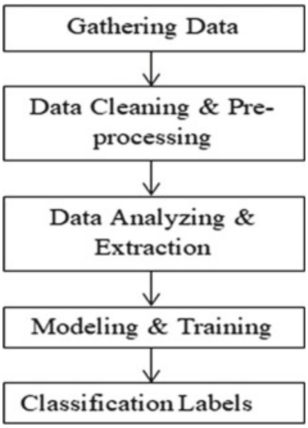


Fig. 2 An overview of the proposed system

Answers to relevant questions are fetched within the data collection process and may determine results from a long-time system by collecting and estimating information on targeted variables. The decision is taken to use YouTube data within the proposed system which is out there on the Web. Alternatively, another method is to develop your database using the YouTube API v3 developed by Google itself for various programmers to speak with YouTube.

In the first step, we process missing data. Since missing values are alleged to be text data within the proposed system, we have to mandatory replace them. The tongue processing clean technique is then used to process text data. Numerical values and punctuation are removed. White spaces are separated by one white space, for tokenization. Eliminating non-alphabetic words and "stop words," ask for words like and, that is, etc., which are important words in learning sentence structure, but which are not useful in predictive analysis. Lemmatization is then used to find the actual root word.

Data is then extracted, cleaned transform, and modeled. Classification labels are methods of producing labels that are the result of the classified model. The labels are going to be associated with the content, so, from labels, it is possible to find what the content is all about.

Anti-Vice Recommendation System Based on Facial Images

The proposed system for the anti-vice recommendation system supported facial images is projected as shown in Fig. 3.

The six modules of the system include image extraction, compression, pixel extraction, and a three-layered neural network, a neural network comparator to match with the info set, and a messenger to point vice detection. The extractor extracts the first facial images from the input; it then compresses the pictures and converts them into an array of pixels. These are then fed to the initial layer of a three-layer neural

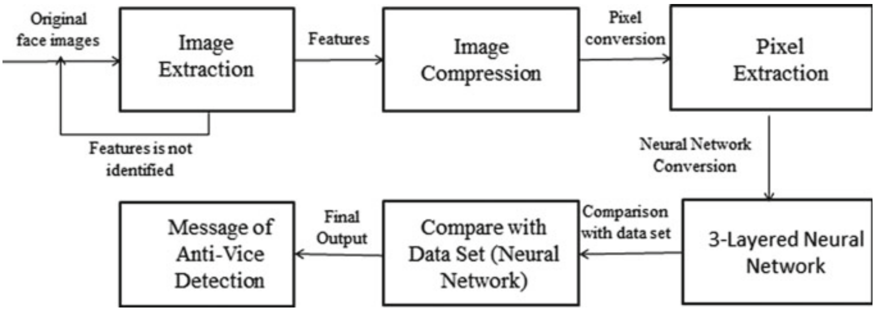


Fig. 3 System architecture

network for training. The final result of this three-layered network is compared with the testing data to get a prediction in the form of vice detection messages.

Detailed Step in Methodology

The process starts by giving facial images as input to our system. This database is created having images of drunk and non-drunk subjects. The images are captured with a high-resolution camera. Then, extract features from facial images like cheeks, neck, chin, ears, etc., as parts to be used for classification.

Figure 4 shows an example of facial images and reduced cheek images extracted from the facial images.

After extracting these features, the compression technique is applied to compress images to a hard and fast size. After the face image is compressed to the fixed size, extract the RGB values from the face images and then normalize the RGB values.

The RGB values of input images are normalized in their matrix form and given to the initial layer of the neural network. Here, we use a backpropagation algorithm to coach neural networks by giving facial images before and after the vice.

We trained a connected three-layer neural network for alcohol consumption detection. We trained a neural network with training patterns and target signals. Once we input a training pattern, the output units of the neural network show some values. We modify the values of the weights that connect the units, so that the output units can display the target signal. An adjustment in the form of feedback from the output layer regarding the weights is dealt with using the popular backtracking algorithm. The training is repeated until the error energy is below a threshold. The knowledge for the input layer contains normalized color values from a facial image. Output units indicate no, light, and heavy drinking and similar results for various vices. Finally, the output level shows vice or novice. Check the detection accuracy using test data.

(a) a face image (b) the resized cheek image

Fig. 4 An example of face images and resized cheek images

Experimental Setup

Generating Recommendations Using Association Rule Mining and Genetic Algorithm

The system was developed in Java since later networking needs would be satisfied. It had been their interface to the R package since only generic interfacing code might be written, and facilities already available within the R package might be used. The Éclat algorithm in R was connected to the straightforward genetic algorithm which was developed in Java to get optimized association rules.

Generic programming was written to form the connectivity. Verification of the implementation was done using an agricultural dataset that possessed data of various crops.

Advertisement Recommendation Related to the Content of the Video

The system was developed on hardware GPU having RAM of 12 GB and a processor AMD Ryzen 5. The software required to develop the system was Python along with its libraries pandas, NumPy, Matplotlib, seaborn, beautiful soup (Web scrapping), OpenCV and TensorFlow 2. AWS EC2 instance is employed to coach deep learning models on the cloud using the Amazon S3 bucket as a public storage.

Anti-vice Recommendation System Based on Facial Images

To develop the system, OpenCV is employed for face detection using image processing. A robust computational system, i.e., a GPU AMD Radeon, is employed for processing. A random-access memory of a sizable amount of 8 GB and above is adequate. A 32 or 64-pixel camera is required to capture the pictures. Python is employed for processing alongside its libraries. The dataset used was created in real-time containing 1052 images of faces of individuals.

Result and Discussion

Generating Recommendations Using Association Rule Mining and Genetic Algorithm

A sample of results of the Apriori algorithm and Éclat-Genetic algorithm was noted, and graphs were plotted as shown in Fig. 5. It is often visualized from the graph that the prediction accuracy is about 28.31%.

Based on the above graph, it is seen that the Éclat-Genetic algorithm performs well because the prediction accuracy is 28.31%. Positive and negative rule generated comparisons also are shown between Éclat-GA and Apriori algorithm in Figs. 6 and 7. The proposed model gives better results than an Apriori in terms of accuracy. We believe that our proposed system is often modified to handle generally, any database. It is also seen, and for given parameters of soil, there is a recommendation suggested for a specific crop to be grown by a farmer.

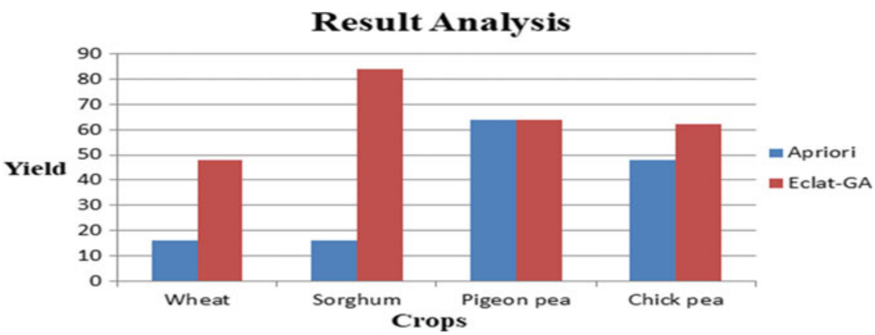


Fig. 5 Result analysis of using Éclat in place of Apriori

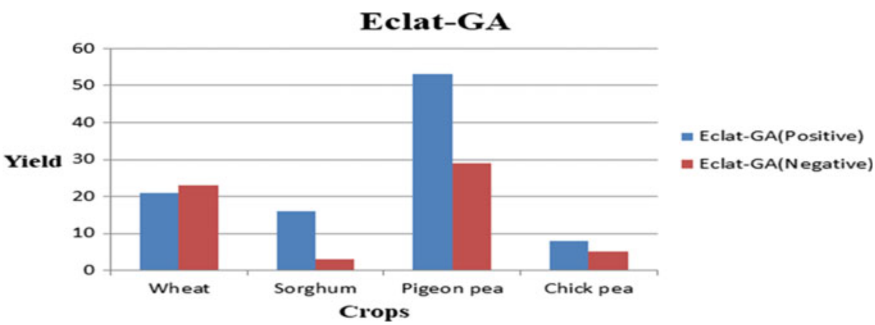


Fig. 6 Éclat-GA rule analysis

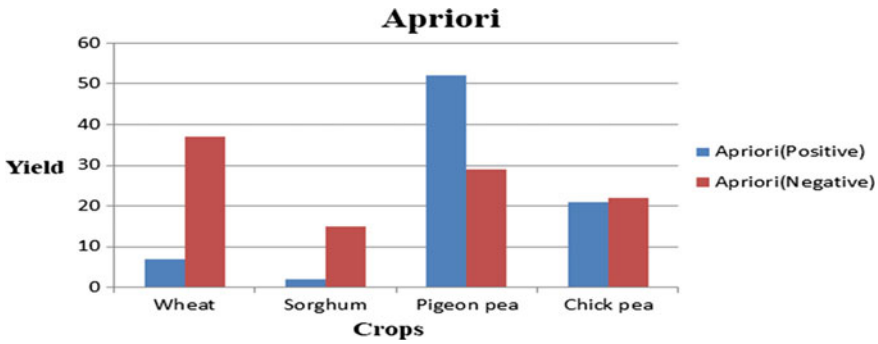


Fig. 7 Apriori rule analysis for positive and negative rules

Advertisement Recommendation Related to the Content of the Video

Text Classification

Testing is performed taking a sample of six categories from varied domains like Science and Technology, Art and Music, Travel, Manufacturing, Food, and History. The LSTM classifier is used for experimentation.

Figure 8 shows a confusion matrix of the result of an LSTM classifier. The rows correspond to the category, and the columns show instances of the category computed by the classifier. For instance, the NB classifier has correctly classified 394 instances of the health class (true positives, TP) while 25 instances are false positives (FP) (Fig. 9).

When we consider precision, recall, f1-score, and support, it is often seen that the LSTM classifier consistently gives the best performance. For all the six categories, thus, the ads recommendation system dynamically recommend the relevant ads up to 98%.

	art and music	food	History	manufact uring	science and technology	travel
art and music	394	2	3	3	0	1
food	5	425	0	1	6	1
history	6	0	411	0	2	2
manufacturing	2	0	0	395	0	0
science and technology	5	1	2	4	391	0
travel	7	6	0	0	0	425

Fig. 8 Confusion matrix

	precision	recall	f1-score	support
art and music	0.94	0.98	0.96	403
food	0.98	0.97	0.97	438
history	0.99	0.98	0.98	421
manufacturing	0.98	0.99	0.99	397
science and technology	0.98	0.97	0.98	403
travel	0.99	0.97	0.98	438
accuracy			0.98	2500
macro avg	0.98	0.98	0.98	2500
weighted avg	0.98	0.98	0.98	2500

Fig. 9 Accuracy by class

	precision	recall	f1-score	support
football	0.88	0.95	0.92	196
tennis	0.90	0.91	0.90	179
weight_lifting	0.97	0.84	0.90	143
accuracy			0.91	518
macro avg	0.92	0.90	0.91	518
weighted avg	0.91	0.91	0.91	518

Fig. 10 Accuracy by class

Video Classification

Because of the exorbitant cost of time used and processing power required by most of the classifiers, and since the experiment possesses is to be repeated 10–100 times to identify the simplest classifier to be used, experiment was done stage by stage (Fig. 10).

With efficiency parameters like precision, recall, f1-score, and support, it is often seen that the video classifier consistently performs 91% accuracy which is appreciable. We show a comparison of the result in Fig. 11.

In Table 3, the labels along with a respective computed score of the pretrained model are presented.

The results of the recommendation system show typically the predicted advertisement representing the keyword identified by the designed convolutional neural network model. Sample keywords representing advertisements are retrieved from the dataset containing advertisements stored with columns indicated by the title and description of the source video. Only the matched data is taken into account for the CNN to identify particular sports. Thus, the classification of videos is achieved by CNN being used as a classifier.

Fig. 11 Output of inception model

Table 3 Pretrained model results comparison tennis

Tennis (score 0.09849)	Fencing (score 0.05321)	Kabaddi (score 0.03406)	Wrestling (score 0.01801)
Table tennis (score 0.09332)	Hockey (score 0.04715)	Baseball (score 0.03229)	Basketball (score 0.01588)
Gymnastics (score 0.09309)	Chess (score 0.04510)	Shooting (score 0.03163)	İce hockey (score 0.01523)
WWE (score 0.09008)	Football (score 0.04218)	Volleyball (score 0.02345)	Motogp (score. 0.01435)
Formula 1 (score 0.07433)	Swimming (score 0.03859)	Weight lifting (score 0.02197)	
Boxing (score 0.06044)	Cricket (score 0.03547)	Badminton (score 0.02169)	

Anti-vice Recommendation System Based on Facial Images

Figure 12 indicates that data is been trained with the epoch of 20 iterative steps, it is used to train the model for detection of the drunken faces, and in the below graph, we can see that the accuracy is been pointing at a 74% score level.

Figure 13 tells about the accuracy counting average, whereas precision tells us that measure feature area of the valid and drunk faces where 0.27 is precision condition differentiation found for the valid faces that we have 0.67 precision differences in

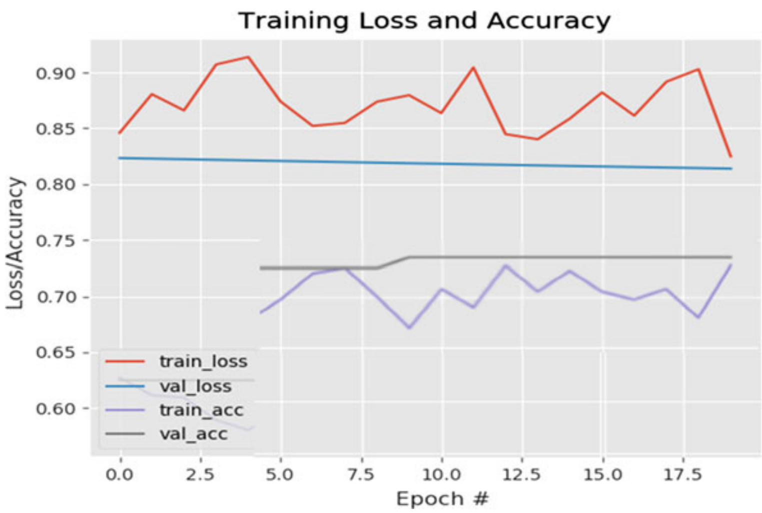


Fig. 12 Drunken detection training graph

	precision	recall	f1-score	support
drunk	0.27	0.08	0.13	36
valid	0.67	0.89	0.77	76
accuracy			0.74	112
macro avg	0.47	0.49	0.45	112
weighted avg	0.54	0.63	0.56	112

Fig. 13 Accuracy by class

the face for recall needed to be 0.08 times for images. F1 score tells about the true positive values of accuracy. Support will indicate that train and test measure data.

Figure 14 explains the training module for detection of drug or not drug with their accuracy score. Ingenerate this graph, we have given the 20 epoch iterative steps for finding the desisting features of the two categories that are personal having drug or not detection purpose. From the collected images, we have trained with binary classification, which helps to detect the drug feature of the faces from the module.

Figure 15 tells about the accuracy counting average, whereas precision tells us that measure feature area of the valid and drug faces where 0.00 is precision condition differentiation found for the valid faces us having the precision condition of 0.86 differences for recall needed to 0.00 times for images. F1-score tells about the true positive values of accuracy. Support will indicate that train and test measure data.

Figure 16 explains the training module for detection of tobacco consumed or not with their accuracy score. From the collected images, we have trained with binary classification, which helps to detect the tobacco feature of the faces from the module.

Figure 17 tells about the accuracy counting average, whereas precision tells us that measure feature area of the valid and no_tobacco faces where 1.00 is precision

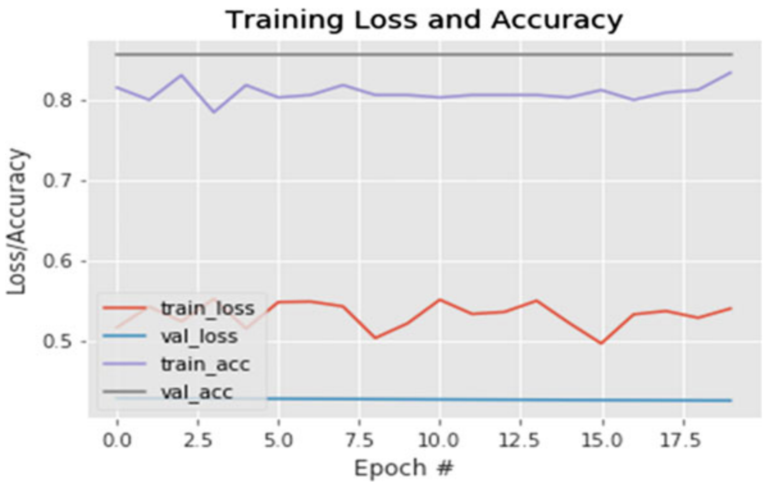


Fig. 14 Drug detection training graph

	precision	recall	f1-score	support
drugs	0.00	0.00	0.00	13
valid	0.86	1.00	0.92	77
accuracy			0.86	90
macro avg	0.43	0.50	0.46	90
weighted avg	0.73	0.86	0.79	90

Fig. 15 Accuracy by class

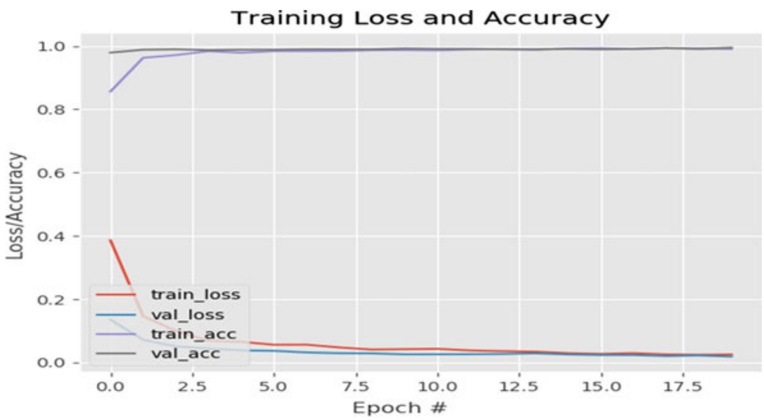


Fig. 16 Tobacco training graph

	precision	recall	f1-score	support
no_tobacco	1.00	0.99	0.99	384
tobacco	0.99	1.00	0.99	383
accuracy			0.99	767
macro avg	0.99	0.99	0.99	767
weighted avg	0.99	0.99	0.99	767

Fig. 17 Accuracy by class

condition differentiation found for the tobacco faces which we are having the precision condition of 0.99 differences for recall needed to 0.99 times for images. F1 score tells about the true positive values of accuracy. Support will indicate that train and test measure data.

Conclusions

The application of genetic algorithms (GA) to association rule mining demonstrates that GA outperforms the Apriori algorithm and produces accurate rules. Even with fewer rules the GAs generated, needless to say, the accuracy was higher compared to RAGA. The principles produced comparative accuracy with the work of Indira and Kamini. The principles produced more accuracy as compared thereto returned by the OLEX-GA plug-in of WEKA.

With the presented system for recommending advertisements that provide a customer with simpler methods to campaign for their product generating more revenue. The model presented uses context to make things more relevant and generating results with better placement duration and format. The context-based recommendation system used for advertisement is placed in a better position providing the user with the knowledge to select an advertisement.

In the system for vice detection recommendations, a way by which alcohol use, drug abuse, and facial changes of other vices might be captured with a camera is presented. The tactic uses parts of the face from the camera-clicked images of the face, and results are generated to detect vices. Within the experiments, we use a three-layer neural network to record alcohol consumption and determine face parts to be identified as defective due to consequences of vices.

Acknowledgements I would wish to say a "Big Thank You" to my Institution for all the assistance extended to me during my work and documentation of it.

References

1. Thankachan S, Kirubakaran S (2014) E-agriculture information management system. Int J Comput Sci Mob Comput 3(5):599–607
2. Agrawal R, Imielinski T, Swami A (1993) Database mining: a performance perspective. IEEE Trans Knowl Data Eng 5:914–925
3. Xu L, Liang N, Gao Q (2008) An integrated approach for agricultural ecosystem management. IEEE Trans Syst Man Cybern Part C Appl Rev 38(4)
4. Bhargavi P, Jyothi S (2009) Applying naive bayes data mining technique for classification of agricultural land soils. Int J Comput Sci Network Secur 9(85):117–122
5. Abdullah A, Hussain A (2006) Data mining a new pilot agriculture extension data warehouse. J Res Pract Inf Technol 38(3):9
6. Gaikwad ST, Desai SB, Kolekar AB (2016) Adoption of information and communication technology (ICT) for development of Indian agriculture. Int J Res Appl Sci Eng Technol 4(4):761–765, 98
7. Swaminathan R Analysis of self organizing maps using visual dm techniques in agro database for prediction of yield. Int J Adv Comput Sci 3(10):508–511, 201
8. Dave K, Lawrence S, Pennock DM Mining the peanut gallery: opinion extraction and semantic classification of product reviews. ACM
9. Prasad JR, Prakash PR, Kumar SS, Babu MS, Rani KS (2012) Identification of agricultural production areas in Andhra Pradesh. Int J Eng Innov Technol (IJEIT) 2(2):51–55
10. Nasira GM, Hemageetha N (2012) Vegetable price prediction using data mining classification technique. In: International conference on Pattern Recognition, Informatics and Medical Engineering (PRIME)
11. Veenadhari S, Misra B, Singh C (2011) Data mining techniques for predicting crop productivity a review article. IJCST 2(1)
12. Thankachan S, Kirubakaran S (2014) E-agriculture ınformation management system. Int J Comp Sci Mobile Comput (IJCSMC) 3(5):599–607
13. Jaiswal A, Dubey G (2013) Identifying best association rules and their optimization using genetic algorithm. IJESE 1:91–96
14. Dey A (2016) Machine learning algorithms—a Review. Int J Comput Sci Inf Technol 7(3):1174–1179
15. Wang J, Yang Y, Mao J, Huang Z, Huang C, Xu W (2016) CNN-RNN: a unified framework for multi-label ımage classification 2285–2294. https://doi.org/10.1109/CVPR.2016.251
16. Shin K, Jeon J, Lee S, Lim B, Jeong M, Nang J (2018) Approach for video classification with multi-label on YouTube-8M dataset
17. Szymański P, Kajdanowicz T (2017) A scikit-based Python environment for performing multi-label classification. J Mach Learn Res 20
18. Hermosilla G, Verdugo J, Farias G, Vera E, Pizarro T, Francisco G, Machuca M (2018) Face recognition and drunk classification using ınfrared face ımages. J Sens 1–8. https://doi.org/10.1155/2018/5813514
19. Tolba A, El-Baz A, El-Harby A (2005) Face recognition: a literature review. Int J Signal Process 2:88–103
20. Zhao W-Y, Chellappa R, Jonathon Phillips P, Rosenfeld A (2003) Face recognition: a literature survey. ACM Comput Surv 35:399–458.https://doi.org/10.1145/954339.954342
21. Tekkam Gnanasekar S (2019) Facial attribute recognition and its application in drug abuse detection (Unpublished master's thesis). University of Calgary, Calgary, AB
22. Yadav D (2019) On matching faces with temporal variations using representation learning. Graduate Theses, Dissertations, and Problem Reports. 3939. https://researchrepository.wvu.edu/etd/3939
23. Pandey K, Lilani R, Naik P, Pol G Human face recognition using ımage processing. İnt J Eng Res Technol (IJERT) IJERT www.ijert.org ICONECT' 14 Conference proceedings
24. Radzi F, Khalil-Hani M, Liew SS, Bakhteri R (2014) Convolutional neural network for face recognition with pose and ıllumination variation. Int J Eng Technol 6

25. Pranav KB, Manikandan J (2020) Design and evaluation of a real-time face recognition system using convolutional neural networks. Eng Technol Appl Res 10(3):5608–5612. https://doi.org/10.48084/etasr.3490
26. Manojkrishna M, Neelima M, Mane H, Matcha VGR (2018) Image classification using deep learning. Int J Eng Technol (UAE) 10(1)
27. Xin M, Wang Y (2019) Research on image classification model based on deep convolution neural network. EURASIP J Image Video Process Volume 2019, Article number: 40

Systematic Review on Churn Prediction Systems in Telecommunications

Gireen Naidu, Tranos Zuva, and Elias Mmbongeni Sibanda

Abstract Customer churn is common in all industries; however, in the telecommunication sector, churning occurs more frequently since the sector is extremely competitive. Customer churn is an essential aspect that needs to be monitored in all industries, specifically in the telecommunication sector as it directly influences customer retention, revenue as well as margin profit or loss. Failure to predict customer churn can result in revenue loss; therefore, this paper reviews machine learning algorithms used to predict customer churn in the telecoms industry. Twenty (20) papers were reviewed from 2017 to 2021. SVM, neural network and random forest have shown more accuracy with the accuracy of above 85%, while logistic regression is the mostly used algorithm on prediction on customer churn in telecommunication, although the accuracy rate is below 85%. Other researchers use different evaluation metrics which makes it difficult to compare algorithms. Therefore, it is necessary to find out how it can be modified to achieve a higher accuracy rate in order to predict customer churn. Logistic regression is the preferred algorithm even though it has a lower accuracy, while the algorithm needs to be modified to provide higher accuracy.

Keywords Churn prediction · Customer churn · Telecommunications · Machine learning

Introduction

The telecommunications industry in all markets across the globe is dealing with a serious burden on its revenue due to the aggressive market competition [1]. Telecommunications operators are in a state of constant evolution in an attempt to maintain

G. Naidu (✉) · T. Zuva · E. M. Sibanda
Vaal University of Technology, Gauteng, South Africa

T. Zuva
e-mail: tranosz@vut.co.za

E. M. Sibanda
e-mail: eliass@vut.co.za

V. Bindhu et al. (eds.), *Proceedings of Third International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 844,
https://doi.org/10.1007/978-981-16-8862-1_64

and grow their revenue streams. With telecommunication service providers aggressively attempting to differentiate themselves by offering more competitive rates and services, customer tenures are rapidly declining and brand loyalty is deteriorating.

One aspect that has consistently impacted revenues is that of customer attrition or customer churn. Churn is explained as the occurrence of a customer that terminates services with an organization. As with any subscription and non-subscription company, telecom providers save costs in retaining customers for repeat sales rather than acquiring new ones.

A mechanism that has been developed for addressing the problem of customer churn is the application of machine learning algorithms and data mining to determine the probability of a customer churning. Since telecommunication providers have detailed customer data in their records, they can use this information to predict churn behavior. Various churn prediction algorithms have been applied in this area, but it is critically important to understand their behavior and outcomes [2].

This systematic review paper focused on churn prediction systems in telecoms using machine learning algorithms. Open issues and challenges are also discussed.

This paper is arranged as follows: Methodology, Results, Discussion, Conclusion and Future Work.

Methodology

The systematic review process shown in Fig. 1 was followed to conduct research into the area of machine learning applications to reduce customer churn in the telecommunication industry.

Results

The 20 articles that were selected for review, were published between the years 2017 and 2021. This was to ensure that the latest algorithms presented in the research was used for the purposes of this study are still relevant as well as to ensure that there is an adequate representation that will depict the current telecommunications industry landscape. Figure 2 indicates the articles and the year in which they were published.

Discussion

This research aims to provide an understanding of logistic regression, random forests, decision trees, SVM and neural network classifiers to assist in understanding the most effective models of churn prediction. A hybrid approach was also applied utilizing a

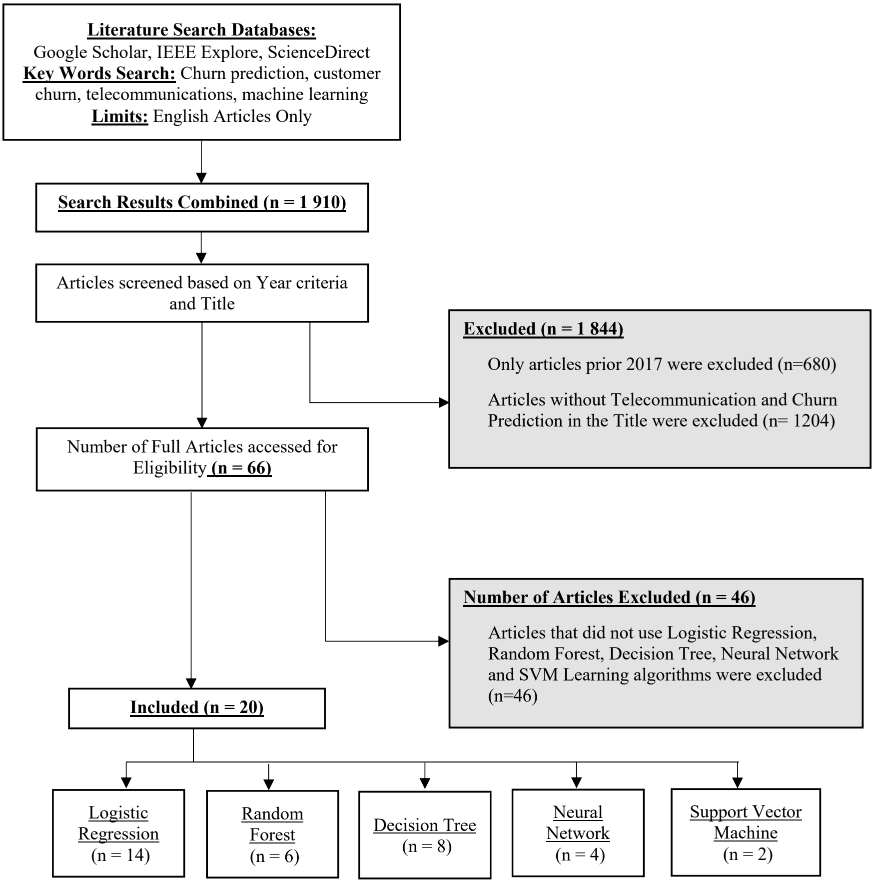


Fig. 1 PRISMA systematic review research methodology. Note: The total number of algorithms exceeds 20 as certain papers may have used multiple algorithms

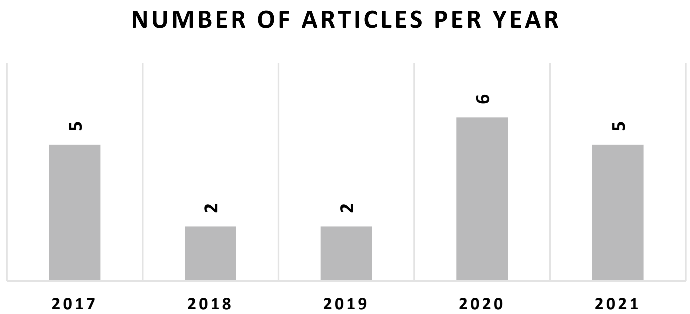


Fig. 2 Annual representation of the articles under review

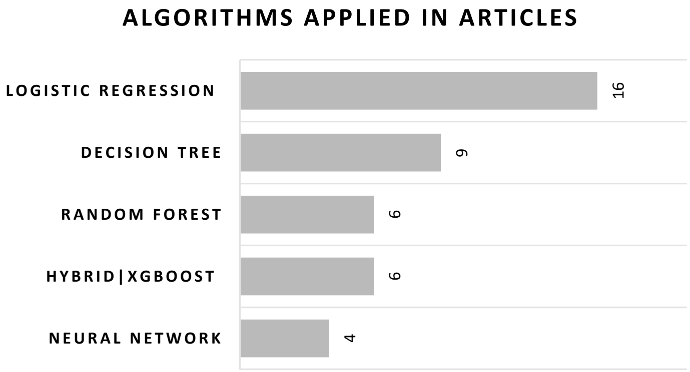


Fig. 3 Application of algorithms in telecommunications churn prediction. Note: The total number of algorithms exceeds 20 as certain papers may have used multiple algorithms

combination of logistic regression, decision tree and XGBoost. Figure 3 represents the number of algorithms that were modeled within the research under review.

Multiple researchers have made comparative studies of the algorithms and models to enable the prediction of customer churn in telecommunications. Various experiments were conducted with varying types of datasets to deliver the most accurate prediction system. Table 1 indicates the various prediction systems used in research as well the multitude of evaluation metrics used to measure prediction performance.

The table above indicates that there are varying types of evaluation mechanisms being utilized to measure model performance and effectiveness. The evaluation metric of accuracy seems to the preferred evaluation mechanism. Although accuracy does indicate the correct level of prediction accuracy, it can sometimes be misleading and produce results with high accuracy. In cases where the dataset is imbalanced toward a specific class variable, the accuracy will be skewed more toward the output with the highest imbalance.

Out of the 20 articles reviewed, 16 (80%) articles measured accuracy as an indication of model performance. 4 (20%) articles measured model performance using a combination of evaluation metrics that excluded accuracy. Table 2 indicates the average accuracy scores for the algorithms under review.

The logistic regression has the lowest average accuracy of 82.33%. This is the algorithm that is most used in research due its dichotomous nature and ease of implementation. [5] used Stratified k-fold cross validation and principle component analysis techniques on a random forest algorithm to achieve the highest level of accuracy.

Table 1 Indicates the various types of machine learning algorithms used in research

References	Model/Method used	Evaluation matrix
[3]	XGBoost, logistic regression	Accuracy
[4]	Neural network	Accuracy, error rate, recall, precision and f-score
[5]	Logistic regression, random forest, K-nearest neighbors	Accuracy precision, recall, f-score
[6]	Gradient boosting, random forest, logistic regression, XGBoost, decision tree, K-nearest neighbors	Accuracy
[7]	XGBoost, logistic regression, decision tree, and naïve Bayes	Accuracy, precision, recall, f-score
[8]	Logistic regression, SVM, logit boost, random forest	Accuracy, precision, recall, f-score
[9]	Logistic regression	Accuracy, precision and recall
[10]	Logistic regression, K-nearest neighbors	Accuracy, recall, precision and f-score
[11]	Logistic regression	Accuracy, error rate, f-score
[12]	Random forest, logistic regression, XGBoost	Precision, recall and f-score
[13]	Logistic regression, random forest, neural networks	Confusion matrix
[14]	Decision tree, naïve Bayes, and decision rules, Logistic regression, random forest, and neural network, SVM	Accuracy, recall, precision and f-score
[15]	XGBoost, random forest, decision tree	AUC
[16]	K-Nearest neighbors, random forest, XGBoost	Accuracy, f-score
[17]	Decision tree, logistic regression, logit leaf model	AUC and TDL
[18]	Decision tree, logistic regression, and naïve Bayes	Accuracy
[19]	Logistic regression, decision tree	Accuracy, mean errors
[20]	Logistic regression	Accuracy
[21]	Decision tree	Accuracy
[22]	Random forest, logistic regression, SVM, multilayer perceptron, K-nearest neighbors, naïve Bayes, neural network	Accuracy

Discussion

Churn prediction models use machine learning and data mining to determine the probability of a customer churning. Since telecommunication providers have detailed customer data in their records, they can use this information to predict churn behavior.

Table 2 Representation of average accuracy scores for the algorithms being researched

Evaluation metrics	Number of articles	Accuracy intervals (%)	Average accuracy score (%)
Logistic regression	13	74.58–91	82.33
Decision tree	6	72.54–91	83.29
Random forest	6	77.38–96	85.25
Neural network	3	77.5–93	85.78
Support vector machine	3	79.2–90	85.73
Hybrid/other	4	78–94	82.96

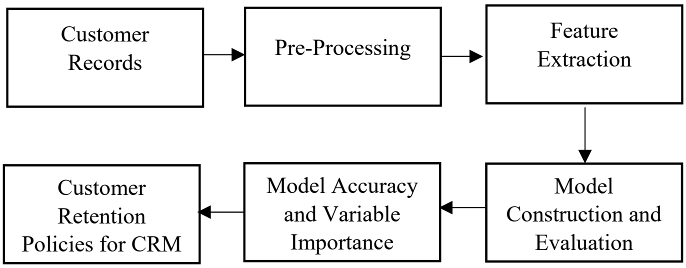


Fig. 4 Phases of a churn prediction system from [23]

A large number of churn prediction models are already in use among companies. Some of which are more sophisticated than others. The different phases of a model churn prediction system proposed by Umayaparvathi and Iyakutti [23] is shown in Fig. 4.

This model consists of five phases:

(1) Preprocessing the input customer records—This phase consists of customer data extraction, cleansing, and formatting. Missing values, outliers, and inconsistencies are removed and/or cleaned. This step is required to enhance the quality of the dataset.
(2) Extracting the required features for developing churn models—This phase consists of identifying the variables that will be used in the prediction model. Various variables such as customer revenue and calling patterns are utilized to predict churn.
(3) Construction of models using different classifiers and cross-validate the models—This phase consists of developing the prediction model.
(4) Calculation of prediction accuracy and variable importance report—This phase consists of measuring the accuracy of the prediction model.
(5) Providing customer retention policies to CRM executives—This phase consists of recommending treatment plans to prevent customers from churning.

Machine Learning Methods in Churn Prediction from Retrieved Articles

This research provides an understanding of logistic regression, random forests, decision trees, SVM, and neural network classifiers to assist in understanding the most effective models of churn prediction.

Logistic Regression

This approach is a statistical model highly used to acquire dichotomous results in the variable predictor. The method is used in research to deduce whether an event either happened or failed to occur [3]. Logistic components are used to acquire log-odds (B) using logistic functions. Therefore, the model quantifies outcomes between results and predictor variables. Multiple regressions which consist of various input features or independent variables are used to build the relationship with respective dependent variables or output. The general equation of multiple regressions is shown in Eq. 1:

$$y = \alpha + b_1 x_1 + b_2 x_2 + \ldots b_i x_i \tag{1}$$

where y is defined as a dependent variable,

$a =$ intercept and is the value of y when all the independent variables are equal to zero,

$x_i = x_1, x_2, x_3 \ldots x_n$ are the independent variables or the predictors,

$b_i = b_1, b_2, b_3 \ldots b_n$ are known as the coefficient for the respective predictors.

In dealing with the binary result, logistic regression is a widely used tool that predicts whether churn will occur or not. Multiple logistic regressions can be expressed in Eq. 2, which replaces the dependent variable with logit function to produce the outcome of the binary result:

$$\log\left(\frac{p}{1-p}\right) = \alpha + b_1 x_1 + b_2 x_2 + \ldots b_i x_i \tag{2}$$

where $p =$ probability of churn,

log = the logarithm of the ratio of the odds to produce the categorical binary result that a customer will churn or not churn.

Decision Trees and Random Forests

Decision trees categorize the observation by arranging in descending order from the root node to the tree node to some leaf/terminal node. Each node in the tree acts as a test observation for some parameter, and each edge resulting from the node corresponds to the possible outcomes to the test observation [21].

The purpose of a decision tree in the organization includes pointing customer's specific class. Technically, most votes from a class classify the customer. Since decision trees are susceptible to their trained data, telecom avoids the menace through bagging. Bagging is, therefore, a random sample from the dataset used in training

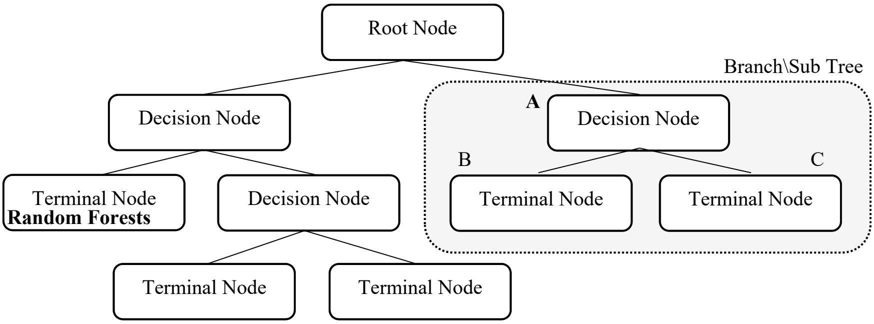


Fig. 5 Representation of decision tree model [24]

decision trees. Figure 5 illustrates how decision tress use nodes to determine the target variable.

Random forest is a collection of decision trees based on majority voting to generate improved accuracy in churn prediction. A good performance is achieved when using this method because it has the capability of combining outputs through majority voting of decision trees. Figure 6 illustrates how an ensemble random forest uses multiple decision tress to determine the target variable.

SVM

SVM is a supervised learning method that makes predictions using linear combinations of kernel functions [25]. The concept is an implementation of structural risk minimization principle. It seeks to minimize the upper bond of generalization error. Therefore, it maps the input data to high-dimensional spaces. The accuracy of an SVM model links its functionalities with kernel parameters selection. The parameters, therefore, contain adverse effects on kernel method performances. The method when employed in churn prediction can help to create high retention campaigns and help improve marketing efforts by a company to help reduce customer churn. Figure 7

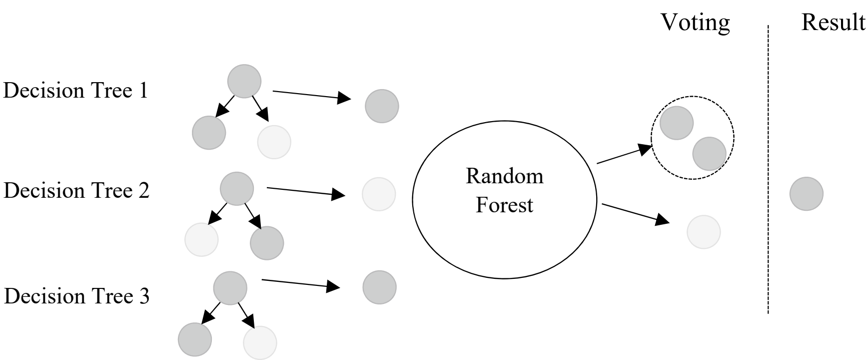


Fig. 6 Representation of random forest structure [8]

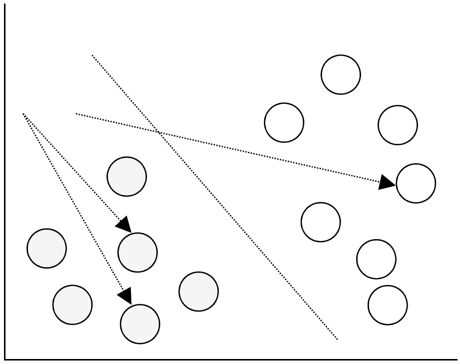


Fig. 7 Representation of SVM structure [8]

represents an SVM model mapping training data points into a higher dimensional space to determine the target variable.

Support vectors are the coordinates of the variables. These data points help in building SVM model. While building SVM, the margin between the coordinates and hyperplane tries to be increased [2]. The loss function is calculated by Eq. 3.

$$c(x, y, f(x)) = \{0, ify * f(x) \geq 11 - y * f(x), \text{else}\} \tag{3}$$

The performance of the SVM model is very dependent on the appropriate model parameters. Several parameters need to be tuned by cross validation; these parameters include error/margin trade-off parameter and controls the width of the intensive zone [25]

Neural Networks

A neural network simply reflects human brain behavior, thus allowing computer programs to solve common problems in deep learning, machine learning and AI. The patterns recognized by the neural networks are numerical and can help predict outcomes from a large dataset. Neural networks rely on training data. The neural networks are comprised of node layers with an input and output layer and several hidden layers between the two. All the nodes are interconnected and pass data from one node to another. Below is a diagram depicting how a neural network works.

The neural networks classifier works by training data that improves over time. The nature and organization of the connection between neuron are the factors that determined the structure. Subsequently, in the learning process, it is crucial to train these connections in order to achieve a desired overall outcome of the algorithm [26]. Figure 8 represents an overview of a neural network.

Evaluation Metrics

See Table 3.

