ISSN (ONLINE) : 2045-8711
ISSN (PRINT) : 2045-869X

# INTERNATIONAL JOURNAL OF INNOVATIVE TECHNOLOGY & CREATIVE ENGINEERING

August 2023
Vol - 13 No - 8
@IJITCE Publication

**UK: Managing Editor**
International Journal of Innovative Technology and Creative Engineering
1a park lane,
Cranford
London
TW59WA
UK

**USA: Editor**
International Journal of Innovative Technology and Creative Engineering
Dr. Arumugam
Department of Chemistry
University of Georgia
GA-30602, USA.

**India: Editor**
International Journal of Innovative Technology & Creative Engineering
36/4 12th Avenue,
1st cross St,
Vaigai Colony
Ashok Nagar
Chennai, India 600083

Email: editor@ijitce.co.uk

**www.ijitce.co.uk**

**IJITCE PUBLICATION**

# *International Journal of Innovative Technology & Creative Engineering*

Vol.13 No.08

**August 2023**

IJITCE
INNOVATE TO CREATE

**www.ijitce.co.uk**

Dear Researcher,

Greetings!

In this issue discusses about recent trends on mathematics concepts and computer science.

We look forward many more new technologies in the next month.

Thanks,
Editorial Team
IJITCE

# Editorial Members

# Review Board Members

# Contents

# Generative AI: An overview

**Gowri A**
Assistant Professor, Department of Information Technology,
Hindusthan College of Arts and Science, Coimbatore, Tamil Nadu, India
**Bala Ganesh B**
M.Sc (Information Technology) Department of Information Technology,
Hindusthan College of Arts and Science, Coimbatore, Tamil Nadu, India
**Chandana Pradeepa**
M.Sc (Information Technology) Department of Information Technology,
Hindusthan College of Arts and Science, Coimbatore, Tamil Nadu, India

**ABSTRACT-Generative Artificial Intelligence (AI) has emerged as a powerful paradigm within the field of machine learning, allowing machines to create original content and display creative behaviors. The paper provides an overview of generative AI, covering its algorithms and applications. The outline of the paper is the key principles and methods employed in generative AI, including prominent models like Generative Adversarial Networks (GANs) and Variational Autoencoders (VAEs). The first section of the paper elucidates the concept of generative AI and its associated algorithms, while the second section explains the practical applications of this technology.**

Keywords: GANs, VAEs, AI

## I. INTRODUCTION

Generative artificial intelligence (AI) is a form of AI that can autonomously generate new content, such as text, images, audio, and video. Generative AI provides innovative approaches for content production in the metaverse, filling gaps in the development of the metaverse. Products such as ChatGPT have the potential to enhance the search experience, reshape information generation and presentation methods, and become new entry points for online traffic. This is expected to significantly impact traditional search engine products, accelerating industry innovation and upgrading. Generative AI refers to a branch of artificial intelligence that focuses on generating new content, such as text, images, or even music, based on patterns and examples from existing data. It utilizes machine learning techniques, particularly deep learning models, to create original content that resembles human-generated data.

One popular approach to generative AI is the use of generative adversarial networks (GANs). GANs consist of two components: a generator network that generates new samples, and a discriminator network that tries to distinguish between the generated samples and real data.

Through an iterative process, the generator and discriminator improve their performance, resulting in more realistic and high- quality output.

Generative AI has various fields, including art, design, entertainment, and even scientific research. It can be used to generate realistic images for video games, assist in the creation of synthetic training data for machine learning models, aid in creative tasks such as painting or music composition, and even support research in areas like drug discovery or molecular design.

However, it's important to note that generative AI also poses certain challenges and ethical considerations. There are concerns about the potential misuse of generative AI for creating fake or misleading content, deepfakes, or infringement of intellectual property rights. Ensuring responsible use and appropriate regulation of generative AI technologies is crucial to mitigate these risks and promote beneficial applications of this powerful technology.

## II. EVOLUTION

Artificial intelligence (AI) may have a first-rate impact on the way modern societies respond to the tough disturbing conditions they face. Properly harnessed, AI can create an extra fair, healthy, and inclusive society. Today, AI has come to be a mature technology and a more and more important part of the modern life fabric. AI is already deployed in one-of-a-type software domains, e.g., recommendation systems, direct mail filters, photograph recognition, voice recognition, virtual assistants, etc. It spans many sectors, from treatment to transportation, and all through decades, for the purpose that term was modified in-between the 1950s. The strategies moreover evolved, from the foundational AI algorithms of the 1950s to the paradigm shift in symbolic algorithms and expert machine development in the 1970s, the appearance of machine-gaining knowledge in the 1990s, and the deep gaining knowledge of algorithms of the 2010s.

The thought of generative models, in any case, commenced to require shape over a long time and revolutionized several businesses, collectively with discourse acknowledgment, photo handling, and characteristic dialect preparing (NLP).

New generative models like Bayesian systems and Markov models have developed as accessible at the turn of the thousand years, significantly progressing mechanical technology and computer vision. Profound learning's innovation and ensuing headway, be that as it may, raised the bar for generative AI. The profound neural arrangement period has enabled researchers and makers to form exceptional propels in generative AI. The improvement of generative dangerous systems (GANs) in 2014, which created instep genuine looking pictures, motion pictures, or possibly sounds taken after genuine human-generated substance, adjusted into one exceptional breakthrough.