TP: Correctly categorized as the class variable

TN: Correctly categorized as the class variable

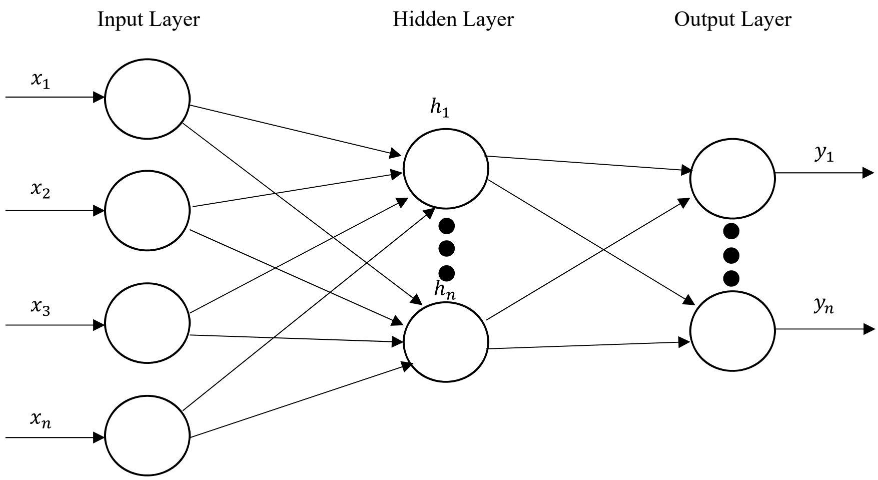


Fig. 8 Representation of neural network [27]

Table 3 Representation of recurring evaluation metrics

<table>
<tr><th></th><th>Evaluation metrics</th><th colspan="4">Formula</th><th>References used</th></tr>
<tr><td rowspan="4">1</td><td rowspan="4">Confusion matrix</td><td colspan="2"></td><td colspan="2">Actual Results</td><td rowspan="4">[11]</td></tr>
<tr><td rowspan="3">Predicted Results</td><td></td><td>Positive</td><td>Negative</td></tr>
<tr><td>Positive</td><td>TP</td><td>FP</td></tr>
<tr><td>Negative</td><td>FN</td><td>TN</td></tr>
<tr><td>2</td><td>Precision (P)</td><td colspan="4">$\frac{TP}{TP+FP}$</td><td>[4, 5, 7, 8, 10, 12]</td></tr>
<tr><td>3</td><td>Recall (R)</td><td colspan="4">$\frac{TP}{TP+FN}$</td><td>[4, 5, 7, 8, 10, 12]</td></tr>
<tr><td>4</td><td>Accuracy</td><td colspan="4">$\frac{TP+TN}{TP+TN+FP+FN}$</td><td>[3–12, 14, 16–20]</td></tr>
<tr><td>5</td><td>F-Score</td><td colspan="4">$\frac{2*P*R}{P+R}$</td><td>[4, 5, 7–10, 12, 14]</td></tr>
<tr><td>6</td><td>Error rate</td><td colspan="4">$\frac{FP+FN}{TP+TN+FP+FN}$</td><td>[4, 7, 9]</td></tr>
</table>

FP: Incorrectly categorized as the class variable

FN: Incorrectly categorized as the class variable

Area Under Curve (AUC)—Receiver operating characteristics (ROC) is a probability curve and AUC indicates the grade or size of separability. It highlights to what value the model can distinguish between categories. The increase in the AUC value, the more the model is capable at predicting the correct classes.

Jain et al. [2] also indicated that logistic regression is the algorithm that is most used in research as it performs well where there is limited training data to build the model. This is the algorithm that is most used in research due its dichotomous nature and ease of implementation.

The research highlighted that although accuracy was primarily utilized as the statistical metric for the evaluation of model performance, Accuracy alone should not be utilized for the measurement of performance. Although accuracy does indicate the correct level of prediction correctness, it can sometimes be misleading and produce results with high accuracy. In cases where the dataset is imbalanced toward a specific class variable, the accuracy will be skewed more toward the class with the highest imbalance. Although all the methodologies researched have produced decent levels of predictions and will have a less margin of error and can help in predicting the churn rates, the acknowledgment that the lack of standardization of the measurement of model performance needs to be addressed. The implementation of a consistent and standardized approach should be explored. Furthermore, a standardized approach for model measurement will ensure that future models can be accurately benchmarked. There is a definite requirement for future work that will strategize the implementation on the guidelines to ensure transparency and standardization of quality of the evaluation metrics and protocols utilized to measure prediction model performance.

Processing times and computational overhead when executing prediction algorithms need to be taken into consideration when creating the algorithm. Training and testing prediction algorithms on subsets of data may be adequate for the purposes of research; however, considerations need to be made when applying similar models on real-world datasets. Telecommunications datasets have grown exponentially over time, so processing times and computational overhead becomes a contributing factor to the effectiveness of the prediction model.

Additionally, the average accuracy score across the models researched was 84%. Stratified k-fold cross validation and principle component that was used to achieve the highest accuracy needs to be considered to improve model performance. Model parameter and hyperparameter tuning are another concepts not consistently performed. These will control how the model behaves when being trained and produce more accurate outcomes. Hyperparameters optimization techniques include grid search, random search and Bayesian optimization. Fine-tuning the parameters of the algorithm will define the architecture of the model and assist in improving accuracy scores.

Conclusion and Future Work

In any service industry, customers are their primary assets as they bring revenue to the company. In the telecom industry, customers play vital roles in the growth and profit of a company. A mechanism that has been developed for addressing the problem of customer churn is the application of machine learning algorithms and data mining to determine the probability of a customer churning. Many models and schemes have been produced in the telecom market based on machine learning to predict churn. Most of the models exhibit different characteristics based on complexity, interpretability levels, and model type [18]. The research has been able to determine the five methods of predicting churn. The methods include logistic regression, random forests, decision trees, neural networks and support vector machine. Future work can be conducted in the areas of standardizing the approach to performance measurement, architecture relating to processing times, and computational overhead can be improved by balancing computing power with model performance and enhancements in model parameter and hyperparameter tuning.

References

1. Umayaparvathi V (2012) Applications of data mining techniques in telecom churn prediction 42(20):5–9
2. Jain H, Khunteta A, Srivastava S (2021) Telecom churn prediction and used techniques, datasets and performance measures: a review. Telecommun Syst 76(4):613–630. https://doi.org/10.1007/s11235-020-00727-0
3. Parmar P (2021) Telecom Churn Prediction Model using XgBoost Classifier and Logistic Regression Algorithm May:1100–1105
4. Mishra A, Reddy US (2018) A novel approach for churn prediction using deep learning. 2017 IEEE International conference on computational intelligence and computing research. ICCIC 2017, Jan 2019, pp. 1–4. https://doi.org/10.1109/ICCIC.2017.8524551
5. Senthilnayaki B, Swetha M, Nivedha D (2021) Customer Churn Prediction. Iarjset 8(6):527–531. https://doi.org/10.17148/iarjset.2021.8692
6. Singh D, Jatana V, Kanchana M (2021) Survey paper on churn prediction on telecom. SSRN Electron J 27(5):395–403. https://doi.org/10.2139/ssrn.3849664
7. Xu T, Ma Y, Kim K (2021) Telecom churn prediction system based on ensemble learning using feature grouping. Appl Sci 11(11). https://doi.org/10.3390/app11114742
8. Jain H, Khunteta A, Srivastava S (2021) Telecom churn prediction using seven machine learning experiments integrating features engineering and normalisation
9. Jain H, Khunteta A, Srivastava S (2020) Churn prediction in telecommunication using logistic regression and Logit boost. Procedia Comput Sci 167(2019):101–112. https://doi.org/10.1016/j.procs.2020.03.187
10. Joolfoo K, Jugurnauth M, Joofloo R (2020) Customer churn prediction in telecom using big data analytics. IOP Conf Ser Mater Sci Eng 768(5). https://doi.org/10.1088/1757-899X/768/5/052070
11. Arivazhagan B, Sankara Subramanian DRS, Scholar R (2020) Customer churn prediction model using regression with Bayesian boosting technique in data mining. Ijaema.Com XII(0886):1096–1104, [Online]. Available: http://www.ijaema.com/gallery/132-ijaema-may-3946.pdf
12. Kavitha V, Hemanth Kumar G, Mohan Kumar SV, Harish M (2020) Churn prediction of customer in telecom industry using machine learning algorithms. Int J Eng Res V9(05):181–184. https://doi.org/10.17577/ijertv9is050022
13. Kavita M, Sharma N, Aggarwal G (2020) E-commerce industry using machine learning 17(9):6–15
14. Ben A (2020) Enhanced churn prediction in the telecommunication industry. SSRN Electron J 2:6–15. https://doi.org/10.2139/ssrn.3577712
15. Ahmad AK, Jafar A, Aljoumaa K (2019) Customer churn prediction in telecom using machine learning in big data platform J Big Data 6(1). https://doi.org/10.1186/s40537-019-0191-6
16. Pamina J et al (2019) An effective classifier for predicting churn in telecommunication. J Adv Res Dyn Control Syst 11(1) Special Issue: 221–229
17. De Caigny A, Coussement K, De Bock KW (2018) A new hybrid classification algorithm for customer churn prediction based on logistic regression and decision trees. Eur J Oper Res 269(2):760–772. https://doi.org/10.1016/j.ejor.2018.02.009
18. Chathuranga L, Rathnayaka R, Arumawadu New customer churn prediction model for mobile telecommunications industry

19. Kau FM, Masethe HD, Lepota CK (2017) Service provider churn prediction for telecoms company using data analytics. Lect Notes Eng Comput Sci 1:385–388
20. Sebastian HT, Wagh R (2017) Oriental J Churn Anal Telecommun Logistic Regression
21. Saini NM, Garg K (2017) Churn Prediction in telecommunication industry using decision tree. Int J Eng Res V6(04):439–443. https://doi.org/10.17577/ijertv6is040379
22. Sabbeh SF Machine-learning techniques for customer retention: a comparative study. Int J Adv Comput Sci Appl 9(2). https://doi.org/10.14569/IJACSA.2018.090238
23. Umayaparvathi V, Iyakutti K (2016) A survey on customer churn prediction in telecom industry : datasets , methods and metrics
24. Chauhan NS (2020) Decision tree algorithm, explained—Kdnuggets. https://www.kdnuggets.com/2020/01/decision-tree-algorithm-explained.html (Accessed 09 Aug 2021)
25. Rodan A, Faris H, Alsakran J, Al-kadi O (2014) A Support vector machine approach for churn prediction in telecom industry. Aug 2014
26. Rupawon NA A review of churn predictive in telecommunication
27. Bre F, Gimenez JM, Fachinotti VD (2018) Prediction of wind pressure coefficients on building surfaces using artificial neural networks. Energy Build 158(April):1429–1441. https://doi.org/10.1016/j.enbuild.2017.11.045

Machine Learning-Based Approach for Depression Detection Using PHQ-9 and Twitter Dataset

Sumitra Motade, Alima Hassan, Fiza Mir, and Karan Parikh

Abstract Depression is a severe mental health issue that affects people of all age groups, sexes, and races. People are more relaxed about expressing their opinions on social networking sites (SNS) almost every day in this age of modern connectivity and technology. This work aims to propose a model based on data analysis and machine learning to detect depression in a person. Two types of depression detection tests are conducted on the same Web site. Two tests are based on the PHQ-9 dataset approved by the National Institute of Health (NIH) and Twitter Dataset. In the PHQ-9 dataset, the maximum accuracy achieved using SVM is 99.74%. In the Twitter dataset, comparatively, a little less accuracy is obtained because of the informal language used in tweets. The outcome of this work will help the researchers with an insight into the automatic techniques to detect depression. Proposed work significantly improves the accuracy as SVM algorithm gives the highest accuracy with correct data preprocessing.

Keywords Depression detection · Machine learning · Support vector machine · Twitter · PHQ-9 · Decision tree classifier · Random tree classifier · Naive Bayes classifier

Introduction

Depression is a mental health disorder characterized by the steadily depressed frame of mind or loss of attention in daily life activities, causing noteworthy harm to a person. The depression causes comprise a blend of biological, psychological, and social sources of suffering. Progressively, many research suggests that these factors may reason changes in brain function, including the distorted activity of certain neural circuits in the brain. It is a major psychiatric disorder that can lead to suicide. It is also a leading cause of injury worldwide [1]. More than 300 million people suffer

S. Motade (✉) · A. Hassan · F. Mir · K. Parikh
School of Electronics and Communication Engineering, Dr Vishwanath Karad MIT World Peace University, Pune, India
e-mail: sumitra.motade@mitwpu.edu.in

V. Bindhu et al. (eds.), *Proceedings of Third International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 844,
https://doi.org/10.1007/978-981-16-8862-1_65

from depression per annum all over the world. Depression is typically assessed using face-to-face psychiatric depression guidelines. As far as medical science is concerned, it relies on asking the patients a few questions. However, that does not detect depression in a precise manner. If depression increases fast, it will be the second prominent disease worldwide after heart-related diseases. Depression has to be treated properly by medication, and medical advice is required. In the early stages of depression, 70% of patients would not seek medical help, leading to severe mental illness because people feel shy to admit that they are facing mental illness in the starting stages [2].

The proposed work relates to the person experiencing depression detects it at the earliest before it gets worse and serious through our intensely personal depression detection test. With this test, mental health professionals would already know the patient's mental condition and give better treatment to patients. Data are collected from two sources: tweets and questionnaires prepared from NIH. With the growing cases of depression, there is an increase in the demand for more automatic methods for depression detection. Machine learning algorithms and natural language analysis are used to classify whether the data or tweets are pessimistic or not [3]. Exploratory data processing is conducted on two datasets before creating a model to comprehend it fully. In this article, various strategies for detecting depression have been explored from social media data, such as questionnaires, Twitter data, and how these techniques vary from one another.

The remaining of the paper is structured as: Sect. "Literature Survey" contains Literature Survey. Section "System Design and Methodology" explains the System Design and Methodology. In Sect. "Results," Results are discussed. Section "Conclusion" summarizes the Conclusion and Future Work.

Literature Survey

Depression studies are becoming very significant and are in the spotlight of the Internet. Detecting depression from papers and online data has become an increasingly vital research area, with fascinating methods and outcomes reported for Questionnaires, Twitter, and other debate posts. With the help of a worldwide questionnaire survey, research and surveys have been done by many researchers. In one of Lenore Sawyer Radloff's works, a CES-D scale contains 20 questions about mental health conditions, such as sad feelings and different sleep conditions of users. Some questions have many options lined up, and then, scores are labeled to the options, and thus, depression is detected with the ultimate score [3]. The stage of depression is diagnosed based on the score scale. Twenty-one categories of users' mental health and physiological states are expressed like the frame of mind, sentiment of failure, displeasure, petulance, disgrace, sense reprimand, self-hate, self-accusation, and inhibition work. The efficiency characteristics of the Patient Health Questionnaire-9 (PHQ-9) Items are used as a tool for adolescent depression [6]. They were inspired by deep learning in detecting sentiments and provided a proper method to understand the person's emotion and depression based on the deep learning technique. They have used a deep 50-layer residual network in the visual data and a convolution neural network (CNN) in the audio. Recently, another work was conducted [7], which reflected the collection of Facebook data and proposed the machine learning model to detect depression efficiently and effectively.

Numerous machine learning techniques can be used to detect depression. However, each has its advantages and disadvantages that can affect the accuracy and performance of a system. It depends on how much data are used for training and testing purposes, and which methods are used for data preprocessing. For example, in related work [6], they used two datasets—one from the questionnaire and the other from Reddit. For dataset I, the XG Boost classifier provided the greatest accuracy, i.e., 83.87%. The logistic regression provided the lowest accuracy, unlike dataset II, where the logistic regression gave the highest accuracy, 86.45% because dataset I consists of 18 characteristics which were more than the words. Support vector machine, k-nearest neighbor (k-NN), Naive Bayes, decision tree, and other grading algorithms of supervised machine learning exist in machine learning. Such algorithms have been used based on the various tasks and data available. By using ensemble learning techniques, the accuracy of the model can be increased. In addition, the NIH approved PHQ-9 questionnaire is used to get good accuracy, and for processing the dataset, the mode operation is used in the proposed work so that the data are in the binomial distribution. It will also help to increase accuracy and efficiency [8]. The proposed model uses a questionnaire similar to PHQ-9, and it covers all aspects and symptoms that lead to depression. Also, the Twitter dataset is used to train the model as Twitter contains more negative sentiments than positive ones [9] (Table 1).

System Design and Methodology

This research work aims to diagnose depression using various categories, life issues and physical wellbeing. It can ask questions to the user based on external datasets from the PHQ-9 approved by NIH and Twitter and then detect the type of depression into five stages of depression. The depression detection test using two datasets is completely based on client–server architecture. The front end needs to access the data stored in the backend, and the backend gives access to all the databases. In the front end, UI designing, user interaction is done, and the front end displays data. A setup connection between the database providers at the backend is maintained. Whatever, Webpage has been designed that is on the front side. On the backend, Python and machine learning modules are used. On the front end, HTML, CSS, Java, and Bootstrap are used. Flask is used to maintain connectivity between the front end and back end.

A user can use the function to respond to the system's questions and provide text answers. The system sends the operation to the user to see if they are depressed. The machine is fully under the control of the administrator. The framework analyzes

Table 1 Literature review

Sr. No.	Author	Paper Name	Method	Result
1	Jain S. Narayan, S.P., and kumar, R. (2017) [6]	Depression detection using machine learning using Twitter and questionnaire	PHQ-9, Praw	XG Boost classifier accuracy-83.87%
2	Krishna Shrestha (2018) [9]	Depression diagnosis using machine learning and Twitter data	Twitter for prediction and machine learning used	Random forest classifier worked better
3	Islam, M. R., and Kabir, M. A. (2018) [5]	Depression analysis using social network data and machine learning techniques	Machine learning technique as an scalable method	Social network, emotions, depression, sentiment analyses
4	Alsagri, H., and Akhlef, M. (2017) [2]	Machine learning-based depression detection using Twitter	Method is a data-driven, predictive approach for early detection	Exploration part of the features
5	Huiye Lin (2011) [10]	Psychological stress detection	Using neural network	SAE got the highest accuracy of 89%
6	Jie Tang (2011) [11]	Quantitative study of emotional states in social sites	Mood cast with different algorithms	62.18% accuracy
7	Mummun De Chowdhur (2013) [12]	Social media as a tool for depression detection	SVM classifiers, (CES-D)	SMDI detects depression
8	Budhaditya Saha (2015) [13]	A framework for classification of depression	LIWC, LDA	Accuracy 87.6%
9	Elvis (2016) [14]	MIDAS: Mental Illness detection and analysis via social media	Epidemiologic studies, also used TF-IDF, PLF	Extracted feature by considering two mental disorders
10	Keumhee (2016) [15]	Identification of depressive users in twitter	SVM-based work, used visual sentiment ontology	Proves multimodel has higher accuracy and can analyze user's mood
11	Maryam Mohammed [16]	Prediction of depression using social media	RapidMiner, SVM and Naïve Bayes classifier used	Best precision and minimal accuracy
12	Adrian (2017) [17]	Multitask learning for mental Health using social media	Multitask and single task learning approach used	Performs better than other LR models

(continued)

Table 1 (continued)

Sr. No.	Author	Paper Name	Method	Result
13	Munmunn De chadur [18]	Predicting postpartum anxiety depression	PHQ	Postpartum depression detection
14	Hong-Ha net (2020)	Study on social media mental disorder	Tensor model	Accuracy CR 80.5
15	Nafız Al Asad (2020)	Analyze social media posts of user	NLP	Accuracy 74% and precise 100%

the user's feedback using machine learning and generates a result about the type of depression observed. Consequently, different machine learning algorithms are used based on their accuracy and compatibility with datasets. Figure 1 represents the proposed model of depression detection.

Data are collected from two different sources, i.e., PHQ-9 questionnaires approved by NIH and Tweets from Twitter, a social media site. The below section provides a detailed explanation of used datasets, data preprocessing method, proposed algorithms, and model.

Dataset 1 from Questionnaire

Patient health questionnaire is a reliable depression module and clinical research tool, and it scores each of 9 DSM-IV criteria as "0" (not at all) to "3" (nearly every day). PHQ-9 is a readymade dataset completed by around 6,000 patients in 15 mental care clinics, as shown in Fig. 2. It is a trustworthy and convincing measure of depression severity. It makes it a constructive clinical and research tool for detecting depression as the National Institute also approves it for health. If there is a question, e.g., How often do you feel depressed? then there are four choices, and the choices are encoded by the label encoder, as shown in Table 2. Table 2 shows the values for the four options of the quiz-based depression detection test. The results of the test are evaluated as per the value held by the options.

For the PHQ-9 dataset, the following steps are required for data preprocessing:

- Column extraction (take columns from q1–q10 and the class column).
- Handling missing data using mode (mode because need the data in binomial form).
- Data splitting, 60% for training, and 40% for testing purposes.

Users fill a test similar to the PHQ-9 questionnaire. Various counselors have been consulted to make questions accurate and useful in determining the level and type of depression. The mode method is used to handle missing data. Features of focus are mentioned below.

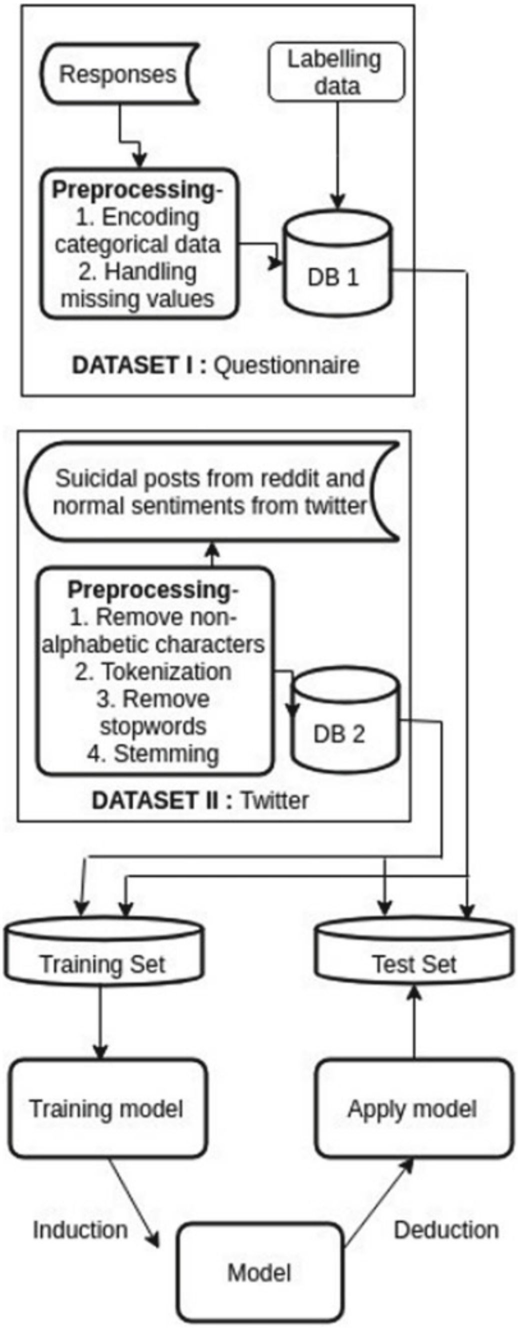


Fig. 1 Proposed model of depression detection

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1	id	q1	q2	q3	q4	q5	q6	q7	q8	q9	q10	score	class	time	period.na	start.time	
2	1	3	2	2	2	3	0	0	0	0	3	15	3	########	evening	########	
3	2	0	3	0	0	0	3	0	0	0	0	6	1	########	evening	########	
4	3	0	0	0	0	0	0	0	0	0	0	0	0	########	morning	########	
5	4	2	1	1	2	0	0	2	3	0	3	14	2	########	midday	########	
6	5	1	3	1	1	2	1	2	3	0	1	15	3	########	midday	########	
7	6	3	2	1	0	1	2	2	1	2	3	17	3	########	morning	########	
8	7	2	1	2	3	2	3	2	1	3	2	21	4	########	midday	########	
9	8	0	3	0	0	0	3	0	0	0	0	6	1	########	evening	########	
10	9	0	1	2	3	2	1	3	2	2	2	18	3	########	evening	########	
11	10	0	1	1	3	2	3	3	3	0	1	19	3	########	morning	########	
12	11	2	3	0	1	0	1	1	1	2	0	11	2	########	morning	########	
13	12	2	2	0	1	3	1	0	0	3	3	15	3	########	midday	########	
14	13	0	0	0	0	0	0	0	0	0	0	0	0	########	evening	########	
15	14	1	1	1	0	0	3	3	0	3	2	14	2	########	morning	########	
16	15	0	3	0	0	0	3	0	0	0	0	6	1	########	midday	########	
17	16	0	1	0	2	3	2	1	2	3	0	16	3	########	evening	########	
18	17	3	1	3	3	1	3	2	3	2	1	22	4	########	morning	########	
19	18	2	0	3	0	2	3	2	1	0	3	16	3	########	evening	########	
20	19	1	1	2	1	0	1	3	2	2	3	16	3	########	evening	########	
21	20	3	3	1	1	2	2	0	0	2	1	15	3	########	midday	########	
22	21	0	1	1	1	3	3	2	0	3	1	17	3	########	evening	########	

Fig. 2 PHQ-9 Dataset

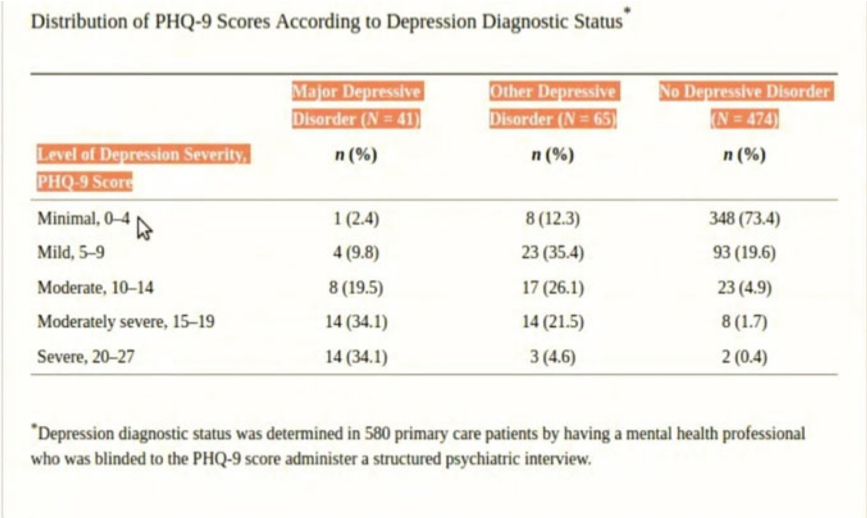
Table 2 Values of options

Options	Value
Not at all	0
Several days	1
More than half the days	2
Nearly everyday	3

- Gender
- Age group
- Punctuality in educational institutions
- Feeling fatigued/having less motivation and energy
- Feeling low or discouraged.
- Insomnia severity
- Appetite loss
- Difficulty in focusing
- Suicidal thoughts
- Any earlier experience of abuse

There are ten questions in the dataset, e.g., if the question is little interest or pleasure in doing things, then there are four options mentioned in Table 2. Then, there are labels encoded to them. All these labels give us the score. After that the class of depression is determined with the help of a score. With the column extraction method, all unnecessary columns are removed (Fig. 3).

PHQ-9-based depression detection test covers all aspects and symptoms related to depression. A total of ten questions are asked to the user, and each question has four options. It is a quiz-based test. The highest accuracy is achieved, which is 99.74%, with the help of SVM in this test. 60% data are used for training, and 40% data are used for testing the model. Model is trained using the training set and tested using the testing set. Hence, the accuracy is calculated.



Distribution of PHQ-9 Scores According to Depression Diagnostic Status*

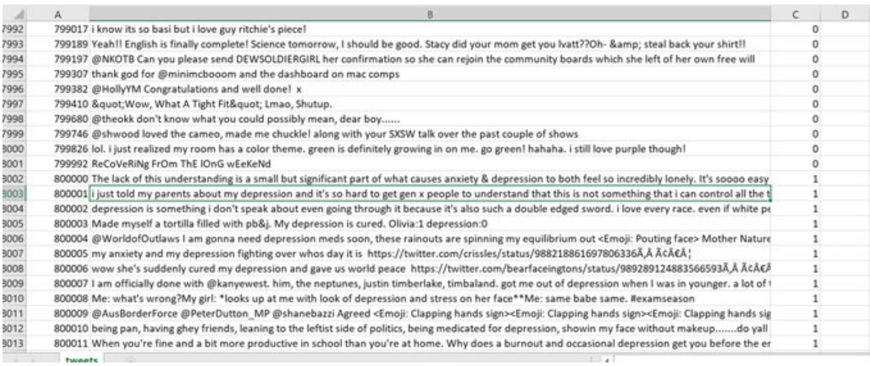
Level of Depression Severity, PHQ-9 Score	Major Depressive Disorder (N = 41) n (%)	Other Depressive Disorder (N = 65) n (%)	No Depressive Disorder (N = 474) n (%)
Minimal, 0–4	1 (2.4)	8 (12.3)	348 (73.4)
Mild, 5–9	4 (9.8)	23 (35.4)	93 (19.6)
Moderate, 10–14	8 (19.5)	17 (26.1)	23 (4.9)
Moderately severe, 15–19	14 (34.1)	14 (21.5)	8 (1.7)
Severe, 20–27	14 (34.1)	3 (4.6)	2 (0.4)

*Depression diagnostic status was determined in 580 primary care patients by having a mental health professional who was blinded to the PHQ-9 score administer a structured psychiatric interview.

Fig. 3 Distribution of PHQ-9 scores

Dataset 2 from Kaggle

Kaggle is used for obtaining labeled datasets that contain Twitter posts. As shown in Fig. 4, this dataset contains both positive and negative tweets (related to anxiety and depression). So, it has both positive and negative sentiments. Twitter is selected over any other social media site (e.g., Facebook); because, as per a survey, Twitter contains many negative tweets regarding depression. After collecting the dataset, you need to extract columns and keep the message columns and the corresponding label column containing 0 and 1. After that, data cleaning is done by removing all non-alphabetic characters then removed all the special characters (e.g., #, @) and



	A	B	C	D
7992	799017	i know its so basi but i love guy ritchie's piece!	0	
7993	799189	Yeah!! English is finally complete! Science tomorrow, I should be good. Stacy did your mom get you lvatt??Oh- & steal back your shirt!!	0	
7994	799197	@NKOTB Can you please send DEWSOLDIERGIRL her confirmation so she can rejoin the community boards which she left of her own free will	0	
7995	799307	thank god for @minimcbooom and the dashboard on mac comps	0	
7996	799382	@HollyYM Congratulations and well done! x	0	
7997	799410	"Wow, What A Tight Fit" Lmao, Shutup.	0	
7998	799680	@theokk don't know what you could possibly mean, dear boy......	0	
7999	799746	@shwood loved the cameo, made me chuckle! along with your SXSW talk over the past couple of shows	0	
8000	799826	lol. i just realized my room has a color theme. green is definitely growing in on me. go green! hahaha. i still love purple though!	0	
8001	799992	ReCoVeRiNg FrOm ThE lOnG wEeKeNd	0	
8002	800000	The lack of this understanding is a small but significant part of what causes anxiety & depression to both feel so incredibly lonely. It's soooo easy	1	
8003	800001	i just told my parents about my depression and it's so hard to get gen x people to understand that this is not something that i can control all the t	1	
8004	800002	depression is something i don't speak about even going through it because it's also such a double edged sword. i love every race. even if white pe	1	
8005	800003	Made myself a tortilla filled with pb&j. My depression is cured. Olivia:1 depression:0	1	
8006	800004	@WorldofOutlaws I am gonna need depression meds soon, these rainouts are spinning my equilibrium out <Emoji: Pouting face> Mother Nature	1	
8007	800005	my anxiety and my depression fighting over whos day it is https://twitter.com/crissles/status/988218861697806336Ã‚Â Ã¢Â€Â¦	1	
8008	800006	wow she's suddenly cured my depression and gave us world peace https://twitter.com/bearfaceingtons/status/989289124883566593Ã‚Â Ã¢Â€Â	1	
8009	800007	I am officially done with @kanyewest. him, the neptunes, justin timberlake, timbaland. got me out of depression when I was in younger. a lot of t	1	
8010	800008	Me: what's wrong?My girl: *looks up at me with look of depression and stress on her face**Me: same babe same. #examseason	1	
8011	800009	@AusBorderForce @PeterDutton_MP @shanebazzi Agreed <Emoji: Clapping hands sign><Emoji: Clapping hands sign><Emoji: Clapping hands sig	1	
8012	800010	being pan, having ghey friends, leaning to the leftist side of politics, being medicated for depression, showin my face without makeup.......do yall	1	
8013	800011	When you're fine and a bit more productive in school than you're at home. Why does a burnout and occasional depression get you before the er	1	

tweets

Fig. 4 Twitter dataset

stop words. For that, NLTK stop word corpus is used. Then, tokenization is done by collecting a bag of depressive words. It was done by stemming the words and have used porter stemmer for this step.

Term frequency–inverse document frequency (TF-IDF) is used; it is a statistical measure that shows how relevant a word is to a document in a set of various documents. It is completed by multiplying two metrics: how many instances a word comes in a document and the inverse document frequency across a set of documents. Through the test, the text is entered, and the result is generated by searching the keywords of the text in the two bags. Two bags will be formed, one having depressive words and one with non-depressive words. The labeling is done; the bag with depressive words has label 1, and the bag with non-depressive words has label 0.

A depression detection test based on Twitter allows the user to write anything about his sentiments and feelings to predict depression. Two drawbacks of the Twitter-based test are that its accuracy is less as compared to PHQ-9-based test as it is a textual test, and it only shows whether a person is suffering from depression or not, unlike PHQ-9-based test, which shows the class of depression. When a keyword is received in the text, the word's probability or frequency of repetition is searched, and it should pass the threshold value.

The probability depends on the dataset, and both the results and probability will change with the change in the dataset.

Model of Depression Detection Using Machine Learning

Figure 5 shows the block diagram of the depression detection model executed using machine learning. The first step, data collection, includes downloading the datasets from two different sources. The first PHQ-9-based dataset is downloaded from online health source Web sites, and the Twitter-based dataset is downloaded from kaggle.com. Data cleaning includes various steps that help to eliminate unnecessary data, like for PHQ-9-based dataset, column extraction is done, and for Twitter-based dataset, stop words are removed. Data normalization in PHQ-9-based dataset includes filling missing data in cells by using the mode method. In Twitter, the dataset includes word tokenization for forming a bag of words.

Feature extraction is a step carried out to decrease the features and make new features from the existing ones. Training data include the splitting of data into two sets for training and testing, respectively. For PHQ-9-based dataset, 60% data are trained, 40% is tested, and for the Twitter-based data set, 70% is used for training and 30% for testing. Depression classification involves the classification of depression. The PHQ-9 based dataset classifies depression into five classes (no depression, mild depression, moderate depression, moderately severe, and severe), whereas the Twitter-based test classifies in only two ways, either yes or no depression. Finally, detection outcome involves the generation and display of results.

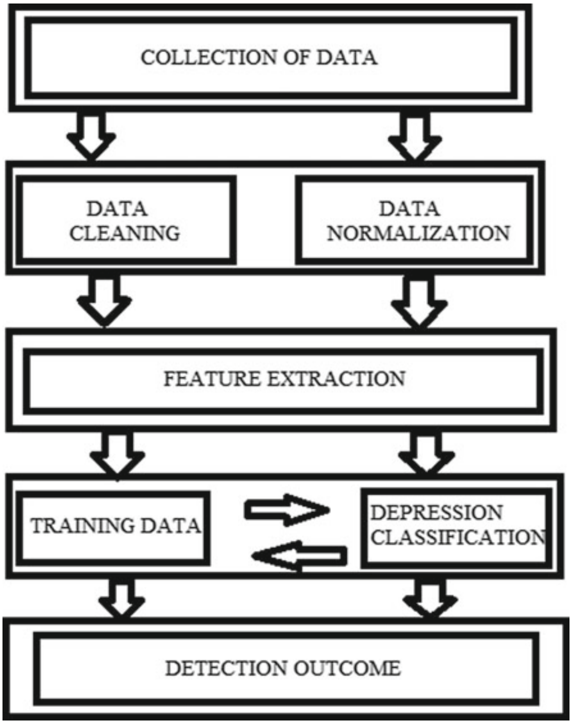


Fig. 5 System block diagram

Results

Machine Learning Algorithms

In this proposed work, many previous studies have been analyzed to detect depression using machine learning. It has been observed that for conducting such similar works, most of the data are collected from social media sites, eg., Facebook and Twitter. This huge amount of data from social media sites has many informal languages, and hence, accuracy gets compromised. However, to extend the model's accuracy, efficiency, and applicability, it is necessary to use data from some professional health approved sources. In the proposed model, data are collected from two sources. One is the PHQ-9 dataset approved by the National Institute of Health. It is a numerical type of dataset and hence improves the efficiency and accuracy. Another one is the Twitter dataset collected from Kaggle.com. Two Web-based depression detection tests are developed, one from the PHQ-9 questionnaire, quiz-based, and the other from the Twitter dataset, which is sentiment based.

The front end uses HTML, CSS, JavaScript, and Bootstrap in the proposed work, and the back end uses machine learning, Python, and flask. The Web site, which contains two depression detection tests, works on the front end with the help of the

Table 3 Accuracy obtained

Classifier	Accuracy (%)
Support vector machine	99.74
Decision tree classifier	99.37
Random tree classifier	99.59
Naive Bayes classifier	96.62

back end. Whenever an option is selected on the Web site, the front end sends a request to the back end, and then, the backend sends back the response.

For the PHQ-9-based test, 60% data are trained, and 40% data are tested, which gives us a very high accuracy of 99.74% when SVM is used. The accuracy obtained is remarkably higher than any of the previous works done. This is because of the data preprocessing methods and the machine learning algorithm used in the proposed work.

For the Twitter-based test, 70% of data is trained, and 30% is tested. Even after putting a large section of data for training, the desired accuracy is not obtained; this is because of the textual nature of the dataset. There is an enormous use of informal language in the tweets, which causes problems in getting accurate results. Also, the accuracy will vary with the change in the dataset.

The accuracies of the proposed work are compared with that of previous works in this section, and insight of the Web site and tests are given.

Table 3 shows accuracy obtained by training 60% data and testing 40% data in the proposed system. SVM gives the highest accuracy of 99.74%, significantly higher than any of the accuracies achieved in any of the previous works.

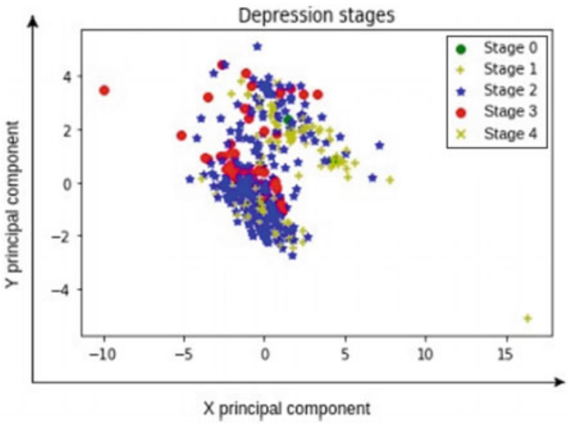
Table 4 shows the accuracy for various machine learning algorithms achieved in the literature review's similar work [6]. The highest accuracy in results is attained by implementing the logistic regression classifier. However, the accuracy is still less than the proposed work because of a different method used for the data preprocessing.

- Support Vector Machine: The SVM algorithm plots a dataset in two dimensions. There are linear and non-linear separators depending on the results. The kernel was needed for non-linear separators. SVM is used in the proposed work as it gave the highest accuracy for the PHQ-9-based test. An accuracy of 99.74% is achieved by using SVM. Random Forest Classifier: It fits a figure of decision tree classifiers on different sub-samples of the dataset. It increases exactness. An accuracy of 99.59% is achieved using this classifier.

Table 4 Accuracy result obtained in the related work [6]

Classifier	Accuracy (%)
Random forest tree classifier	82.05
XG Boost classifier	84.02
Logistic regression classifier	86.45
Support vector machine	85.44

Fig. 6 PCA plot is used to lessen a big set of dimensions to a little set which still gives all the information



- Decision Tree Classifier: The decision tree classifier poses a sequence of watchfully created questions about the attributes of the test record. An accuracy of 99.37% is achieved.
- K-Nearest Neighbors Algorithm (k-NN): It is used for classification and regression. An accuracy of 99.51% is achieved for the k-NN classifier.
- Naïve Bayes: These classifiers are tremendously measurable, through loads of parameters linear in the figure of variables in a learning problem. An accuracy of 96.62% is achieved for this algorithm.

PCA for Plot Training Dataset of Questionnaire

40% data is used for testing and 60% for training. Figure 6 shows the training of the dataset by a PCA plot in which a dataset has more than nine features is plotted. A new coordinate system is discovered, with each point having a different (x, y) value. The axes are combinations of features known as "principal components," which give one axis lots of distinction. PCA plot is used to lessen a big set of dimensions to a little set which still gives all the information.

Correct Representation of Author Names

Figure 7 shows the login page of the Web site. After entering the username and password credentials, the Web site's home page or landing page will appear. It is a three-minute quiz-based depression detection test which is based on the PHQ-9 questionnaire.

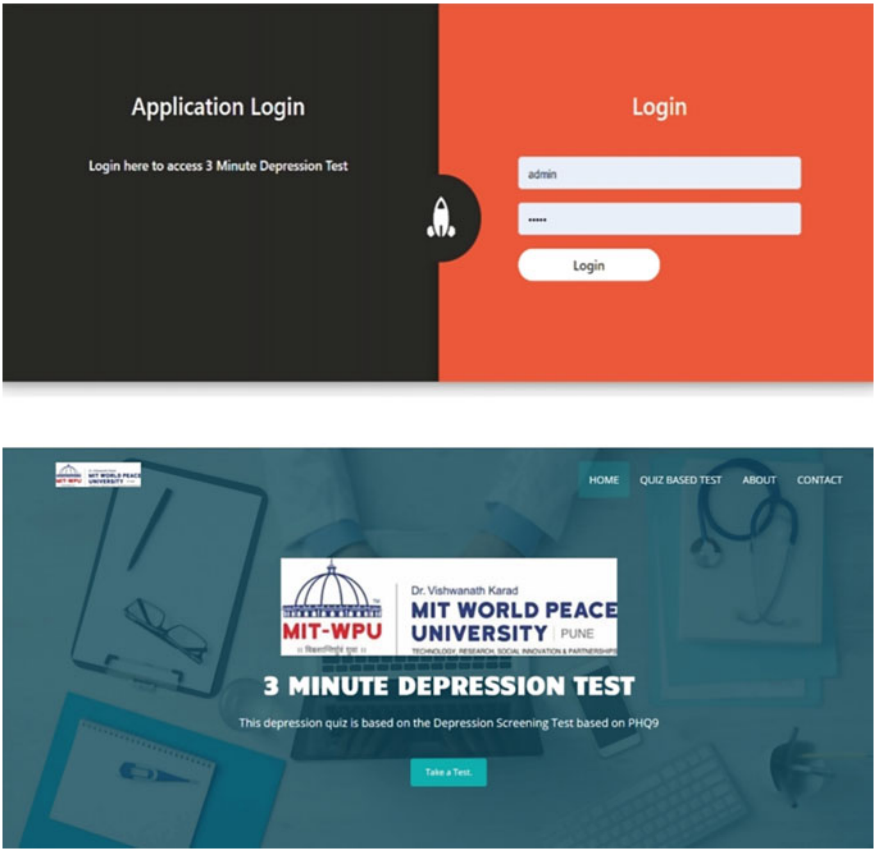


Fig. 7 Login and home page

Figure 8 shows the PHQ-9 quiz-based depression detection test. There are ten questions in the test, and for each question, there are four options. The input will be given by selecting the radio buttons.

Figure 9 illustrates the test result generated for a clinically depressed patient who had taken this test. The results show that the patient is suffering from moderately severe depression, which is accurate.

Figure 10 represents the sentiment-based test. This test is based on the dataset acquired from Twitter. The test taker gives a textual input, and then, keywords are identified in the sentence, and the frequency is checked from the bag of words stored in pickle files if the frequency value of the keyword is greater than the threshold value, then the result is generated. The result is either Yes or No for depression.

The graph in Fig. 11 illustrates the cases of depression found in different age groups ranging from 15 to 25. It can be deduced that the people ranging from the 19–21 age group have the highest cases of depression recorded.

The pie chart from Fig. 12 illustrates that 45% of people feel depressed and hopeless on several days.

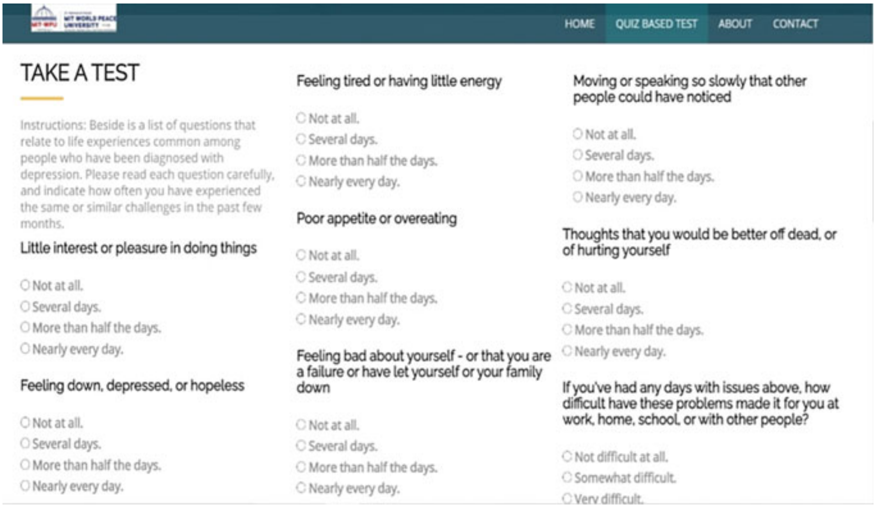


Fig. 8 PHQ-9-based depression detection test

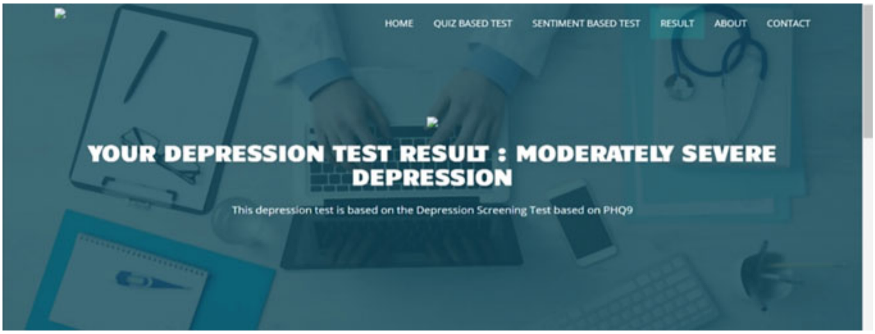


Fig. 9 Result generated for the test

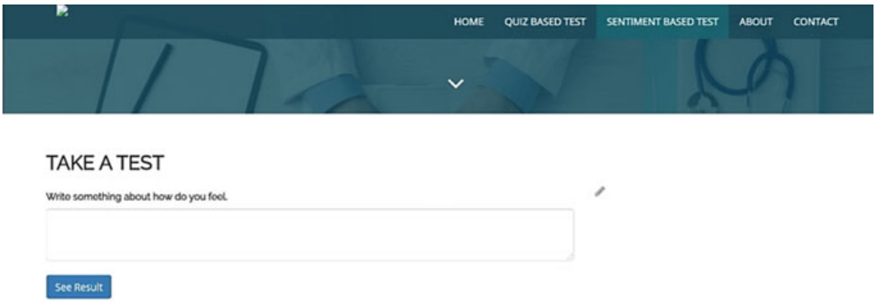


Fig. 10 Twitter-based depression detection test

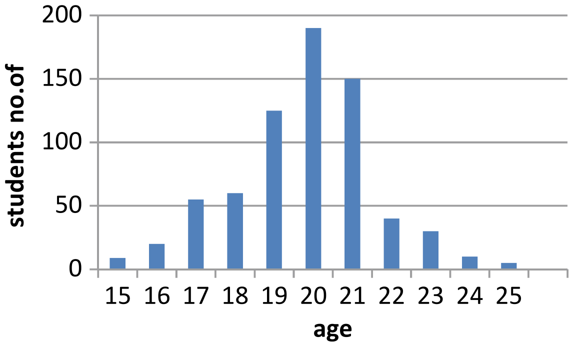


Fig. 11 Depression in different age groups

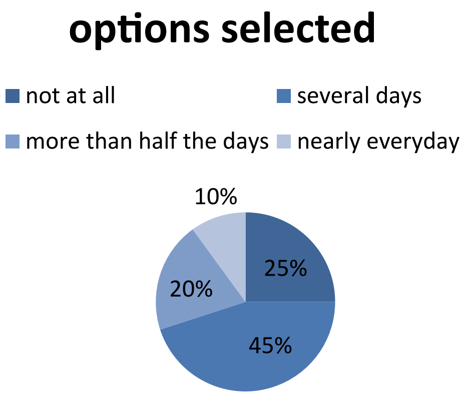


Fig. 12 Comparison of option selection

The pie chart in Fig. 13 shows a survey conducted for the proposed work on a social networking site. The survey was conducted for various people ranging in the

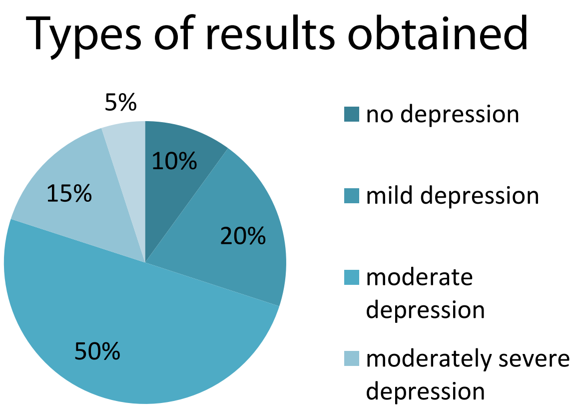


Fig. 13 Results obtained based on survey

age group of 20–25. The results clearly show that a majority of people were detected as having moderate depression.

Conclusion

In the proposed work, quiz-based and sentiment-based tests have been successfully implemented. The PHQ-9 questionnaire is used for the quiz-based test, and for the sentiment-based test, the Twitter dataset from Kaggle is used. The proposed work accuracy of 99.74% is achieved for the PHQ-9-based test, which is higher than any previous works. The accuracy is so high because the mode method is used for handling the missing data during the data preprocessing and by using SVM for predicting results. However, the desired accuracy is not achieved from the Twitter-based sentiment test because informal language is vast in a text-based dataset and input.

Surveys have been conducted. The depression is accurately classified, and correct results are generated, and this helps in the early detection of depression.

Major depression causing factors among the age group of 15–29 found through test are non-accomplishment, parental force, aggravation, body shaming, inferiority complex, examination tension, peer pressure, physical, and sexual abuse. With this work, we want to ensure that the person experiencing depression detects it at the earliest before it gets worse and serious through our intensely personal depression detection test. With this test, mental health professionals would already know the patient's mental condition and can give better treatment to patients.

References

1. Victor E, Aghajan ZM, Shewart AR, Christian R (2019) Detecting depression using a framework combining deep multimodel neural networks with a purpose-built automated evaluation. Psychol Assess 31(8):1019–1027
2. Hatoon S, Alsagri H, Ykhlef M (2020) Machine learning-based approach for depression detection in Twitter using content and activity features. Cornell University E103.D(8):1825–1832
3. Arun V, Prajwal V, Shyam V (2018) A boosted machine learning application for detection of depression. IEEE symposium series on computational ıntelligence (SSCI), pp 41–47
4. Suyash D, Pramila C (2020) Depression detection on social media using machine learning techniques: a survey. Int Res J Eng Technol (IRJET) 7(11):97–100
5. Islam MR, Kabir MA Depression detection using machine learning techniques and social network data. Health Inf Sci Syst 1–12
6. Jain S, Narayan SP, Kumar RA A machine learning based depression detection using questionnaire and Twitter. In: 5th students conference on engineering and systems. National Institute of Technology, Allahabad, vol 5. Pp 56–61
7. Nalinde PB, Shinde A A Machine learning analysis for detection of psychological disorders at OSN. Int J Innov Technol Exploring Eng (IJITEE) 8:3293–3298

8. Ramalingam D, Sharma V, Zar P Study of Detection Analysisusing MLT. In: International Conference on Multidisciplinary Perspective in Engineering and Technology (ICMPET_19) Int J Innov Technol Exploring Eng (IJITEE) 187–191
9. Shreshthra K Machine learning for depression analysis using Twitter. Int J Comput Eng Res Trends (IJCERT) 5
10. Lin H, Jia J, Guo Q (2011) Psychological stress detection. Preceedings on the 22nd ACM International Conference on multimedia, pp 507–516
11. Tang J, Zhang Y (2011) Quantitative study of emotional states in social networks. IEEE transaction on affective computing, pp 132–144
12. Saha B, Phing D (2016) Framework for classifying online mental health related communications with an interest in depression. IEEE Biomedical Health Information
13. Savaria E, Chang C, De Lorenzo RJ, Chen Y (2016) MIDAS: mental Illness detection and analysis via social media. IEEE/ACM International Conference on Advances in Social Network Analysis and Mining (ASONAM), pp 1418–1421
14. Adrian, Mitchell M, Houg D (2017) Multitask learning for mental health conditions with limited social media. Preceeding of 15th Conference of the EACL, pp 152–162
15. Kang K, Yoon C, Kim EY (2016) Identifying depressive users multimodal analysis. International Conference on Big Data and Smart Computing (Big comp), pp 231–238
16. Aldarwish MM, Ahmad HF (2017) Prediction depression levels using social media posts. IEEE 13th International Symposium on Autonomous Decentralized System (ISADS), pp 277–280
17. Shuai H-H et al. A comprehensive study on social network mental disorders detection via online social media mining. IEEE Transactions on knowledge and data engineering, pp 1212–1225
18. Asad NA, Mahmud Pranto MA, Afreen S, Islam MM (2019) Depression detection by analyzing social media posts of user. IEEE International Conference on Signal Processing, Information, Communication and Systems (SPICSCON), pp 13–17

Early Detection of Autism Spectrum Disorder (ASD) Using Machine Learning Techniques: A Review

Shreyas Kalikar, Amol Sinha, Sumit Srivastava, and Gaurav Aggarwal

Abstract Autism spectrum disorder (ASD) is a complex psych-growing condition that affects an ability to perform socially by individuals, in verbal and non-verbal exchange, and also includes repetitive/anomalous behavior. Though ASD can become irreversible and a lot of people diagnosed with this condition do continue to lead normal complete lives, it is also important to address this condition and early as possible to minimize hindrances, as it only gets more rigid as the individual grows older. Earlier, a lot of progress on this has been made using machine learning techniques through various algorithms that process a variety of datasets. The motivation behind this review is to bring to light an array of results from different methods already implemented to give perspective on possible ways they can complement each other. The concept of computer vision and its effectiveness, however, is something less investigated and is being looked into.

Keywords Autism spectrum disorder · Machine learning · Generative adversarial networks · Linguistic

S. Kalikar (✉) · S. Srivastava · G. Aggarwal
Department of Information Technology, Manipal University Jaipur, Jaipur, India
e-mail: shreyas.189302036@muj.manipal.edu

S. Srivastava
e-mail: sumit.srivastava@jaipur.manipal.edu

G. Aggarwal
e-mail: gaurav.aggarwal@jaipur.manipal.edu

A. Sinha
Department of Computer Science, Manipal University Jaipur, Jaipur, India
e-mail: amol.189301080@muj.manipal.edu

V. Bindhu et al. (eds.), *Proceedings of Third International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 844,
https://doi.org/10.1007/978-981-16-8862-1_66

Introduction

Today, according to the Center for Disease Control and Prevention [1], about, 1% of the world's entire population has autism spectrum disorder, considering all types and all kinds of severity. This number has only substantially risen over the years, the WHO declaring 1 in every 270 people have ASD, the approximate coming to 1.8 million people. It is unknown how common ASD is in countries with lower- or middle-income strata. It is a neurodevelopmental condition that affects subjects at a very early stage, typically at 2–3 years old. However, there have also been cases where the subject loses their learning ability at the toddler stage, and either finds it hard to learn new skills further on, or tends to forget skills that have already been learned. Factors such as genetics also prominently come into play; autism has been seen to be prevalent among almost three times boys in ratio to girls. There is no single symptom that can give an affirmative diagnosis of autism, instead, there is an array of signs like behavioral deterioration in interaction, repetitive behavior, and restrictive linguistic skills.

The aim of this review paper is to discuss, through sections, areas of research that have been explored using non-clinical methods, or clinical methods that have complimented machine learning techniques for an accurate diagnosis. The directive is also to reinstate the importance of an early diagnosis, which can be made possible through the use of technology, making its accessibility tenfold, while still being precise. Further, it is planned to emphasize on some drawbacks and developments that can be made on currently existing remedial projects which can be implemented to increase efficiency and accuracy of the diagnosis (Fig. 1).

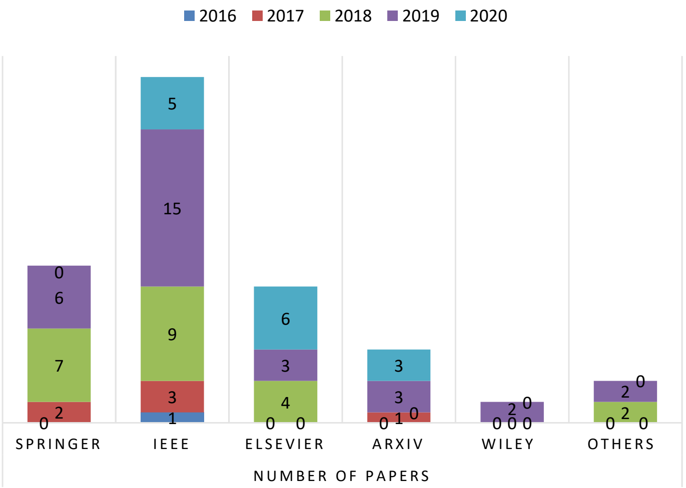


Fig. 1 Number of papers published since last five years on ASD detection

Signs and Symptoms of ASD

- Repetitive behavior [2] can be of many forms: motor stereotypy, sameness orders, obsessive behavior, repetitive self-injury, etc. This section, however, has been one of the least explored sections that is a major "tell" of this condition. Some of the more prominent signs that are under discourse are hand-flapping, teeth clenching, head banging, fist-clenching, and repetitive gestures and postures.
- Behavioral changes in interaction and social skills are some of the most prominent symptoms of the early onset of ASD. Children may not be able to make eye contact, look or listen attentively, and may have difficulties in holding conversations. They may also give slow responses to any sort of stimuli that may require their attention, which can be as simple as calling them by their name.
- ASD might affect motor behavior. Clenching, movement, or even basic actions such as a head tilt can have delayed or involuntary responses and response times, respectively, which was observed in a "head tilt reflex" test [3]. This also resulted in an observation that imbalanced motor skills might have an impact on developing linguistic skills as the child grows older.
- A lack of response specific stimuli is also observed, which can be analyzed through comparison of gaze or facial features. Undiagnosed [4], ASD can also morph into other neurological disorders such as depression, anxiety, disturbed sleep, prolonged sleep, all due to physiological and neurological connection irregularities.

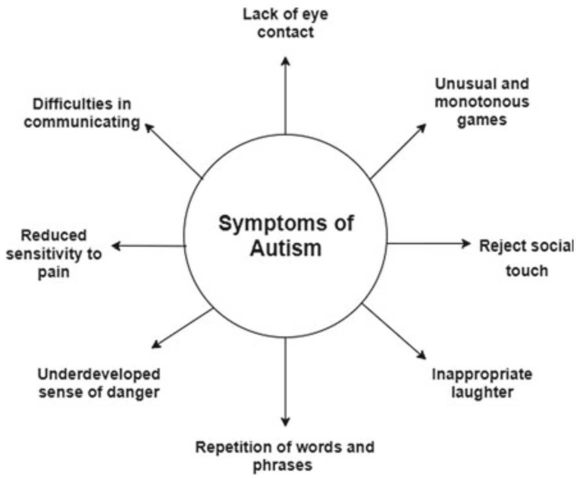
Detecting ASD can be a challenging task. There is no such medical test or quantitative test that can conclude on an ASD diagnosis. Professionals observe developmental stages and general behavior to gauge and make a diagnosis. Getting an early diagnosis is extremely essential to intercept any further severity. The process of detection, however, can be a long observational period, and the option of early detection may eventually no longer be an option (Fig. 2).

Causes of ASD

In specific, there are no causes for ASD. However, some factors [5] that can be taken into account are genetics and environmental factors. For a lot of individuals, genetic disorders like the Rett syndrome may be associated with ASD. Similarly, a lot of people may have genetic mutations that may affect brain development and may determine the severity of ASD. Possible environmental causes such as medications or complications during the pregnancy period, infections, drug or alcohol abuse, and pollutants are being researched on and explored further. Also, babies that are born prematurely (before 26 weeks) are at a much higher risk of ASD.

Another suggested cause could be illness and antibodies during pregnancy [6]. Studies have shown that mothers that were exposed to illnesses like rubella have had a 37% statistic on having children with the neurological disorder being a learning

Fig. 2 Sıgns and symptoms of ASD



impairment. Fortunately, rubella is an obsolete infection, and since then, there hasn't been an impactful epidemic that risked the possibility of ASF through maternity.

Cure for ASD

As of today, there is no set-in-stone cure for ASD, only mitigation methods that can control the severity. However, it is also very important that this should be detected at a malleable age group as these behavioral temperaments may continue to grow into different directions of the spectrum. Research suggests [7] that the development of fluent speech before the toddler phase (around the age of 5) shows positive signs of cognitive development, as the chance of mitigation at an older age is exponentially lower than that at a younger age (around the age of 3).

The idea behind an early diagnosis is that younger children have greater malleability with regards to their thought processes, which greatly increases the odds of improvement in areas where there is developmental lacking. Similarly, applied behaviour analysis (ABA training) interventional methods that promote a correct response to prompts.

Methodology for Classification of ASD

Dataset Collection

Numerous experiments have been conducted over the years to develop effective strategies for detecting autism. The most commonly used data collection tools are AQ-10, QCHAT-10, ADOS, ADI-R, and the social responsiveness scale (SRS) [8].

Further, to build a useful predictive model, the AQ-10 dataset is used, consisting of three different datasets based on the problems in the AQ-10 testing tool. These three datasets contain data for the age groups of 4–11 years old (children) and 12–17 years old (teenagers) [9], and over 18 years old (adults). AQ-10 (or quotient tool for patients with autism) is used to determine whether to recommend a person for a comprehensive assessment of autism. These themes mainly involve the fields of imagination, communication, and social interaction.

The best-known data sources used in the studies include the Boston Autism Consortium (AC), the Autism Genetic Resource Exchange (AGRE), and the National Autism Research Database. It is necessary to have regulated ASD data repositories for machine learning (ML) studies to achieve standardized comparative results [10].

Feature Extraction

Data preprocessing is one of the most important tasks in the application of machine learning models. Initially, the data collected for training machine learning tasks is not suitable for training purposes. Therefore, to prepare the data for input into machine learning processes, performing data preprocessing is crucial. Processing includes techniques for identifying noise and managing missing data. It involves the transformation of data in order to improve its reliability and thus the efficiency of algorithms in machine learning, such as predictive accuracy and reduced training time.

One of the major tasks of data preprocessing is the collection of features; of which, only, a few functions are selected and used in the training process of the machine learning algorithm. The goal is to find the best feature subset that can improve the efficiency of the learning algorithm. Selection of suitable features involves defining relevant features and using them in machine learning applications, as well as ignoring the remaining features that have little or no predictive information.

After reviewing various research papers on ASD, it is observed that the features selected to detect autism are based on physical characteristics such as facial features [11], textural and statistical analysis features [12], and asymmetry-based features. In the application of coordinate-based image processing techniques, the positions of objects are detected as x and y coordinates in an image [13]. Other extraction techniques include skin color-based and Haar digital image processing for increased accuracy of object detection.

- Face detection

  Histogram of oriented gradient (HOG) is used for frontal face detection. Eye and mouth markings are checked within this region of interest (ROI). A shape predictor object is applied to detect those landmarks, which accurately identifies the point locations of prime facial landmarks such as the corner of the mouth and eye that defines the shape of the object in the input frame sequence [13].
- Hand detection

  Before initiating the hand detection process, preprocessing techniques are applied to the image data frame. After successful detection of the face, the bottom of the frame is cropped. Then, hands and legs are identified using skin detection techniques.
- Preaggression detection

  There is more white color detected in adjacent image frames of the mouth; then, we can assume that teeth are presented in the frame. So, if the mouth is opened and teeth are presented, then we can assume it as tightened teeth detected. By considering these variations in mouth and eyes, the system indicates that autistic child tends to show preaggressive behavior.

  There are many forms of aggression, such as physical aggression, neurasthenia, and self-harm, which can prove to be injurious to children with autism as well as others around them. Before the aggressive state, children with autism usually exhibit pre-aggressive behaviors.

  In various models, the angry emotions of autistic children are used to recognize the facial expressions before the attack. Therefore, there is a need to successfully acknowledge the changes in the eyes and mouth distinctions [14]. Autistic children usually open their mouths with clenched teeth for 60 seconds or more. The image processing system returns the value and reports it as an open mouth. Therefore, if the mouth is opened and the teeth are exposed, it can be assumed that the teeth have been tightened, and these mouth and eyes variations successfully indicate if the child is prone to aggressive behavior or not.

A compact and essential collection of features not only simplifies the representation of the data pattern but also assists the classifier in the model training process. When the number of training patterns is minimal, fewer features reduce the curse of dimensionality. Extracting and selecting features are two crucial steps to create a precise representation of the pattern.

Evaluating Classification Techniques

A standard classification model using machine learning algorithms, wherein the cases have been categorized previously, is best suited for ASD detection [15].

Various algorithms have been developed and verified to predict the characteristics of autism. Classification techniques such as KNN, linear regression, Naïve

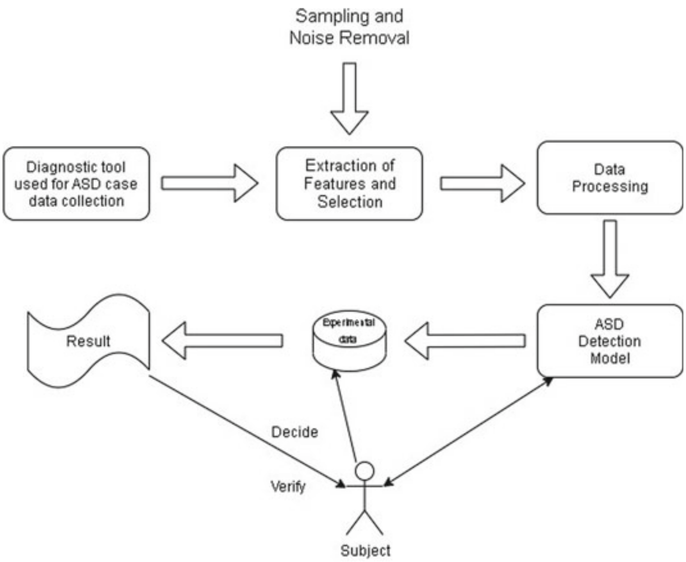


Fig. 3 ADS diagnostic test using contemporary feature extraction methods

Bayes, SVM, random forest have been constantly built and applied for ASD detection [16, 17].

In the next section, we will review some recent work on classification techniques and feature extraction used for ASD (Fig. 3).

Literature Review

This section discusses the technologies and detection systems that are already in place or are undergoing research currently. All these changes were brought upon only after years of clinical diagnosis methods, explaining how important it is to diagnose ASD at the earliest, and its effects and severity as the subject ages. Earlier [18], it was conventionally a parametric diagnosis consisting of two levels. In the first level, the subject is screened to meet the criteria of "delayed development," which is followed by an ASD-specific screen if the criteria in the first diagnostic checks out. If they have signs of ASD-specific conditions, the subject is referred to level 2, which would be the official clinical diagnosis and evaluation of ASD by a professional. Further, neurological and specific evaluations are made to estimate the child's developmental status, and this implementation is considered essential.

There have been credible developments in the field of machine learning when it comes to ASD diagnosis. This is due to the detection of different types of diseases in the same field using machine learning garnering a lot of implementation such as [19] Cruz et. al. attempted to diagnose cancer using machine learning, and there was also an effort made to calculate the odds of the subject being diabetic. ASD detection has also gotten a fair amount of global contribution such as [20] Wall et. al. who implemented reduced screening times, which resulted in quicker detection of ASD symptoms by making use of alternating decision tree (ADTree). With data of 891 subjects, they achieved a substantial amount of precision with the use of Autism Diagnostic Interview and Revised, or the ADI-R method. However, the test was restricted between the ages of 5–17 and failed to predict ASD for the major different age demographics, i.e., children, adolescents, and adults. [20] Wall et. al. also worked toward autism classification with validation and reduced screening time test and discovered that ADTree, as well as the functional tree, had performed with great precision, specificity, and sensitivity.

Liu et al. [21] studied the possibility of face scanning and recognition patterns in the identification of children with ASD by making use of machine learning algorithms to analyze a gaze movement dataset for the categorizing purpose. The study concluded with an accuracy of 88.51%, the sensitivity of 93.10%, the specificity being 86.21%, and AUC being 89.63%. Allison et. al. [22] considered the usage of "Red Flags" as a tool for ASD detection with autism spectrum degrees for children and adults, then went on to nominate them to AQ-10, achieving more than 90% accuracy.

With the procedural advent of technology, however, the end goal of early diagnosis is possible using machine learning techniques that can be modeled and train to provide very specific results as per the need to meet a more elaborate criterion which provides more accurate data and conclusions on developmental profiles of a child.

- *Generative Adversarial Networks in Speech Recognition and Correction*: There has been great progress on the speech recognition aspect of early detection of ASD using machine learning [23]. Considering there has been a lot of in-depth research and implementation on more approachable elements of diagnosis, being visual cues or biological evidence, there is also the facet of speech impairments that can be characterized in ASD. The paralinguistics of speech by a child can provide a great amount of information on any linguistic impairments or underlying developmental barriers, and using generative adversarial networks (GAN), core markers used for accurate data collection can be obtained. Speech and linguistic attributes such as pitch, loudness, rhythm, and speed have been found to be useful when using comparative algorithms such as GAN, as audio cues also have intricate comparisons that can be made. Research shows that the use of GAN can enhance speech skills, but only if used on a deep learning level [24], for the purpose of automatic diagnosis. Here, the GAN periodically takes in sample training data and then is made to predict outcomes from either the training speech data or the generated result. The discriminator here makes the differentiation from the source and the generated output, again and again, predicting the outcome more accurately after each cycle. This, eventually, allows the discriminator to compare the generated output to the initial pre-training input (which is now more refined), and gives the outcome based on how accurately it gauges the similarity or dissimilarity.

- *Modified Grasshopper Optimization Algorithm (MGOA)*: The grasshopper optimization algorithm is an algorithm derived from the mathematical nature of "swarming," a trait followed by organisms, particularly grasshoppers in this conceptual model [25]. It has been modeled after mathematical deliberations were made looking at the interactive capabilities of grasshoppers traveling in massive groups, or swarms. Some characteristics that have been implemented in this model inspired by this event are movement pattern tracking between nymphs and grasshoppers at adulthood, and their food-seeking capabilities and desires that cause them to display such behavior. The GOA has proven to be very effective in feature extraction methods, by searching in anonymous search spaces looking for entries. The implementation of GOA along with support vector machines have been used in the past for the detection of seizures and epilepsy, as it provides almost completely accurate data comparison with those which have an already provided list of clinical data pertaining to these disorders. MGOA (being GOAs supported at an MLP level) can be extremely effective for the detection of the early onset of neurological disorders that are and may follow the diagnosis of ASD.
- *AS Detect*: Josephine Barbaro [26], along with her team came up with AS Detect, which is a project that has made use of the accessibility and versatility of families possessing mobile phones. AS Detect is a mobile phone application that is made to allow parents of children in the demographic of up to 3 years to use reaction stimuli to gauge if the subject is at risk of ASD or not. It uses visual cue detection and assesses the visual cues of the subject using small videos that demand expression prompts. This is designed to cut down on the long observational period that an initial clinical diagnosis might take, which may exceed that critical 3-year-old mark and make a lot of adversities harder to mitigate. Following this, parents who find the objective results of this app diagnostic alarming or concerning can share these results with a professional clinician who can try to eliminate some barriers (the "wait-and-see" approach) which can be molded only at a very young age. The main aim of this was not only to aid the early detection and prevention of early onset of ASD, but it is also to encourage more accessible demographics, parents, or caregivers, for example, to take the opportunity to help their child have a more comfortable future by attempting to eliminate and streamline autism symptoms and their consequent neurological disorders.
- *Combining gaze and demographic feature descriptors for autism classification*: In this paper, Canavan and other authors [27] proposed a novel method to classify autism using gaze and demographic descriptors (including subject's age and gender). They create feature descriptors that contain the age and gender of the subject and features based on eye gaze data, which are further tested against three different classifiers—random regression forest, C4.5 decision tree, and PART.

  The study showed promising results as it proved that gaze can be widely used as a marker for autism spectrum disorder, complementing various other approaches. It was also observed that the validated medium risk classifiers, out of the low, medium, and high-risk ones, have the highest classification level with only one error. However, since the test was performed mostly on subjects over the age of 60, this model leaves a lot of scope for detection of ASD predominantly in children, as well on larger datasets.

- Identifying predictive features of autism spectrum disorders in a clinical sample of adolescents and adults using machine learning: In previous strides of new developments in the field, SVM has been extensively used as these models showed the most accurate and promising results. The frequent use of SVM for the classification of ASD data is contributing to the fact that SVMs have great predicting abilities. As compared to the usage of the random forest [28], results were more promising using the SVM classifier. First, feature extraction is done on a given trainable dataset to shortlist a smaller set of profiles that have the same predictive power level. Next, this reduced feature model is now trained on the trainable dataset again over and over, to achieve more suitable, stable predictive results. Then, the next step is to train the same reduced feature model with a dedicated dataset, to allow the model to be trained to differentiate between non-ASD and ASD subjects and classify their traits accurately. From the whole set, five main features were extracted for classification: gesture features, unusual eye contact, social facial expressions, quality of social response, and amount of reciprocal social communication.
- *Analysis and Detection of Autism Spectrum Disorder Using Machine Learning Techniques*: Suman Raj and team [29] discuss the exploration of usage of algorithms such as Naïve Bayes, support vector machine, and logistic regression to evaluate three datasets. The first dataset deals with ASD in children; the second deals with ASD in adults, and the third dataset consists of adolescent subjects with ASD. The result strongly suggested that CNN algorithm works most effectively on all three datasets.
- *Mobile detection of autism through machine learning on home video*: Qandeel Tariq et al. [30] hypothesized that it takes from 20 to 100 behaviors and several hours to evaluate, and is a process that is a long, procedural process. They have analyzed records from two standard records which consist of classifiers that provide extractable information for nonexpert examiners from merely 3 min long home videos examining behavior. They use eight independent classfiers, with 116 home videos of children diagnosed with ASD and 46 videos of the natural development of children. All classifiers showed an accuracy of greater than 94% across all ages.
- *Machine learning approach for early detection of autism by combining questionnaire and home video screening*: Halim Abbas and team [31] incorporate two algorithms to train the model, one dealing with a short questionnaire which is parent-reported and the other handling analysis of home videos. This was performed on a collective set of 162 children and are quantified on the basis of sensitivity and specificity. This has been come into inception to compare machine learning processes with clinical settings. Some screeners they compare to are M-CHAT and CBCL.

Future Scope

In the medical area, data scarcity is usually a major issue that may be handled rather easily with the aid of deep GANs. In addition to raw data, handmade features may be derived from data and fed to DL networks, which can assist improve performance by combining the potential of traditional approaches with the potential of DL-based models. To help ASD patients, several researchers have developed different DL-based rehabilitation techniques. Designing a low-power consumption DL algorithm-based gadget that is dependable, accurate, and wearable is the future aid for ASD sufferers. Wearing smart glasses to assist youngsters with ASD is a feasible rehabilitation technique.

Conclusion

The paper reviews the various causes of autism spectrum disorder (ASD) and the machine learning approaches involved in predicting ASD at various levels. The work mostly covers the image-based recognition algorithms; however, it is interesting to study the generative adversarial networks (GAN) approach for speech disorder and the mathematical nature of swarming for the detection of ASD. The research is further aligned to extend the existing image and signal-based approaches to identify the ASD in the video samples and attempt to develop the hybrid approaches acceptable for both images and video-based ASD detection.

References

1. Beary M, Hadsell A, Messersmith R, Hosseini MP (2020) Diagnosis of autism in children using facial analysis and deep learning. arXiv preprint arXiv:2008.02890
2. Bodfish JW, Symons FJ, Parker DE, Lewis MH (2000) Varieties of repetitive behavior in autism: comparisons to mental retardation. J Autism Dev Disord 30(3):237–243
3. Samango-Sprouse CA, Stapleton EJ, Aliabadi F, Graw R, Vickers R, Haskell K, Sadeghin T, Jameson R, Parmele CL, Gropman AL (2015) Identification of infants at risk for autism spectrum disorder and developmental language delay prior to 12 months. Autism19(3):327–337
4. Bhat S, Acharya UR, Adeli H, Bairy GM, Adeli A (2014) Autism: cause factors, early diagnosis and therapies. Rev Neurosci 25(6):841–850
5. Ratajczak HV (2011) Theoretical aspects of autism: Causes—a review. J Immunotoxicol 8(1):68–79
6. Folstein S, Rutter M (1978) A twin study of individuals with infantile autism. In: Autism. Springer, Boston, MA, pp 219–241
7. Woods JJ, Wetherby AM (2003) Early identification of and intervention for infants and toddlers who are at risk for autism spectrum disorder
8. Shahamiri SR, Thabtah F (2020) Autism AI: a new autism screening system based on artificial intelligence. Cogn Comput 12(4):766–777

9. Cavus N, Lawan AA, Ibrahim Z, Dahiru A, Tahir S, Abdulrazak UI, Hussaini A (2021) A systematic literature review on the application of machine-learning models in behavioral assessment of autism spectrum disorder. J Personalized Med 11(4):299
10. Thabtah F (2019) Machine learning in autistic spectrum disorder behavioral research: a review and ways forward. Inform Health Soc Care 44(3):278–297
11. Katuwal GJ, Cahill ND, Baum SA, Michael AM (2015) The predictive power of structural MRI in Autism diagnosis. In 2015 37th annual ınternational conference of the IEEE engineering in medicine and biology society (EMBC). IEEE, pp 4270–4273
12. Zhou Y, Yu F, Duong T (2014) Multiparametric MRI characterization and prediction in autism spectrum disorder using graph theory and machine learning. PloS One 9(6):e90405
13. Gamaethige C, Gunathilake U, Jayasena D, Manike H, Samarasinghe P, Yatanwala T (2017). “SenseA”-autism early signs and pre-aggressive detector through ımage processing. In: 2017 Asia modelling symposium (AMS). IEEE, pp 125–130
14. Xanthopoulou M, Kokalia G, Drigas A (2019) Applications for children with autism in preschool and primary education. Int J Recent Contributions Eng Sci IT 7(2):4–16
15. Bone D, Bishop SL, Black MP, Goodwin MS, Lord C, Narayanan SS (2016) Use of machine learning to improve autism screening and diagnostic instruments: effectiveness, efficiency, and multi-instrument fusion. J Child Psychol Psychiatry 57(8):927–937
16. Büyükoflaz FN, Öztürk A (2018). Early autism diagnosis of children with machine learning algorithms. In: 2018 26th signal processing and communications applications conference (SIU). IEEE, pp 1–4
17. Goh KL, Morris S, Rosalie S, Foster C, Falkmer T, Tan T (2016) Typically developed adults and adults with autism spectrum disorder classification using centre of pressure measurements. In: 2016 IEEE ınternational conference on acoustics, speech and signal processing (ICASSP). IEEE, pp 844–848
18. Pinto-Martin J, Levy SE (2004) Early diagnosis of autism spectrum disorders. Curr Treat Options Neurol 6(5):391–400
19. Cruz JA, Wishart DS (2006) Applications of machine learning in cancer prediction and prognosis. Cancer inform 2:117693510600200030
20. Wall DP, Dally R, Luyster R, Jung JY, DeLuca TF (2012) Use of artificial intelligence to shorten the behavioral diagnosis of autism. PloS One 7(8):e43855
21. Liu W, Li M, Yi L (2016) Identifying children with autism spectrum disorder based on their face processing abnormality: a machine learning framework. Autism Res 9(8):888–898
22. Allison C, Auyeung B, Baron-Cohen S (2012) Toward brief “red flags” for autism screening: the short autism spectrum quotient and the short quantitative checklist in 1,000 cases and 3,000 controls. J Am Acad Child Adolesc Psychiatry 51(2):202–212
23. Deng J, Cummins N, Schmitt M, Qian K, Ringeval F, Schuller B (2017) Speech-based diagnosis of autism spectrum condition by generative adversarial network representations. In: Proceedings of the 2017 ınternational conference on digital health, pp 53–57
24. Zhang Z, Deng C, Shen Y, Williamson DS, Sha Y, Zhang Y, Song H, Li X (2020) On loss functions and recurrency training for gan-based speech enhancement systems. arXiv preprint arXiv:2007.14974
25. Goel N, Grover B, Gupta D, Khanna A, Sharma M (2020) Modified grasshopper optimization algorithm for detection of autism spectrum disorder. Phys Commun 41:101115
26. Barbaro J, Yaari M (2020) Study protocol for an evaluation of ASDetect-a Mobile application for the early detection of autism. BMC paediatrics 20(1):1–11
27. Canavan S, Chen M, Chen S, Valdez R, Yaeger M, Lin H, Yin L (2017) Combining gaze and demographic feature descriptors for autism classification. In: 2017 IEEE ınternational conference on ımage processing (ICIP). IEEE, pp 3750–3754
28. Küpper C, Stroth S, Wolff N, Hauck F, Kliewer N, Schad-Hansjosten T, Roepke S (2020) Identifying predictive features of autism spectrum disorders in a clinical sample of adolescents and adults using machine learning. Sci Rep 10(1):1–11.
29. Raj S, Masood S (2020) Analysis and detection of autism spectrum disorder using machine learning techniques. Procedia Comput Sci 167:994–1004

30. Tariq Q, Daniels J, Schwartz JN, Washington P, Kalantarian H, Wall DP (2018) Mobile detection of autism through machine learning on home video: a development and prospective validation study. PLoS Med 15(11):e1002705
31. Abbas H, Garberson F, Glover E, Wall DP (2018) Machine learning approach for early detection of autism by combining questionnaire and home video screening. J Am Med Inform Assoc 25(8):1000–1007

Virtual System for Restoration of Ancient Architecture Heritage Based on Remote Communication Technology

Kaihong Wang

Abstract Virtual system for restoration of ancient architecture heritage based on remote communication technology is studied in this paper. Although the successful realization of parallel processing of remote sensing image enhancement in a grid environment has explored an effective way for the rapid processing of massive remote sensing images, there are still many problems that need to be resolved. Hence, we consider the theoretical system, basic framework and technical framework of quantum remote sensing have been greatly improved and improved on the original basis, and phased results have been achieved in the quantum remote sensing fine spectroscopy experiments and also quantum remote sensing information mechanism research. The model is applied to the restoration of ancient architecture heritage, and the simulation has proven the performance.

Keywords Remote communication · Virtual system · Restoration model · Ancient architecture heritage · Data transmission

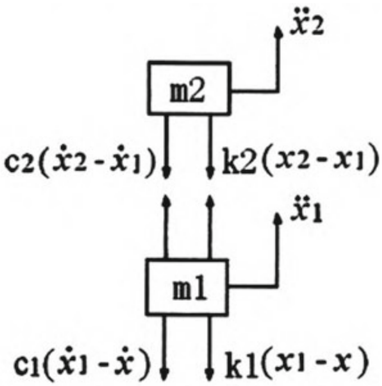
Introduction and Theoretical Basis

Quantum remote sensing is a theory and method that reflects the motion law of remote sensing technology at the quantum level, and it is also a theory that reflects the motion law of remote sensing microscopic particles; it uses multi-space–time and the dynamic earth surface systems as the research object and uses quantum remote sensing information and quantum remote sensing technology. Quantum remote sensing communication, quantum remote sensing computing, quantum remote sensing network and computer information processing are the main technical systems [1, 2]. The amplitude amplification factor from the frame to the optical element is only affected by the dynamic parameters of the optical element assembly itself and has nothing to do with the dynamic parameters of the system composed of the bottom support and the frame [3, 4].

K. Wang (✉)
Guangdong University of Foreign Studies, Guangzhou 510006, Guangdong, China
e-mail: wkh18802054567@163.com

V. Bindhu et al. (eds.), *Proceedings of Third International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 844,
https://doi.org/10.1007/978-981-16-8862-1_67

Fig. 1 General fusion model



Quantum fusion is a method of image display enhancement and information fusion. It has the advantages of flexibility and applicability. It is a fusion method with strong practicability and obvious effect. The fusion method can basically maintain the spectral characteristics of the multi-spectral image before fusion and at the same time incorporate high-resolution geometric information. But while enhancing the geometric resolution, this method will lose a certain amount of the spectral information of the original image [5–7].

In Fig. 1, the fusion model is defined. After selecting a certain fusion method for image processing, it is necessary to then evaluate the fusion effect scientifically and reasonably. At present, there is no unified method and standard for the evaluation of the fusion effect of multi-source remote sensing images [8, 9].

Generally, a combination of subjective qualitative evaluation and objective quantitative evaluation is adopted. With the emergence of parametric modeling and the concept of the model features embedded in the sequential history-based architecture, a large number of the modern commercial CAD applications have adopted parametric, feature and history-based design methods. This traditional design system is changing. The characteristic must accept the performance loss caused by regenerating the entire model from the historical edit point.

The matching fusion algorithm based on the direct positioning is essential for the modelling, due to the inherent error of positioning sensors such as odometers, directly using them for positioning, the accuracy of the results obtained is poor. In order to obtain better positioning results, it is usually combined with Kalman filter to correct the error or combined with laser ranging radar for the multi-sensor information fusion processing. In the next sections, we will be based on the mentioned models to then finalize the proposed methodology.

Literature Review

In our review, the focus will be based on the re-constructions. The generated geometric data of building models such as houses are too simple, the texture is too rough, the generated urban 3D scene is then not realistic, and the user's visual experience is poor. The advantage of the latter is that the 3D model can better reflect the characteristics and details of building structure, realistic texture map, realistic terrain and landscape and good visual effect for users. The disadvantage is that there are many manual interventions, a large amount of the data and low efficiency. The model does not strictly refer to the real geographical information of features, so it is difficult to ensure the accuracy of the model [10, 11].

It can be imagined that the performance impact and design flexibility brought by synchronous modeling technology can be edited without regenerating the entire model, because synchronous modeling technology discovers, locates and resolves dependencies in real time. Designers no longer need to study and reveal complex constraints in order to understand how to edit models, they do not have to worry about the downstream involvement of editing, and they can imagine the positive benefits brought about by the complexity of product development. The rendering of clouds now generally adopts rendering technology based on lighting model, texture mapping technology and also volume rendering technology. The rendering method based on lighting model is to apply various lighting models and use ray tracing technology in computer graphics to generate cloud images. Texture mapping is to generate a two-dimensional texture map of the cloud through the color setting of the data field obtained by the modeling and then use the texture mapping to generate the cloud model.

The Proposed Framework

The Remote Sensing Communication Model

From the beginning of the development of remote sensing to the present stage, the key issue is still further practical application. However, in general, the overall level of remote sensing applications cannot meet practical requirements. It is prominently manifested in two aspects: The real-time detection and processing capabilities are not enough to meet the synchronization requirements such as the natural disasters, and data acquisition and also analysis are completed in quasi-synchronization. The ability and accuracy of quantitative inversion of geological parameters from remote sensing data cannot meet the practical requirements yet. In summary, the features of remote sensing are as follows [12, 13].