Today, generative AI is being connected in huge sort of bundles through numerous distinctive businesses. It has long gone past what is considered to be conventional depictions and inventiveness, thinking the entry of intriguing melodic compositions and captivating works of art. Also, it has empowered businesses to make smooth products, improve healthcare results, and essentially modify how we utilize period. The region of generative AI is balanced to extend indeed more as time creates and insights accessibility rises, sorting out endless openings for inventiveness and disclosure. The thought of generative models, in any case, commenced to require shape over a long time and revolutionized several businesses, collectively with discourse acknowledgment, photo handling, and common dialect handling (NLP).

## ALGORITHM

### 1. Generative Adversarial network

A Generative Adversarial Network (GAN) is a type of Machine Learning system that can create realistic images by playing a game between two parts: a Generator and a Discriminator. The Generator tries to produce fake data that looks similar to real data, while the Discriminator tries to tell the difference between real and fake data. GANs are used to generate complex data like images, audio, and videos. They were first used in 2017 to make humanfaces look more realistic in pictures.

To train a GAN, the Generator creates fake data, like counterfeit money, and the Discriminator is given both the real data and the fake data. The Discriminator's job is to figure out which is real and which is fake. The GAN is trained by adjusting the Generator and Discriminator based on the mistakes the Discriminator makes.

Example of GAN Mathematical Equation

The mathematical equation for training a GAN is:

G = Generator
D = Discriminator
Pdata(x) = distribution of real data
p(z) = distribution of the generator
x = sample from real data
z = sample from generator
D(x) = Discriminator network
G(z) = Generator network

In the study of GANs' ability to convert 2D images to 3D, the process involves collecting live data, creating a benchmark dataset with important features, merging images to determine suitability, and pre-processing the images through segmentation and cleaning. The GAN is trained to analyze patterns and generate accurate 3D images.

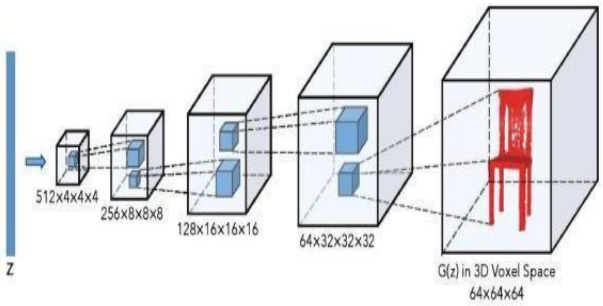



**Figure 1: Generative Adversarial networks**

### Variational autoencoders

In the last few years, deep learning-based generative models have gained more and more interest due to (and implying) some amazing improvements in the field. Relying on huge amounts of data, well- designed network architectures, and smart training techniques, deep generative models have shown an incredible ability to produce highly realistic pieces of content of various kinds, such as images, texts, and sounds.

A variational autoencoder (VAE) offers a probabilistic approach to representing an observation in latent space. Instead of constructing an encoder that produces a single value for each latent attribute, we design our encoder to describe aprobability distribution for each attribute.

To illustrate this concept, let's consider a scenario where we have trained an autoencoder model on a large dataset of faces, with an encoding dimension of 6. A well-trained autoencoder will learn distinctive characteristics of faces, such as skin color or whether the person is wearing glasses, aiming to capture an observation in a condensed form.

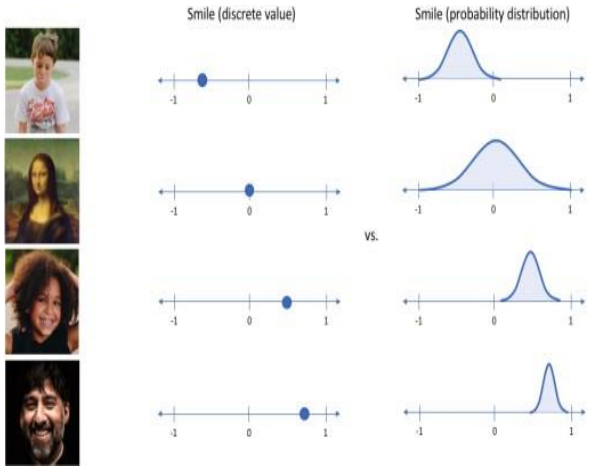



**Figure 2: Variational auto encoders**

In the given example, we represented the input image by utilizing a single value for each attribute in its latent representation. However, it might be preferable to represent each latent attribute as a range of potential values. For instance, if we input a photo of the Mona Lisa, assigning a single value to the smile attribute would be challenging. By employing a variational autoencoder, we can describe latent attributes in terms of probabilities or likelihoods.

## 1. RECURRENT NEURAL NETWORK

During the 1990s, there was significant research and development focused on recurrent neural networks. These networks were specifically designed to learn patterns that occur in a sequence or change over time. A recurrent neural network is a type of neural network that includes feedback connections, forming a closed loop [Fausett, 1994]. Notable examples of recurrent neural networks include BAM, Hopfield, Boltzmann machine, and recurrent backpropagation nets [Hecht-Nielsen, 1990]. These techniques have been applied to a wide range of problems. In the late 1980s, researchers such as Rumelhart, Hinton, and Williams introduced simple partially recurrent neural networks to learn strings of characters [Rummelhart, 1986]. Additionally, recurrent neural networks have been used in various applications involving dynamical systems with sequences of time-based events.