(1) There are few restrictions on access to information. In many harsh areas hard to reach by humans, remote sensing technology can be used to obtain various valuable data.
(2) There are many ways to obtain information, and the amount of the information is large. According to the different tasks, different bands and remote sensing instruments can be selected to obtain information.

Since the edge detail information is reflected as coefficients with large absolute value in the high-frequency part after the SIDWT decomposition, for the high-frequency part, the fusion is based on the maximum absolute value criterion. After obtaining the new low-frequency part and high-frequency part, the inverse discrete small transformation with translation invariance is used to finally obtain the fusion result. The algorithm leads to a large number of redundant wavelet representations, which is an over-complete wavelet transform. The source image is decomposed into one layer to obtain an approximate image and $3n$ detail images representing the detail information at different resolutions. The size of the image is the same as the source image defined as Eq. 1 [14, 15].

$$g(x, y) = \sum\nolimits_{\text{general}} f(i, j)/nn \tag{1}$$

Remote sensing electromagnetic waves are greatly limited in the microscopic field, mainly because its description of charged matter only reflects its particle nature. The description of electromagnetic waves only reflects its volatility. In fact, charged particles have volatility, and electromagnetic fields also have particle properties. When charged materials mainly show particle properties and the electromagnetic fields mainly show volatility and the calculation results of remote sensing electromagnetic waves can approximately reflect the objective reality.

In the interior of atoms, the fluctuation of electrons is obvious, and the wave function must be used to describe the motion state of electrons instead of classical orbits. This fusion method plays an important role in image compression, image enhancement, image change detection and multi-temporal image fusion. In Fig. 2, the demonstration is presented.

Virtual System for Restoration and 3D Modeling

The model first selects simple functions to describe the basic structural features of natural textures and then uses turbulence functions to perturb some parameters in the basic functions to produce complex texture details. Turbulence functions can also be used to express the irregularities of various surface color attributes, such as bump texture and transparency. Now, analyze the structural results to take the CSG model as an example. The CGS model makes the stratum form a whole. Using the above two steps, all stratum bodies can be formed, and the adjacent stratum is used as a common surface, and all stratums are integrated to form a complete i-dimension as

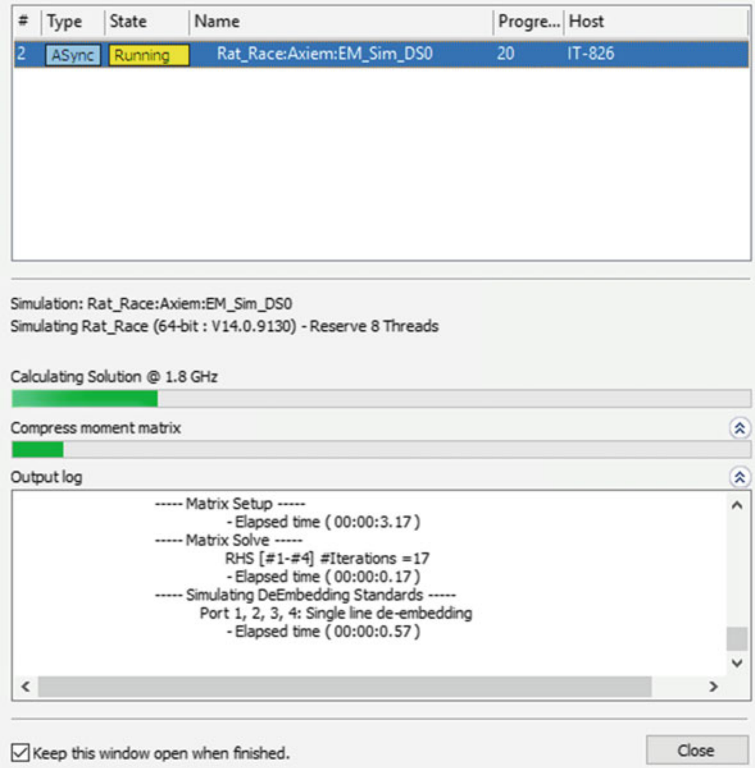


Fig. 2 Remote sensing communication UI demonstration

sectional body. The fusion of the three modeling ideas is used for the structure of the section volume and the division of basic grid volume units and the data organization method, so that the town can then integrate the structural characteristics of the vector and raster models during the construction of the geological section volume with the simple voxel is formed by points, lines, rings, and faces, and also a complex stratigraphic body is formed through a union operation. In Eqs. 2 and 3, the processes are presented [16, 17].

$$px_i = (\pi/2 \sin(f y_{i-1} y)) + \pi \sin(f x_i z/2) \tag{2}$$

$$py_i = (\pi/2 \sin(f y_{i-1} x)) + \pi \sin(f y_i z/2) \tag{3}$$

Physical process-based technology directly starts from the physical factors of the cloud generation, obtains parameters of some different states to simulate its meteorological dynamic change process by solving the physical equation of object movement and can show the detailed process of cloud formation and movement change very truly. Figure 3 shows the framework.

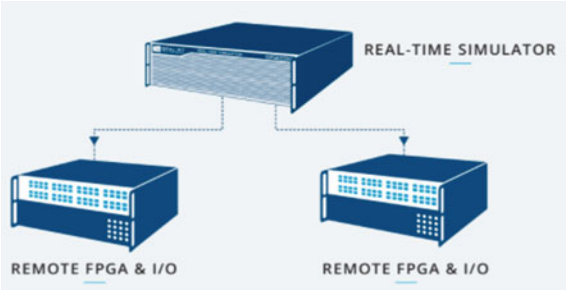


Fig. 3 Framework considered

There is no order in the production of various models in the plot, they can be completed independently, but they must be integrated together completely. After the classification modeling is completed, based on the established terrain and road model, the starting point elevation of buildings, sketches and ancillary facilities is slightly adjusted to make the core spatial relationship of the model in the whole planning plot more accurate and real. The intersection of terrain and road model also needs to be adjusted and edited to ensure the integrity and rationality of the plot model. When the number of the singular values is large, in order to avoid excessive loss of original data information, the method of limiting extra-high neighborhood values can be then adopted; that is, when the absolute value of the difference of a point pair is greater than the given neighborhood limit, the sum of the squares of the difference of the point pair is not included. This method does not eliminate singular values in the original data defined as below [18, 19].

$$Z_{movement}(x) = m(x) + R(x) \tag{4}$$

Detailed phase zoning can better deal with the influence of this factor. If it is not easy to distinguish, the censored distribution can be used to weaken the mutual influence between different populations. In Fig. 4, model is presented.

Restoration of Ancient Architecture Heritage

Before the actual ancient building surveying and mapping process, the relevant information must first be collected and sorted out to find out the commonality between the single ancient building and the ordinary ancient building and the individual personality [20]. Only when we fully understand the personality of the ancient building, we can use the correct ideology to guide our work [21–24].

Many correct consciousnesses are preserved through intangible cultural heritage. If we do not pay attention to these intangible cultural heritage and also use some common things to guide the research of individual units, we will often get results that are inconsistent with the facts and affect our protection of ancient buildings.

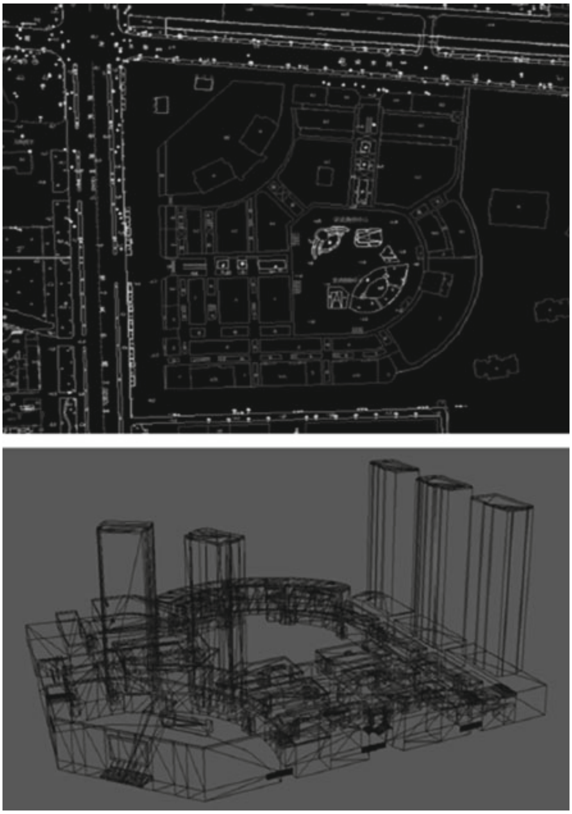


Fig. 4 Model framework presentation

The so-called reconstruction actually covers two phenomena of “restoration” and “reconstruction” in the protection of Chinese architectural heritage. The main difference between restoration and reconstruction lies in the ability to accurately and credibly reproduce the disappeared historical buildings. Restoration refers to the accurate reproduction of the appearance and technical techniques of historical buildings on the original site with traditional materials and techniques based on exact historical pictures and documents. Specific problems are analyzed in detail, and construction plans that reflect the principles of ancient building restoration are adopted to achieve the purpose of properly protecting ancient buildings.

This is the correct guiding ideology for the ancient building restoration. As a Chinese ancient building protection worker, we must carefully study the damage characteristics and repair methods of Chinese ancient buildings, and we must understand the basic difference between stone and wooden buildings. The western stone buildings are rigid, the eastern wooden buildings are flexible, and also, the damage process and rules of ancient Chinese buildings are completely different from those of the western stone buildings. In view of these characteristics of the wooden buildings, effective protection measures and restoration methods have long been summarized in ancient my country, and this is to constantly carry out targeted restorations of different scales.

Experiment

In this section, the experimental analysis is further conducted, and we focus on the communication systems. The image size used in the test is the standard size. Adjust the number of nodes, and the calculation time initially decreases as the number of nodes increases, when the number of nodes increases to 8. The calculation time increases as the number of nodes increases. The window picking mode to correct remote sensing images uses the polynomial model. The polynomial correction avoids the spatial geometric process of the general imaging and directly performs mathematical simulations on the image deformation itself.

The geometric deformation of remote sensing images is caused by many factors, and the law of the change is very complicated. For this reason, the overall deformation of remote sensing images is regarded as the result of translation, scaling, rotation, affine, deflection, bending and higher-level basic deformation. It is difficult to describe with a core strict mathematical expression, but with the appropriate polynomials to then describe the coordinate relationship between the corresponding points of the image before and after correction. In Fig. 5, the systematic framework is presented, and in Fig. 6, the simulation is done.

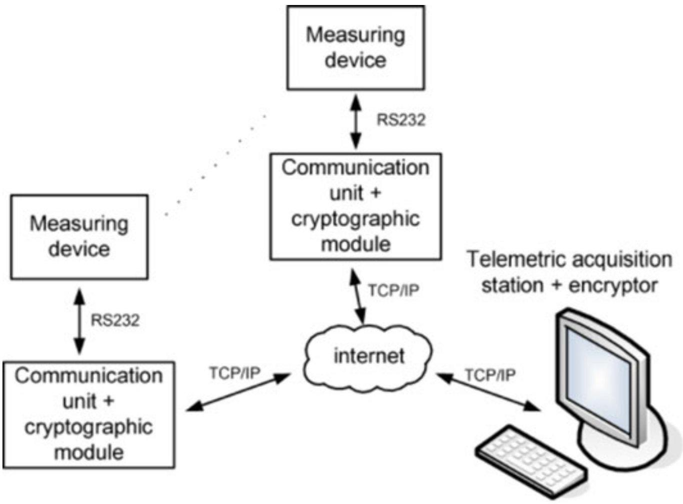


Fig. 5 Experimental structure model

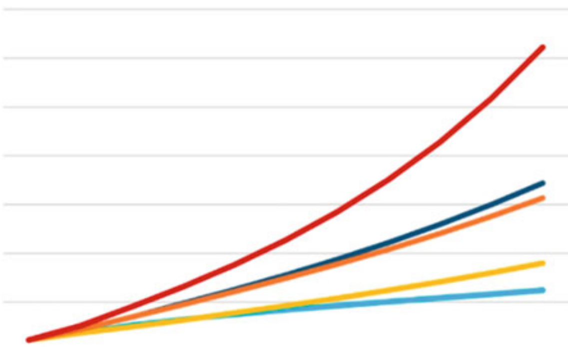


Fig. 6 Experimental curve model

Conclusion

Virtual system for restoration of ancient architecture heritage based on the remote communication technology is studied in this paper. The computing resources of the cluster are scheduled to realize the parallel processing of remote sensing image enhancement. The core experimental results prove that compared with the serial program, this mode makes full use of various computing resources in the grid environment, gives play to the characteristics of grid computing resource sharing, and obtains a higher speedup. In view of the characteristics of the image to be fused and the purpose of fusion, this article mainly deals with the low-frequency part, realizes the complementarity of the two types of images in the low-frequency basic information, and adds the edge detail information of the high-frequency part. The proposed model is simulated under different scenarios. In the future study, we will consider some more aspects of the model to validate the robustness.

Acknowledgements This work is funded by the characteristic innovation project of colleges and universities of Guangdong Provincial Department of Education. It is a phased achievement of "little Taishan, big world—Research on the promotion of architectural culture of Taishan Qiaowei", No. 2019wtscx018. And get Guangdong University of foreign studies to introduce talents research start-up project funding.

References

1. Higo J, Kusaka A, Kasahara K, Kamiya N, Fukuda I, Mori K, Hata Y, Fukunishi Y (2020) GA-guided mD-VcMD: a genetic-algorithm-based method for multi-dimensional virtual-system coupled molecular dynamics. arXiv preprint arXiv:2006.06950
2. Smys S, Basar A, Wang H (2020) Artificial neural network based power management for smart street lighting systems. J Artif Intell 2(01):42–52
3. Herdt V, Große D, Pieper P, Drechsler R (2020) RISC-V based virtual prototype: an extensible and configurable platform for the system-level. J Syst Archit 109:101756
4. Chen JIZ, Yeh L-T (2020) Analysis of the ımpact of mechanical deformation on strawberries harvested from the farm. J ISMAC (3):166–172

5. Covaciu F, Pisla A, Iordan A-E (2021) Development of a virtual reality simulator for an intelligent robotic system used in ankle rehabilitation. Sensors 21(4):1537
6. Poux F, Valembois Q, Mattes C, Kobbelt L, Billen R (2020) Initial user-centered design of a virtual reality heritage system: applications for digital tourism. Remote Sens 12(16):2583
7. Fung A, Irvine M, Ayub A, Ziabakhsh S, Amed S, Hursh BE (2020) Evaluation of telephone and virtual visits for routine pediatric diabetes care during the COVID-19 pandemic. J Clin Trans Endocrinol 22(2020):100238
8. Dewar S, Lee PG, Suh TT, Min L (2020) Uptake of virtual visits in a geriatric primary care clinic during the COVID-19 pandemic. J Am Geriatr Soc
9. Zhao Y, An X, Sun N (2020) Virtual simulation experiment of the design and manufacture of a beer bottle-defect detection system. Virtual Reality Intell Hardware 2(4):354–367
10. Arora PK, Haleem A, Kumar H, KhanSA (2020) Recent development in virtual cellular manufacturing system. In Recent advances in mechanical engineering. Springer, Singapore, pp 1–7
11. Fu Y, Huang G, Liu L, Li J, Zhang X, Zhai M, Pan X (2021) Multi-hierarchy virtual-water management–a case study of Hubei Province, China. J Cleaner Prod 293(2021):126244
12. Nayak S, Patgiri R (2021) 6G communication technology: a vision on intelligent healthcare. In: Health ınformatics: a computational perspective in healthcare. Springer, Singapore, pp 1–18
13. Mensah EA, Blankson IA, Loglo FS, Freeman E (2021) Emergency remote teaching and learning (ERTL) during COVID-19: the Ghana Communication Technology University Experience. In: Handbook of research on ınequities in online education during global crises. IGI Global, pp. 321–339
14. Yu K, Tan L, Lin L, Cheng X, Yi Z, Sato T (2021) Deep-learning-empowered breast cancer auxiliary diagnosis for 5GB remote E-health. IEEE Wirel Commun 28(3):54–61
15. Hafermalz E, Riemer K (2021) Productive and connected while working from home: what client-facing remote workers can learn from telenurses about 'belonging through technology.' Eur J Inf Syst 30(1):89–99
16. Embrett M, Liu RH, Aubrecht K, Koval A, Lai J (2021) Thinking together, working apart: leveraging a community of practice to facilitate productive and meaningful remote collaboration. Int J Health Policy Manag 10(9):528–533
17. Byaruhanga J, Paul CL, Wiggers J, Byrnes E, Mitchell A, Lecathelinais C, Tzelepis F (2020) Connectivity of real-time video counselling versus telephone counselling for smoking cessation in rural and remote areas: an exploratory study. Int J Environ Res Public Health 17(8):2891
18. Liu Q (2021) Application of remote sensing and GIS technology in urban ecological environment investigation. Arab J Geosci 14(17):1–12
19. Kniel J, Comi A (2021) Riding the same wavelength: designers' perceptions of shared understanding in remote teams. SAGE Open 11(3):21582440211040130
20. Desnitsky, Vasily, and Igor Kotenko. "Enhancing technology of producing secure IoT devices on the base of remote attestation." In MATEC Web of Conferences, vol. 329. EDP Sciences, 2020.
21. Amelia R, Istianah TN (2021) Teaching strategies: How do teachers in remote area survive during remote learning? In ELT Forum J Engl Lang Teach 10(2):146–153
22. Dendani B, Bahi H, Sari T (2020) Speech enhancement based on deep AutoEncoder for remote Arabic speech recognition. In: International conference on image and signal processing. Springer, Cham, pp 221–229
23. Scott ME (2020) Identifying barriers to organizational identification among low-status, remote healthcare workers. Commun Stud 71(4):685–698
24. Islam MA, Igwe PA, Rahman M, Saif ANM (2021) Remote working challenges and solutions: ınsights from SMEs in Bangladesh during the COVID-19 pandemic. Int J Qual Innov 5(2):119–140

Integration of University Digital Resources Based on Cloud Computing and Electronics Systems

Wei Zhao

Abstract Integration of university digital resources based on the cloud computing and electronics systems is studied in this manuscript. Coordinated computing objects occupy a lot of system resources and have high requirements for network resources after creation. The activation module can control the creation and release time of coordinated computing objects, and hence, the novel computing method is considered. With the wide application of the Internet, digital resources can integrate information through the network with high-speed and large-capacity computer and network system, hence, the integration model is designed. To verify the results of the proposed model, we use the public database with the integration of different parameters to verify the model. The comparison results have proven that the designed model is efficient.

Keywords Electronics systems · Cloud computing · Digital resources · Integration system · Smart platform

Introduction

Although the distributed computer network structure has more advantages than the conventional computer network structure in practical applications, there is still a lot of room for optimization and improvement to meet the higher demands of the society for network services [1–3]. Including the terminal characteristics of the load, temperature, humidity, wind speed, sunshine and other meteorological data related to the load operating state, aggregated to obtain the operating state of the load group, and transmitted to the dispatch center [4]. According to the current operating status of the power grid, the dispatching center issues load control instructions to the load aggregation layer, and the load aggregation layer parses the control instructions and issues control signals to the loads under its jurisdiction.

W. Zhao (✉)
Kunming Railway Vocational And Technical College, No. 16 Xiaoshiba, Economic and Technological Development Zone, Kunming 650208, China
e-mail: kmtdzyjsxyzw@163.com

V. Bindhu et al. (eds.), *Proceedings of Third International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 844,
https://doi.org/10.1007/978-981-16-8862-1_68

Fig. 1 The cloud data analytic model

Distributed computer interlocking system can configure system substations freely according to user requirements. For example, the entire station yard can be divided into several areas, and the equipment in each area (such as the turnouts, signals, track sections, etc.) are concentrated in one extension for control, and the extensions in each area exchange data with the interlocking host through the bus. The advantage of this is that when the system is maintained and reformed, the corresponding sub-controller can be closed for the reformed area without affecting other areas; in addition, the same type of the equipment (such as switch, signal machine and track section) are concentrated in several sub-controllers for control, and the sub-controllers exchange data with the interlocking host through the bus.

In the network connectivity problem, the connectivity of the network is expressed in the form of probability [5]. If the number of summary points and the number of error nodes of the network are known, the total connectivity is obtained by then analyzing whether the network is connected under each error condition frequency. Due to the large number of the nodes distributed in the distributed computer network structure, when a certain node has a problem, it may cause a part of the structure to be interrupted, which is not conducive to the improvement of the overall stability of the distributed network, Fig. 1 shows the model.

For our study, the digital resource is essential for applications. For traditional paper resources, they can only be provided to designers through the catalog and physical media [6, 7]. At the same time, with the continuous development of most enterprises, many technical data have been shelved for a long time and cannot be well used. With the wide application of the Internet, digital resources can integrate information through the network with high-speed and large-capacity computer and network system, which is convenient for designers to use. Regardless of whether the digital resources are databases, e-books, or e-journals, libraries and users have some of the same requirements for them [8–10].

Therefore, this indicator system will formulate some general indicators to evaluate the above-mentioned resources. For example, the life of the storage body, the obsolescence of the operating system and the application system, etc. In fact, the probability of hardware and the software failure is relatively high, such as the damage of the hard disk, the paralysis of the system, the loss of digital resources or being infected by viruses, and the negligence of digital formats. It even needs to deal with immunity and disasters and the immunization is a preventive measure. Generally, libraries will consider virus protection and hacker intrusion, but they often lack consideration in handling disasters [11, 12].

Once a disaster occurs, it seems helpless. Although disaster management is an emergency measure, it is essential to protect digital resources, and we should attach great importance to it when formulating safety management measures. In our next sections, the solutions will be proposed for the analysis.

The Theoretical Review

In the process of the effect evaluation theory acting on the electronic resource use evaluation work of university libraries, its corresponding value connotation is shown as follows: through the reasonable application of the effect evaluation theory, it is determined that the electronic resource services provided by university library are relative to the customers with the level of value, impact and degree of benefit. Electronic archives management also has certain risks in the process of storage and use. Firstly, when using the traditional archives management mode, there may be a clerical error or a small error in the process of archives processing, which has a small impact [13, 14]. If we use the computer to process a large number of electronic documents, because the computer can process and accept a large amount of data, mistakes in the processing process will then cause unimaginable consequences. In addition, if electronic archives are used to process documents, if the documents cannot be updated, it may cause the loss of document information. In addition, when using electronic archives for management, due to the complex internal structure of the computer, some information may be lost in the general transmission and transmission of electronic archives, making the data integrity of archives incomplete.

Due to the diversity of network databases, the scope and search question styles included in each search system are not the same, so. When a search fails, don't give up easily. We can make new permutations and combinations of search terms, add or delete search terms, add or change search logic, or replace them with the synonyms, and search again. Generally, you will get better search results. The first step in using electronic archives for resource management is to ensure the safety of storage devices, and to clean up, maintain, update and upgrade software and hardware regularly. Ensure that the electronic storage device can operate normally and keep it in a good condition.

During the storage of archives, the archives should be backed up regularly to prevent the loss or destruction of archive data resources. The location of electronic resources is then transmitted from the browser to the server through the resource information represented by these URL parameters. Using the script to run URL parameters to call the requested web page. Because dynamically generated web pages are indispensable when constructing truly large-scale websites, electronic resources in the network often use dynamic URLs to locate resources.

The Designed Pipeline

The Cloud Computing and Electronics Systems

In order to study these scheduling and also fault-tolerant models, system simulations are needed [15]. However, it is difficult to then develop and test a real distributed computing platform on the Internet. It is very difficult to set up the operating scenarios and input data. Therefore, this article proposes and describes a simulator suitable for application evaluation in a distributed computing platform.

The main goal of the simulator is to provide the accurate analysis of the execution process, using a tracking-driven approach to collect task-related data and results in the computing platform. In Fig. 2, we demonstrate the model.

The client uses an active strategy to obtain subtasks to avoid tedious and complicated tasks such as discovering client resources and task scheduling by the server, thereby reducing the load on the server. Whether it is simulation system or control system, reasonable system and data planning is very important [16, 17].

Through system and data planning, the system design can be optimized to solve the problems of task balance and resource utilization of each machine, and give full play to the overall efficiency and performance of the system, while giving consideration to reliability, security and fault diagnosable, and allowing a certain scalability. We consider 2 aspects as follows.

(1) The operating status of a single load is transmitted to the dispatching center, which directly sends control instructions to a single load based on the current operating status of the power grid.
(2) Distributed control does not need to communicate with the dispatch center. The load controller automatically responds to system frequency changes, and the implementation cost is low. However, the reserve capacity that the load

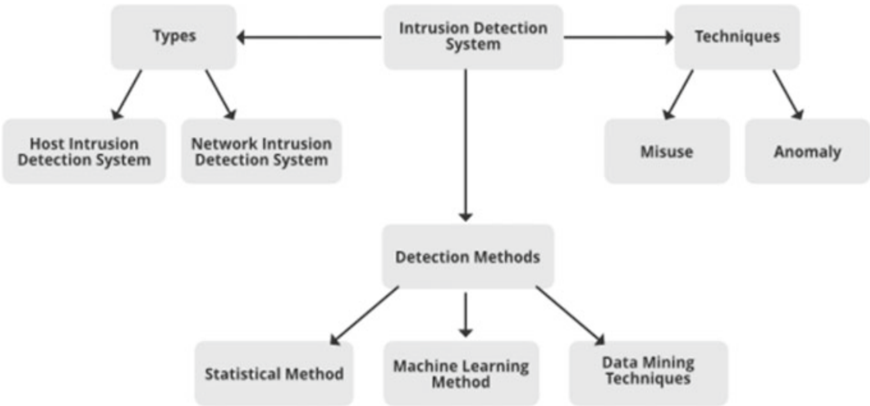


Fig. 2 The model structure details

response frequency can provide cannot be determined, and it is easy to cause over-control or under-control [18].

The designed subsystem includes the two parts: the subsystem management module and the calculation examples; it also creates. The subsystem management module is responsible for receiving various instructions from users, and meeting user needs by interacting with the coordination layer; the calculation examples of the subsystem cooperate with the general coordination calculation module of the coordination layer to jointly complete the calculation of the distributed dynamic power flow. Distributed computing requires that the distributed subsystems can start participating in the calculation at the same time. The function of the startup module is to coordinate the local calculation of the start time of each subsystem. When a subsystem applies to the coordination layer to start distributed computing, the coordination layer uses asynchronous call method to issue commands to all subsystems uniformly to meet the requirements of each subsystem to start local computing at the same time. The lemma 1 defines the node model [19–21].

$$E(A) \leq T \tag{1}$$

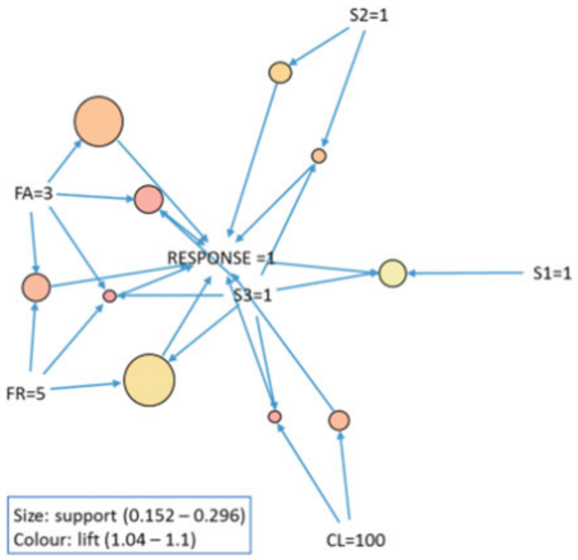
The simulator can support the setting of some parameters, such as the size of the checkpoint file, the usage ratio of the checkpoint, and the selection of the scheduling strategy of the subtasks. Two checkpoint strategies are supported: local checkpoint and centralized checkpoint. The essence of SQL optimization is to use statements that can be identified by the optimizer to minimize type conversion and calculations, make full use of indexes, reduce the number of I/Os of table scans, and try to avoid table searches. In fact, the performance optimization of SQL is a complicated process, and the above content is only some manifestations at the application level. Figure 3 shows the pattern.

Digital Resources Modeling

The construction of digital library resources is to meet the requirements for the rational allocation of resources and systematic dispatch, gradually improve the functions of the shared service platform, and continuously enhance the three-dimensional service functions between users and the digital library, so as to then achieve the goal of providing digital information services to the general public purpose. Digital preservation is suitable for the preservation of the native digital documents, the storage of online documents, and the preservation of analog-digital conversion products. It involves many programs for long-term preservation of electronic documents and for long-term access by people [22, 23].

Compared with paper-based documents, the preservation process of the digital resources is more complicated. It is necessary to organize the information content, but also to then ensure long-term preservation, and can adapt to the continuous

Fig. 3 Cloud computing and electronics system pattern



development of technology to ensure that it can still be retrieved and obtained by users after many years. Because the preservation of digital resources depends on the network, and information technology changes rapidly, hardware and software are aging unpredictably, so the storage media has poor stability, and it may be caused by the inability to access the physical storage media for storing the digital information after a few years and the digital information is lost. In Fig. 4, the steps are presented with the detailed procedures.

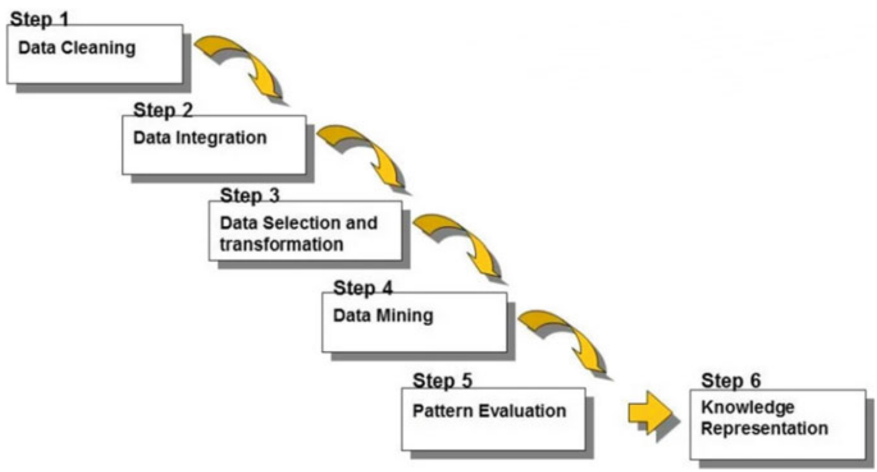


Fig. 4 The detailed steps from the data analysis perspectives

The integration of digital resources is relative to a certain range of users, and its personalized services require on-demand integration of information resources on the basis of public integration. That is to realize integration and personalization of information resources and services according to user needs and characteristics that considers the following steps.

(1) With the continuous development of information security technology, the current identity authentication technology can also be applied as a module to security products, such as the firewalls, three-layer routing switches, databases, servers or other third-party technologies.
(2) Message authentication technology is used to authenticate messages to prevent hackers from tampering information. The use of encryption technology is an active and also open preventive means of network information security. By encrypting the information transmitted on the network, the information service is more secure and reliable, ensuring the integrity and unity of information. Thus, achieve the purpose of ensuring the information security. Digital watermarking technology is adopted to prevent the illegal dissemination of digital resources.
(3) The professional digital resource library is the information collection of all the characteristic resources provided by each library in the digital service, and it is also a sharable document database established by the digital library in the electronic network environment according to the needs of target users.

We have built a large number of constantly updated and constantly emerging knowledge elements into a general knowledge meta-database, so that the various knowledge contents generated in this way are internally networked, so as to reflect the characteristics of cross penetration of various disciplines and specialties. By embedding the knowledge meta-database into various databases, it will closely link the content of the knowledge information within the database and between the databases, and through the knowledge discovery process of the literature, all the databases will be in the knowledge meta-database under the domination of a network, and it becomes an organic whole with highly connected content. In Fig. 5, the parameters are defined.

Fig. 5 Digital resources modeling parameters

CDS–Invenio	Dspace	Eprints
DublicCore MARC21	DublicCore	DublicCore

Integration of University Digital Resources

University libraries should fully consider the needs of the school's guiding and scientific research, and select the type and content of electronic resources according to the school's subject structure and specialty and existing collections. At the same time, it is best to try it out before subscribing to electronic resources, and do a good job of usage statistics and feedback during the trial period, so as to provide references for subscribing to general electronic resources. This requires discrimination and analysis of the URL addresses of electronic resources in the database integration work. Although the information content of the URL is not much and just enter a URL information in the address bar of the IE browser [24].

Internet can link to the target page, but its operation mechanism and stability largely affect the utilization and maintenance of the electronic resources. In the process of evaluating the use of electronic resources in academic libraries, we always adhere to the user-centered approach. The advantages of constructing a user-centered model are: First, it can fully reflect the nature of university libraries and its education function in the entire university education and teaching activities; second, it can realize the evaluation mode of university libraries from the book-based model. The transformation of the human-centered model not only improves the reliability and pertinence of evaluation data, but also encourages university libraries to further improve the application and configuration of the electronic resources based on the evaluation results.

The static URL usually contains the path and file name corresponding to the visited inside the website. In the use of the Internet, sometimes the web pages to be located are hidden in several levels of subdirectories. This kind of electronic resources with static URLs has slightly changed website structure, and the URL is changed accordingly. The simulation will be studied in the next section.

The Experiment

For the systematic simulation of the proposed model, the communication efficiency is the essential factor. The entire simulation process was carried out in the participation of 64 desktop PCs in a scientific research institution, and the external P address was assigned to simulate a large-scale real Internet environment. The running time is about 12 days, and the data of the tracked task is collected every 12 min. The master station system has the advantage of exchanging data with each slave station to complete collection and drive. The startup process is completed after 12 cycles. If the data can be exchanged with the slave station correctly, the connection is considered normal, the sealing of the mapping area is lifted, and other programs are allowed to access the mapping area. In Fig. 6, the estimation result is presented.

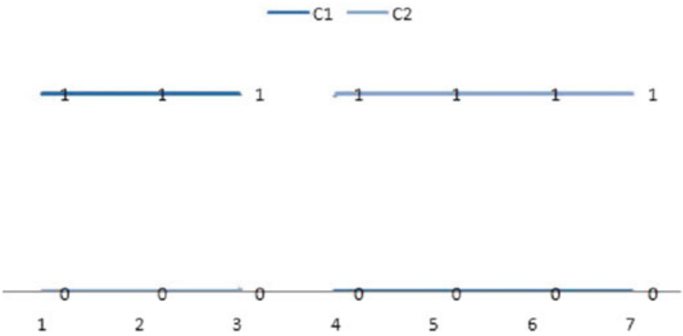


Fig. 6 The efficiency estimation result

Conclusion

Integration of the university digital resources based on the cloud computing and electronics systems is studied in this manuscript. The user machine we designed adopts an active strategy to obtain subtasks to avoid the tedious and complicated tasks of the core server from discovering client resources and performing task scheduling, thereby reducing the load on the server. The static URL usually contains the path and file name corresponding to the visited inside the website, and the model is based on the URL model. In the next stage of the study, we will consider the different application models to verify the robustness.

References

1. Nalajala S, Moukthika B, Kaivalya M, Samyuktha K, Pratap NL (2020) Data security in cloud computing using three-factor authentication. In: International conference on communication, computing and electronics systems. Springer, Singapore, pp 343–354
2. Swarnakar S, Kumar N, Kumar A, Banerjee C (2020) Modified genetic based algorithm for load balancing in cloud computing. In: 2020 IEEE 1st ınternational conference for convergence in engineering (ICCE). IEEE, pp 255–259
3. Harika GL, Chowdary H, Satya Kiranmai T (2020) Cloud-based ınternet of things for smart water consumption monitoring system. In: 2020 5th ınternational conference on communication and electronics systems (ICCES). IEEE, pp 967–972
4. Mehraj S, Tariq Banday M (2020) Establishing a zero trust strategy in cloud computing environment. In: 2020 ınternational conference on computer communication and ınformatics (ICCCI). IEEE, pp 1–6
5. Padmavathi M, Basha SM, Jaya Rama Krishnaiah VV (2020) Load balancing algorithm to reduce make span in cloud computing by enhanced firefly approach. In: 2020 ınternational conference on electronics and sustainable communication systems (ICESC). IEEE, pp 896–900
6. Shukur H, Zeebaree S, Zebari R, Zeebaree D, Ahmed O, Salih A (2020) Cloud computing virtualization of resources allocation for distributed systems. J Appl Sci Technol Trends 1(3):98–105

7. Haji LM, Ahmad OM, Zeebaree SR, Dino HI, Zebari RR, Shukur HM (2020) Impact of cloud computing and internet of things on the future internet. Technol Rep Kansai Univ 62(5):2179–2190
8. Prakash V, Bawa S, Garg L (2021) Multi-Dependency and time based resource scheduling algorithm for scientific applications in cloud computing. Electronics 10(11):1320
9. Alnajrani HM, Norman AA, Ahmed BH (2020) Privacy and data protection in mobile cloud computing: a systematic mapping study. Plos One 15(6):e0234312
10. Goyal S, Bhushan S, Kumar Y, Bhutta MR, Ijaz MF, Son Y (2021) An optimized framework for energy-resource allocation in a cloud environment based on the whale optimization algorithm. Sensors 21(5):1583
11. Kim K-C, Zhao Y (2021) OneNet cloud computing based real-time home security system. J Korea inst Electr Commun Sci 16(1):101–108
12. Singh S, Singh D (2020) Live virtual machine migration techniques in cloud computing. In: Data security in ınternet of things based RFID and WSN systems applications. CRC Press, pp 99–106
13. Raković RM (2020) Application of cloud computing in electric power utility systems: advantages and risks. In: Cyber Security of Industrial Control Systems in the Future Internet Environment. IGI Global, pp 229–247
14. Demigha O, Larguet R (2021) Hardware-based solutions for trusted cloud computing. Comput Secur:102117
15. Pires IM, Garcia NM (2020) Identification of warning situations in road using cloud computing technologies and sensors available in mobile devices: a systematic review. Electronics 9(3):416
16. Cho S, Hwang S, Shin W, Kim N, Peter H (2021) Design of military service framework for enabling migration to military SaaS cloud environment. Electronics 10(5):572
17. Mardia J, Jiao J, Tánczos E, Nowak RD, Weissman T (2020) Concentration inequalities for the empirical distribution of discrete distributions: beyond the method of types. Inform Infer J IMA 9(4):813–850
18. Hussam M (2021) New lightweight hybrid encryption algorithm for cloud computing (LMGHA-128bit) by using new 5-D hyperchaos system. Turk J Comput Math Educ (TURCOMAT) 12(10):2531–2540
19. Khan MK, Shuja J, Jararweh Y, Yu G, Guizani M, Verikoukis C, Ahmad RW (2020) IEEE access special section editorial: mobile edge computing and mobile cloud computing: addressing heterogeneity and energy ıssues of compute and network resources. IEEE Access 8:163769–163774
20. Abdul-Jabbar SS, Aldujaili A, Mohammed SG, Saeed HS (2020) Integrity and security in cloud computing environment: a review. J Southwest Jiaotong Univ 55(1)
21. Malar BA, Prabhu J (2020) Data storage security in the cloud environment using the honey pot system. Int J e-Collab (IJeC) 16(4):1–14
22. Chitturi AK, Swarnalatha P (2020) Exploration of various cloud security challenges and threats. In: Soft computing for problem solving. Springer, Singapore, pp 891–899
23. Hassija V, Chamola V, Saxena V, Chanana V, Parashari P, Mumtaz S, Guizani M (2020) Present landscape of quantum computing. IET Quantum Commu 1(2):42–48
24. Kim S-C, Choi T, Song S, Strinati EC, Chung J-M (2020) Special issue on 5G & B5G enabling edge computing, big data and deep learning technologies. ETRI J 42(5):639–642

Intelligent Auxiliary System for Sports Public Service Information Platform Based on Communication and Computing Methods

Jinyuan Zhu and Jijun Chu

Abstract In the recent time, the combination of the information system with computational models is the essential trend, hence, this paper studies the intelligent auxiliary system for the sports public service information platform based on communication and computing methods. Computer communication network security technology plays a very important role in the development and application of computer technology. The development of various new technologies also makes network applications more convenient and safer. With this integration, this paper adopts the novel information system to finalize the framework. The application scenario of the sports public service information platform is integrated. The test on the performance has shown the satisfactory performance.

Keywords Computing methods · Communication system · Intelligent auxiliary model · Sports public service · Information platform

Introduction and Background

With the rapid development of the Internet, more and more information is presented to users, but at the same time, the problem is that it is more and more difficult for the users to obtain the information they need most [1, 2]. Intelligent measurement technology is the application of intelligent control technology in measurement, which aims to solve the intelligent control problems involved in measurement. For the measurement of the three-dimensional medium and large workpieces, the main solution is to improve the measurement efficiency without reducing the measurement accuracy [3, 4]. The method is to increase the measurement step length and reduce the number of some measurement points. The optimization parameter can be then selected as the radius

J. Zhu
Guangdong Industry Polytechnic, Guangzhou 510300, China
e-mail: ariel_zhujy@163.com

J. Chu (✉)
Guangzhou Sports University, Guangzhou 510500, China
e-mail: ariel_zhujy@163.com

V. Bindhu et al. (eds.), *Proceedings of Third International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 844,
https://doi.org/10.1007/978-981-16-8862-1_69

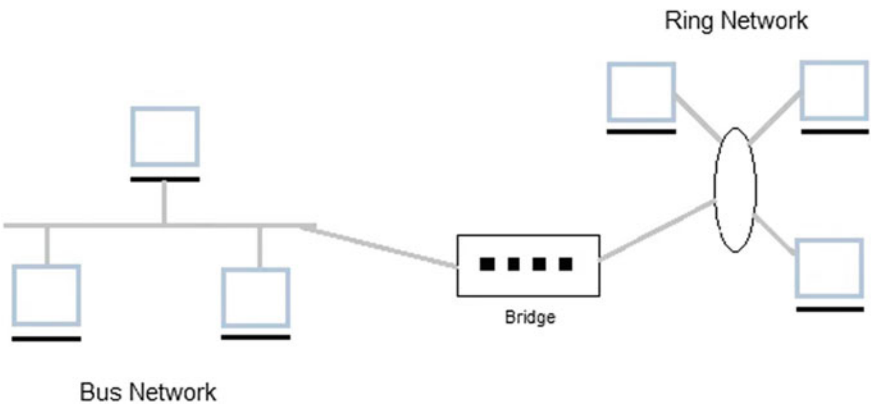


Fig. 1 The service information platform structure

of curvature of the workpiece. The optimization method is to automatically adjust the measurement step size in real time according to the change of the radius of curvature, so as to achieve "slowly sparsely dense, large sparse and small dense" [5].

Through the database interface and data exchange, the system obtains the part feature information, process information and resource information required by the process design that realizes the reconstruction and reuse of the software, and can support a variety of process designs and dynamic process design process control. In Fig. 1, we denote the service information platform structure.

In this system, a task is a download request submitted by a user. The system provides a powerful task management function in addition to task management, the system also provides user management function, we will system users into the administrators, advanced users and general users three levels. In our designed system, the ideas can be considered from listed aspects [6, 7].

(1) User module, the main function for the user management. Including user login verification, user add, enable, disable, user information view, user password modification, etc.
(2) Command management module, the main function is then command management. There are two types of commands, one is a single command, and the other is a combination command. This module includes the management of all the commands, including adding, viewing, and removing combinations, etc.

Literature Review

The communication receiving system based on cluster computing adopts the radio frequency conversion gating unit and the intermediate frequency conversion gating unit to then realize the isolation and dynamic mapping between the signal resource and the processing platform, which fully reflects the flexibility of the software radio receiving system [8, 9]. In addition, it also created triple play technology, realized the effective integration of data information, voice and video, made the broadband network fully carry rich traffic, and realized the integration of three networks. It effectively solves the problems of data delay, picture jitter and system redundancy, and improves the bearing level and access energy efficiency of the network system [10]. With the comprehensive application of the multimedia network technology, the market demand for it continues to rise, which will further promote the integration of network technology and communication industry, and meet the modern needs of users for distance learning, the medical treatment and rich e-commerce services [11]. In addition, in the new era, broadband network technology has also achieved leapfrog development.

The filtering algorithm, demodulation algorithm and other parameters of the data demodulation module can be then dynamically changed and loaded to meet different demodulation quality requirements [12, 13]. The core channel decoding module receives and caches the demodulation operation results of the node server, and according to the time stamp and channel information carried by the data block, the result data of each channel submitted by each node server is sorted and spliced in chronological order to form a continuous data stream post-decoding [14, 15]. In this way, the balance of computing tasks among heterogeneous nodes can be realized, which not only ensures that all computing resources of the system are fully utilized, but also can ensure that the system has a higher throughput rate and a smaller delay [16].

The Proposed Methodology

The Communication and Computing Methods

Modern communication system is a very complicated engineering system, and the design and research of communication system are also a very complicated task. Due to the complexity of technology, in modern communication technology, more and more attention is paid to the use of computer simulation technology for system analysis and design. The essential working feature of data confidentiality technology is that only under a designated user or network can the password be unlocked and the original data can be obtained. This requires the data sender and receiver to use some special information for encryption and decryption. This is the so-called data encryption technology that uses keys for security protection [17].

Therefore, the total baseband transmission signal is defined as below.

$$s[n] = \sum\nolimits_{k=1}^{K} A_k b_k s_k[n] \tag{1}$$

The algorithm has the worst load balancing effect. This is because the task is assigned to the node with the latest start time, which causes some nodes to be heavily loaded, while some nodes may be then idle, resulting in uneven load and negatively affecting global performance influence [18, 19].

At the same time, communication network operators pay more attention to FTTH and GEPON and other means and technologies, both of which have good development prospects. Of course, based on people's comprehensive needs for wireless network applications, future wireless network access technologies will become the focus of research. In this regard, efforts should be made to improve the high-quality mobility of broadband network technologies, thereby conforming to the characteristics of the times and meeting the comprehensive needs of the public. In the lemma 2, the calculation process is defined [20, 21].

$$y(m) = r[m] \times s_k[m] \tag{2}$$

The goal of network isolation technology is to ensure the isolation of harmful attacks, and to complete the secure exchange of data between networks on the premise that information outside the trusted network and the internal information of the trusted network is not leaked, shown in Fig. 2.

In the actual environment, the wireless channel is easy to produce multipath effect and bring multipath interference to the receiver. Moreover, the actual system needs correction coding to reduce the bit error rate. These two factors are considered in the following system simulation. In a cluster-based communication receiving system, multiple Gigabit Ethernet ports with PCI Express interfaces are used for high-speed parallel distribution of data, which can ensure that the data channel from the system memory to each network port is exclusive, and there will be no network. The situation of competition between ports for the use of the bus. At the same time, the 2-level dynamic load balancing technology is adopted to ensure that the data distribution is relatively balanced among the multiple network ports of the collection server and between the multiple node servers managed by each network port, which can then also greatly improve data distribution and also the processing efficiency. After the node server completes the current computing task and sends the result data to the result collection server, it can then exit the clustered computing system. Through the

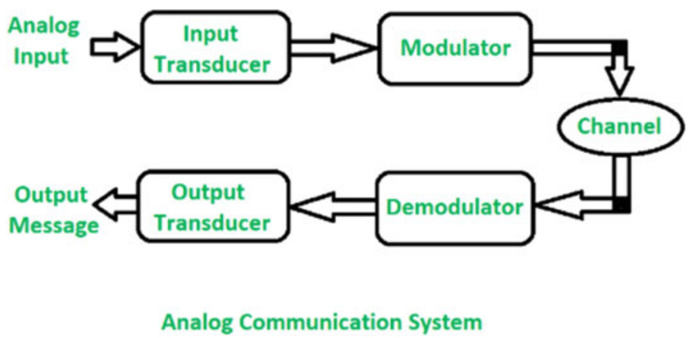


Fig. 2 The communication channel model

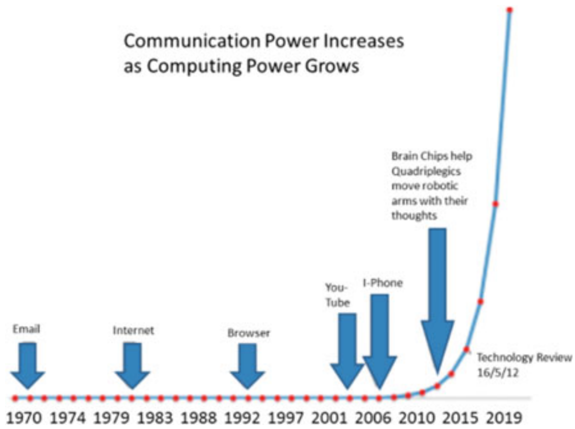


Fig. 3 Communication and computing method developmental trend

mechanism of the joining and quitting heterogeneous computing nodes, the communication receiving and processing platform based on cluster has the capacity of scale and processing and arbitrary reduction and also expansion. Accordingly, in Fig. 3, we then consider the developmental trend of the technology.

Service Information Platform

The task management module also provides a mechanism for sending messages to the users. When a user submits a new task, if the administrator performs an operation on the task, such as confirming or rejecting it, the system will send a message to the submitter of the task to notify the user. Similarly, when a task is stopped or completed, the system will also send a message to the submitter of the task. In the final analysis, the realization of the various functions of the system is completed by various operations on the database. Therefore, to realize the function of system planning and improve the performance of the system, it is necessary to choose a database management software with reasonable design and complete functions, and design a database that is strictly prohibited by the organization and has no redundant data and is easy to maintain.

For the platform efficiency modeling, the process is defined as below.

$$M = (y_1 - y_0)(z_2 - z_0) - (z_1 - z_0)(y_2 - y_0) \tag{3}$$

$$N = (z_1 - z_0)(x_2 - x_0) - (x_1 - x_0)(z_2 - z_0) \tag{4}$$

We propose a model of system filtering recommendation system based on item rating prediction, which predicts unrated items by then calculating the similarity between users, thus increasing the number of items that users have scored together, thereby effectively alleviating the sparsity problem. The bottom layer of this system will also operate the equipment in the form of commands.

If the configuration file is modified, it will cause compatibility problems, and the configuration files of different brands and models of devices will be very different. In addition to the feature of online classification, another feature of this system is that users can then customize the hierarchical classification system by themselves, that is, by specifying the class samples or class keywords and their weights, customizing the classification system, before classification must first set a classification system for the system, and the system can classify the downloaded information according to the classification system as follows (Fig. 4).

Control requires a program algorithm to realize the control algorithm is also one of the core technical research contents of the intelligent prestressing system. It needs to transform the displacement or load change sensed by the intelligent prestressing system into the amount of control information [22–24].

Research can be nonlinear information for multivariable input and outputand the intelligent control algorithm for some correct processing has the characteristics of the material creep and nonlinear load randomness and diversity, and the control algorithm needs to be calculated in real time. The system has listed components.

(1) The data processing Agent is responsible for processing each user's purchase information, updating the user's shopping record, and then generating a user preference information model from the shopping record, calculating the user's similarity based on the model, and preparing data for the product recommendation Agent.

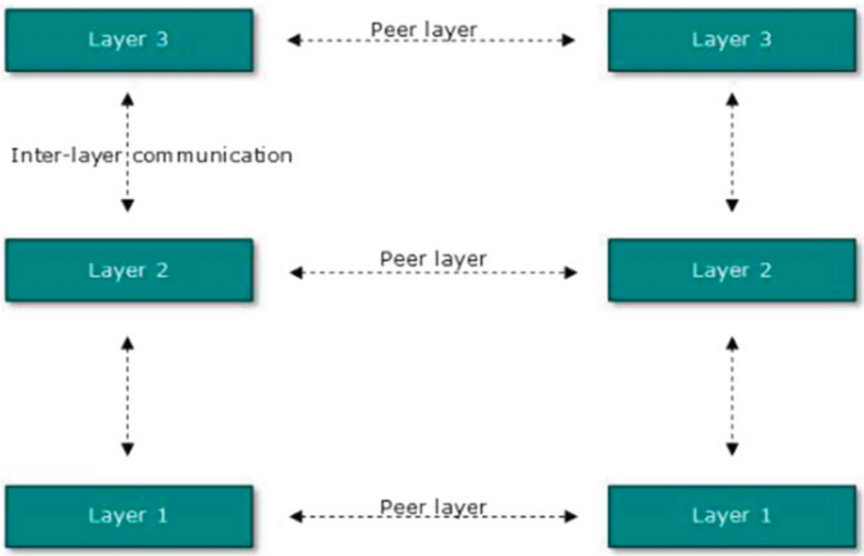


Fig. 4 The information platform layers

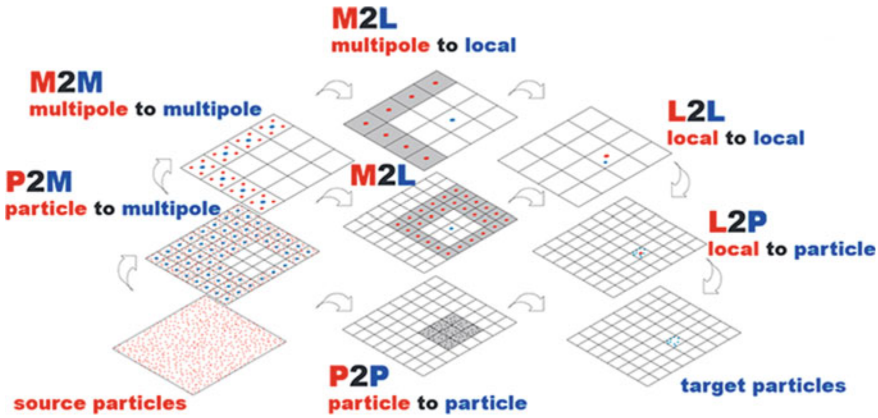


Fig. 5 Service ınformation patterns

(2) The user Agent is automatically generated when the user logs in to the system. It can read the user's personal information and the results generated by the product recommendation Agent, generate user records, record the user's webpage links and visits, forward the search requests submitted by the user to the product search Agent, and return the search results.

(3) The expression of precision features is based on the shape features and geometric elements as carriers. Management features, material features, and technical features describe the overall information of parts, and they are combined into overall information. Through the above-mentioned hierarchical structure of part information, a part information model based on part class and feature class can be constructed.

The system also provides a function that allows users to customize the search results, such as selecting the sorting method of the search results. We provide four sorting methods, which are relevance, time, and attention. The default is to sort by relevance. The system can also allow users to choose the maximum number of search results they want to return and the number of results displayed on one page. By default, all results are returned, and 20 results are displayed on each page. Figure 5 shows the patterns.

Intelligent Auxiliary System for Sports Public Service

The rich expressiveness of the digital learning resources can attract students' interest in learning, and its shareability and interactivity are also conducive to the development of classroom learning activities. Through mobile smart terminals, students in the classroom can receive learning resources distributed by teachers, learners and teachers, learners and learners can communicate and share with each other, and communicate at any time. With the development of general global informatization and networking, the scope of intelligent management is expanding, and the development of foreign intelligent buildings shows obvious trends in two aspects. First, smart buildings have expanded from office buildings to hotels, hospitals, stadiums, residences, factories and the other fields; second, with the expansion and increase in the number of smart buildings, they are developing toward smart cities. In the long run, it not only meets the needs of people-oriented and solves the problems of the sustainable development of gymnasiums and cities, but also enriches, improves, renews and expands the cultural living environment. Combining the two concepts of green and intelligence, that is, insisting on the concept of green intelligence, can truly achieve the general goal of sustainable development of the stadium.

There are various media forms and types of mobile learning resources. Different media have the different characteristics and functions, and their roles in presenting information are also different. All kinds of information related to stadiums are acquired, stored, managed and reproduced in digital form, through comprehensive analysis and effective use of the stadium information provide decision-making support for improving venue management, improving service levels, enhancing competitiveness, and the healthy and sustainable development of venues.

The Experiment

The performance of the proposed model will be evaluated in this section. In a working environment with diverse tasks, each data block that enters the system for processing must add data classification features such as time stamp, block size information, A/D sampling information, and the channel information. At the same time, according to the type of service carried by each signal, the priority of data block processing is determined, giving higher priority to voice and video services, and lower priority to data services. In Fig. 6, the performance is presented.

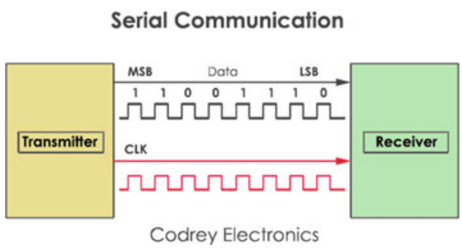


Fig. 6 The performance evaluation model result

Conclusion

In the recent time, the combination of the information system with computational models is the essential trend, hence, this paper studies the intelligent auxiliary system for sports public service information platform based on communication and computing methods. Today's communication objects are mainly aimed at terminals rather than individuals. For example, in fax services, the object of data transmission is a fax machine, not a certain person. If there is no fax machine at the time of transmission, then the data may not be received. Hence, we consider the proper combinations of the different models. The performance is simulated under different scenarios. In our future study, the comprehensive performance will be tested and evaluated.

References

1. Kuang L, Gong T, OuYang S, Gao H, Deng S (2020) Offloading decision methods for multiple users with structured tasks in edge computing for smart cities. Futur Gener Comput Syst 105:717–729
2. Han Y, Jiao J, Weissman T (2020) Minimax estimation of divergences between discrete distributions. IEEE J Sel Areas Inform Theory 1(3):814–823
3. Wan S, Zonghua G, Ni Q (2020) Cognitive computing and wireless communications on the edge for healthcare service robots. Comput Commun 149:99–106
4. Liu Y, Yuan X, Xiong Z, Kang J, Wang X, Niyato D (2020) Federated learning for 6G communications: Challenges, methods, and future directions. China Commun 17(9):105–118
5. Peixoto MLM, Maia AHO, Mota E, Rangel E, Costa DG, Turgut D, Villas LA (2021) A traffic data clustering framework based on fog computing for VANETs. Veh Commun 31(2021):100370
6. Chen S, Li Q, Zhou M, Abusorrah A (2021) Recent advances in collaborative scheduling of computing tasks in an edge computing paradigm. Sensors 21(3):779
7. Liu Y, Zhang W, Zhang Q, Norouzi M (2021) An optimized human resource management model for cloud-edge computing in the internet of things. Cluster Comput 1–13
8. Meng Q, Zheng H, Guo H (2020) Classification of abnormal traces in a collaborative fog computing architecture. In: International conference on security, privacy and anonymity in computation, communication and storage. Springer, Cham, pp 420–429
9. Yang M, Ma H, Wei S, Zeng Y, Chen Y, Yuemei H (2020) A multi-objective task scheduling method for fog computing in cyber-physical-social services. IEEE Access 8:65085–65095
10. Besta M, Kanakagiri R, Mustafa H, Karasikov M, Rätsch G, Hoefler T, Solomonik E (2020) Communication-efficient jaccard similarity for high-performance distributed genome comparisons. In: 2020 IEEE ınternational parallel and distributed processing symposium (IPDPS). IEEE, pp 1122–1132
11. Akdeniz BC, Egan M (2020) A molecular communication scheme to estimate the state of biochemical processes on a lab-on-a-chip. In: Proceedings of the 1st ACM ınternational workshop on nanoscale computing, communication, and applications, pp 8–13
12. Islam MSU, Kumar A, Hu Y-C (2021) Context-aware scheduling in Fog computing: a survey, taxonomy, challenges and future directions. J Network Comput Appl 103008
13. Vimal S, Suresh A, Subbulakshmi P, Pradeepa S, Kaliappan M (2020) Edge computing-based intrusion detection system for smart cities development using IoT in urban areas. In: Internet of things in smart technologies for sustainable urban development, pp 219–237

14. Wagh DP, Fadewar HS, Shinde GN (2020) Biometric finger vein recognition methods for authentication. In: Computing in engineering and technology. Springer, Singapore, pp 45–53
15. Ageed Z, Mahmood MR, Sadeeq M, Abdulrazzaq MB, Dino H (2020) Cloud computing resources impacts on heavy-load parallel processing approaches. IOSR J Comput Eng (IOSR-JCE) 22(3):30–41
16. Raj JS (2020) Machine learning based resourceful clustering with load optimization for wireless sensor networks. J Ubiquit Comput Commun Technol (UCCT) 2(01):29–38
17. Armingol E, Officer A, Harismendy O, Lewis NE (2021) Deciphering cell–cell interactions and communication from gene expression. Nat Rev Genet 22(2):71–88
18. Banjarey K, Sahu SP, Dewangan DK (2021) A survey on human activity recognition using sensors and deep learning methods. In: 2021 5th international conference on computing methodologies and communication (ICCMC). IEEE, pp 1610–1617
19. Ullo SL, Sinha GR (2020) Advances in smart environment monitoring systems using IoT and sensors. Sensors 20(11):3113
20. Salimitari M, Chatterjee M, Fallah YP (2020) A survey on consensus methods in blockchain for resource-constrained IoT networks. Internet Things 11 (2020):100212
21. Ulidowski I, Lanese I, Schultz UP, Ferreira C (2020) Reversible computation: extending horizons of computing: selected results of the COST action IC1405. Springer Nature
22. Kashani MH, Ahmadzadeh A, Mahdipour E (2020) Load balancing mechanisms in fog computing: a systematic review. arXiv preprint arXiv:2011.14706
23. Kaur A, Singh P, Nayyar A (2020) Fog computing: building a road to IoT with fog analytics. In: fog data analytics for IoT applications. Springer, Singapore, pp 59–78
24. Al Hayani B, Ilhan H (2020) Image transmission over decode and forward based cooperative wireless multimedia sensor networks for Rayleigh fading channels in medical internet of things (MIoT) for remote health-care and health communication monitoring. J Med Imaging Health Inform 10(1):160–168

Key Technologies of English Document Grammar Recognition System with Cloud Computing and Electronics Systems

Wei Guo and Cong Wang

Abstract This paper studies the key technologies of English document grammar recognition system with the cloud computing and electronics systems. The security of the platform relative to the user refers to the user's information security guarantee on the cloud computing platform; hence, at the initial design of the proposed model, the electronics system is combined for the overall performance enhancement. Characters have a special line structure, which can basically be classified as the horizontal, vertical, and oblique line combination. Hence, we consider the text mining model and the neural networks model to finalize the framework. The experimental results compared with the latest ones have shown the robustness.

Keywords Electronics systems · Cloud computing · Key technologies · English document · Text analysis · Recognition system

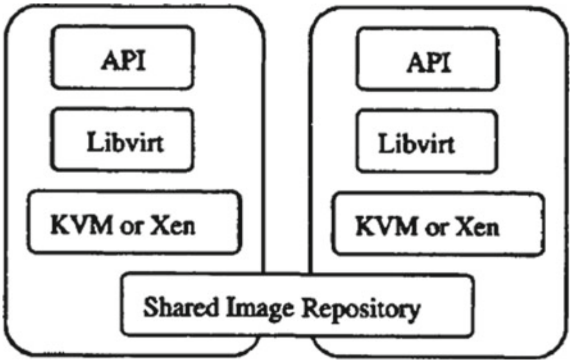
Introduction

The latest developments in the open source virtualization have led to the development of many open source IaaS vendors [1, 2]. Consumers can obtain services from a comprehensive computer infrastructure through Internet. This type of the service is called infrastructure as a service. Internet-based services are part of IaaS. The node controller is the final computing node, which starts and shuts down the virtual machine by calling the virtualization technology of the operating system layer. All the computing nodes in the same cluster must be in the same subnet. A storage server usually needs to be deployed in a cluster to provide data storage services for the computing nodes in the cluster [3, 4]. The security of the platform relative to the user refers to the user's information security guarantee on the cloud computing platform. One is the security of information between users and users who have the right to access the information they own, but they do not have the right to access the information of other users; the second is the security of users in the process of

W. Guo · C. Wang (✉)
School of English Language & Culture, Xi'an Fanyi University, Xi'an 710105, Shaanxi, China
e-mail: wangcong@xafy.edu.cn

V. Bindhu et al. (eds.), *Proceedings of Third International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 844,
https://doi.org/10.1007/978-981-16-8862-1_70

Fig. 1 Cloud computing and electronics system pipeline



accessing information. Information is required to be able to be stored and transmitted to ensure the confidentiality and integrity of data; the third is the security of user information on the platform, which is the most dangerous [5–7].

In Fig. 1, the pipeline is presented. Studying the load balancing service mode of virtual resources, and solve the problem that the server configuration does not need to be configured according to the peak period of the application, but only needs to be configured according to usual configuration to solve the problem that the system automatically adds machines for load according to the strategy during the peak period of the application, and can also temporarily use the services of a third-party cloud provider to start the machine for load to save costs and improve the horizontal scalability of the application. To this idea, the model will be applied to the recognition applications [8, 9].

As far as the named entity recognition method is concerned, on the one hand, the rule-based method is highly subjective and has poor portability, and it is difficult for rules to cover all language phenomena; on the other hand, the use of natural language is not purely a random process, while the use of statistical-based methods alone will make the state search space very large.

Therefore, early pruning with knowledge of grammatical rules is a more effective method. The current method with higher recognition accuracy is based on the combination of statistics and rules. Code knowledge points refer to the program design knowledge points included and implied in the code written by the program design. As a whole, it also refers to all the knowledge points in all programming languages, data structures, discrete mathematics, algorithms, and some other programming fields involved.

This paper proposes an automatic recognition method for knowledge points based on the syntax tree. First, the code is lexically analyzed, and the code is converted into a syntax tree. Then, the feature tree is used for feature matching and recognition, and the program design language, knowledge points in the field of data structure. In the next parts, the model will be discussed.

Review Section

Conjunction grammar retains some important characteristics of context-free grammar, but it has stronger generation ability than context-free grammar and can better describe the grammatical structure of the language [10].

Nowadays, the development of the computers tends to be parallel, and the parallel processing of the multiprocessors is an irreversible era. Therefore, the analysis and research on the processing process of random connection grammar in a parallel environment has theoretical and practical significance [11, 12]. In the rule-based method, named entity recognition uses not only the composition rules of various named entities, but also the entity itself, context, and word usage. Because the rule-based method integrates the prior knowledge of human experts, the recognition accuracy is high, but the rule-based method is highly subjective, it is difficult to cover all language phenomena and has poor portability, and the requirements for corpus are relatively strict; on the other hand, the state search space of statistical method alone is huge, and early pruning with rules is a more effective method [13, 14].

Therefore, at present, there is almost no named entity recognition system using the statistical model alone. In many cases, hybrid methods are used, that is, statistical model combined with syntax rule knowledge for recognition [15]. The dot matrix input network extracts features according to the connection relationship between the character dot matrix and the nodes of the network input layer.

Different connections have different features extracted, and global connection network only extracts the characteristics of the overall change of the character, while the local connection network also extracts the characteristics of each local change of the character. Although byte streams provide sufficient functions to handle any type of the input/output operations, they cannot directly manipulate Unicode characters. Therefore, since one of the purposes of Java is to support "write once, run everywhere," direct character input/output support is necessary. In the next sections, the details will be discussed.

The Proposed Model

The Intelligent Cloud Computing and Electronics Systems

Using the virtualization technology of cloud computing can build diversified virtual resources, virtualize and also instantiate personal desktop systems, server systems, and other operating systems and devices, and provide different users with flexible computing system architectures and devices to meet user needs. In the system design link under the cloud computing platform, there will be certain security vulnerabilities, and it is easy to form a certain backdoor program.

When these backdoor programs or security vulnerabilities are exploited, the program will show unreliability in the execution link. For the system, we consider the following aspects should be referred [16, 17].

(1) Users will get a special-purpose client that allows users to remotely access via the Internet and charge fees based on usage.
(2) Users on the network access data in a variety of the formats and sources through services in a transparent and rational way [18].

How to build a heterogeneous virtualization platform on the basic platform based on automated intelligent scheduling and localized elastic resource allocation is also a major problem to be solved in the future infrastructure cloud.

Virtualization uses logical isolation to provide the feeling of the physical independence and isolate each group of the virtual machines in their respective network segments, that is, different VLANs, so as to minimize the risk of data leakage from one virtual machine partition to another through the network [19, 20]. For the information security of users in the cloud platform, enterprises can divide information and users into several levels from low to high according to the degree of information security and users' different information needs, and strictly control users' access to information. In Fig. 2, the structure is presented.

During the installation process, we need to set the cloud controller's host name, virtual machine intranet mode, network segment, subnet mask and DNS, and cluster name for cluster management.

The cloud controller is the communication interface between the client and the cloud computing platform and implements scheduling management for the working nodes of the entire platform. Its components roughly include the cloud interface, platform component manager, scheduler, monitor, user manager, storage manager, and network manager. The storage system is used to store the image files used in the

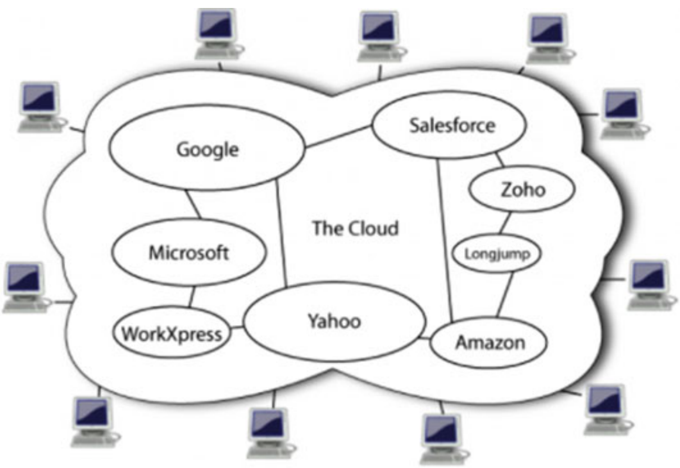


Fig. 2 Presented structure model

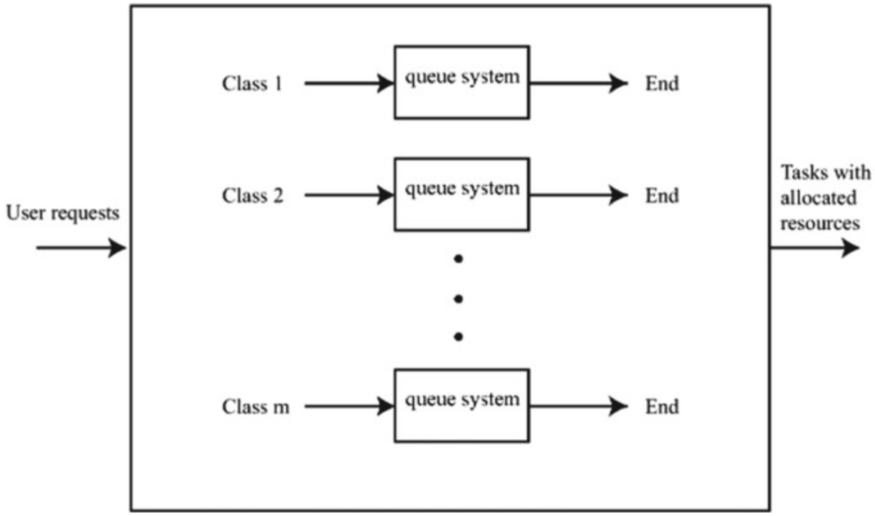


Fig. 3 Intelligent cloud computing and electronics system structure

platform. Clients (users and cloud computing platform administrators) can access the cloud computing platform through command lines and browser interfaces. The cloud access then converts the commands from the client into a mode that is uniformly recognized by the entire platform. Agent cloud is divided into two categories, one is the front-end cloud, and the other is the back-end cloud. In the cloud, each operation of the agent is regarded as a service. If a service's response time is not long, this article will design it as a front-end cloud, otherwise it will be a back-end cloud. In the back-end cloud, technologies such as multicast, workflow, and job scheduling are used.

This makes the originally complicated technology completely transparent to developers, reduces the difficulty of work, and improves development efficiency. The system can realize the fine-grained marking and tracking of program data object level. Even the data of different tenants in the process address space can be effectively tracked and isolated.

This part of the work is reflected in the information flow control at the programming language level. In terms of programming model, developers only need to add mark-based strategies to existing applications with a fixed program structure. Figure 3 presents the structure.

The English Character Text Recognition

At present, there are many methods for automatically locating characters in images. They are basically divided into two categories: connected area analysis method and the texture analysis method unique geometric constraint positioning character area. In the structural method, the text is regarded as a collection of the primitives and positional relations, and simple strokes are extracted as primitives, and then, the strokes are analyzed by the strokes, and finally, the recognition results are given. But no matter which method is used, a complex clustering algorithm can be used to achieve the goal.

With the deepening of research and the development of the signal processing technology, artificial neural networks, wavelet analysis, and stochastic models have penetrated into the field of handwriting recognition and injected new vitality into its development. Characters have a special line structure, which can basically be classified as the horizontal, vertical, and oblique line combination. Statistical characteristics, such as the Chinese characters and English characters are meaningful for the analysis. The line structure of characters in the image has special texture characteristics, and the gray color is different from the background.

The edge changes more drastically show the obvious vertical, oblique, and horizontal edge features, strong high-frequency information. In the lemma 1, the calculation model is defined.

$$y_i' = \frac{y_{i+1} - y_i}{x_{i+1} - x_i} \tag{1}$$

The above whole process is called the classification process. The purpose is to use the information of the horizontal, vertical, and diagonal sub-bands to classify each point of the original image at this scale as the character point or a non-character point according to the unique texture characteristics of the character. In Fig. 4, the node framework is defined.

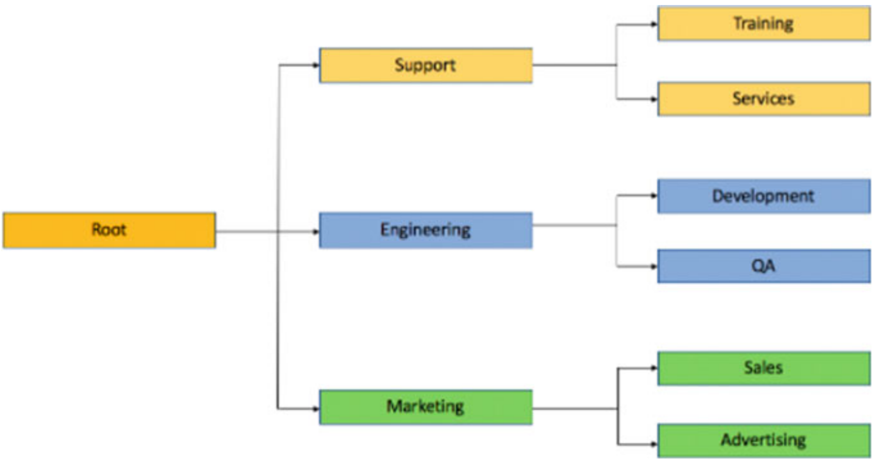


Fig. 4 Node framework

Learning samples are crucial to the success or failure of the neural network applications. Poor learning samples will not only lead to incorrect mapping of the network, but also may make the learning process of the network not converge; hence, we consider listed steps.

(1) The sample should be extensive, and it is best to provide the input of the BP network in various situations and give the corresponding expected output. The extensiveness of the samples can make the trained BP network which have better adaptability [21].
(2) If the learning sample contains a large number of redundant components, the convergence of network learning process will be difficult or non-convergence; the trained network will produce error mapping, making the network output too much biased to the output direction formed by redundant learning components.

Although this paper provides a good algorithm for the handwritten English character recognition, it needs to then reduce the computational complexity of the algorithm in practical application. For the re-estimation of parameters, we should further find more effective methods [22]. The calculation is defined as below.

$$P_{bl} = \left\{ p_i | b_i \in B_{mp} \right\} \tag{2}$$

Because the edges of characters change drastically and show obvious vertical and diagonal edge features, the first step in classification is to extract edge feature information. The judging function is defined as below.

$$Judging_k(i, j) = \begin{cases} 1, & E_k(i, j) > \tan \\ 0, & others \end{cases} \tag{3}$$

The thick density distribution of the character area appears in slices and has a certain height and width, while the background edge regularity is poor. It is shown as a dot-line distribution. According to these characteristics, a combination of the projection and threshold is used to detect the line boundary of the character area to generate candidate character lines. Figure 5 shows the projection [23].

English Document Grammar Recognition System

The reduced grammar rule model takes the preprocessed file as the input that recognizes the named entity according to the grammar rules, and outputs the file containing the recognition results. The reason why the reduced grammar rules are adopted can reduce the complexity of the design of the grammar rule model on the one hand, and on the other hand, the intermediate results obtained by partial matching can effectively reduce the core feature space of the maximum entropy model and reduce the time complexity of the maximum entropy model.

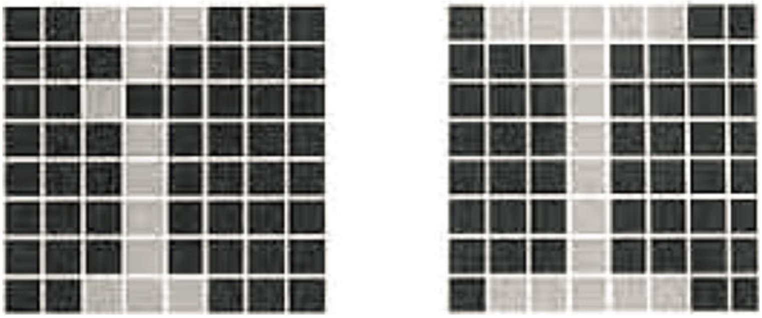


Fig. 5 English character text recognition projection

The problem of low recall caused by the use of reduced syntax rules can also be compensated by partial matching. Search for each gray from low to high, and find the first gray node. If the node is not the parent node, use a dotted line to connect; if it is a parent node, use a solid line to connect; if the connected node is not the next node corresponding to the original tree, use a hollow arrow, otherwise use a solid arrow. After partial matching, the result is passed to the maximum entropy model for further and also final processing. The basic idea of the principle of maximum entropy is: use the given training data to select an appropriate statistical model to make it satisfy all known facts without making any assumptions about the unknown facts. The effective way to realize this idea is to use entropy. Entropy defines the uncertainty of a random variable. When the entropy is the largest, the random variable is the most uncertain, random, and distributed most uniformly. The statistical model at this time satisfies our requirements.

$$p(w_T|w_{T-n+1....}w_{T-1}) \tag{4}$$

There are two conversion modes to convert grammars that are not in Chomsky canonical form. By rewriting the grammar, the grammar can be converted into the Chomsky canonical form. According to the rewritten grammar, the second and third steps of the above algorithm can be then continued. Step, as the grammatical analysis and recognition are successfully completed, so the above algorithm is modified. In order to realize the automatic generation of scheduling components, a special factory object SchedulerFactory can be also implemented in the server-side ORB, which is responsible for analyzing component descriptions and generating instances of scheduling components. When a customer needs a new scheduling component, first obtain the SchedulerFactory reference, and then further send the component description to the core SchedulerFactory, requesting the creation of a scheduling object, and after obtaining the object reference, the real-time task can be monitored. Figure 6 shows the steps.

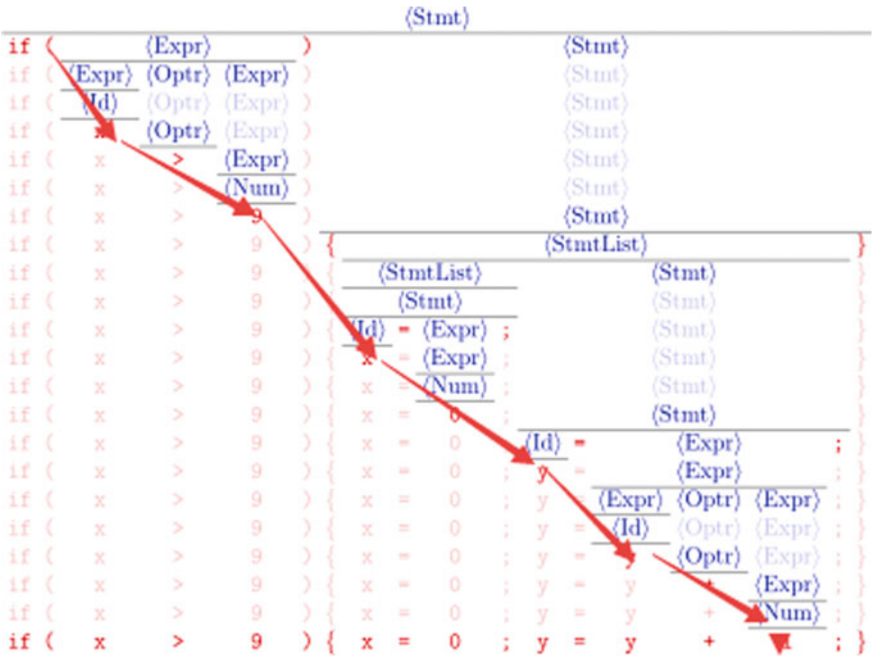


Fig. 6 Document grammar recognition steps

The Simulation and Verification

Using the named entity recognition method combining the reduced rules and the maximum entropy model proposed in this paper, we have implemented a hybrid named entity recognition system named DDCS system. The distribution of character pixels in each region of the dot matrix changes greatly, which brings difficulties to the learning of the network. When the number of samples increases, that is, when the change of the same character font increases, network learning is often difficult to converge, showing that “learning the later samples, we forget the previous samples.” To solve this problem, you must strengthen the memory ability of the network, which means increasing the scale of the network, especially the number of nodes in the first hidden layer. Figure 7 shows the data.

U + 00000000 ~ U + 0000007F	0xxxxxxx
U + 00000080 ~ U + 000007FF	110xxxxx 10xxxxxx
U + 00000800 ~ U + 0000FFFF	1110xxxx 10xxxxxx 10xxxxxx
U + 00010000 ~ U + 001FFFFF	11110xxx 10xxxxxx 10xxxxxx 10xxxxxx
U + 00200000 ~ U + 03FFFFFF	111110xx 10xxxxxx 10xxxxxx 10xxxxxx 10xxxxxx
U + 04000000 ~ U + 7FFFFFFF	1111110x 10xxxxxx 10xxxxxx 10xxxxxx 10xxxxxx 10xxxxxx

Fig. 7 Simulation data considered

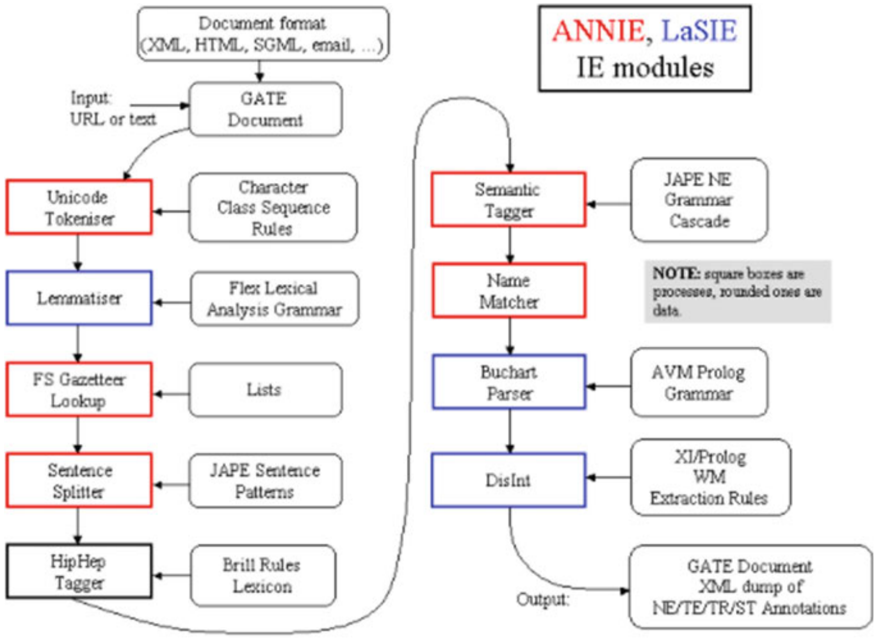


Fig. 8 Recognition process simulation

Count the number of projection values greater than two-thirds of the maximum value N. If N is less than 8 or the maximum value is less than half of the line height, the candidate character is judged to be a false line and removed. In Fig. 8, the simulation process is simulated.

Conclusion and Prospect

This paper studies the key technologies of English document grammar recognition system with the cloud computing and electronics systems. The biggest feature of this method is to ensure that the correct detection rate is high and the false detection rate is low, and the detection and positioning time is significantly shortened, which is particularly beneficial to the general semantic feature-based image/video retrieval and query on the Internet. The proposed model is simulated on the different platforms to validate the performance. In the future study, the ideas will be improved for implementations.

Acknowledgement This paper is supported by: Business English Provincial First-class Major Construction Project Sponsored by Xi'an Fanyi University.

References

1. Nalajala S, Moukthika B, Kaivalya M, Samyuktha K, Pratap NL (2020) Data security in cloud computing using three-factor authentication. In: International conference on communication, computing and electronics systems. Springer, Singapore, pp 343–354
2. Li X, Dang Y, Aazam M, Peng X, Chen T, Chen C (2020) Energy-efficient computation offloading in vehicular edge cloud computing. IEEE Access 8:37632–37644
3. Subha S, Nagalakshmi S (2021) Design of cloud computing-based control algorithm for hybrid power system in smart grid applications. IEEE Can J Electr Comput Eng 44(1):22–29
4. Butt UA, Mehmood M, Shah SBH, Amin R, Waqas Shaukat M, Raza SM, Suh DY, Piran M (2020) A review of machine learning algorithms for cloud computing security. Electronics 9(9):1379
5. Sheikh MS, Liang J, Wang W (2020) Security and privacy in vehicular ad hoc network and vehicle cloud computing: a survey. Wirel Commun Mob Comput 2020
6. Xiaoyu W, Zhengming G (2020) Research and development of data security multidimensional protection system in cloud computing environment. In: 2020 ınternational conference on advance in ambient computing and ıntelligence (ICAACI), IEEE, pp 67–70
7. Cotroneo D, De Simone L, Liguori P, Natella R (2020) Fault ınjection analytics: a novel approach to discover failure modes in cloud-computing systems. IEEE Trans Dependable Secure Comput
8. Butpheng C, Yeh K-H, Xiong H (2020) Security and privacy in IoT-cloud-based e-health systems—a comprehensive review. Symmetry 12(7):1191
9. Al-Malah DKAR, Aljazaery IA, Alrikabi HTS, Mutar HA (2021) Cloud computing and its ımpact on online education. IOP Conf Ser Mater Sci Eng 1094(1):012024. IOP Publishing
10. Abdelsamea A, Nassar SM, Eldeeb H (2020) The past, present and future of scalable computing technologies trends and paradigms: a survey. Int J Innov Appl Stud 30(1):199–214
11. Stergiou, CL, Plageras AP, Psannis KE, Gupta BB (2020) Secure machine learning scenario from big data in cloud computing via internet of things network. In Handbook of computer networks and cyber security. Springer, Cham , pp 525–554
12. Siddiqui S, Darbari M, Yagyasen D (2020) An QPSL queuing model for load balancing in cloud computing. Int J e-Collab (IJeC) 16(3):33–48
13. Khalil H, Osman T, Miltan M (2020) Extracting arabic composite names using genitive principles of Arabic grammar. ACM Trans Asian Low-Res Lang Inform Process (TALLIP) 19(4):1–16
14. Xin Y (2021) Design and application of English Chinese translation system based on feature extraction algorithm. In The ınternational conference on cyber security ıntelligence and analytics. Springer, Cham, pp 778–784
15. Sidnyaev NI, Butenko YI, Bolotova EE (2021) Formal grammar theory in recognition methods of unknown objects. Autom Documentation Math Linguist 54(4):215–225
16. Sage C, Aussem A, Eglin V, Elghazel H, Espinas J (2020) End-to-end extraction of structured information from business documents with pointer-generator networks. In: Proceedings of the fourth workshop on structured prediction for NLP, pp 43–52
17. Balaha HM, Ali HA, Badawy M (2021) Automatic recognition of handwritten Arabic characters: a comprehensive review. Neural Comput Appl 33(7):3011–3034
18. Xu N, Ma L, Wang L, Deng Y, Ni G (2021) Extracting domain knowledge elements of construction safety management: rule-based approach using Chinese natural language processing. J Manag Eng 37(2):04021001
19. Zavorin I, Bills A, Corey C, Morrison M, Tong A, Tong R (2020) Corpora for cross-language ınformation retrieval in six less-resourced languages. In: Proceedings of the workshop on cross-language search and summarization of text and speech (CLSSTS2020), pp 7–13
20. Nguyen HD, Huynh T, Hoang SN, Pham VT, Zelinka I (2020) Language-oriented sentiment analysis based on the grammar structure and ımproved self-attention network. ENASE 339–346
21. Demner-Fushman D, Elhadad N, Friedman C (2021) Natural language processing for health-related texts. In: Biomedical Informatics. Springer, Cham, pp 241–272

22. Premananada R, Jambukesh HJ, Shrıdhar H, Rajashekar U, Harısha KS (2020) Design and ımplementation of automated ımage handwriting sentences recognition using hybrid techniques on FPGA. In: 2020 third ınternational conference on multimedia processing, communication & ınformation technology (MPCIT). IEEE, pp 152–161
23. Kanwal T, Moqurrab SA, Anjum A, Khan A, Rodrigues JJPC, Jeon G (2021) Formal verification and complexity analysis of confidentiality aware textual clinical documents framework. Int J Intell Syst

Changes in Commodity Circulation Channels Based on International Trade E-commerce Platform from Computing and Electronics System Perspectives

Wenji Wei

Abstract Changes in the commodity circulation channels based on the international trade e-commerce platform from computing and electronics system perspectives are studied in this paper. From the perspective of the overall service functions and links of logistics, logistics infrastructure not only meets the needs of logistics organization and management, but also provides important places or organizations to give full play to the logistics service functions. Hence, the RFID technology is optimized to be combined for the designed channels and the intelligent platform is constructed through the integration of the theoretical modelling. For the systematic verification, the public data sets are considered, and the state-of-the-art methods are simulated to validate the overall performance.