## 2. TRANSFORMERS

Transformers have had a significant impact on generative AI. They have emerged as a powerful architecture for a range of generative tasks, including natural language processing, image synthesis, and music generation.The introduction of Transformers, exemplified by models like GPT (Generative Pre-trained Transformer), has revolutionized language generation. Unlike earlier methods that relied on recurrent neural networks, Transformers use self-attention mechanisms to capture the relationships between words in a sequence. This enables them to produce coherent and contextually relevant sentences. In the realm of image synthesis, Transformers have been applied to tasks such as image captioning, text-to-image synthesis, and style transfer. By leveraging the attention mechanism, Transformers can effectively understand the overall context and generate visually appealing images based on textual descriptions.

Moreover, Transformers have also demonstrated promise in music generation. By treating musical notes or events as sequential input, Transformers can learn long-term dependencies and create consistent musical compositions. This advancement has led to the development of Transformer-based models capable of composing melodies, harmonies, and even complete musical arrangements.

## 3. DEEP CONVOLUTIONAL GENERATIVE ADVERSARIAL NEURAL NETWORK

Deep Convolutional Generative Adversarial Neural Network is a popular and successful network design for Generative Adversarial Networks. Unlike traditional Convolutional Neural Networks, DCGAN does not utilize max pooling or fully connected layers. Instead, it employs convolutional stride and transposed convolution for down sampling and up sampling. The generator in DCGAN follows a specific network design, as shown in the figure below. DCGAN is an implementation of GANs using convolutional neural networks. In traditional CNNs, feature extraction and down sampling are accomplished through convolutional and pooling layers. However, DCGAN eliminates the pooling layer in both the discriminative and generative models Gong, Wang and Lazebnik (2014). This allows the network to learn spatially up sampling and down sampling on its own. The discriminative model in DCGAN is a convolutional neural network without fully connected layers, using LeakyReLU activation functions and either Sigmoid or SoftMax functions for binary problems.

Its purpose is to compress an image into a feature vector. On the other hand, the generative model involves a deconvolution process.

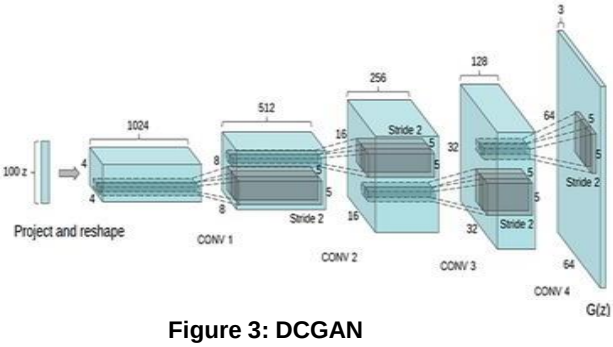



**Figure 3: DCGAN**

Compared to traditional GANs, DCGAN stands out by replacing the multilayer perceptron with a CNN. The pooling and sampling layers are removed, and the discriminator employs convolutional layers for image discrimination, while the generator utilizes deconvolution layers for image generation. The specific structure of the DCGAN generator includes an input layer, batch normalization layer for convergence acceleration, and reshaping layer for data normalization, followed by upsampling, Conv2DTranspose, and batch normalization layers for sampling, deconvolution, and data normalization, respectively.

## 4. DEEP REINFORCEMENT LEARNING

Deep Reinforcement Learning (DRL) combines deep learning techniques with reinforcement learning principles to train agents in complex environments. By interacting with the environment and receiving rewards or penalties as feedback, agents learn optimal behaviours. The key components of DRL include the agent (an intelligent entity learning to interact based on observations and policies), the environment (representing the external world and providing feedback), the state (capturing relevant information for decision-making), actions (choices made by the agent based on its policy), and rewards (scalar values indicating the desirability of actions to maximize cumulative rewards).

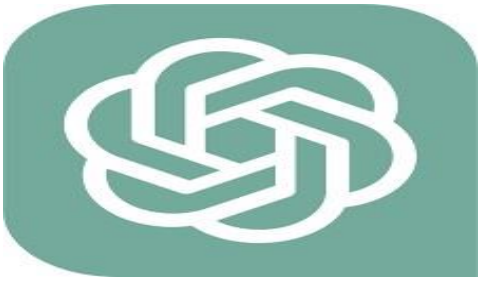


**Figure 4: Chat GPT**

DRL algorithms utilize deep neural networks such as CNNs or RNNs to approximate the agent's policy or value function, enabling them to learn complex representations from high-dimensional inputs. Some popular DRL algorithms include Q-Learning (which learns the action-value function, Q-function) combined with deep neural networks in Deep Q- Networks, Policy Gradient Methods (optimizing the agent's policy through estimating gradients of cumulative rewards, e.g., Proximal

Policy Optimization, Trust Region Policy Optimization), and Actor-Critic Methods (utilizing both an actor for action selection and a critic for estimating value or action-value function, e.g., Advantage Actor-Critic, Asynchronous Advantage Actor-Critic).

DRL has demonstrated successful applications in various domains, including game playing (e.g., AlphaGo), robotics control, autonomous driving, and natural language processing. Its ability to learn complex behaviors and make decisions in challenging environments has proven effective.

## III. REAL-WORLD APPLICATIONS

Generative AI is applied in various ways, such as in the development of chatbots like ChatGPT, image generation models like DALL-E, and language models like BART.

### CHAT GPT

ChatGPT, developed by OpenAI, is a language model that can be useful in all fields. It can generate text that is similar to human writing, which can help individuals and communities make well-informed decisions. While ChatGPT is a valuable tool, it should be used alongside human analysis and judgment. We should embrace this new technology while recognizing the importance of human involvement in the scientific process. Planning for how AI can contribute to science in the future is worth considering.

ChatGPT is an AI model that can engage in conversations with users. It goes through two main steps: pre-training and fine-tuning.

During pre-training, ChatGPT learns from a vast amount of internet text data. It tries to predict the next word in a sentence based on the words that came before it. This helps the model understand patterns, grammar, and relationships between words. After pre-training, ChatGPT goes through fine- tuning. In this phase, human reviewers provide feedback and rate the model's responses to different prompts. The model is trained to generate responses that make sense in a conversation and are relevant to the given context.

When a user interacts with ChatGPT, the model takes their input and generates a response based on what it has learned. It tries to provide coherent and contextually appropriate answers by using the knowledge gained from pre-training and fine- tuning.

However, it's important to remember that ChatGPT is not perfect. It can sometimes give incorrect or nonsensical answers. It may also be sensitive to how the user phrases their input and can show biases based on its training data. Ongoing research aims to improve these limitations and make models like ChatGPT even better activities based on individual students' objectives and preferences.

Virtual office hours: Professors can use ChatGPT to conduct virtual office hours, answering students' questions and providing real-time assistance.

Student engagement: ChatGPT increases student engagement in online classes by offering interactive activities and questions aligned with the course material.

Currently, self-directed learners in various educational applications are using ChatGPT. ChatGPT contributes to education in the following ways:

1. Tutoring and assistance: ChatGPT provides personalized guidance, helping students with their coursework and addressing their questions.
2. Research support: ChatGPT assists students in their research by providing relevant resources, articles, and papers related to their topics.
3. Class scheduling and reminders: ChatGPT helps students manage their class schedules and reminds them of upcoming assignments and tests.
4. Personalized learning: ChatGPT offers tailored recommendations for learning resources.

ChatGPT enhances the learning experience by providing personalized and interactive support. It offers customized exercises and games that cater to learners' specific needs, along with recommendations for learning materials and resources. It can act as a tutor or mentor, giving feedback and assistance throughout the learning journey. Additionally, ChatGPT empowers self- directed learners to set their own learning goals and strategies, serving as a tool for self-reflection and evaluation. This helps learners take ownership of their education, develop necessary skills, and succeed as independent learners.

## IV. CONCLUSION

Generative AI holds immense potential across various fields, such as art, music, image synthesis, and text generation. It empowers machines to autonomously produce original and diverse outputs, often surpassing traditional rule-based or deterministic approaches. The implications of generative AI extend to transforming industries like entertainment, design, and scientific research. We have explored the fundamental principles and methodologies underlying generative AI, including well-known algorithms like Generative Adversarial Networks (GANs) and Variational Auto encoders (VAEs). Furthermore, we have investigated the wide-ranging application of generative AI.

Overall, generative AI represents a revolutionary and exciting frontier in the realm of artificial intelligence, with the capacity to reshape how we create and engage with content. As advancements in this field progress, it is vital to strike a balance between innovation, ethics, and societal impact to fully harness the benefits of generative AI.

## REFERENCES

1. https://www.tandfonline.com/doi/abs/10.1 080/02564602.2014.987328
2. Generative adversarial network: An overview of theory and applications
3. Author links open overlay panelAlankrita Aggarwal a, Mamta Mittal b, Gopi Battineni
4. NVAE: A Deep Hierarchical Variational AutoencoderArash Vahdat, Jan KautzNVIDIA{avahdat, jkautz}@nvidia.com
5. https://www.eweek.com/artificial- intelligence/ generative-ai-and- cybersecurity/ #:~:text= Generative %20AI %20is%20an% 20emerging, imagery%2C %20audio%2C%20and%20video.
6. Education in the Era of Generative Artificial Intelligence (AI): Understanding the Potential Benefits of ChatGPT in Promoting Teaching and Learning
7. Generative Adversarial Network (GAN): a general review on different variants of GAN and applications
8. A Chat(GPT) about the future of scientific publishingElisa L. Hill-Yardin a,b,1, MarkR. Hutchinson c,d,1Robin Laycocka,1, Sarah J. Spencer a,*,1
9. https://www.linkedin.com/pulse/evolution -generative-ai-its-applications-tagx#:~:text= The% 20exciting %20journey %20of%20 generative, thought%20and%20 decision%2 Dmaking % 20processes.
10. https://generativeai.net/

INTERNATIONAL JOURNAL OF INNOVATIVE TECHNOLOGY AND CREATIVE ENGINEERING (ISSN:2045-8711)
VOL.13 NO.08 AUG 2023

# ECMDS: PLANT LEAF IMAGE DISEASES USING SEGMENTATION AND IMPROVED GLCM FEATURE EXTRACTION

**Dhana Priyadharsini K[1]**
Research Scholar, Avinashilingam University for Women, Coimbatore, Tamil Nadu, India
**Kalpana B[2]**
Professor, Avinashilingam University for Women, Coimbatore, Tamil Nadu, India