Keywords Computing and electronics · Systematic modelling · E-commerce platform · Commodity circulation channels · Online system

Introduction

The achievements of the information technology have brought unlimited possibilities to all countries and industries in the new century. E-commerce based on information networks, a brand-new form of international trade, has brought huge development opportunities to international trade [1–3].

If we can solve the problems faced by e-commerce, we can make good use of e-commerce to bring great benefits to international trade. The advancement of science and technology will inevitably bring about changes and reforms. What is innovated is the part of the production relationship that does not adapt to the development of productive forces. Advanced technology brings about higher-level production relationships. The same is true for the development of e-commerce. New forms and virtualized trading models have a great impact on the supervision of international

W. Wei (✉)
School of Information Technology, ShangQiu Normal University, ShangQiu 476000, HeNan, China
e-mail: sqwj_wei@163.com

V. Bindhu et al. (eds.), *Proceedings of Third International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 844,
https://doi.org/10.1007/978-981-16-8862-1_71

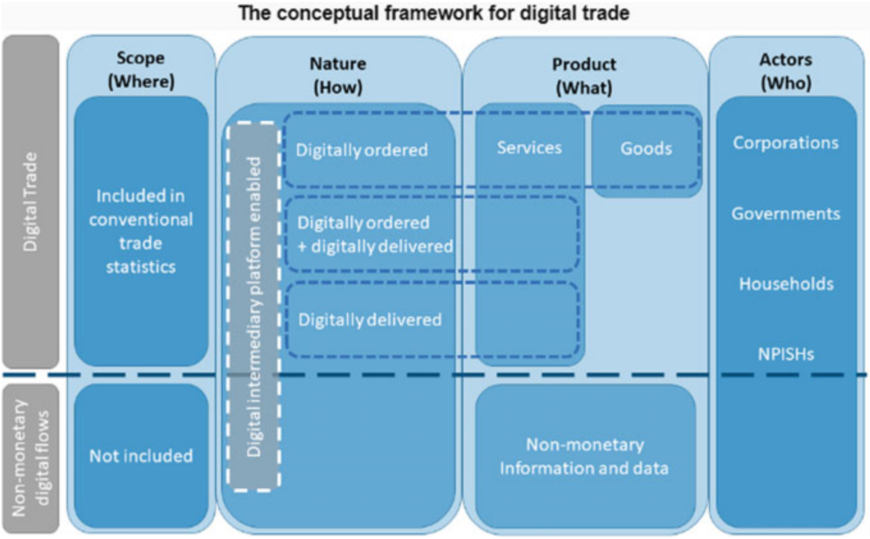


Fig. 1 The framework of the modern digital trade

trade. At the same time, it also creates new opportunities for the innovation of international trade macromanagement. The Chinese government's understanding of e-commerce has been around for a long time. Many government departments have used e-commerce technology to conduct macromanagement of international trade, and they have then effectively promoted the development of e-commerce. As the sample, in Fig. 1, we demonstrate the framework of the modern digital trade [4, 5].

Choosing a practical e-commerce implementation model is essential. The e-commerce model is a way for companies to provide value and profit for the target market at the certain position in the value chain system [6]. Digital time stamp technology refers to the use of computer technology to digitize traditional time stamps. A trusted time stamp service provided by a third party, and the digital time stamp data obtained through this service can prove that the data already exists at a certain moment [7, 8]. At present, since the country has not issued a related digital source calibration standard for digital electric energy measurement systems, the digital signal source cannot be used as a calibration benchmark for accuracy judgment. In order to ensure the traceability of the system to be verified, a high-precision analog-to-digital signal method is now used to transfer and convert data from an analog standard power source to solve the problem that the digital signal cannot be traced. Hence, in our designed methodology, we will adopt it from the following two core aspects. (1) The operation status of the remote time server is monitored through the network, and if it is abnormal, the relevant person in charge will be notified by means of short messages to deal with it in a timely manner. (2) The difference between the two is used as the phase discriminator, and an ideal second-order digital phase-locked loop is used to track the source so that the output of the local crystal oscillator has high precision and stability. If one source is abnormal, switch to the other source to ensure system reliability.

Literature Review

In this literature review, we are focused on two aspects, namely: electronic international trade and digital commodity circulation [9, 10].

The main body of international trade has also changed a lot, from the previous entity company to the virtual company of the e-commerce platform. Network information technology integrates the core technologies of some different entity companies in their fields of expertise, unites a large number of companies in the virtual world, forms a rope, weaves into a group company network, completes the market functions that a single company cannot achieve, and also provides more valuable goods and higher-level services [11]. The development of e-commerce has greatly promoted the international division of labor [12].

The deployment of production factors such as capital, raw materials, and talents on the network platform has promoted the internationalization of production. This promotes the international division of labor, and ultimately accelerates the global flow of products and semi-finished products [13, 14]. For example, production companies accept orders through the network platform, and follow the principle of proximity to produce on-site or cooperate with local companies to produce, and the finished products are sold and distributed on-site.

After the system sends out the data frame according to the test requirements, the measured electric energy meter calculates the measured value through the internal measurement chip, and the measured value is sent to the pulse acquisition module of the tester. The core test system compares this pulse value with the data of the standard electric energy meter [15]. Perform comparative processing and analysis, and obtain a more accurate measurement error through the method of the cumulative averaging. Commodity barcode reading is realized by distinguishing the border and width of the bar and space. Therefore, it is required that the color contrast between the bar and the space is as large as possible. The strip color is dark, and the empty color is light. Black strips and white shorts are the most ideal color combination. Since the working light source of the bar code reading instrument is the color of the red, the bar code with red bar cannot be read.

The Designed Methodology

The RFID System Model

The system hardware design adopts the principle of modular design. The hardware platform is divided into microprocessor core module, GSM wireless communication module, voice navigation module, GPS positioning acquisition module, and RFID positioning acquisition module [16, 17]. These modules are the core part of the system and determine the terminal system as the important indicator of the performance. The basic principle of using RFID technology for communication is: after the tag enters the magnetic field, it receives the radio frequency signal from the reader and uses the energy obtained by the induced current to send out product information stored in the chip, or actively send a signal of the certain frequency. After the information is read and decoded, it is sent to the central information system for data processing [18, 19].

Considering the communication distance, transmission frequency, cost, and power consumption required by this design, active electronic tag is used in the system. The electronic tag carries user information. When it receives the trigger signal sent by the reader, the tag will be activated, and within a specified time, it will continuously send request commands to the reader until it then receives the response command from the reader [20]. Then the tag will send the message carried, and the reader will receive and process the message. In Fig. 2, the model structure is presented with following focuses.

1. Carry on the development of driver program, complete the initialization of the extended device, provide an interface for the application program to access the device, and realize the data reading and transmission of the device and the user space [21].

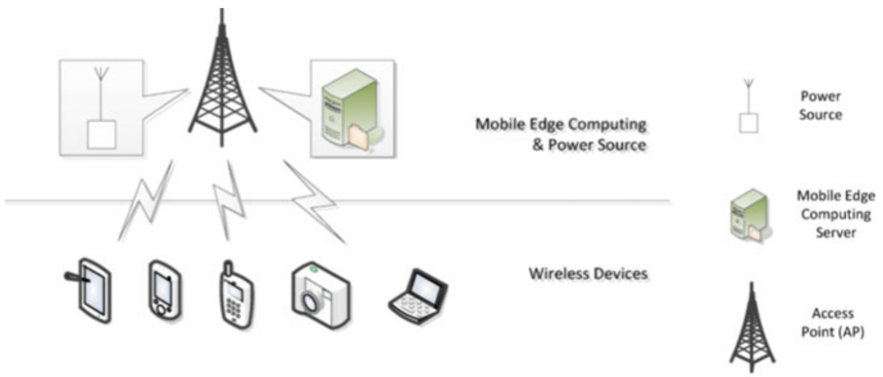


Fig. 2 RFID system model structure

2. GPS acquisition process: In order to realize the system positioning and tracking function, a series of data acquisition and processing are required. The system main core process first completes the initialization, waits for the GPS positioning request command, and calls the general GPS acquisition process when receiving the GPS positioning request command.

Electronic International Trade and Forecast Model

The international trade model of vertically differentiated products can be divided into two main categories: the first category still retains the assumption of perfect competition. After incorporating the vertical differentiation of products, the conclusion obtained is that intra-industry trade still conforms to the traditional proposition of the H–O theorem. Another type of model assumes an oligopoly market. In the competition, some companies are then eliminated, and the surviving companies can realize economies of scale and reduce production costs and sales prices, thereby benefiting consumers. Based on evolution mechanism extraction of the above empirical analysis, this paper integrates the growth of nodes, the optimal connection of nodes based on fitness, the creation of edges, the deletion of edges, and the mechanism of node deletion. This paper establishes a weighted directed network evolution model as follows [22].

$$\Pi_{\text{est}} = \frac{\eta(i)s^{\text{in}}(i)}{\sum_{\text{est}} \eta(l)s^{\text{in}}(l)} \tag{1}$$

The Π is the calculation model considered, because the data required by the gravity model has the characteristics of strong availability and high credibility. The application of the gravity model of trade has become more and more extensive, and it has become the main empirical research tool for the current international trade flow. It is set according to the main factors affecting the bilateral trade flow. Different explanatory variables to analyze the influence direction and size of these factors, and calculate and analyze the trade potential. In Fig. 3, the node distribution model is considered.

From the formation of the initial theory of the international trade to the establishment of the extended gravity model, it reflects the process of people's exploration of the root causes of the emergence and continuous evolution of international trade, that is, the process of progress from ideological experiment to demonstration. Summarizing the evolution law of various research methods in this process will be of great help to the future economic and statistical research. In the lemma 2, the matrix model is defined [23].

$$G_{\text{matrix}} = \begin{pmatrix} H & Y \\ X & O \end{pmatrix}_{(2nn+n)\times(2nn+n)} \tag{2}$$

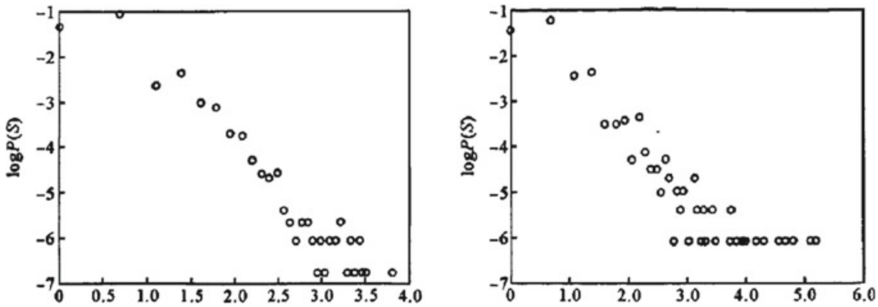


Fig. 3 The distribution model

The H, Y, X, O are the different dimensions. Trade function is an important tool of modern economic analysis, which not only objectively reflects trade and economic relations, but also provides the accurate reference data for economic regulation. Therefore, its estimation will be one of the important contents of future statistical work, so it is necessary for us to do some systematic research on its construction. The development and operation of e-commerce depend on complete information network facilities. Although China's investment and construction in information network facilities have achieved certain results, and also the network architecture has basically taken shape, compared with foreign countries, China's investment in information network facilities is significantly insufficient and the foundation is relatively weak. In order to promote the healthy and also rapid development of e-commerce in China, it is more necessary to make effective use of existing resources on the basis of accelerating the construction of information network facilities. Traditionally, the technology used to exchange electronic data between international business partners is EDI based on value-added services. Its infrastructure investment is huge, which is the patent of large enterprises. Small- and medium-sized enterprises have been squeezed out of traditional EDI because they cannot afford their investment outside. In addition, traditional EDI still has a problem of unifying global general technical standards, and all current standards are not suitable for Internet EDI or Web-EDI, which means that it is not suitable for data search, positioning and display through the Internet. Internet EDI and Web-EDI are the latest developments of EDI. They take advantage of the characteristics of the Internet, such as fast transmission speed and low cost, to meet the basic requirements of modern small- and medium-sized enterprises when they use the Internet to conduct international trade with some uncertain business partners. Companies and large companies have the opportunity to stand on the same starting line. They are a symbol of modern EDI and a breakthrough for the widespread application and development of traditional EDI. In Fig. 4, the layer model is presented.

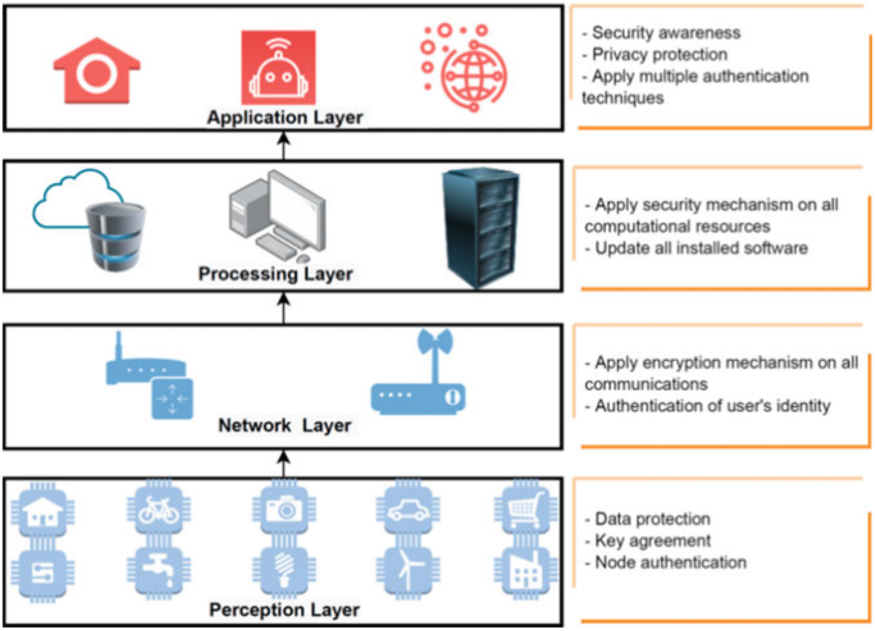


Fig. 4 Electronic international trade and forecast model layer model

Commodity Circulation Channel Modelling

ARM maintains the local time in the driver layer through interrupt processing functions, where the interrupt interval is Ires. Therefore, the minimum time resolution obtained through the hardware is then lms, and the resolution below milliseconds is maintained by the ARM's own counter, and its time resolution is configured as 40 ns. Hence, we consider listed aspects.

1. Uniqueness is the basic principle of commodity barcode coding, which means that commodities with the same basic characteristics should be assigned the same commodity identification code, and commodities with different basic characteristics must be assigned different commodity identification codes. The basic characteristics of commodities include commodity name, trademark, type, specification, packaging type, etc.
2. Stability refers to the fact that the product identification code has been assigned, so long as the basic characteristics of the product do not change, the bar code should remain unchanged. Even if the same commodity is discontinued, its identification code cannot be used on other commodities.
3. Meaninglessness means that each digit in the product identification code does not represent any specific information related to the product.

In general, the above four ways are combined with each other. The transfer of commodity ownership is accompanied by the movement of commodity entities.

Commodity circulation mainly depends on the organic unity of business flow and logistics, and this process is inevitably combined with the transfer of capital flow and information flow. Commodity exchange and social division of labor are the basis for the emergence and development of commodity circulation. Commodity circulation industry continues to develop with the changes of market system and commodity exchange system. At the same time, the level and scope of commodity circulation industry have been continuously improved with the intervention of various exchange media and the reform of some commodity circulation industry structure and system, while giving birth to more and more developed commodity circulation. The PC terminal starts the data transmission service with the tester, and immediately sends a specified number of data messages. The PC terminal and the tester enter the status synchronization protection interaction, and regularly obtain the message receiving status of the embedded receiver.

The use of a programmable device FPGA solves the space problem of discrete devices, and makes the hardware more flexible. It can easily implement various logic functions, such as frequency division, phase shift, encoding, and decoding, which is particularly important in time application systems. After the test system, collector converts the measured electric energy meter signal, the high-pass filter filters out the DC component, and then the core instantaneous active power is calculated by the relevant algorithm, and then the active energy and also reactive power are obtained through the steps of low-pass filtering, error correction, and energy integration. Electric energy and electric energy value of each phase. There are usually two ways to obtain electrical energy. The first is to use a dedicated metering chip to achieve it through hardware. The chip integrates AD acquisition and hardware integration. This method does not require software calculation. The second is AD sampling software calculation method.

During electric energy calculation, the sampling points convert each harmonic into orthogonal waveform within the designed harmonic range, and the power value is obtained by direct multiplication and integration of the sampling points. The reactive power value calculated by Hilbert algorithm is more accurate than Fourier algorithm. Therefore, system adopts Hilbert algorithm to realize electric energy measurement and conversion. In Fig. 5, the model is defined.

The Experiment

As the source of vitality of the regional logistics organizations, logistics innovation capabilities are naturally important. Constantly seeking reforms and implementing innovations are the requirements of the development of the logistics industry and also market changes, as well as the guarantee for the sustainable development of logistics capabilities. Any innovation ability is inseparable from the investment of talents and energy, and regional logistics innovation ability is no exception. The signal received from the analog standard power source is converted into a digital

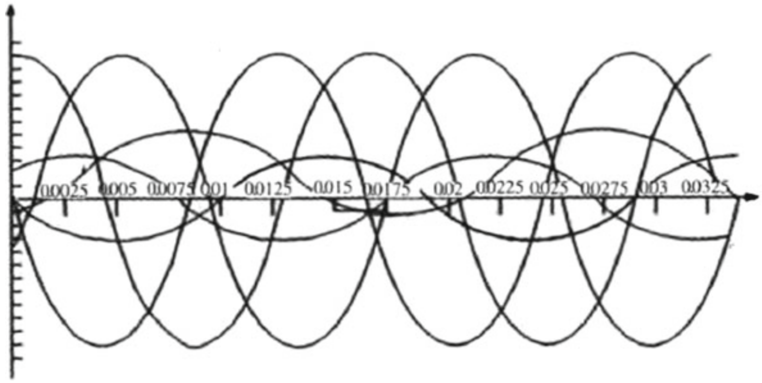


Fig. 5 Commodity circulation channel modelling

signal through a high-precision transformer and NI collector. The transformer and collector are embedded in the digital meter tester. In Fig. 6, the result is presented.

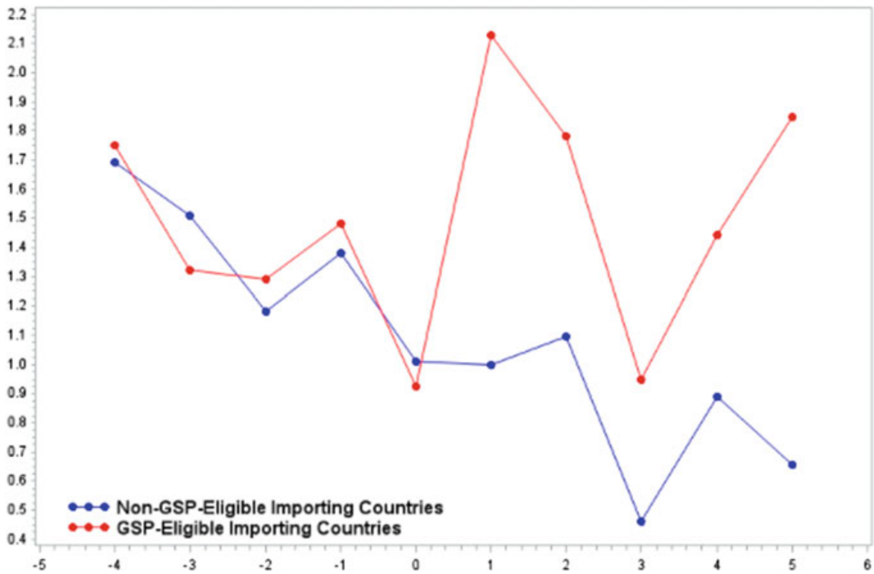


Fig. 6 The estimation performance result

Conclusion and Prospects

Changes in the commodity circulation channels based on the international trade e-commerce platform from computing and electronics system perspectives are studied in this paper. The rationality of the logistics network and node layout also has an important impact on the regional logistics carrying capacity. In the development process of the logistics, logistics networks and nodes have spread throughout various urban areas. The more reasonable distribution of logistics networks and nodes, the higher the efficiency and carrying capacity of logistics operations. This paper integrates the computing and electronics system to construct the efficient systematic model. The simulation has shown the robustness, and in the future study, we will consider some further combinations on applications.

References

1. Do TN, Visell Y (2017) Stretchable, twisted conductive microtubules for wearable computing, robotics, electronics, and healthcare. Sci Rep 7(1):1–12
2. Watkinson N, Zaitsev F, Shivam A, Demirev M, Heddes M, Givargis T, Nicolau A, Veidenbaum A (2021) EdgeAvatar: an edge computing system for building virtual beings. Electronics 10(3):229
3. Song X, Pickert V, Ji B, Naayagi RT, Wang C, Yerasimou Y (2017) Questionnaire-based discussion of finite element multiphysics simulation software in power electronics. IEEE Trans Power Electr 33(8):7010–7020
4. Neamtu OM, Gordan CE, Gordan IM (2021) A computer control for electronics actions of smart voice ınteractive devices. In: 2021 16th ınternational conference on engineering of modern electric systems (EMES). IEEE, pp 1–4
5. Ahmad A (2017) Challenges for test and fault-tolerance due to convergence of electronics, semiconductor systems and computing. In: 2017 ınternational conference on ınfocom technologies and unmanned systems (Trends and Future Directions)(ICTUS). IEEE, pp 64–68
6. Abu Al-Haija Q, Zein-Sabatto S (2020) An efficient deep-learning-based detection and classification system for cyber-attacks in IoT communication networks. Electronics 9(12):2152
7. Patra B, Incandela RM, Van Dijk JPG, Homulle HAR, Song L, Shahmohammadi M, Staszewski RB et al (2017) Cryo-CMOS circuits and systems for quantum computing applications. IEEE J Solid State Circ 53(1):309–321
8. Akmandor AO, Jha NK (2017) Smart health care: an edge-side computing perspective. IEEE Consum Electr Mag 7(1):29–37
9. Shastri BJ, Tait AN, Ferreira de Lima T, Pernice WHP, Bhaskaran H, David Wright C, Prucnal PR (2021) Photonics for artificial intelligence and neuromorphic computing. Nat Photon 15(2):102–114
10. Liu C, Chen H, Wang S, Liu Q, Jiang Y-G, Zhang DW, Liu M, Zhou P (2020) Two-dimensional materials for next-generation computing technologies. Nat Nanotechnol 15(7):545–557
11. Sun Z, Pedretti G, Ielmini D (2019) Fast solution of linear systems with analog resistive switching memory (RRAM). In: 4th IEEE international conference on rebooting computing, ICRC 2019. Institute of Electrical and Electronics Engineers Inc., pp 1–5.
12. Lacka E, Chan HK, Wang X (2020) Editorial on "Technological innovation in B2B international trade". Ind Market Manag 91:174
13. Toleuuly A, Yessengeldin B, Khussainova Z, Yessengeldina A, Zhanseitov A, Jumabaeva S (2020) Features of e-commerce risk management in modern conditions. Acad Strateg Manag J 19(1):1–6

14. Zhang YY, Zheng Q, Wang H (2021) Challenges and opportunities facing the Chinese economy in the new decade: epidemics, food, labor, e-commerce, and trade. Chin Econ, 1–3
15. Gu Y, Tan J, Zeng Q (2020) Measurement of cross-border e-commerce convenience level in countries along the belt and road initiative. In: Innovative Computing. Springer, Singapore, pp 997–1007
16. Xi Y, Ming L (2020) Research on the talent training mode of application-oriented undergraduate cross-border e-commerce innovation and entrepreneurship education. In: 2020 international conference on big data and informatization education (ICBDIE). IEEE, pp 119–122
17. Dai Y (2020) Data protection laws—one of the most important sources of competitive advantage in the context of international trade. J Data Protect Priv 4(1):72–80
18. Cao X, Deng M, Li H (2021) How does e-commerce city pilot improve green total factor productivity? Evidence from 230 cities in China. J Environ Manag 289(2021):112520
19. Liu YY (2020) Empirical analysis of the impact of China's cross-border electronic commerce on foreign trade based on VAR model. J Phys Conf Ser 1533(2):022064 (IOP Publishing)
20. Wang N (2021) The ınfluence of computer network technology on trade development. In: International conference on applications and techniques in cyber security and ıntelligence. Springer, Cham, pp 22–27
21. Dahsh FA, Kazim NA (2020) Legal measures to curb fraudulent e-commerce websites during the COVID-19 period. Int Res J Eng Technol (IRJET) 7(09)
22. Savinov Y, Taranovskaya E, Gavryushin O (2021) On determining the effectiveness of digital marketing in transnational e-commerce. In: Digital strategies in a global market. Palgrave Macmillan, Cham, pp 173–183
23. Pavicevic D, Novicevic R, Abramovic N, Crvenica D (2021) The importance of applying the principle of sustainable development in trade. In: SHS web of conferences, vol 111. EDP Sciences, p 01003

Ultrasonic Sensor-based Canopy Height Measurement and Root Depth Estimation

S. Gilbert Rozario and V. Vasanthi

Abstract Today, agriculture is plagued because of low yield. Good crop management practices lead to good yield. There is some observation required for monitoring crop growth in different growth stages such as canopy height, and root depth. Traditionally, canopy height is measured manually. This method is low, labor-intensive, and lacks accuracy. IoT and wireless sensor networks help to monitor crop growth in different stages. Canopy height has been used to predict yield, fertilizer application in the crop, water requirement, or biomass. The ultrasonic sensor emits sound waves to the plant area. The sound waves are reflected from the plant and absorb by the receiver. The time difference between emitting and absorb was calculated to measure the canopy height and transfer the data to the server using IoT and wireless sensor network architecture. Canopy height is highly correlated with root depth. Using artificial intelligence (AI) technics, easy to estimate root depth. Root depth information helps in irrigation scheduling. During photosynthesis, the crop converts carbon dioxide to carbohydrates, this cycle is known as the Calvin cycle. Depend on canopy height, Calvin cycle will varieties. So canopy height is directly related to crop yield.

Keywords Ultrasonic sensor · Canopy height · Root depth estimation · Wireless sensor network · IoT · Linear regression

Introduction

The precision farming system required the basic crop growth information with reliable and fast acquire accurate data at low cost. Acquired crop growth information will get through destructive sampling in the agricultural field. In destructive sampling

S. Gilbert Rozario (✉)
Department of Computer Science, Rathinam College of Arts and Science, Coimbatore, India

V. Vasanthi
Sri Krishna Adithya College of Arts and Science, Coimbatore, India
e-mail: vasanthiv@skacas.ac.in

V. Bindhu et al. (eds.), *Proceedings of Third International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 844,
https://doi.org/10.1007/978-981-16-8862-1_72

method, it will take more time and more expensive to get a reliable result. These approaches are labor-intensive and subjective, therefore constraining the number of hectares that have to cover by the same person. In the rainy season, it is very difficult to enter the field and collect samples.

Agronomists and plant breeders used to measure canopy height are by using a meter scale. Multiple samples are taken in a single plot and use a rule to average a canopy height and record the topmost part of the plant. However, this method is a time-consuming, costlier, and individual decision for the highest part of the crop. Therefore, the quality might vary from person to person. Nowadays, agriculture required non-destructive sampling technologies for basic crop growth monitoring information in real-time.

Wireless sensor networks place an important role in collecting basic crop growth information in the field. An ultrasonic sensor uses to find the actual height of the plant. The ultrasonic sensor emits high frequency (40 kHz) sound waves at time intervals. If they strike a plant, then the waves are reflected as echo waves to the ultrasonic sensor, the time taken between emitting sound waves and receiving the echo to the sensor helps to calculate the canopy height. Sound waves are hammerless waves, it will not affect the crop or human. Ultrasonic sensors were used for many applications in the field of agriculture for example object dictions, water level management, wind speed, and direction estimation. Canopy height is the basic and important key parameter for most agricultural research.

The ultrasonic sensor is a part of a wireless network. Inbuilt program is written in the base station in time interval base station trigger the ultrasonic sensor and receive canopy height and sent to the server. In a field experiment, we observed more than 250 groundnut plants with basic parameters like canopy height, root depth, planted date, observed date, etc. Once we received new canopy height from the base station server estimate root depth with existing data using artificial intelligence. In the real-world agricultural sector, we can destructive plant sampling every time and take canopy height, root depth, it is a very tedious process and the very farmer cannot do this kind of observation. But all the parameters are highly correlated with yield. So crop growth information has to be collected without destructive plants and with zero knowledge in observation.

Canopy height is an important parameter for estimation yield prediction. Yield does not come in overnight. It will take a plant life cycle. So in every stage of the plant has to be maintained properly. Water requirement, fertilizer requirement, and the light intensity will vary from every stage of the plant without knowing the exact plant requirement if we apply water or fertilizer, it may be deficient or extra fertilizer or water will be wasted. It also directly affected yield. To estimate plant water and fertilizer requirement, crop growth information has to be observed in a fixed time interval and also in every stage of the crop. Canopy height is the important parameter for estimate water requirement and fertilizer requirement. In every growth stage, canopy height has to be increased to a certain height. If not fertilizer or water is deficient.

Root depth is an important parameter for estimating water requirements. If water below the root depth was wasted. Depend on root depth, irrigation has to be scheduled.

If canopy height is more, it inhales more carbon dioxide, and during the photosynthetic process, carbon dioxide converts to more carbohydrates. This process is directly related to yield. All the above parameters are more important for crop production. So wireless sensor network collects all this information and transfer to server. Crop growth information monitors at every stage of the crop and calculate the difference in all parameter values. This kind of information helps to take proper decision in right time.

Materials and Method

Field Experiments and Height Measurement

The field experiments were conducted from month of July 2019 for 4 months of groundnut crop. Randomized block design as followed and more than 250 crops wear destructed and observed major parameters such a canopy height, root depth, sowing date, and observation date. Canopy growth monitoring will be observed weekly once using an ultrasonic sensor, therefore we can analyze canopy growth every week. Around 15–20 groundnut, plants were taken to observation in a week. An ultrasonic sensor will measure canopy height in centimeters. Ultrasonic sensor fixed in sensor bar 1 m from the ground. To calculate the distance between plant and sensor,

$$\text{Distance} = (\text{Time taken to receive echo} \times \text{Speed of sound})/2 \tag{1}$$

$$\begin{aligned}\text{Canopy height} =& (\text{Distance between ground and sensor}) \\ &- (\text{Distance between plant and sensor}).\end{aligned} \tag{2}$$

Using an ultrasonic sensor, we will get the exact height of the canopy in centimeters.

Canopy distance measuring code

```
const int trigPin = 4;
const int echoPin = 3;
long duration;
int distance;
void setup() {
```

```
pinMode(trigPin, OUTPUT);
pinMode(echoPin, INPUT);
Serial.begin(9600);
}
void loop() {
digitalWrite(trigPin, LOW);
delayMicroseconds(2000);
digitalWrite(trigPin, HIGH);
delayMicroseconds(1000);
digitalWrite(trigPin, LOW);
duration = pulseIn(echoPin, HIGH);
height = 200-duration*0.034/2;
Serial.print("Canopy Height: ");
Serial.println(height);
}
```

Hardware Details

Arduino UNO R3 microcontroller user to process the sensor data and transfer to a base station. Ultrasonic sensor HC-SR04 used for canopy height measurement. This sensor non-contact range detection with high accuracy. It can detect from 2 to 400 cm range and 30° measuring angle from the sensor. In Fig. 1, it clearly explains about ultrasonic sensor connected in Arduino UNO R3. The configuration pin of HC-SR04 is VCC (red wire), TRIG (yellow wire), ECHO (green wire), and GND (black wire). The supply voltage of VCC is +5 V and you can attach TRIG and ECHO pin to any digital I/O in your Arduino board. To create the ultrasound, we need to set the trigger pin on a high state for 10 μs. That will convey a 8 cycle sonic burst which will go at the speed sound and it will be gotten in the echo pin. The echo pin will yield the time in microseconds the sound wave voyaged.

Data Transmitted to Server

Wireless sensor-based monitoring system is cost-effective and durable at outdoor operations. Ultrasonic sensor node receives echo signal and calculates the canopy

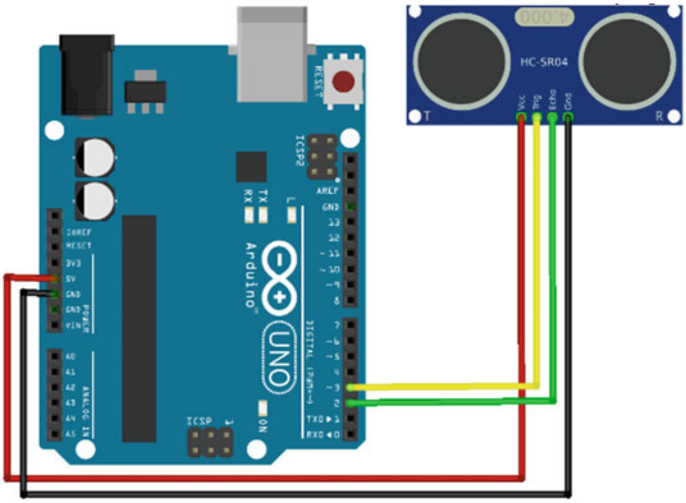


Fig. 1 Sensor connection to Arduino UNO microcontroller

height, then data are transmitted to the base station using wireless sensor network technic. The data are processed and transmitted from the base station to the remote server for further analysis.

Estimating root depth

In the field experiment conducted in the groundnut plot, more than 250 plants were taken in a basic observation like canopy height, root depth in various growth stages. The complete observed data were stored in the server. Canopy height is correlating with root depth. Every time, it is difficult to take destructive plant samples to find root depth and lead are index. So machine learning technic help to estimate root depth with help of canopy height. The sensor node transmits canopy height to the base station and the base station sent the data to the server. Root depth can be estimated using canopy height in the server.

Linear regression is a statistical method for estimating the relationship between variables. The linear regression model is mainly focused on the relationship between a dependent variable and one or more independent variables. In this paper, canopy height is an independent variable and root depth is a dependent variable. Once we give canopy height as an input variable to the linear regression modal, it will predict the root depth.

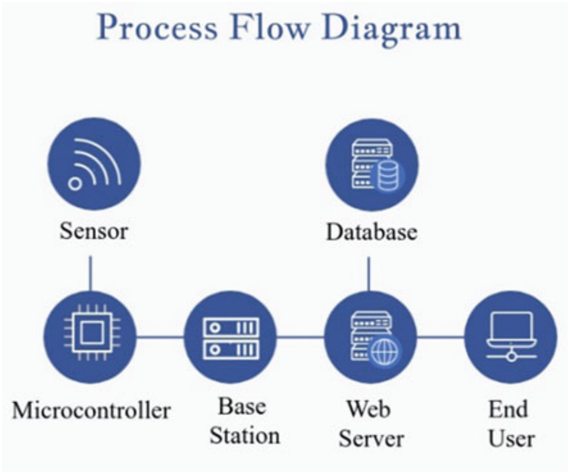


Fig. 2 Process flow diagram

Decision Support System

Decision support systems help to decide irrigation, fertigation, and other operations in agriculture. Canopy height helps to calculate water demand for the crop, and root depth is helps to define irrigation type. So these two pieces of information help the decision support system to take proper decisions on irrigation scheduling (Fig. 2).

Results

Ultrasonic sensor groundnut canopy height calculation showed the expected curvilinear growth pattern. Groundnut crop canopy height measured by both manually and ultrasonic sensor. The comparison study has been done both manually, and sensor-based using various statistical methods such as mean absolute Error is 1.11. Therefore, the error rate between manually and sensor-based is 1.11 cm. Same like that mean squared error is 2.01 and root mean squared error is 1.41. It means the error rate between sensor-based and manual measurement difference where 1 or 2 cm that also depends upon wind speed and air temperature.

Ultrasonic sensor data transfer to server. In a field experiment conducted in the groundnut plot, more than 250 plants observed data are stored in the server. Groundnut canopy height and root depth have the strongest linear relationship of 0.9 correlation coefficient value. We used a linear regression model to estimate root depth as a dependent variable and canopy height as an independent variable. Data preparation for the simple linear regression model, we split 70% data as train data and 30% for test data, then we build the model.

In Fig. 3, it clearly explains the linear regression summary. This model has a 0.845 R-squared value, it represents the goodness of fit for the model. The data are perfectly fit to the model. The F-statistic is a statistics test overall significance of the model. Therefore, this model has very good results.

In Fig. 4 overall linear regression model prediction in graphical representation. The plant height and root depth are highly correlated in Fig. 4, *y*-axis shows the root depth and the *x*-axis shows the plant height then each point represents a plant. Figure 4 clearly explains that every growth stage plant height is correlated with root depth. Depending on plant height, the root depth also increasing accordingly.

In Fig. 5, overall picture explains plant growth in a different growth stage. The observation was taken in the weekly interval for the whole plant life cycle. Depend on each growth stage, the plant water and nutrition requirements also change. Initial stage plant height and root depth are smaller the water and nutrition requirement also less. For every growth stage, plant height and root depth will increase depend on that plant requirement also increase. This study helps us to estimate plant requirements based on the plant growth stage. This study helps the decision support system to

OLS Regression Results

Dep. Variable:	root_length	**R-squared:**	0.845
Model:	OLS	**Adj. R-squared:**	0.844
Method:	Least Squares	**F-statistic:**	876.1
Date:	Fri, 07 May 2021	**Prob (F-statistic):**	5.11e-67
Time:	16:00:53	**Log-Likelihood:**	-226.79
No. Observations:	163	**AIC:**	457.6
Df Residuals:	161	**BIC:**	463.8
Df Model:	1		
Covariance Type:	nonrobust		

	coef	**std err**	**t**	**P>|t|**	**[0.025**	**0.975]**
const	3.4371	0.270	12.746	0.000	2.905	3.970
plant_height	0.2702	0.009	29.600	0.000	0.252	0.288

Omnibus:	1.157	**Durbin-Watson:**	1.904
Prob(Omnibus):	0.561	**Jarque-Bera (JB):**	1.236
Skew:	-0.146	**Prob(JB):**	0.539
Kurtosis:	2.690	**Cond. No.**	104.

Fig. 3 Simple linear regression modal evaluation

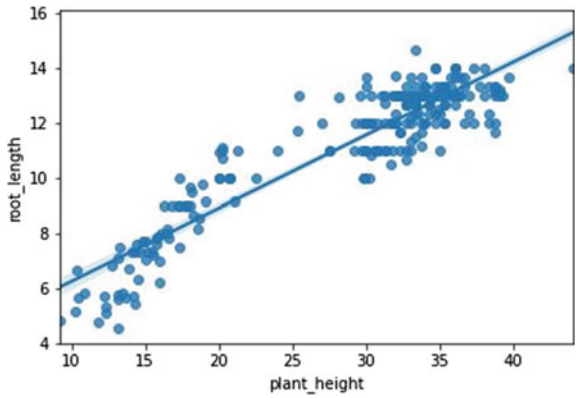


Fig. 4 Simple linear regression model visualization

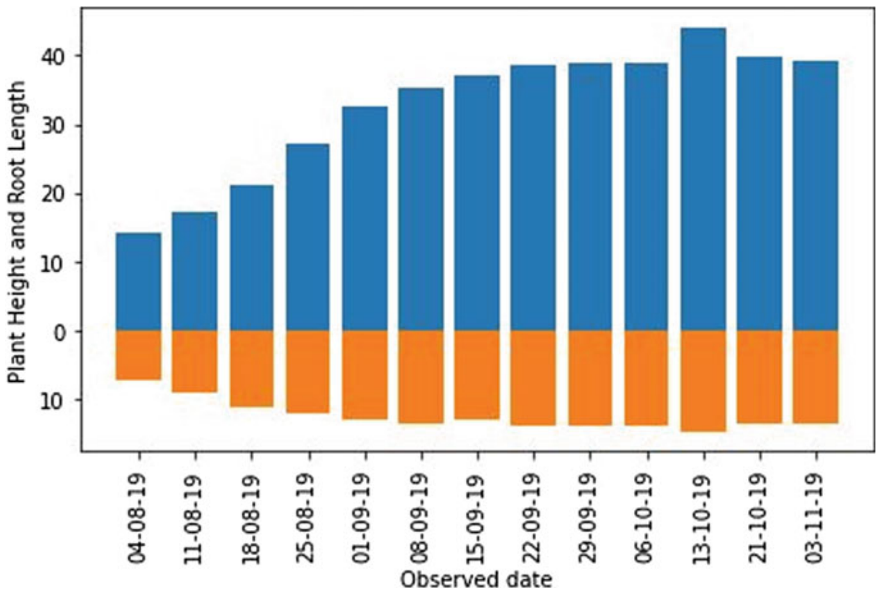


Fig. 5 Observation of plant growth stages

take proper decision irrigation scheduling and fertigation. Wireless sensor networks place an important role in process automation in the agricultural field.

Conclusion

This study confirms the utility of an ultrasonic sensor for plant height measurement that can be used in water and fertigation management. Manual observation takes more time and observation various from person to person. It required more manpower but in the case of ultrasonic sensor-based observation is accurate and cost-effective. Root depth observation is taken by destructive the plant. Machine learning technic helps to estimate root depth with the help of previously observed data and without a destructive plant. It is not actual root depth, it cannot be used for research work, but it is very useful for irrigation scheduling how much mm water has to be irrigated based on plant requirement and plant growth stage.