**Abstract — Productivity in agriculture is important for economic expansion. The presence of disease in plants is very widespread, which is one of the reasons why plant disease detection is important in the agricultural industry. It is possible to identify disease from a plant's leaves using a technique called plant disease detection. Analyzing data augmentation, segmentation and segmentation features data are some of the phases in the process of detecting plant diseases. This paper, introduced the Enhanced Color based Mean shift Disease Segmentation (ECMDS) and Improved GLCM Feature Extraction model for effective plant disease identification. In recent years, deep learning has substantially improved the accuracy of object detection and image categorization systems. The experimental results show that the proposed segmentation algorithm can attain an accuracy of 96.232% . The trials made use of the well-known Plant Village dataset, which contains 54,305 images of various plant disease varieties classified into 38 classes.**

Keywords: ECMDS, Plant Disease, GLCM

## I. INTRODUCTION

Due to the fact that India is an agricultural nation, agriculture is the most significant sector of the Indian economy. Agriculture is a major source of income for more than 80% of individuals. The work of identifying leaf diseases is crucial since the crop's overall production is decreased by a variety of illnesses. Additionally, it can be challenging for farmers to diagnose a particular disease. The automatic identification of plant diseases is a crucial area for research since it could be useful for monitoring vast fields of crops and automatically identifying diseases from symptoms that occur on plant leaves. Because of this, robot guidance for disease management is more accurate and reliable when plant illness is automatically detected with the aid of image processing technology [1].

Visual diagnosis of plant diseases is more time-consuming, less accurate, and only practicable in a few locations. However, using an automatic detection method will require less work, less time, and result in a higher degree of accuracy.

Brown and yellow spots, early and late scorch and other common bacterial, viral, and fungal diseases are all present in plants. Image processing is utilized to quantify the disease-affected region and ascertain how the hue of the afflicted area differs [2], [3], and [4].

Image segmentation is the process of dividing or classifying an image into various components. Various techniques exist today for doing picture segmentation, ranging from the straightforward thresholding method to sophisticated color image segmentation techniques. Normally, these components correlate to something that is simple for people to separate and see as a separate item. Many alternative techniques have been developed to segment photos since computers lack the ability to recognize items intelligently. Several features in the image are used in the segmentation process. This could be a section of a picture [5], [6], or color information. The segmentation of the plant village image dataset using Enhanced Color-based Mean shift segmentation is shown in this paper.

The shortcomings of manual segmentation methods are addressed by this automated solution. The image may have been downloaded from a web repository, taken with a high-resolution mobile phone camera, or with a normal digital camera. The system receives this image as input in order to acquire the leaf features. The system includes various processes, including segmentation, feature extraction, denoising and data augmentation [7].

Identification of different plant diseases is extremely important and a major concern. Making better management decisions for agricultural production may be aided by early diagnosis of plant diseases. There are frequently observable indications or spots on the flowers of diseased plants. In particular, every virus and pest problem leaves behind distinctive patterns that can be used to spot anomalies. Identification of a plant disease requires staff and expertise. Furthermore, manual examination might be biased and time-consuming when identifying the type of plant infection, and the

disease reported by farmers or specialists occasionally isn't real [8].

Researchers have created many solutions to the aforementioned issues. Machine learning can classify plant diseases using a variety of feature sets. The most popular feature sets among them are the classic hand-crafted and deep learning (DL)-based feature sets. Before characteristics, such as segmentation, can be successfully extracted, preprocessing is necessary [9]. Feature extraction may be followed by the usage of several classifiers. prominent classifiers include Deep CNN [10], K-nearest Neighbor [11], Support Vector Machine [12], rule generation [13-14], and Artificial Neural Networks (ANNs).

By separating a leaf from its background, segmentation can be useful in this situation. Additionally, this method can be applied when the classifier needs to be aware of the scene. This paper presents Enhanced Color-based Mean shift segmentation with Improved GLCM feature extraction technique is proposed here.

## II. REVIEW OF LITERATURE

In order to enable disease diagnosis from leaf images, Islam M and Dinh A, Wahid K, 2017 [15] suggested a strategy that combines image processing and machine learning. This automated process analyses a publicly accessible plant image database named “Plant Village” to identify illnesses (or their absence) on potato plants. They exhibit disease categorization over 300 photos with a 95% accuracy using their segmentation method and support vector machine.

According to (Boulent et al., 2019) [16], the applications that were presented might be used as a foundation for the creation of tools that provide specialist assistance or automatic screening. These resources could support more dependable food production and sustainable agriculture methods. They reviewed 19 researches that used CNNs to automatically identify crop illnesses in order to gauge the potential of these networks for such applications. They discussed their implementation strategies, performance, and profile details. Their study enables us to pinpoint the main problems and shortfalls of the works in this field of study.

Deep learning-based automatic image identification systems, particularly those based on convolutional neural networks, were thoroughly examined by Lee et al. (2020) [17]. The authors first investigated whether it would be more useful identifying plants as opposed to a test for recognizing objects in general. They specifically demonstrated, using visualization approaches, how the features taught vary depending on the strategy used and how they do not always concentrate on the area afflicted by the disease. In order to deal with diseases involving crops that are not represented in the training database, they proposed a more natural strategy that takes diseases into account separately from crops and shown that it is more effective than the conventional crop-disease pair approach.

Systematically examining the issue of visual illness detection for plant disease diagnosis was done by Liu et al. in 2020 [18]. Photos of plant diseases typically have randomly scattered lesions, a wide range of symptoms, and complex backgrounds, making them more difficult to distinguish from other types of images. They approached plant disease recognition using this dataset by reweighting visual regions and loss to emphasize sick sections.

The use of deep learning in plant disease recognition was discussed by Lili Li et al. in 2021 [19]. This method can avoid the drawbacks associated with the artificial selection of disease spot features, make the extraction of plant disease feature more objective, and accelerate the pace of research and technological advancement. In the past few years, deep learning technology has made significant strides in the field of identifying crop leaf diseases, according to their review. The authors discussed the existing methods for detecting plant leaf disease utilizing deep learning and cutting-edge imaging techniques, as well as the difficulties that remain.

Restructured residual dense network, which combines the benefits of deep residual networks and dense networks, was proposed by Changjian Zhou et al. in 2021 [20] for the identification of tomato leaf diseases. In order to increase calculation accuracy and the flow of information and gradients, this hybrid deep learning model can reduce the number of training process parameters. Since the original RDN model was developed for image super resolution, the network architecture for classification tasks must be restructured using modified input image characteristics and hyper parameters.

According to Punitha Kartikeyan and Gyanesh Shrivastava [21], classifying and identifying plant diseases are an important undertaking in the world of agriculture and have a substantial impact on crop yield and quality. Early plant disease diagnosis can help to lessen losses and boost crop output. To improve agricultural cultivation and raise crop production yield, accurate diagnosis and categorization of plant diseases were required; for this, an image-processing approach might be applied. To a greater or lesser extent (95.16 to

98.38%) than another model, the proposed hybrid feature extraction technology, which combines Discrete Wavelet Transform decomposition and Grey Level Co-Occurrence Matrices feature extraction with Support Vector Machine classifier, could identify and categories plant diseases.

## III. PROPOSED METHODOLOGY

The research approach takes into account the order in which all of the trials were carried out by implementing the plant village disease dataset [22]. The proposed Enhanced Color based Meanshift Disease Segmentation (ECMDS) and improved GLCM (Gray-Level Co-occurrence Matrix) Feature Extraction method successfully predicts the plant leaf diseases from plant images, yielding enhanced suitable for additional processing such as data augmentation. Figure 1 depicts the overall proposed ECMDS with Improved GLCM feature extraction flow diagram.

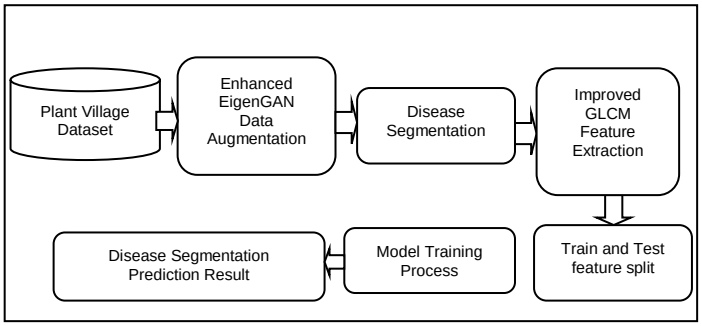



**Fig.1: Proposed Flow**

### Data Acquisition

The Plant Village dataset served as the source for the used data [22]. A database opens to the public that contains around 54,323 images of 14 different crops and 30 distinct plant diseases. Figure 2 depicts each example class. In order to better fit the input size of the network, the images were subsequently shrunk to 512 x 512 pixels. Every plant has at least two classes of 512 x 512 images—one class for sound leaves and the other for harmed leaves. The images were separated into training and testing groups according to the class divides that were decided upon for the train-test set.

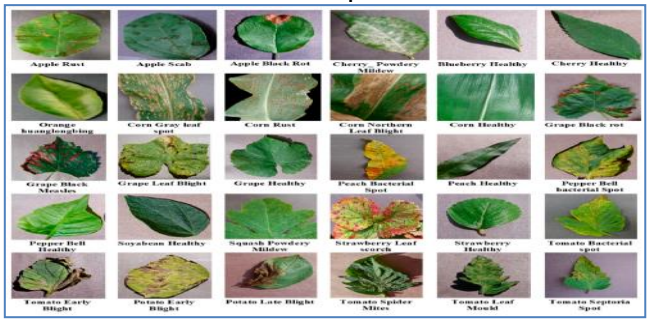


**Fig.2: From the Plant Village collection, examples of 38 distinct types of leaf diseases**

### Data Agumentation

In order to increase the amount of images in the training set for every of the plant village image datasets described in the preceding section, nine different data-augmentation methods are investigated. The Enhanced EigenGAN Data Augmentation was already we presented (Dhana Priyadharsini K and Kalpana B, 2022 []).

The Discriminator and Generator model input parameters of image size, noise dimension, base channels, and max channels will be used to train the proposed enhanced EigenGAN approach on a single CPU device with Eigenblock. ConvLayer is used in the Discriminator section to process the base and max channels with kernel size, and the blocks are then returned as output. The chunks are linearly appended using a sequential layer. According to the Generator model, the image size will grow by a factor of two dependent on the maximum channels and the quantity of blocks. The proposed Data augmentation results are illustrated in figures 3 and 4.