References

1. Bronsona KF, Frencha AN, Conleya MM, Barnesb EM (2021) Use of an ultrasonic sensor for plant height estimation in irrigated cotton
2. Trevisan RG, de SV Junior N, Portz G (2015) Use of crop height and optical sensor readings to predict mid-season cotton biomass
3. Chang YK, Zaman QU, Rehman TU (2017) A real-time ultrasonic system to measure wild blueberry plant height during harvesting. Biosys Eng 157:35–44
4. Yuan W, Li J, Bhatta M, Shi Y, Baenziger PS, Ge Y (2018) Wheat height estimation using LiDAR in comparison to ultrasonic sensor and UAS. Sensors 18:3731
5. Barmeier G, Mistele B, Schmidhalter U (2016) Referencing laser and ultrasonic height measurements of barley cultivars by using a herbometre as standard. Crop Pasture Sci 67:1215–1222
6. Fricke T, Richter F, Wachendorf M (2011) Assessment of forage mass from grassland swards by height measurement using an ultrasonic sensor. Comput Electron Agric 79:142–152
7. Li Z, Chen Z (2011) Remote sensing indicators for crop growth monitoring at different scales. IGARSS 2011
8. Ni J, Zhan J, Rusong W (2018) Development of an apparatus for crop-growth monitoring and diagnosis. Sensors 18:312
9. Srivastava R, Sharma V, Jaiswal V, Raj S (2020) A research paper on smart agriculture using IoT. Int Res J Eng Technol (IRJET) 07(07)
10. Dash D, Agarwal A, Agarwal K, Misra G (2021) Post catastrophe fallouts and challenges to swim to safety. J Inf Technol 3(01):12–17
11. Sivaganesan D (2021) Performance estimation of sustainable smart farming with blockchain technology. IRO J Sustain Wirel Syst 3(2):97–106
12. Manoharan JS (2020) Supervised learning for microclimatic parameter estimation in greenhouse environment for productive agronomics. J Artif Intell 2(3):170–176
13. Changa A, Junga J, Maeda MM, Landivar J (2017) Crop height monitoring with digital imagery from Unmanned Aerial System (UAS). Comput Electron Agric 141:232–237
14. Zhang L, Grift TE (2012) A LIDAR-based crop height measurement system for Miscanthus giganteus. Comput Electr Agric 85:70–76

Food Assessment Model for Indian Elderly Persons Using CNN and Image Processing Techniques

Leena K. Gautam and Vijay S. Gulhane

Abstract Nutrient requirements differ from person to person and are influenced by a variety of factors like age, physical activity, diseases, etc. In Indian elderly persons, loss of apetite, isolation, depression, weak muscles are some of the important factors which not only lead to malnutrition but also raise the risk of getting affected with other major diseases results in the worsening of chronic conditions, frequent hospitalization causing burden to caregivers. In this paper, a user-friendly and intelligent system is proposed that can accurately classify and estimate the food region size of mixed/cooked food images, which can be further used to obtain calories and nutrient information utilizing the Indian Food Nutrient Dataset (IFND). The proposed system uses convolution neural network also known as CNN trained and tested on Indian Food Image Dataset (IFID) and integrated with refine Image processing techniques for better classification accuracy. Experimental results show that the adapted method leads to an accuracy of 97%.

Keywords Convolutional neural network · Refine image preprocessing techniques · Estimation of food region

Introduction

Food and nutrition play a crucial role in our daily lives. The amount of nutrients required in a certain food item changes depending on age, physical factors, and other disorders [1]. In elderly persons, isolation, depression, low immunity level, and weak muscles are added factors determining nutrients requirements [2]. The main issue among elderly people is a loss of appetite and low immunity, which not only leads to malnutrition but also increases the chance of contracting other serious

L. K. Gautam (✉) · V. S. Gulhane
Sipna College of Engineering and Technology, Amravati, Maharashtra, India
e-mail: kgautam@sipnaengg.ac.in

V. S. Gulhane
e-mail: vsgulhane@sipnaengg.ac.in

V. Bindhu et al. (eds.), *Proceedings of Third International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 844,
https://doi.org/10.1007/978-981-16-8862-1_73

diseases, worsening chronic conditions, and delaying recovery [3, 4]. The malnourished elderly are more likely to require hospitalizations, cause a burden to caregivers, and have a gigantic monetary expense to society [5]. To lead a healthy and active life, an efficient food assessment system is highly recommended which classifies and estimates nutrient requirements from the given food image. The method which can automatically assess, i.e., can classify and estimate the nutrient as well as calorie contents from a food image allowing a person to learn from what food contains and how healthy it can be. Many researchers [6–8] had already proved the accepted classification accuracy of the given food images using convolutional neural network (CNN) using existing databases, while others had performed calorie estimation of a given food using different tools and techniques [6, 9–14]. In all of these methods, raw food was segmented using simple segmentation algorithms, which were then assessed for content estimation using an existing, limited, and/or calorie annotated dataset featuring dishes belonging to that particular region. Some methods [8, 15] used pre-trained CNN models to classify the food image and a vector embedding technique to estimate food properties from the image. Estimation of attributes for its ingredients contents and nutritional values can be efficiently applied to the raw foods but not on mixed and cooked food. To reduce the computational time of convolutional neural network, region-based classification called R-CNN was proposed [16]. R-CNN classifies images using a selective region-based approach but for some data, it might consume more training time. The above mentioned strategies increase complexities and cause usability issues. Moreover the classification and estimation were performed using a preexisting monotonous database with some region-based added dishes. Indian cuisine on the other hand is diverse, consisting of different ingredients with different cooking styles for preparing the same food dish item. For example, "*kanda-poha*" is similar to "*Indori-poha*" and "*batata-poha*" with a few minor ingredient differences. Similarly, "*upma*", a popular Indian breakfast, is served differently in different states, with varying ingredient variations and cooking methods and the same is with "*idli*s". Estimation of these commonly cooked Indian food items requires accurate classification, which can certainly improve its estimation accuracy. Correct classification requires fine-tuned segmentation approaches, as well as effective training and testing of CNN on a set of food images. In this work, we propose a CNN-based food assessment system using refine preprocessing techniques for accurate classification and estimation.

The process begins by capturing an image of a food dish from both, front and side views using a pre-registered reference object and a specified distance, as shown in Fig. 1. To measure the real size of the food: its area, and depth, commonly used pre-registered reference objects such as a coin, a cube, a card, and so on can be used. The reference object and the dish regions are isolating and segmented. After segmentation, the top view of the food shot is utilized to classify it using CNN, which has been trained and tested on a variety of Indian food images to ensure accurate classification. Real size area and depth of the food dish are acquired from the top and side views, respectively, with the help of the registered size of the reference object, and are then examined for calorie and nutrient estimation.

Fig. 1 Top view and side view images with reference objects

This paper is organized as follows: The second section delves into the relevant research in this field. System overview, which includes datasets, preprocessing techniques, CNN training, and testing with the nutrient estimation approach, is described in Sect. 3. Section 4 depicts the preprocessing techniques and CNN model in action with results obtained followed by a conclusion in Sect. 5.

Related Work

Classification and estimation are the two main categories of dietary assessment. Author Attokaren et al. [7] proposed a CNN-based food classifier model utilizing IMAGENET and FOOD-101 datasets. To ensure that any image obtained from any perspective could be categorized, basic preprocessing techniques were used. When compared to other less accurate methodologies, the approach yields roughly 87% classification accuracy. Calorie estimation, on the other hand, was done by scratching the web. User-friendly software that uses CNN to classify food with an accuracy of 85% was proposed by author Yunus et al. [8]. For real-time attribute and ingredient estimate, vector space embedding was utilized, which semantically links the nearby food items. An expansion of the FOOD 101 dataset to include subcontinental cuisine was one of the paper's significant contributions. Similar concept was adopted by author Shen et al. [15]. The key contribution of this work is to improve the accuracy of CNN by fine tuning pre-trained inception v3 and v4 models resulting in classification accuracy of 85%. Relating a portion of the food with geometrical shapes and computation of their sizes with reference objects was proposed by some of the researchers. Author Khanna et al. [10] proposed a geometrical shape-based estimation approach. The volume of the food dish was calculated by fitting it into some geometrical shapes. The strategy was effective for most of the food items, with an error rate of 10–11%, but for specific things, such as vegetables, salads, and bread slices, the error rate increases tremendously. Author Pallavi Kuhad et al. enhanced the work done by author Parisa et al. [11] (where SVM was used for classification) and presented a fully automatic and user-friendly calibration of the dimension of portion sizes using deep learning [12]. The collected image is submitted to the cloud for classification and estimation (using the distance calculation technique). Deep learning was used to classify the images, and the calorie details matching the image were obtained from a cloud-based database. The technique works well for single food items, but mixed foods and dissolving components remain an issue. Using a smartphone with no external recognition servers, Author Napoletano et al. [6] proposed an automatic calorie estimation approach. GrabCut segmentation and CNN were used to segment and recognize a photo of a food image and a reference object. Calorie estimation was performed using the size obtained from the top view food image only. Quadratic curve estimation from the 2D size of foods to their calories was employed, rather than simple linear estimation. Each food category's quadratic curve is computed independently based on training data annotated with real food calories giving a 21.3% error rate. By applying segmentation masks to the food images, author Ege et al. [17] approximated the nutrients using a grain of rice as a reference object (which is found in practically every Japanese cuisine). A new food image dataset with bounding boxes and segmentation masks was added to the traditional UEC-100 dataset. In yet another work author Liu et al. [16] proposed food volume estimation of irregular food based on the reference object. To precisely predict the volume of the food, R-CNN is used in conjunction with an optimization technique (a combination of GrabCut and edge filtering methods). The experiment was conducted on five different types of fruits, yielding a 4.5% mean absolute inaccuracy. Author Siriwan and Sunitha proposed improved Mobilenet architecture for food classification [18]. Instead of average pooling, Global average pooling was used. Moreover, fully connected layer was replaced by batch normalization, ReLu, and dropout layers to improve performance. When compared to a typical mobilenet design (46%), the results reveal a 67% higher accuracy.

System Overview

Proposed System Flow

Figure 2 depicts the proposed system's general design as well as its flow. As a first of this process, the user registers the reference object with its dimensions, such as a wallet, a cube, or an expired debit card. For safety reasons one can also use the thumb, a coin, etc. This is a one-time registration that is used to calculate the food dish region size. Users must click two photos: a top view and a side view with a reference object at a specified distance in order to receive the calorie or nutrient information for a specific food dish image. Both pictures will be subjected to image preprocessing techniques such as edge detection and GrabCut segmentation, as described in Sect. 3.3. After that extracted foreground top view image is passed to a convolution neural network for recognition. In the meantime, the weight is calculated by analyzing the photos of the food region and the reference object. The identified image is then used to extract nutrient and calorie information from their calculated sizes.

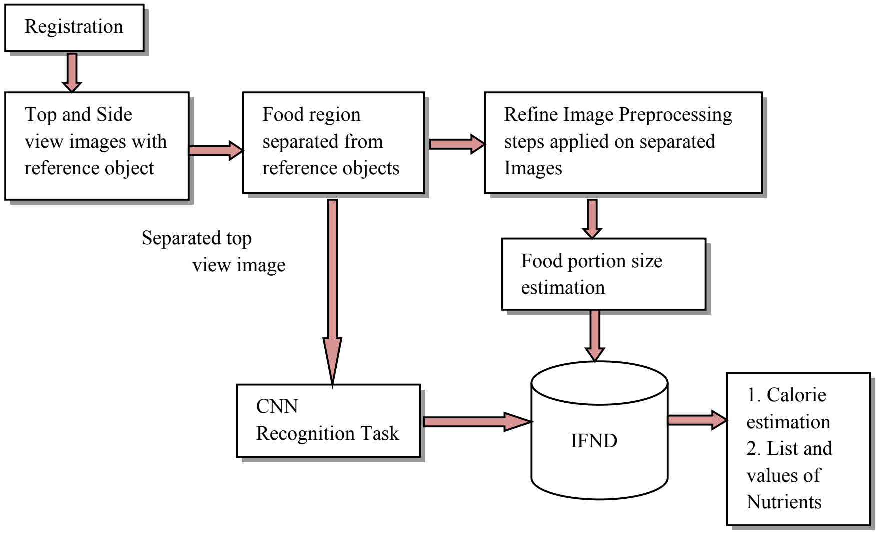


Fig. 2 Proposed system flow chart

Datasets

Nutrients are food-based chemicals that provide humans with energy, the building blocks for repair and growth, and substances that help control chemical processes. When the food is absorbed or digested it releases an amount of energy called calories. Carbohydrates, Fats, Proteins, Calcium, Vitamins, Iron, Dietary Fiber, Sodium, and Potassium are among the necessary nutrients for the elderly [3]. These nutrients values along with their calories are stored in Indian Food Nutrient Dataset (IFND). For this work, a partial dataset was created manually containing nutrient information of different varieties of Indian snacks poha, upma, idlis, etc. from the recognized websites: Myfitnesspal [19], Chronometer [20]. In order to train and test CNN on food images, another dataset IFID was used. This dataset consists of varieties of Indian food dish images with their labels that are divided into snack, meal, dessert, fruit, and dairy categories. Total 155 images were collected. Some of the images were collected manually while most were collected and extracted online through BeautifulSoup [21]. A Python package which parses HTML and XML documents. The search words like [‘rava upma’, ‘kanda poha’, ‘oats idli’ …] are set and the path is specified where we want to download our images. The process is repeated for every word. Request is made by creating a URL of each word and then the images are extracted using the BeautifulSoup method. Text and watermarking present in images were removed, whereas unwanted and blur images were discarded manually. Sample images with their training labels are shown in Fig. 3.

Fig. 3 Sample images with their training labels

Refine Image Preprocessing Techniques

To make food image suitable for classification and food region size estimation following preprocessing techniques are applied.

Edge detection from food dish image. In the image processing framework, various edge detection approaches have been widely used. In 1986, Canny [22] came up with the concept of detecting a wide variety of edges in images using a multi-stage algorithm. A Canny edge detector technique identifies a wide variety of edges in images using a Gaussian filter, suppression method, and threshold settings [22, 23]. The bounding box that surrounds the plate region of the meal is extracted using an edge detection algorithm applied to the top view and side view of the food image.

Segmentation. After detection of edges, we need to segment out the food dish region from the background using GrabCut segmentation technique [24], which pulls foreground images from a complicated background with minimal user interaction segmentation is performed by specifying bounding box coordinates of the plate region.

Convolutional Neural Network (CNN) Configuration

Yann LeCun [25] initially presented convolutional neural networks, often known as ConvNets or CNN, as the most widely used neural network model for image classification tasks. CNN mainly comprises of convolutional layers, pooling layers, subsampling layers and fully connected layer. We have implemented a CNN model which recognizes the name of the food image by extracting its color, size, and shape. For training CNN model images of "idlis" (ragi, rava, spinach, oats, etc.)were collected online. Unwanted labels, watermarking were removed and the images were down sampled to (65 * 65). Data augmentation with parameters rescale = 1./255, shear_range = 0.2, zoom_range = 0.2, and horizontal_flip = true were performed on the collected image set. Out of 306 total images 241 images were used for training and

Table 1 CNN model summary

Layer (type)	Output shape	Param #
conv2d (Conv2D)	(None, 65, 65, 8)	224
max_pooling2d (MaxPooling2D)	(None, 32, 32, 8)	0
dropout (Dropout)	(None, 32, 32, 8)	0
conv2d_1 (Conv2D)	(None, 32, 32, 16)	1168
max_pooling2d_1 (MaxPooling2)	(None, 16, 16, 16)	0
dropout_1 (Dropout)	(None, 16, 16, 16)	0
conv2d_2 (Conv2D)	(None, 16, 16, 32)	4640
max_pooling2d_2 (MaxPooling2)	(None, 8, 8, 32)	0
dropout_2 (Dropout)	(None, 8, 8, 32)	0
flatten (Flatten)	(None, 2048)	0
dense (Dense)	(None, 256)	524,544
dropout_3 (Dropout)	(None, 256)	0
dense_1 (Dense)	(None, 5)	1285

Total params: 531,861
Trainable params: 531,861
Non-trainable params: 0

rest for validation. Functions used for building sequential model: **Conv2D-** kernel size (3 * 3) to extract image features, padding = "same", input size (65 * 65 * 3), activation was performed using an effecient, non linear activation unit called "ReLu". **MaxPooling2D function**—with pool_size (2 * 2) to downsample the image by taking a maximum value over an input window, shifted by strides (2, 2) along each dimension. **Dropout**—to overcome the problem of overfitting, with a scale 0.35 is used to drop the hidden and visible units in neural network. To configure training process **Compile function**—with three arguments: optimizer with learning rate 0.001, loss function (Categorial_crossentropy),and list of metrics. **Flatten**—converting the data into 1D for inputting it to the following layer. **Dense layer**—with 256 neurons and relu as activation. **Softmax function**—used as an activation function in last dense layer to interpret the result as a probability distribution in range 0–1 and sum to 1. Table 1 describes the model summary.

Estimation of Food Region/portion Size

To calculate the weight of the food, we must first compute the area and the volume. To calculate accurate area the best way is to compare the top view image of the food region with geometric shape say circle, square, etc. In most of the cases we will get

Fig. 4 **a** Top view. **b** Side view

circle or ellipse as a similar geometrical shape when compared. By comparing the number of pixels of the food region with that of a reference object image whose size we already know we can get dimensions. Area can then be calculated by applying the specific formula, here in this case shown in Fig. 4a. formula of circle is applied A = 3.14 * r * r. Volume is then obtained by calculating the depth from the food region of the side view image Fig. 4b and multiplying it by the obtained area.

$$V = A * D \tag{1}$$

where V is the volume, A is the area, and D stands for depth.

To obtain calories we need to figure out the mass of a food item which is a result of volume and the density of food

$$M = V * F_{\text{den}} \tag{2}$$

where M is mass, V is volume, and F_{den} is the density of the food. The density of food can be fetched from INFOODS density database [26]. The dataset includes density and bulk density of around 638 food entries in 20 food groups. When it comes to the food measurement, mass and weights are similar, i.e., we got mass, we got weight. The name of the food with its weight can be searched in IFND to obtain required calories and nutrients.

Experiments and Results

This section summarises the findings and observations made during the experiment.

Extracting Food Region

The food dish region present in an image is obtained by applying canny edge detection method. The method identifies points in a digital image where the image brightness alteration dramatically demonstrates discontinuities, i.e., brightness that is discontinuous. As a first step the input image is converted to grayscale and by applying gaussian kernel (5 * 5) the noise is reduced. The intensity of edges are obtained by

Fig. 5 Top view image **a** edge detection **b** segmentation

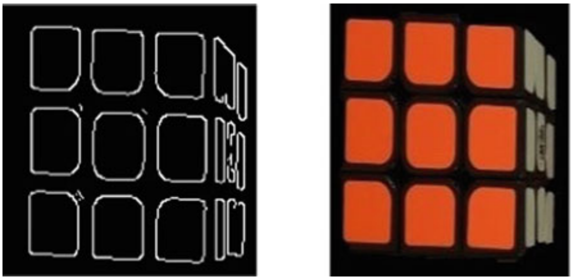


Fig. 6 Reference object image **a** edge detection **b** segmentation

gradient calculations, edges suppression method and, threshold. Tracking and transformation of weak edges into strong is done by hysteresis method. The obtained dish region is then inputted for GrabCut segmentation method which works in an iterative fashion and extracts food region directly. The result obtained after edge detection and GrabCut segmentation is shown in Fig. 5.

Extracting Reference Object

In our method, we had used a calibration object for calculating real size of the food region. In this experiment, cube is used as a reference object whose dimensions are already stored. The method discuss in Sect. 4.1 was applied for obtaining the final reference object as shown in Fig. 6.

CNN Implementation

As an initial level of experiment out of 306 food images 224 images were used for training and rest for validation. Three convolutional layers and one fully connected neural network layer make up the model. The last layer's output is coupled to the softmax classifier, which generates the probability distribution of 5 classes. The first layer filters 65 * 65 * 3 input image by 8 kernels of size 3 * 3.After applying pooling, and dropout the input moves to the second layer and filter the input by 16 kernels

Table 2 Result showing obtained accuracy

Dataset	%Accuracy CNN Model	%Accuracy Preprocessing + CNN
IFID	93	97
FOOD-101	87	89

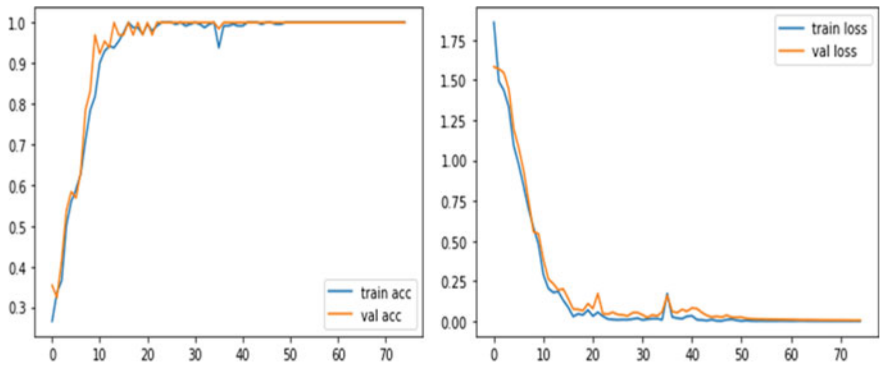


Fig. 7 **a** Training/validation accuracy. **b** Training/validation loss

of size 3 * 3. After applying pooling, dropout and normalization techniques the output of second layer is fed to third convolutional layer with 32 kernels of size 3 * 3.The output undergoes hidden layers consisting of 256 neurons. ReLu is used as an activation function for each of the above mentioned layers. Dropout is used in all the 3 layers of neural network to reduce the complex co-adaptations of neurons. The model is compiled with Adam optimizer having learning rate of 0.001. The model is trained using stochastic gradient method with batch size of 16 and epochs 20 giving around 97% validation accuracy. The present system is also tested on images from FOOD-101 dataset. The model classification accuracy increases from 87 to 89% with refine preprocessing techniques. Table 2 shows the results obtained before and after preprocessing techniques and utilizing both IFID and FOOD-101 datasets and Fig. 7 plots the accuracy/loss.

Conclusion and Future Work

In this paper, we proposed a user-friendly dietary assessment model designed specifically for Indian elders that uses refine image processing techniques to classify images with remarkable accuracy and also predicts the food region size by calculating area and volume, which can then be used to obtain calories and nutrient information from IFND. To enchance the classification accuracy of CNN images were preprocessed with edge detection and GrabCut segmentation techniques. We focused on the recognition task of cooked and mixed foods and our results indicates 97% of accuracy which is highly acceptable.

Limitation observed in our proposed system is in exceptional circumstances, i.e., when the color of plate region and food region are same, edge detection fail to distinguish the edges adequately. In such cases an object detection step might be included before performing edge detection to improve segmentation results. Future work includes implementation as well as expanding the data set to encompass a wider range of Indian food images in each class.

References

1. Gautam LK, Ladhake SA (2018) A mathematical AI-based diet analysis and transformation model. In: Advances in intelligent systems and computing, Singapore. https://doi.org/10.1007/978-981-10-8533-8_1
2. ASN-staff (2020) The role of nutrition in supporting the ımmune system relative to coronavirus (COVID-19). https://nutrition.org/the-role-of-nutrition-in-supporting-the-immune-system-relative-to-coronavirus-covid-19
3. Krishnasamy K (2011) National ınstitute of nutrition. NIN, Hyderabad
4. Taweel A, Barakat L, Miles S et al (2016) A service-based system formalnutrition prevention and self-management. In: Computer science and interfaces, pp 225–233. https://doi.org/10.1016/j.csi.2016.03.005
5. Ramya MS, Ranganath TS, Jadhav J et al (2017) To assess the nutritional status among elderly and factors influencing it, in an Urban Area. Int J Commun Med Public Health 1727–1732. https://doi.org/10.18203/2394-6040.ijcmph20171792
6. Napoletano P, Schettini R, Ciocca G (2018) CNN-based features for retrieval and classification of food images. Comput Vis Image Underst 70–77
7. Attokaren D, Murthy S, Kulagudi S (2017) Food classification from images using convolutional neural networks. In TENCON, Malasyia, pp 2801–2806. https://doi.org/10.1109/TENCON.2017.8228338
8. Yunus R, Ari O, Hammad A et al (2019) A framework to estimate the nutritional value of food in real time using deep learning. IEEE Access 7:2643–2652. https://doi.org/10.1109/ACCESS.2018.2879117
9. Chen M, Yang Y et al (2012) Automatic Chinese food identification and quantity estimation. In: SIGGRAPH Asia Technical Briefs, ACM, pp 1–4. https://doi.org/10.1145/2407746.2407775
10. Khanna N, Carol B, Chang et al (2013) Image-based food volume estimation. CEA-13, pp 75–80. https://doi.org/10.1145/2506023.2506037
11. Puladzadeh P, Shirmohammadi S (2014) Measuring calorie and nutrition from food image. IEEE Trans Instrum Measur 63
12. Kuhad P, Shirmohammadi S, Yassine A (2015) Using distance estimation and deep learning to simplify calibration in food calorie measurement. IEEE Trans. https://doi.org/10.1109/CIVEMSA.2015.7158594
13. Shehzad A, Zafar N, Hassan M, Shen Z (2019) Food ıtem recognition and ıntake measurement techniques. In: Proceedings of the 2019 11th ınternational conference on machine learning and computing, Zhuhai, China, pp 405–411
14. Yingnan S, Jianing Q, Benny L et al (2020) Image-based food classification and volume estimation for food assessment: a review. IEEE J Biomed Health Inform
15. Shehzad A, Chen S, Sun H, Liu J, Shen Z (2020) Machine learning based approach on food recognition and nutrition estimation. In: Elevier B.V, pp 448–453

16. Lai J, Sun W, Liu Y et al (2020) Food volume estimation based on reference. In: 4th international conference on ınnovations in artificial ıntelligence, China, pp 84–89
17. Shimoda W, Yana K, Ege T (2019) A new large-scale food ımage segmentation dataset and ıts application to food calorie estimation based on the grain of rice. In: 5th ınternational workshop on multimedia assisted dietary management (MADiMa'19), Nice, pp 82–87
18. Surinta O, Phiphiphatphaisit S (2020) Food ımage classification with ımproved mobilenet architecture and data augmentation. In: Proceedings of the 2020 the 3rd ınternational conference on ınformation science and system, Cambridge, United Kingdom, pp 51–56
19. MyFitnessPal 2021. http://www.myfitnesspal.com
20. Cronometer 2020 http://www.cronometer.com
21. Richardson L (2020) Crummy.com. https://www.crummy.com/software/BeautifulSoup/bs4/doc/
22. Canny J (1986) A computational approach to edge detection. IEEE Trans Pattern Anal Mach Intell PAMI 8(6):679698
23. Katiyar S, Arun V (2012) Comparative analysis of common edge detection techniques in context of object extraction. IEEE Trans Geosci Remote Sens 50:68–78
24. Rother C, Kalmogorov V, Blake A (2004) GrabCut: interactive foreground extraction using iterated graphcuts. In: ACM SIGGRAPH, pp 309–314. https://doi.org/10.1145/1015706.1015720
25. Sutskever I, Hinton G, Krizhevsky A (2012) Imagenet classification with deep convolutional neural networks. Adv Neural Inf Process Syst 1097–1105
26. Haytowitz D, Stadlmayr B, Charrondiere UR (2012) Food and agriculture organization of United States. http://www.fao.org.

A Review on IOT Response to COVID-19 Pandemic, Challenges and Open Issues

Bryan David Julies and Tranos Zuva

Abstract In the history of humanity, we have never faced an invisible enemy like the COVID-19 pandemic. The medical field has never been overwhelmed so much. Scientists and engineers are working around the clock to develop a suitable vaccine for COVID-19. In order to fight this pandemic technology will need to be utilized to its full potential. A fourth industrial revolution (4IR) technology, called the Internet of things (IOT), is an interconnection of physical devices and the Internet, has been identified to fight the battle against COVID-19. Numerous papers have been written about the Internet of things technology and how the technology can be use in different environments, thereof it is imperative to review how this technology can be utilize in the the response to COVID-19 pandemic. In this paper, we examined literature on the coronavirus (SARS-COV-2) and IOT technologies that can fight COVID-19 to minimize its spread. Various challenges and open issues related to the used of the technology in the fight of COVID-19 where identified and discussed, such as security and privacy issue, limited spectrum and bandwidth, scalability and interoperability as a threat. This technology can be recommended for use in a pandemic period.

Keywords COVID-19 · Internet of things · Smart home technology · eHealth · IOT Smart Hospitals · Privacy and security · Trust · IOT Smart City · Scalability

Introduction

The world has been turned up-side-down with the outbreak of the coronavirus (SARS-COV-2) in December 2019. All through the ages, humankind has encountered numerous pandemics that demolished enormous quantities of the populace. As per ref. [1] these outbreaks includes the 1720 Marseilla pandemic, the global cholera

B. D. Julies (✉) · T. Zuva
Department of Information and Communication Technology, Vaal University of Technology, Johannesburg, South Africa

T. Zuva
e-mail: tranosz@vut.ac.za

V. Bindhu et al. (eds.), *Proceedings of Third International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 844,
https://doi.org/10.1007/978-981-16-8862-1_74

outbreak from 1817 to 1923, cholera pandemic that occurred in India between 1817 and 1823, the Spanish flu that happened from 1918 to 1919, another pandemic called the Asian flu in the year 1957, 1958 and 1968, the Hong Kong flu in 1968 and then lastly Swine flu (H1N1) that occurred in 2009. None of these previous pandemic compares to the coronavirus (SARS-COV-2) which occurred at the end of 2019 in the city of Wuhan of the Hubei Province in China. The coronavirus (SARS-COV-2) has been named COVID-19 (AAP FactCheck 2020) and has spread rapidly around the world. On March 11 2020, the World Health Organization (WHO) has declared the virus a pandemic. The declaration was made in less than 3 months after the first detection [1]. The United States of America (USA) was hit hard by the pandemic as by middle March, all 50 states reported COVID-19 infections and suddenly humanity found itself with a momentous treat that caused the death of over seven million people worldwide in a couple of months. The coronavirus infection lead to human passing as well as set mankind in an uncommon basic circumstance that influenced all circles of life [2].

In order to regulate the transmission of COVID-19, lockdowns, isolation and remain at home orders have been given by countries around the world, which had a devastating effect on the economy [3]. According to ref. [3] restraints on industries, businesses and travel limitation will keep on having a drawn out effect on global supply chains and economy. As to date, there is no vaccine to cure the pandemic and for this reason humanity's safeguard to COVID-19 is protective measures like personal protective equipment (PPE), self distancing, sanitizing of surfaces, wearing of masks, self-confinement, restrictions on travel, lockdowns, quarantine, large crowd limitations and closures of schools.

As per ref. [4] the Internet of things (IOT) has different factors to help nations to reduce COVID-19 affections as ref. [5] declares that IOT has a wide scope of uses which would be successful to ensure that all the rules of security and safety measures given by World Health Organization (WHO) are followed.

According to ref. [3] the Internet of things (IOT) is an intelligent network used interchanged information and communication via sensors with protocols alignment. Well as per ref. [6] the Internet of things (IOT) comprises of complicated smart devices which transfer data over the Internet and has redesigned the world articles into shrewd virtual items. The goal of IOT can be establish to unite all smart device around the world and keep the users up-to-date with information [7].

In other words, IOT is a network of gadgets, appliances and "electronic things" that exchange data and uses low cost Internet connected devices and sensors [8].

The Internet of things (IOT) can be defined as a network of physical objects and can be classified into three things:

- People-to-People.
- People-to-Things.
- Things-to-Things.

IOT forms part of radio frequency identification (RFID) and has evolved with the increase of mobile devices, ubiquitous communication, cloud computing and data analytics [9]. Other protocols and technologies of IOT includes, Wi-Fi, Bluetooth, ZigBee and, 5G cellular, NFC and Neul. The term Internet of things (IOT) was firstly use by Kevin Ashton in 1999 to describe how objects communicates by the use of sensors and coin radio frequency identification (RFID) [9]. In 2003, IOT became prominent in related market examination and its distributions. The advantage of IOT is that the technology is tied with smart objects via a network without any human intervention [10]. If any device has access to the Internet it can be monitored or data can be exchanged and has potential to be an IOT device [11]. With the approach of IOT idea, various organizations zeroed in it and attempted to perceive its noteworthiness and started to distinguish its job and the connected future perspectives [12]. IOT has taken a huge uptake over the last few years and a great number of IOT devices and smart sensors have been presented.

Based on the user need and requirements there is a number of smart devices on the market which includes HDMI sticks, smart cameras, smart watches, smart light bulbs, temperature-sensors, smart health like fitness trackers and blood pressure monitors [3].

Reference [13] states that with the IOT market at $267 million per year, it still forecast to grow to $520 billion by 2021 and 100 billion devices will be connected to the Internet by 2025. Based on these statistics, ref. [3] states that smart devices can possibly be a significant achievement in efforts to control and fight the COVID-19 virus.

The main objective of the IOT is for things to communicate anytime, anywhere with anything. This has the guarantee of better use of accessible assets, decrease cost, and limiting manual interaction. For this reason it is absolutely critical to discuss IOT's role as the COVID-19 virus is spreading across the globe.

Reference [12] claims that IOT has an immense scope of uses which would be successful to ensure that all the rules of well-being and safeguards gave by World Health Organization (WHO) are followed. As declared by ref. [4], the IOT network is very scalable which can possibly manage a enormous amount of information received from IOT device sensors to help the fight against COVID-19. Reference [14] claims that the COVID-19 pandemic present the worst global health crisis since the influenza pandemic in 1918. Reference [14] also continue to say that the statics provide by the World Health Organization (WHO) states that on the July 19 2020, the infection rate passed 14 million people and the death rate stood at 540 thousand people [3].

COVID-19 was declared as a pandemic on March 11 2020 by the World Health Organization (WHO). The COVID-19 virus was first discovered in China, Wuhan City in Hubei Province, at the end of 2019. The sickness has comparable side effects with flu, for example, fever, cough, weariness which are basic to know to analyze at the beginning phase (C. for disease control and prevention). The virus is transmitted by salvia and nasal droplets which are transfered through either coughing or sneezing, affecting anyone less than 6 feet from the infected person. The virus can also be transmitted by contaminated surfaces (example door handles, railings of escalator and public restrooms) and high risk areas. Mild side effects of COVID-19 includes fever, coughing and breath shortness, well more severe symptoms are organ failure. Other newly discovered symptoms, includes chills and fever, muscle ache, headache, inflamed throat, loss of smell and taste [3]. It also possible that a person without any symptoms can be a transmitter of the virus to others. When a person is in isolation, the incubation period is 1–14 days.

Reference [12] define that the coronavirus is 600–1400A in diameter and has proteins spikes bulging from its surface. Reference [14], states that in the past two decades there has been several forms of coronavirus outbreaks, which includes 2002–2004 SARS-CoV outbreak, MERS-CoV infection in 2012, the SARS-CoV pandemic in China's Guandong province. As per (WHO) MERS-CoV was first recognized in Saudi Arabia which had 871 casualties.

To curb the COVID-19 spread, global lockdowns, isolation as well as stay at home mandate has been rolled out by countries around the world, and had devastating effect on the economy [3]. Other measures and mitigation strategies against COVID-19 includes wearing of masks when going in public places, washing of hands with water and soap or hand sanitizer, keeping a social distance (2 m) from people, avoid touching your face and public surfaces. By touching contaminated surfaces, the virus is spread to other areas by your hands. Infections occurs when a person touches his/her face and the virus is transferred through mouth, nose or eyes into the lungs.

As the virus spread with a rapid pace, IOT has a reliable network that decreases the delivery time for essential data that can help safe human lives.

It is important to examine the adequacy of IOT in battling the worldwide pandemic of COVID-19 so that in future it can be utilized approriately to combat future pandenics (Table 1).

Table 1 Attainments of IoT approaches in handling COVID-19 [15]

	Attainments	Description
1	Nusing of COVID-19 patient	Check patients health stats
2	Smart Hospitals	IOT using dedicated network
3	Archieving the data of COVID-19 patient	Storage of COVID-19 data
4	Originate multiple sources and devices	IOT uses multiple sources for analyzing
5	Scientific decision making	Use for accurate and quality decisions
6	Observing levels of COVID-19 patient	Helps monitoring patients health levels
7	COVID-19 disease alerts	Real-time tracking of COVID-19
8	Statictics to the healthcare assistant	Provide data to healthcare workers
9	Accepted medication	Monitors diet and medications
10	Acceptable resources	Exchange data to healthcare workers
11	Inspect levels of glucose	Checks patients glucose
12	Help in isolated areas	Monitor patiets in remote locations
13	Recognition of an asthma skirmish	Identifies asthma symtoms
14	Reminds to a take medication on time	Rememder patients to take medication
15	Emergency instances	IOT analyze patients data upon hostipal arrival
16	Smart bedding	Adjust the bed to patients preferrences

In this paper, we will look at different applications and technology systems that mitigates the transmission of COVID-19.

This paper is arrange in two sections: Sect. 2 looks at the mitigations against COVID-19 like IOT eHealth, IOT Smart Hospital, Remote Patient monitoring, IOT Smart Home, Contact Tracing, IOT Smart City, Smart Tracking, Smart Testing Vehicles, etc. Section III provides the challenges and open issues with IOT. In this section, we look at security and privacy issue, limited spectrum and bandwidth, scalability and interoperability as a threat.

Mitigations Against COVID-19

As COVID-19 was spreading rapidly, there was a need to find ways to stop the spread and different researches were doing research to find out how the spread can be reduced or stopped. Therefore, there were many areas that suggested in literature, for example, IOT eHealth, IOT Smart Hospital, Remote Patient monitoring, IOT Smart Home, Contact Tracing, IOT Smart City, Smart Tracking, Smart Testing Vehicles, etc.

IOT eHealth

As per ref. [16], eHealth is a practice that is backed by electronic processes and communication. According to ref. [17], eHealth not only addresses "Internet medicine" but includes "virtually everything related to computers and medicine". By 2022, according to ref. [3], the Internet of Medical Things (IOMT) industry is anticipated to flourish to 136.8 billion and goes on to say that currently, 3.7 million smart devices are used to monitor patient's vitals by data-driven applications. eHealth can be split on three scenarios or use-cases.

- Smart Hospitals.
- Remote Patient monitoring.
- Test booth.

According to ref. [3] all of their use-cases make use of smart sensors, smart devices, robotics, patients, doctors, employees, etc. and IOT application data driven services.

Clinical staff related with ambulances are normally managing high strain and blunder circumstances [18] and with the outbreak of the COVID-19 pandemic the pressure has mounted even more on medical staff as they need to deal with pandemic. As per ref. [4] IOT-supported ambulances offer a compelling arrangement in which far off clinical specialists recommend important activities to the clinical staff managing the patient in the emergency vehicle as it prompts a convenient reaction and successful treatment of patients.

IOT Smart Hospitals

The remarkable possible development and expansion in abilities of new advances in medication, the quick IT advancement, improvement and execution of the thoughts of the fourth Industrial Era in the course of recent years have denoted the possibility for a arrangement of a totally extraordinary medical care model.

According to ref. [11], the deputy Prime Minster of the Russian Federation O. Yu Golodets once stated that demographic development and health concerns is any government's preferences. Reference [19] declares that the Smart Hospital model alters all layers of health organization administration and re-designs the health organization architecture. As hospitals are overwhelmed with the pandemic crisis like influx of sick people, shortage of hospital bed, PPE shortage, smart hospital can be implemented to regulate actions for COVID-19 patients. This regulations can be actioned by the means of RFID sensors installation to track stockpile masks, face shields, gauzes, disposable patient examination papers, boxes of gloves and plastic bottles and vials [3]. Monitoring and tracking an individual's vital signs, torso-temperature, pulse rate and oxygen saturation levels, smart bracelets with installed RFID sensor, is provided to each patient. This allows doctors to monitor the patient on a regular base by means of remote applications without direct contact with the patient.

Remote Patient Monitoring Through IOT Smart Devices

According to ref. [20] distant well-being checking is one of the best innovation took advantage of in clinical experts to keep a beware of the patient's significant well-being factors intermittently. As per ref. [21] in universal processing climate, administration dependability is normal by the client and services are a fundamental precondition. There are cases when patients are tested positive and do not need to admitted to a hospital but can self-quarantine at home. With these cases the doctors can monitor these patients remotely by checking their body temperature, BP levels, pulse rate, coughing patterns EGG, etc. As per ref. [20] wearable sensors can be use to monitor and analysis an individual's themoregulation process through the achievements of 6LoWPAN and RPL IoT for healthcare applications. These patient's data can be extracted from a smart bracelets with RFID sensor that uploads the data to a central cloud. The doctor can then access this data via their smart phones [3].

IOT Smart Home

According to [3], there has been various studies done on health within the IOT smart homes and how to keep the home clean from deseases. The dawn of the current COVID-19 outbreak the most crucial step to combat the coronavirus is y sanitizing our homes. Reference [3] states that using voice control minimizes people to directly touch objects around the house. These objects includes smartphones, televisions remote controls, light switches, thermostats, doorknobs, etc.

Self-isolation requires an individual to quarantine to prevent the spread of the virus. Reference [3] states that with IOT Smart Homes, it possible for practitioners and family members to remotely monitor the infected individual. This is done through smart sensors that can track an individual's body specifications, like temperature and monitor the patients cough density using a cough detection application stated by [22].

Touching, is part of us as a human being, and in our everyday life, we touch door handles, sinks, cabinet handles, fridge handle, escalator railings and so forth. Through our touching affected surfaces, the COVID-19 virus is spread like wild fire. But this can be mitigated in smart homes with voice control devices. Other smart devices that can be use for home sanitization includes smart disinfect robots, smart UV-floor mats, UV-clean phone sanitize devices, smart UV light boxes, voice-control smart dustbins, voice-control hand sanitize pumps, voice control door knobs and cabinet handles, smart cameras, smart door-bells, etc. [3].

Home Confinement by Means of Smart Application Contact Tracing

Contact tracing is one of the tools in the COVID-19 toolbox to fight the pandemic in order to break the chain of infection. According to refs. [23, 24], in numerous countries like Hong Kong, Singapore, Australia and Unite Kingdom, the contact tracing app has become an essential tool to assist in the fight of COVID-19. Although the contact tracing app is an essential tool, it does come with concerns like privacy and security for location and home. A benefit of the tool is for people to come out of lockdown and able to move past the limits of their home. As ref. [23] at Hong Kong airport, upon arrival, individuals are issued with an electronic bracelet to invoke a two week stay at home quarantine. According to ref. [25] once the individual has setup the device, it links the smart device to his/her home and will detect any movements outside the house. Alerts are then triggered to inform police and health officials of the home violation which can lead to heavy fines, relocation of a new quarantine facility or even time in prison. Reference [9] claims that numerous cities allowed only one individual per household to go out for their household shopping. This occurred in cities like Hubei were more than 50 million people were force to be indoors under lock down.

Surveillance Through Digital Technology

According to ref. [23], digital technology, like Proptech, has been used over time to manage residential tenancies, but with the eruptionof COVID-19 pandemic this technology has accelerate to an extent that it is causing a threat to public health as justification. Under normal conditions surveillance technology would have been contested but because of COVID-19 certain practices and technologies are being approved without following the right protocols. Reference [14] therefore states because of this, power inequality between the landlord and their tenants are developed. One way that the virus is spread is through contaminated surfaces touched by individuals who are affected by the COVID-19 virus. Bioconnect cares is a touchless facial recognition access system, use to combat the spread of the virus and to prevent the second wave [14].

"The virus has the ability to infect through a single direct and non-direct touch, and without proper security, it is very easy for someone who has the virus to enter a building and infect hundreds of previously healthy people" [14].

As per ref. [14] vision "where a person's body and biometric identifiers serve to make their health status legible to large proptech firms and the landlords they serve".

In order to make this practical, it involves access to an individual's health records and integrate to with contact tracing application [14].

Studying and Working from Home

The affordances of home for study designs are significant, however frequently unacknowledged. As per ref. [26], the home has for some time been a work environment, not only for home-grown work crafted by social propagation, children and young adults. According to ref. [23] the current COVID-19 pandemic, the home cannot be seen as a private place as the pandemic has accelerated the use of new digital technologies to fight the new age pandemic. According to ref. [23], the year 2020 has saw working and study from home through a number of digital platforms like ZOOM with a 500% increase and Microsoft Teams saw a user increase of 70%. Reference [23] goes on to say that platforms like Examity and ProctorU specialises is distance-learning and digital exam supervision and due to COVID-19 most universities around the world was compelled to online classrooms.

IOT Smart City

According to ref. [27] a novel framework utilizing a huge measure of information assembled from different sensors and gadgets, for example, reconnaissance camera film and Visa exchanges of confirmed COVID-19 patients to reproduce their developments. As per ref. [28], the Voxel Physical Distancing Index (PDI) is a tracking system that tracks COVID-19 virus on social behavior by means of metric computer vision models and live video stream. The PDI assist major cities to capture the human activity and social distancing behaviors. Smart cities can have a rich exhibit of mechanical items that can aid early recognition of COVID-19 flare-ups by way of IOT sensors, and relieve contamination by way of physical distancing.