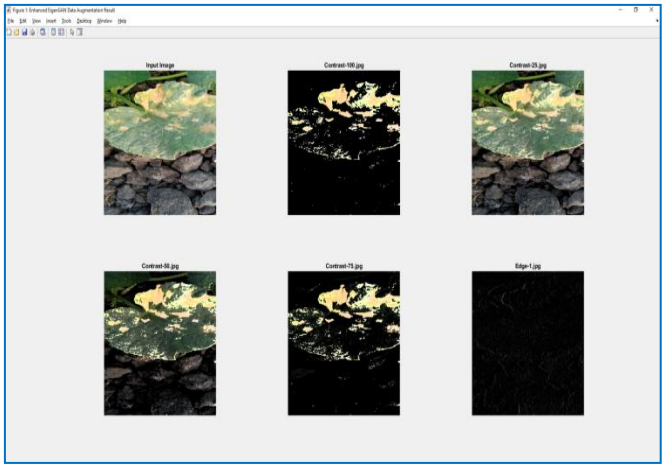


**Fig. 3: Enhanced EigenGAN Result -1**

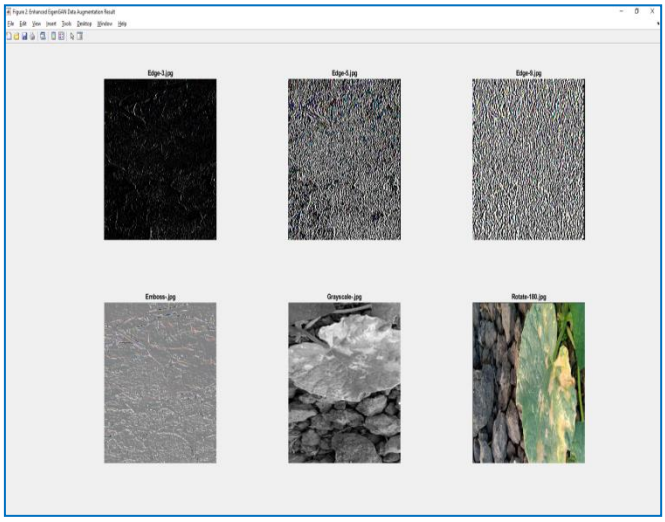


**Fig. 4: Enhanced EigenGAN Result -2**

### Enhanced Color based Meanshift Disease Segmentation (ECMDS)

The proposed enhanced color based Meanshift disease segmentation method is characterized a leaf disease portions extracted are discussed in this section. This technique works In this step, pixels primarily in green are hidden. This system calculated a threshold value for these pixels. Then, largely green pixels are covered up as follows: The red, green, and blue components of a pixel are each given a value of zero if the green component of that pixel's intensity is below a predetermined threshold value. In the case of processed images with 8 bits per channel and pixel, the color range is 0-255. Additionally, to speed up computation, images that are longer than 512 pixels are shrunk to 256 pixels. It is easy to later map the outcomes of the extracted leaf positions from the smaller-sized images to the locations of the original images.

Background removal is simple when using the right color histogram threshold to eliminate non-green backgrounds. The excess green index and excess red index of the RGB color space are computed, respectively. If their textures, like the grass, are extremely intricate, even green areas, like the grass, might become unattractive. The local entropy of grayscale values is used to calculate statistical unpredictability. If a window's local entropy exceeds a predefined cutoff value (>220), it is classified as the background for all three-by-three windows.

To choose color as the primary feature in this study since it is frequently the most prominent and distinctive visual component and suitable for a variety of segmentation tasks. A three-element color vector that describes the average color components is computed over the pixels in a region. According to the Meanshift disease segmentation method, an image is divided into $r$ regions $Rg_m$, $m$ = 1,…, $r$. For each area, the mean vector $V_{Rgm} = v_{1m}, v_{2m}, v_{3m}$ is computed, where $v_{1m}$, $v_{2m}$, and $v_{3m}$ are the mean pixel intensities of the $m^{th}$ region in the three distinct color spaces, respectively. The final ECMDS segmentation result is described in figure 5.

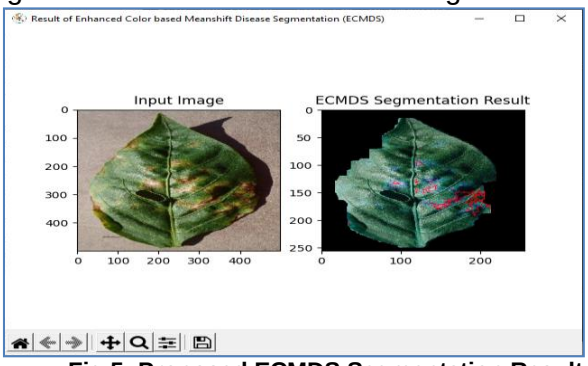



**Fig.5: Proposed ECMDS Segmentation Result**

### Improved GLCM Feature Extraction

The improved GLCM functions determine the frequency with which pairs of pixels with particular values and in a given spatial relationship appear in a picture, forming a matrix from which statistical measurements are subsequently extracted to explain the texture of the image. The statistical features are extracted area, perimeter, physiological_length, physiological_width, aspect_ratio, rectangularity, circularity, contrast, correlation, inverse_difference_moments, entropy, dissimilarity, homogeneity, Angular Second Moment (ASM) and energy. The final improved GLCM feature extraction results is shown in figure 6.