Smart Tracking

As per ref. [29], nations have utilized cellular information to follow residents developments during the pandemic. A smart city can be defined as an improved smart framework execution to assist with the coronavirus pandemic and saving human lives. Drones which is an unmanned aerial vehicle (UAV) has been very affective during the COVID-19 outbreak. According to ref. [30] droves have been update by Spain police to enforce lockdown and to announce messages like "Please respect the safety distance" through speakers attached to the drone. It has been reported that in the Handan Hebei province of China, these droves are also use to discharge disinfectant spray to fight the COVID-19 virus. Other useful features of drones is to single out those individuals how does not follow the pandemic protocols like wearing of mask, social distance keeping of six feet (two arms' length) and measuring individuals body temperature by means of thermal sensors [31].

Smart Testing Vehicle

As per ref. [3] in order to protect healthcare workers, speed testing stations, like drive-tru, are implemented in urban and rural areas. These smart testing vehicles incorporate infrared body temperature, oxygen level sensors, smart test kit, camera, a microphone, and local edge services and help lessens the exposure of those individuals with underlining conditions. The individual will enter the vehicle from one side which is separated by the glass-wall where the patient temperature will be taking by in-build sensors. Healthcare providers will issue the patient with self-test kids, which is return through the car window [3].

Challenges and Open Issues

Within these application of IOT in migitating against COVID-19, there were challenges and open issues identified in the literature such as security and privacy issue, limited spectrum and bandwidth, scalability and interoperability as a threat.

Security and Privacy Issues

As per ref. [23] because of adaptability and energy restrictions of IOT smart devices, the customary cryptographic procedures are not plausible answers to execute security in IOT. On the contrary, ref. [32] states that security arrangements should be energy efficient and calculations defined. This is in order to make sure that IOT networks should have less computational complexities to bring to the table to address data protection, consumer privacy and secure authentication. Reference [4] continues the states that lightweight security calculations should be planned to execute security in IOT. Due to the COVID-19 pandemic, the IOT security prerequisite has escalated. In order to combat the COVID-19 with IOT the following security concerns can be listed:

- the information which is sent from the sensors of COVID-19 patent should be precise
- the information ought to effectively arrive at the objective
- the information ought not be counterfeit
- the information ought not be caught from the correspondence way
- the in memory data of the IOT device ought not be available to everybody.

According to ref. [3] that in eHealth, when a COVID-19 patient gets tested, the individual's test data sent and stored on private cloud which is shared with medical care state or regional government to follow and check the person's movements. According to ref. [3], in order to protect an individual's identity and privacy differential privacy as per [33] and data masking techniques [15] can be used. According to refs. [34] and [35], these techniques are called pseudonymize and anonymize, respectively.

In a pseudonymize procedure, information are traced back into its unique state with high danger of trading off client protection, though it gets difficult to restore information into its unique state in anonymize [3]. Secure authentication is a critical aspect when it comes to security and intelligence control models for cloud IOT devices. Reference [36] states that there has been numerous studies on access control models which includes Amason Web Services (AWS) [37], Google Services [38] and Azure Services [39]. Customary security models arent sufficient in tending to dynamic and advancing access control prerequisites in IOT. As per refs. [4, 40] attribute based access control offers an adaptable and dynamic access-control model for a distributed IOT climate like IOT Smart Homes [41], smart vehicles [28] and wearable IOT devices [32, 42]. As per ref. [24] the data flow between different parts IOT cloud-enable platform must be safeguarded from illegal data access and conversions and for this reason attribute-based communication control has been proposed.

Limited Spectrum and Bandwidth

In today's age, a large number of IOT devices uses spectrum licences provided by mobile operators and as the quantity of IOT smart devices are growing, more bandwidth is needed to transfer data from the IOT smart device to the cloud [4]. As a result, data latency is bound to occur occasionally. As per ref. [4] numerous IOT smart devices makes use of limited spectrum (3G/LTE/4G) and will soon not be enough due to rapid growth in IOT devices. Latency between smart devices is therefore a huge concern during COVID-19 pandemic as this could cause loss of human lives.

Scalability

The biggest challenge in fighting the COVID-19 pandemic with IOT smart devices is scalability as a large number of IOT Smart Homes is needed to gather and analysis vital signs from patients [4]. Currently there are 3.7 million smart devices globally and each device has numerous sensors [4].

According to ref. [43] IOT devices have grown rapidly and the reasoning behind this is that smart devices are not restricted to just a single application however there are numerous utilizations of IOT smart devices in use. Network parameters that can influence the ease of use of IOT smart devices that can be listed as high latency, response time, packet loss and low bandwidth as smart device produce huge volumes of information which brings about a flood in information traffic that causes network blockage and dormancy [3].

Interoperability as a Threat

Interoperability is another huge challenge of IOT devices as smart devices are adaptable with producers of the same vendor and for this reason the need of competition among merchants is essential [4]. According to ref. [3] interoperability has been a challenge from the start of IOT smart device distribution. Reference [4] states as vendors competes among for quality, it increases IOT operations and evolution it contribute to heterogeneity among IOT devices.

Conclusion

The COVID-19 pandemic has forces humanity to live and adapt to a "new normal". The pandemic has spread rapidly around the world and millions of human lives have been affected. In order to combat the COVID-19 virus, IOT assumes a significant part in lessening the dangers of the COVID-19 virus spread by implementing protocols prescribe by the World Health Organization (WHO). The IOT network is very scalable which can possibly manage an enormous amount of information received from IOT device sensors to help the fight against COVID-19. As the virus spread with a rapid pace, IOT has a reliable network that decrease the delivery time for essential data that can help safe human lives.

Acknowledgement This work is supported by the Vaal University of technology under the supervision of Professor Tranos Zuva in Johannesburg (Vanderbijlpark) South Africa.

References

1. Budd J et al (2020) Digital technologies in the public-health response to COVID-19. Nat Med. https://doi.org/10.1038/s41591-020-1011-4
2. Abusaada H, Elshater A (2020) COVID-19 challenge, information technologies, and smart cities: considerations for well-being. Int J Commun Well-Being. https://doi.org/10.1007/s42413-020-00068-5
3. Gupta D, Bhatt S, Gupta M, Tosun AS (2020) Future smart connected communities to fight COVID-19 outbreak. arXiv. 2020
4. Kamal M, Aljohani A, Alanazi E (2020) IOT meets COVID-19: status, challenges, and opportunities. arXiv. 2020
5. Truta TM, Vinay B (2006) Privacy protection: P-Sensitive k-Anonymity property. In: ICDEW 2006—proceedings of the 22nd ınternational conference on data engineering workshops. https://doi.org/10.1109/ICDEW.2006.116
6. Srinivasan CR, Rajesh B, Saikalyan P, Premsagar K, Yadav ES (2019) A review on the different types of internet of things (IoT). J Adv Res Dyn Control Syst
7. Panarello A, Tapas N, Merlino G, Longo F, Puliafito A (2018) Blockchain and IoT integration: a systematic survey. Sensors (Switzerland). https://doi.org/10.3390/s18082575
8. Bajaj RK, Rao M, Agrawal H (2018) Internet of Things (IoT) in the smart automotive sector?: a review. J Comput Eng
9. Rose L, Eldridge K, Chapin S (1987) The internet of things: an overview—understanding the issues and challenges of a more connected world. Mod Plast 64(9):54–55
10. Ali ZH, Ali HA, Badawy MM (2015) Internet of Things (IoT): definitions, challenges and recent research directions. Int J Comput Appl. https://doi.org/10.5120/ijca2015906430
11. HaddadPajouh H, Dehghantanha A, Parizi RM, Aledhari M, Karimipour H (2019) A survey on internet of things security: requirements, challenges, and solutions. Internet of Things. https://doi.org/10.1016/j.iot.2019.100129
12. Oyeniyi J, Ogundoyin I, Oyeniran O, Omotosho L (2020) Application of internet of things (IoT) to enhance the fight against Covid-19 pandemic. Int J Multidiscip Sci Adv Technol 1(3):38–42
13. Axisa F, Schmitt PM, Gehin C, Delhomme G, McAdams E, Dittmar A (2005) Flexible technologies and smart clothing for citizen medicine, home healthcare, and disease prevention. IEEE Trans Inf Technol Biomed 9(3):325–336. https://doi.org/10.1109/TITB.2005.854505
14. Nasajpour M, Pouriyeh S, Parizi RM, Dorodchi M, Valero M, Arabnia HR (2020) Internet of Things for current COVID-19 and future pandemics: an exploratory study. J Healthc Inf Res 4(4):325–364. https://doi.org/10.1007/s41666-020-00080-6
15. Vaishya R, Javaid M, Khan IH, Haleem A (2020) Artificial intelligence (AI) applications for COVID-19 pandemic. Diabetes Metab Syndr Clin Res Rev. https://doi.org/10.1016/j.dsx.2020.04.012

16. Della Mea V (2001) What is e-health (2): the death of telemedicine? J Med Internet Res. https://doi.org/10.2196/jmir.3.2.e22
17. Eysenbach G (2001) What is e-health? J Med Internet Res. https://doi.org/10.2196/jmir.3.2.e20
18. Park E, Kim JH, Nam HS, Chang HJ (2018) Requirement analysis and implementation of smart emergency medical services. IEEE Access. https://doi.org/10.1109/ACCESS.2018.2861711
19. Ilin IV, Levina AI, Iliashenko OY (2017) Reengineering of high-tech and specialized Medical care delivery process for telemedicine system implementation. In: Proceedings of the 29th international business information management association conference—education excellence and ınnovation management through vision 2020: from regional development sustainability to global economic growth
20. Hariharakrishnan J, Bhalaji N (2021) Adaptability analysis of 6LoWPAN and RPL for healthcare applications of Internet-of-Things. J ISMAC 2(2):69–81. https://doi.org/10.36548/jismac.2021.2.001
21. Suma V (2021) Wearable IoT based distributed framework for ubiquitous computing. J Ubiquitous Comput Commun Technol 3(1):23–32. https://doi.org/10.36548/jucct.2021.1.003
22. Zhang L, Alharbe N, Atkins AS (2017) An IoT application for ınventory management with a self-adaptive decision model. In: Proceedings—2016 IEEE international conference on internet of things; IEEE green computing and communications; IEEE cyber, physical, and social computing; IEEE smart data, iThings-GreenCom-CPSCom-Smart Data 2016. https://doi.org/10.1109/iThings-GreenCom-CPSCom-SmartData.2016.77
23. Maalsen S, Dowling R (2020) Covid-19 and the accelerating smart home. Big Data Soc. https://doi.org/10.1177/2053951720938073
24. Bhatt S, Sandhu R (2020) ABAC-CC: attribute-based access control and communication control for internet of things. In: Proceedings of ACM symposium on access control models and technologies, SACMAT. https://doi.org/10.1145/3381991.3395618
25. Florida R, Rodríguez-Pose A, Storper M (2021) Cities in a post-COVID world. Urban Stud, p 004209802110180. https://doi.org/10.1177/00420980211018072
26. Blunt A, Dowling R (2006) Home
27. Raad O, Makdessi M, Mohamad Y, Damaj I (2018) SysMART indoor services: a system of smart and connected supermarkets. Can Conf Electr Comput Eng. https://doi.org/10.1109/CCECE.2018.8447626
28. Gupta M, Benson J, Patwa F, Sandhu R (2019) Dynamic groups and attribute-based access control for next-generation smart cars. In: CODASPY 2019—Proceedings of the 9th ACM conference on data and application security and privacy. https://doi.org/10.1145/3292006.3300048
29. Gupta M, Abdelsalam M, Mittal S (2020) Enabling and enforcing social distancing measures using smart city and ITS ınfrastructures: a COVID-19 use case. arXiv. 2020
30. Li X, Lu R, Liang X, Shen X, Chen J, Lin X (2011) Smart community: an internet of things application. IEEE Commun Mag. https://doi.org/10.1109/MCOM.2011.6069711
31. Uzairue S, Nkordeh N, Matthews VO, Bob-manuel I (2018) IOT and 5G?: the interconnection, October, 2018
32. Aman MN, Basheer MH, Sikdar B (2019) Two-factor authentication for IoT with location information. IEEE Internet Things J. https://doi.org/10.1109/JIOT.2018.2882610
33. Dwork C, Roth A (2013) The algorithmic foundations of differential privacy. Found Trends Theor Comput Sci. https://doi.org/10.1561/0400000042
34. Kobsa A, Schreck J (2003) Privacy through Pseudonymity in User-Adaptive Systems. ACM Trans Internet Technol. https://doi.org/10.1145/767193.767196
35. Thakare A, Lee E, Kumar A, Nikam VB, Kim YG (2020) PARBAC: priority-attribute-based RBAC model for azure IoT cloud. IEEE Internet Things J. https://doi.org/10.1109/JIOT.2019.2963794
36. Gavrila DM, Philomin V (1999) Real-time object detection for 'smart' vehicles. Proc IEEE Int Conf Comput Vis. https://doi.org/10.1109/iccv.1999.791202
37. Bhatt S, Patwa F, Sandhu R (2017) An access control framework for cloud-enabled wearable ınternet of things. In: Proceedings—2017 IEEE 3rd ınternational conference on collaboration and ınternet computing, CIC 2017. https://doi.org/10.1109/CIC.2017.00050

38. Gupta D, Bhatt S, Gupta M, Kayode O, Tosun AS (2020) Access control model for google cloud IoT. In: Proceedings—2020 IEEE 6th international conference on big data security on cloud, BigDataSecurity 2020, 2020 IEEE international conference on high performance and smart computing, HPSC 2020 and 2020 IEEE internatioanal conference on intelligent data and security, IDS 2020. https://doi.org/10.1109/BigDataSecurity-HPSC-IDS49724.2020.00044
39. Tekeoglu A, Tosun AS (2017) A testbed for security and privacy analysis of IoT devices. In: Proceedings—2016 IEEE 13th international conference on mobile Ad Hoc and sensor systems, MASS 2016. https://doi.org/10.1109/MASS.2016.051
40. Hu VC et al (2014) Guide to attribute based access control (abac) definition and considerations. NIST Spec Publ
41. Bhatt S, Tawalbeh LA, Chhetri P, Bhatt P (2019) Authorizations in cloud-based internet of things: current trends and use cases. In: 2019 4th ınternational conference on fog and mobile edge computing, FMEC 2019. https://doi.org/10.1109/FMEC.2019.8795309
42. Bhatt P, Bhatt S, Ko M (2020) Poster: IoT SENTINEL—an ABAC approach against cyber-warfare in organizations. In: Proceedings of ACM symposium on access control models and technologies, SACMAT. https://doi.org/10.1145/3381991.3396227
43. Gupta A, Christie R, Manjula R (2017) Scalability in Internet of Things: features, techniques and research challenges. Int J Comput Intell Res

Hybrid Radio Frequency Identification and Global Positioning System Asset Tracking

Itumeleng Matshego and Tranos Zuva

Abstract Tracking involves information gathering, manipulation and proving information on the location of a set item. Many single or hybrid technologies—Global Positioning System (GPS), radio frequency (RF), Bluetooth (BLT) or Wireless Fidelity (Wi-Fi)—have been used to provide tracking information of an asset of interest. The use of hybrid technology in tracking assets has proven to be effective if the selection of the technologies used is done correctly. This study used a hybrid of GPS and radio frequency technologies to track assets of interest because of their characteristics for use inside and outside a building. In this study, GPS geofencing was used, and time interval was used to receive data from the technology. Heuristic methodology, which enabled us to divide a room into sections, was used, where testing was done in sections in a room with different types of material, such as bricks, wood or metal, and the RF signal degradation, called attenuation, was measured. A straight-line distance and a sum of distances at 30 min intervals were calculated to determine how far the asset had travelled from the point of origin to the new position. A distance of less than 10 m was ignored. Geofencing was used to trigger an event since it indicates that the asset has crossed the permitted boundary. An RF reader was placed at the door to identify when the asset left a building and triggered an event. A model was used for searching for a missing item in a room. The results showed that the system was able to produce two distances, one straight-line distance and the other the approximate sum distance travelled by the asset in 30 min intervals. The RF model was able to find an asset in a room filled with different materials.

Keywords Global Positioning System · Radio frequency tracking · Mobile tracking

I. Matshego (✉) · T. Zuva
Information Communication Technology Department, Vaal University of Technology, Vanderbijlpark, Gauteng Province, South Africa
e-mail: itumelengm@vut.ac.za

T. Zuva
e-mail: tranosz@vut.ac.za

V. Bindhu et al. (eds.), *Proceedings of Third International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 844,
https://doi.org/10.1007/978-981-16-8862-1_75

Introduction

GPS signals do not permeate solid structures like trees, walls, buildings and rocks. This makes it problematic to track assets inside buildings using GPS, as signals reflect, causing multipath reflections which result in major inaccuracy [1]. The utilisation of RFID location-based tracking using triangulation outside in open areas and less densely populated areas is a major issue because of the fact that it utilises beams which require transmitters, towers and receivers [2]. Therefore, using RF to track assets outside buildings requires extensive aerial coverage, and hence, most studies focus on using RF inside buildings, plantations and warehouses [2–4]. This study aims to find an effective way of combining the two technologies in tracking assets inside and outside buildings. The contribution of this study is to also assist organisations in tracing possible routes that stolen assets take and put control systems in place to prevent other items from being moved using the same route.

The purpose of this study is to investigate how to improve effective asset tracking of movable assets both inside and outside buildings. The main research question is as follows:

- How can we accurately track mobile assets outside and inside buildings?

In order to the answer the main question, the following questions were answered:

- What mechanism can be used to track mobile assets inside and outside buildings?
- How can a prototype of a mobile asset tracking system be developed for outside and inside buildings?
- How can the effectiveness of the prototype be evaluated?

In order to answer the research questions, the following objectives were fulfilled:

- To propose a mechanism to track mobile assets inside and outside buildings
- To develop a prototype of a mobile asset tracking system for outside and inside buildings
- To evaluate the effectiveness of the prototype.

The rest of the paper is organised as follows: Sect. 2 addresses the limitations of the study with respect to the design, implementation and test of our asset tracking solution. The background, research literature related to asset tracking is covered in Sect. 3. The methodology and building a new hybrid asset tracking system based on the use of GPS and RF are discussed in Sect. 4 inclusive of the asset travelled distance formulae is shown. The appropriate parameters setting of our hybrid tracking solution is discussed in detail. Section 5 presents data analysis and results of our hybrid system finally, and concluding remarks are made in Sect. 6.

Limitations of the Study

The study was subjected to the following limitations that can be improved upon in future studies:

- The study is geographically limited to South Africa.
- The assets on which the hybrid tracking solution was tested are limited to just one asset.
- The findings came from using two specific technologies, and therefore, findings cannot be generalised across all other technologies.

Related Work

In his paper, [5] indicates that the Global Positioning System is the most popular technology used to establish the outdoor position information of any item as it covers a wide geographical terrain and area. Because the technology utilises latitudinal and longitudinal positions calculated from at least four satellites, locating an asset can be very simple. However, inside buildings, the use of GPS to locate assets has major limitations, mainly due to line of sight interference with satellites orbiting the earth. The use of both RF and GPS solves these limitations. The use and reliability of RF is superb in that it cannot be easily replicated and is also stable. Hence, RF has been used in various industries, such as security, healthcare, retail, transportation, warehousing and distribution. Indoor positioning systems are dependent on the use of various technologies [6]. The use of RFID is also advocated in tracking patients inside hospitals as they are being moved from one ward to another. By doing so, the authors argue that one would have eliminated user-initiated activation and the line of sight problem faced by GPS technology whilst providing one hundred per cent accuracy. In their research, [7] utilises a combination of GPS, Wi-Fi and smart phone to track assets inside and outside buildings. Outside buildings, the GPS works well by utilising trilateration to find the accurate position of the asset; whereas inside buildings, the utilisation of Wi-Fi technology is vital in pinpointing the location of a particular asset. The hardware configuration consists of a Wi-Fi transmitter and receivers situated in various areas of the building. As the transmitter moves from one floor to another, at certain instances, a receiver picks up that a particular unit is within its range, thereby helping in locating the asset. A hypothetical test was conducted in a building with a four floors, where the asset in question was represented by a dot on the mobile phone showing its exact location on the floor. The major limitations of this solution are the costs associated with setting up various Wi-Fi receivers at various points of the building. Secondly, the solution gives an approximate location. So, in case of recovery in a congested building, the unit can be traced to a floor, but its exact location cannot be known, and this is not ideal for recovery.

In [8], a different approach was taken in tracking assets inside congested buildings or areas, especially on construction sites. They use a hybrid model of RFID and

Bluetooth. The model is based on registering all asset information on a database, then tagging each asset and each person with an RFID. The manager is then equipped with a smart phone which notifies him as to which asset has been taken by whom and where the particular asset is in the room, warehouse or construction site. If the asset leave the designated area for any reason, then the manager is notified via the smart phone. The idea behind this process is to eliminate the limitations that come with GPS inside buildings, where signal cannot penetrate or where signal is disturbed due to congestion. These limitations of a GPS-based system are further noted by [9], who discovered that the utilisation of GPS in combination with GSM applications software to track assets is not effective indoors owing to the aforementioned issues. However, their study focussed mainly on tracking assets outside as opposed to both inside and outside. Secondly, the solution was based on the premise that all users had utilised a desktop to log in so as to be able to track their assets, and therefore, it does not accommodate mobile access for those users who are on the go but are still determined to track and locate their assets. According to [10], the solution of asset tracking by a wireless system utilising GPS was based on the premise that a receiver linked to a server with maps integrated provides the location data for to the system. User interface is then used as an application running on machine. The analysis and distribution of data feeds were done by both the client and server application through the internet. The client and server application communicate via cellular modem connection to the internet. In his studies, [4] had the aim of implementing low-cost DGPS to get accurate positioning by using two smartphones, where one phone was used as a receiver and the other as a transmitter. The method entails having the transmitter phone sending signal via cellular network infrastructure using 3G network. The transmitter smartphone is placed at a stationary position or location, and then, an android application is created as the interface for the receiver smartphone. The coordinates are ascertained at four hours' regular intervals at the same location. The limitation of such a solution is that tracking is not done in real time. The other limitation is that the solution does not work inside buildings because the GPS signal is not able to penetrate buildings and solid structures.

This inability of GPS signal to penetrate walls was overcome by the use of RF in locating assets inside buildings by [3]. He implemented a fixed RF reader inside a building acting as a checkpoint which would measure and analyse information about an asset's location change. The reader would then pick up signal from all RF units embedded into all assets transmitting their location within the building. The reader's signal range worked well in larger areas, and signal interference was not an issue. Then, he utilised a handheld RF reader that led the search to the exact or terminal asset being traced. When there is a missing item, the purpose of applying RFID is to minimise search time, effort and investment cost by guiding the detector handheld reader only or handheld and fixed RFID reader. The experiment was conducted only in a rectangular-shaped room with no obstacles or unusual layout. In terms of [11], who developed a real-time vehicle-tracking system which utilised an Arduino Uno R3 module, a SIM 808 module linked to a GPS and a GSM antenna with a battery source. The system was embedded into a vehicle and allowed for real-time monitoring. In the above system, the data are received via the GPS receiver; this data would constitute of the latitudinal and longitudinal positions of the particular vehicle via the satellite constellation, the information being transferred via short message service (SMS) through a GSM modem. The GSM modem is linked to an Arduino Uno R3 microcontroller. At the same time, hypertext transfer protocol is used to transfer information of the vehicle's position to the HTTP server. The specially built Web application interface used PHP, Javascript or jQuery, HTML with embedded Google maps. Application programming interface (API) is used to display the vehicle's position. The real-time monitoring is achieved by having the geopositional data retrieved every two seconds and the maps synchronising at the same time, thus achieving real-time tracking. The system allowed for accurate data in real time which enabled the user to track a vehicle and also assisted in terms of recovery of the vehicle in case of loss or theft. The shortcoming of such a system was the amount of time it took for the GPS module to connect to the network due to poor weather conditions as the GPS module requires a clear line of sight with the GPS constellation. This means that in case of indoor tracking and monitoring such a solution was unable to track the vehicle.

The inability to track is caused by errors to the receiving signal of the GPS receiver owing to the fact that obstacles like tall buildings and solid infrastructure block the view of the sky; the effect of this is multipath error to the receiving signal of the GPS receiver [12]. As a result, on error being initialised, inaccurate results are received as the signal tends to jump from one place to another, sending inaccurate values of latitude and longitude to the server for display on the Google map. In his paper, [13] investigated the use of cellular Internet of Things that can be applied in telematics with the use of low power wide area network (LPWAN) which is low cost and long range.

Research Method and Design

Introduction

The solution comprises of two different mechanisms for tagging the asset we wanted to track, in this case a smart TV, we achieved this by means of both a GPS and RFID module. We developed a C# and SQL DB application that was used to track and monitor the tagged asset outside buildings along with BING API Maps. A geofence was then setup in the area where when the tagged asset exited, a notification was sent to the user informing them of the asset breach. This was done through the designed user interface.

Search an asset outside buildings

In the event that the asset which is attached with a GPS module is missing, a search will be required. When searching for the asset, the GPS location and route that the asset took will be calculated in order for the asset to be required. The location of the asset was denoted as GPS points P_{k-1} at time intervals of 30 min, where P_0 is the

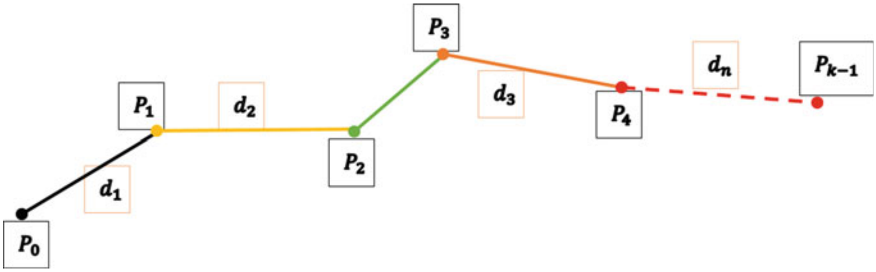


Fig. 1 GPS asset route calculation

initial position of the asset (the location where the asset must be), = 0, 1, 2 … $k-1$, and P_{k-1} represents the last location of the asset. The distance between two points is given as d_i where $i = 1, 2, 3 \ldots n$. The approximate distance travelled by the asset outside buildings is given by Eq. (1). Figure 1 illustrates the model of how the asset travelled.

Total distance adding d_1 to d_n is the total distance covered using Eq. (1) [14].

$$d = \sum_{i=1}^{n} d_i \tag{1}$$

where $i \in [1, n]$,

$$\sum_{i=1}^{n} d_i = \sum_{i=1}^{n} 2r \arcsin \arcsin\left(\sqrt{\sin^2\left(\frac{\varphi_i - \varphi_{i-1}}{2}\right) + \cos(\varphi_i)\cos(\varphi_{i-1})\sin^2\left(\frac{\lambda_i - \lambda_{i-1}}{2}\right)}\right)$$

φ_i *is* latitude, and λ_i is longitude points.

The location of the asset was denoted as GPS points p_{k-1} at a time interval of 30 min, where p_0 is the initial position of the asset (the location where the asset must be), = 0, 1, 2 … k - 1, and p_{k-1} represents the last location of the asset. The distance between two points is given as d_i where $i = 1, 2, 3 \ldots n$. The approximate distance travelled by the asset is given by Eq. (1).

It is very important to note that lat and lon stand for latitude and longitude, respectively. North latitudes and west longitudes are taken as positive, and south latitudes and east longitudes are taken as negative [15, 16].

Looking at Fig. 1, the approximate distance travelled by the asset p_0 to p_{k-1} is calculated in Fig. 1, using the latitude (φ_i) and longitude (λ_i) points. The straight-line distance results using the same Eq. (1), however, we only we only two points for the straight-line distance, p_0 and p_{k-1} using the same Eq. (1). Thus, we will have the true distance covered and the approximate distance travelled by the asset being tracked. The true distance would therefore be calculated by calculating the distance between two consecutive GPS coordinates.

Search an asset inside building

Upon discovering the route that the tagged asset took and locating the asset to a building where GPS does not work due to signal issues inside building or a room. Radio frequency would therefore need to be used to locate the asset. In some cases, the searcher might be faced with a situation, whereby the room being searched has various items such as cupboards, chairs, tables, metal or wooden stands, and sectional walls. Furthermore, we assume that a searcher can predict the location where she/he can find the item location, using GPS coordinates. These items will affect the search time have a bearing on the signal of radio waves known as attenuation. In this study, we used a specific handheld RF reader which will scan the room based on the signal strength detected from the RF module of the tagged asset.

This reader works on a continuous mode, which scans the tagged item strength level continuously in repeating cycles within a specified range. We considered using data from the handheld reader in continuous mode to find the location of a tagged item inside a room. The handheld reader identified the object signal in the minimum distance in order for the exhaustive search to succeed. When the searcher does not have any background about the room, even though the searcher holds the RFID handheld reader in the particular area, it is difficult to suggest a certain heuristic policy.

The object that we are supposed to find is located in a given room. The size of the room is X (width) and Y (length), as given in Fig. 2a. We initially assume that the tagged item is located in any given location inside the room with equal probability. This study used the steps proposed by [3].

In this study, we assume that the room is rectangle. Let the area of the room be XxY, and let the detecting radius of the sensor be r as in Fig. 2b. Let "v" _1 be the searching velocity from left to right as the searcher walks parallel to the room's bottom line. The handheld sensor has then searched an area of 2r"v" "1" t entirely

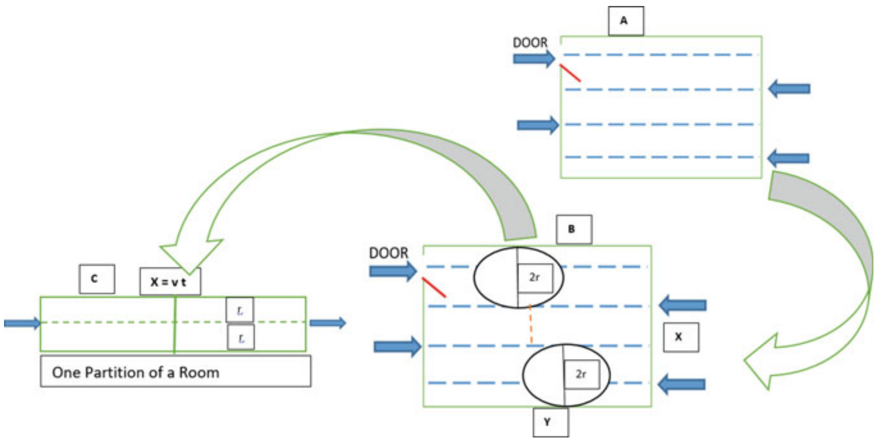


Fig. 2 Exhaustive search routes in a room [3]

when $t = X/$"V" "1." The minimal searching time of walking around the room exhaustively is defined as "T" _"e" = $XY/(2r$"v" _"1").

It takes T_e duration to search the entire room, and under the assumption that the tagged item is in the room, the probability that the tagged item is found within t_p seconds is uniformly distributed in Eq. (2).

$$t_p = \frac{1}{T_e}, \quad 0 \le t_p \le T_e \tag{2}$$

where t_p is time in seconds, and T_e is duration to search a room.

According to the definition of $E[T]$ (anticipated value of T), the expected time to find an object is as follows: $E[T] = 1/$"2""T" _"e." There is one more aspect to consider in this case. Using a handheld reader to detect the tag within a radius r, we must look for the tag manually by opening cupboards, looking beneath tables and so on. The manual time to find the tag is proportional to the radius r. Define the time to find the tagged item manually by $h(r)$. For the expected time $g(r)$, to find the tag in the room, given that tag is in the room, we use Eq. (3).

$$g(r) = \frac{1}{2}T_e + h(r) \tag{3}$$

where r is the radius, $g(r)$ expected time to locate the asset in a room, and $h(r)$ is manual search time.

As r decreases, the time to find the approximate location is increased, but the time to do the fine search increases.

Entering the room, a searcher tours all sections of the room in order, from the left section to the right, and from the upper section location to the lower section. The section that the searcher visits initially is given the lowest index. To avoid detecting the tag in the other section during searching of the current section of a room, we need to restate t_e mentioned before. Thus, with a handheld RF reader, the time to do an exhaustive search of a section when the searching path is given below.

Therefore, the expected time to find the item when doing an exhaustive search of the entire room uses formula (4).

$$\boldsymbol{E}^1[\boldsymbol{T}] = \sum_{i=1}^{N} \overline{p}_i \times \left((i-1) \times t_e + t_e \times \frac{1}{2} \right) \tag{4}$$

where $\overline{p}_i$ is the path in a section of a room.

Data Analysıs and Results

The purpose of this section is to present and discuss the findings of the experiments conducted in the previous section. The discussion is based on data obtained from tracking an asset, using hybrid technologies. The findings are presented in both graphical and narrative formats. The research and experiment focussed on utilising Global Positioning System (GPS) and radio frequency (RF) to track a particular asset inside and outside of buildings, along with the potential benefits of each. The section further contains the results obtained to answer the research questions of the study. A chosen asset was installed with a particular GPS chip and RF chip in order for us to prove our theory. The main objective was to develop an understanding of what has been achieved in the tracking of assets inside and outside buildings by first investigating what has been done in the literature to develop an optimal solution in the adoption of the current technologies and to establish whether hybrid technologies can be adopted, and then to recommend the most suitable solution. The findings presented in this section demonstrate the potential for using multiple technologies to achieve our objective.

Firstly, a Samsung LED TV model C21FCS unit was installed with a GPS module and RF module, as seen in Figure 3.

A user is presented with a login screen which requires him to provide his login credentials in order to access the platform. Controls were primarily designed to allow them to change without altering the source code, which gives us flexibility for different screen orientations to be handled on the fly by the framework. This allows

Fig. 3 GPS AND RF ınterior tagged Samsung LED TV

Fig. 4 Desktop ınterface tracking application login page

for the application to handle differently on msds mobile devices as well as desktop devices. Figure 4 reflect this.

As the tagged asset moved from a set location where it was supposed to be and geofenced, the starting point was indicated by the green flag (Fig. 5), and the point where the asset has ended is indicated as point two in Fig. 4.13. This is where the unit was placed having travelled an approximate sum distance of 60.3 km, which is different from the straight-line distance of 45.82 km. Upon reaching the set destination, this being a block of buildings, the longitude and latitude positions of the GPS module were recorded as 26.31867 and 28.003214, respectively, as the tagged asset entered a room at the set location, as seen in Fig. 5.

Upon reaching the building, the standard RF tag is woken up via a signal from the RF receiver. The receiver signal strength indicator (RSSI) measured in decibel-milliwatts dBm represents the measure of power received from a radio signal. The RF module is then activated using a RF handheld reader, depicted in Fig. 6.

To find an RF-tagged item indoors, a handheld RF reader was used, based on the following experiment. In this case, a tagged item was searched for using a handheld reader. The radius that we used for the RF reader to start searching for the tagged item was from 1 to 4 m. The tagged item was in an enclosed room, which gave us a certain probability of detecting the asset in a subsection of the area, as in the table of probability below.

This means that the time it took to locate the tagged item and the area in which the tagged item was to be found would be divided into parts to make sure that the searched area was not repeated. Searching the room with a handheld RF reader is advantageous

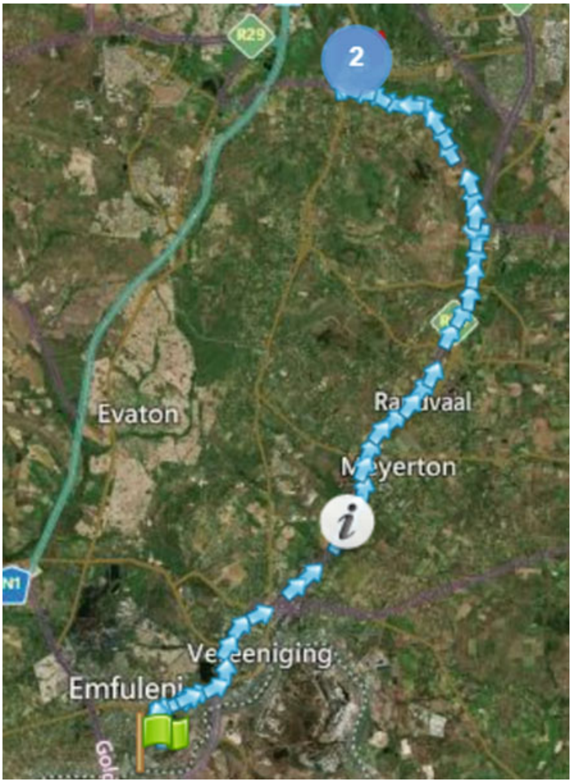


Fig. 5 Unit travel route with destination coordinates

in that detection, and location is conducted without unnecessary repeated calculations. In some instances, such as when a fixed reader is used, heuristics are used in order to locate tagged items.

Three tests were then run inside the building in a room. The first test was conducted by placing the tagged asset in an empty room; the *r* was measured to determine the proximity of the tagged asset. The main aim was to get the RF handheld reader to illuminate a minimum of four bars for each search measuring. The RF reader was placed close to the tagged item and then moved further away. This was done in order to measure the change in attenuation as the distance increased between the RF transmitter and receiver (Fig. 7).

Table 1 above shows the experimental results of the tagged asset. Searching for the asset in a room, we needed a reading of a minimum of -10dbm, which in turn confirmed the existence of the tagged item within the room.

Figure 8 shows how signal strength degradation at various *r* points is affected by the various items. As discussed, the optimal minimum strength required for the handheld reader is -10 dbm, which illuminates four LED lights on the handheld receiver to indicate the presence of the tagged item in the room. Figure 8 shows that this required reading was attained at different *r* values in a room that had various

Fig. 6 RF handheld receiver apparatus

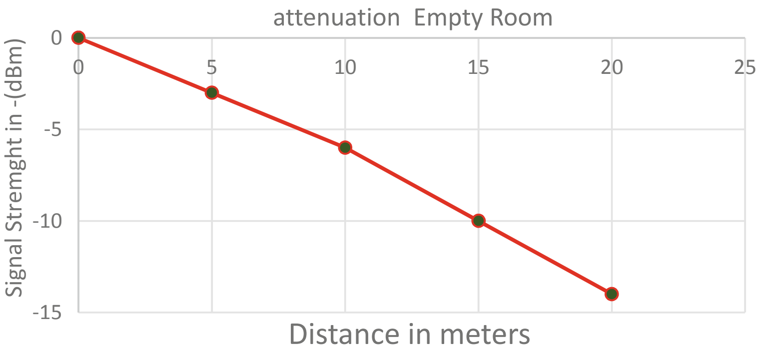


Fig. 7 RF signal attenuation in an empty room

Table 1 Attenuation signal strength results

Distance (m)	Empty room (dbm)	Wooden cupboard (dbm)	Brick wall (dbm)	Under metal table (dbm)
0	0	0	0	0
5	−3	−3.45	−3.66	−4.4
10	−6	−6.9	−7.32	−8.7
15	−10	−11.5	−12.2	−14.5
20	−14	−16.1	−17.8	−20.3

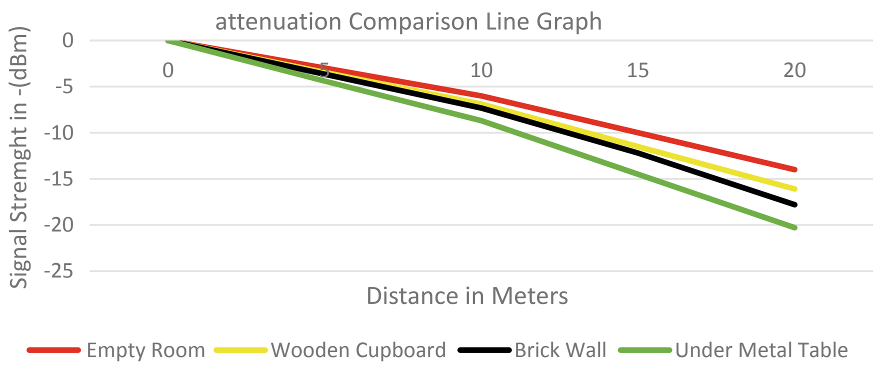


Fig. 8 RF signal attenuation graph line comparison

items. In an empty room, the *r* value required is 15 m; however, in a room that has metal the *r* value was roughly 13 m, and in a room that has wood, the *r* value is 9.6 m. This means that if a room has both metal and wood, the *r* value is the average of the two, which is 11.3 m. The demarcation of a room using heuristic search shown in Figure 3 was influenced by the type of material in the room, that is, depending on what types of obstacles are found in a room, a searcher is able to determine how to search the particular room for a tagged asset.

Conclusion

The monitoring, tracking and location of assets inside and outside of buildings are paramount, not only to protect the safety of assets but also to assist with asset management, be it for personal or business use. Knowing where your assets are and being able to locate them is not only a want but a need. In this study, we looked at a hybrid model in the tracking of an asset, and we also built a model and designed and implemented it to achieve our objectives. An asset was tagged with a GPS and RF module, which was then moved from its actual position to an unknown place. Results showed that the model works in both environments and where one technology lags, the other has the capability and robustness to locate the asset. The research was also able to contribute to the battery life of a live trackable asset because of different time intervals when the modules pinged, also assisting organisations in having in depth knowledge on how assets are moved or taken and possible routes takes due to the history logs. Future work is proposed at looking at IoT solutions with low power wide area networks (LPWA) such as Sigfox and devices using cellular Internet of Things (CIoT) and which will also assist in congestion of access of devices [13].

References

1. Leick A (2004) GPS satellite surveying, 3rd edn. Wiley Publisher, Hoboken, New Jersey, USA
2. Sweeney PJ (2006) RFID basics for dummies. Wiley Publishing, New York
3. Kim GG (2007) Locating and tracking assets using RFID. Master's thesis, Texas A&M University
4. Lawrence LB (2013) Implementation and evaluation of a differential GPS based on smartphones and ınternet technology. Master's thesis, Halmstad University
5. Reddy AR, Basha PA (2012) An efficient model for design of 64-bit high speed parallel prefix VLSI adder. Int J Modern Eng Res 2:2316–2319
6. Fisher JA, Manahan T (2012) Evaluation of real-time location systems in their hospital contexts. Int J Med Inform 81:705–712
7. Sultana S, Tahsin M, Reza T, Hossam-E-Haider MD (2016) An innovative implementation of indoor positioning system using GPS. In: 2016 3rd ınternational conference on electrical engineering and ınformation communication technology (ICEEICT), Dhaka, 2016, pp 1–4
8. Bisio I, Sciarrone A, Zappatore S (2015) Asset tracking architecture with bluetooth low energy tags and ad hoc smartphone applications. In: 2015 European conference on networks and communications (EuCNC), Paris, pp 460–464
9. Verma P, Bhatia JS (2013) Design and development of GPS-GSM based tracking system with GOOGLE map based monitoring. Int J Comput Sci Eng Appl (IJCSEA) 3:33–40
10. Geske K (2000) Development of a wireless position tracking system. Undergraduate thesis, Manitoba University
11. Ahmed S, Rahman S, Costa SE (2015) Real-time vehicle tracking system, Master's Thesis, BRAC University
12. Samrand E (2016) Automated monitoring system designing a laboratory equipment tracking system
13. Suma V (2021) Power efficient time-division random-access model based in wireless communication networks. IRO J Sustain Wirel Syst 2(4):155–159
14. Chopde NR, Nichat M (2013) Landmark based shortest path detection by using A* and haversine formula. Int J Innov Res Comput Commun Eng 1(2):298–302
15. Zuva T, Olugbara OO, Ojo SO, Ngwira SM (2012) Image content in location-based shopping recommender systems for mobile users. Int J 3(4):1–8
16. Rose C, Britt J, Allen J, Bevly D (2014) An integrated vehicle navigation system utilizing lane-detection and lateral position estimation systems in difficult environments for GPS. IEEE Trans Intell Transp Syst 15(6):2615–2629

Author Index

V. Bindhu et al. (eds.), *Proceedings of Third International Conference on Communication, Computing and Electronics Systems*, Lecture Notes in Electrical Engineering 844,
https://doi.org/10.1007/978-981-16-8862-1