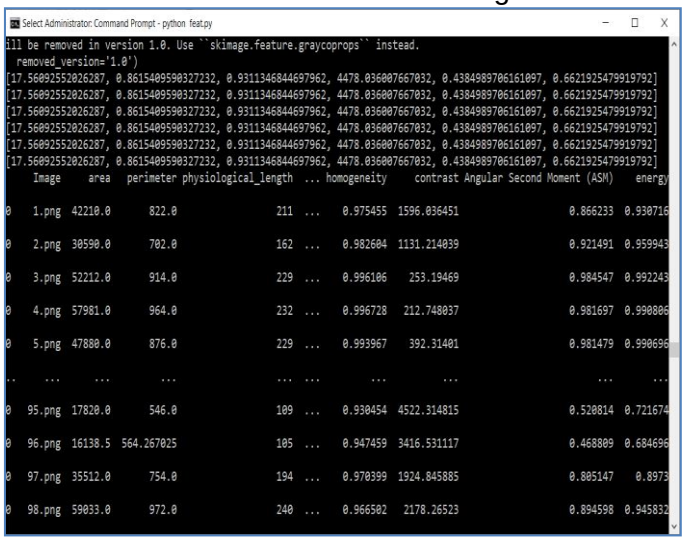



**Fig.6: Proposed Improved GLCM Feature Extraction Result**

## IV. RESULTS AND DISCUSSION

Results have been estimated using the method suggested. On a PC running Windows 10 with Python 3.8 simulations, the findings were implemented using an Intel I5-6500U series processor running at 2.71 GHz and 8GB of main memory. A dataset on plant diseases called Plant Village was released by Pennsylvania State University [22]. The trials made use of the well-known Plant Village dataset, which contains 54,305 images of various plant disease varieties classified into 38 classes. To evaluate segmentation accuracy with existing indices-based histogram segmentation [23] with proposed ECMDS segmentation method. The mean segmentation accuracy rate for the datasets from the Plant Village dataset is shown in Table 1.

**Table 2: Rates of mean Segmentation accuracy for the Plant Village dataset**

| Methods | Indices-based histogram segmentation | Proposed ECMDS Method |
|---|---|---|
| Accuracy | 92.06 | **96.231** |

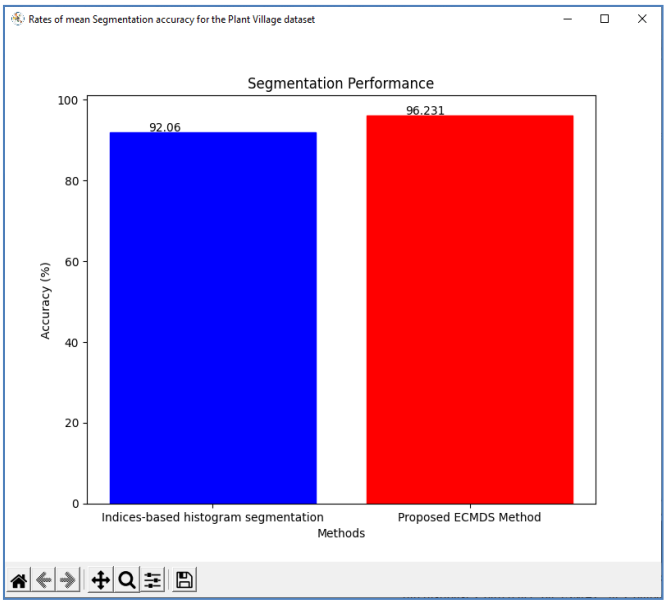



**Fig.7: Segmentation Accuracy Performance Chart**

According to Figure 7 the Proposed ECMDS method produced the best results for the Plant village dataset; with a mean segmentation accuracy rate is 96.231%.

## V. CONCLUSION

In the present work, an Enhanced Color based Meanshift Disease Segmentation (ECMDS) and Improved GLCM Feature Extraction is proposed for the detection and identification of plant village diseases. In order to achieve improved accuracy in terms of its segmentation and feature extraction capabilities performance, the model combines the advantages of proposed segmentation technique. The test results show that the model performs admirably on publicly accessible datasets from plantvillage that are of various sizes and contain images taken against a variety of backgrounds. Proposed ECMDS method obtains 96.23% overall mean accuracy on the PlantVillage datasets in the trial for the detection of plant diseases. In future, In order to produce reliable Transform based neural network models, to examine more tasks and take into account various intricate neural network and information extraction configurations.

## REFERENCES

[1] Sandesh Raut and Amit Fulsunge, “Plant Disease Detection in Image Processing Using MATLAB” International Journal of Innovative Research in Science, Engineering and Technology (IJIRSET), Volume 6, Issue6, June 2017.

[2] Ghaiwat Savita N, Arora Parul. Detection and classification of plant leaf diseases using image processing techniques: a review. Int J Recent Adv Eng Technol 2014;2(3):2347–812. ISSN (Online).

[3] Dhaygude Sanjay B, Kumbhar Nitin P. Agricultural plant leaf disease detection using image processing. Int J Adv Res Electr Electron Instrum Eng 2013;2(1).

[4] Bashir Sabah, Sharma Navdeep. Remote area plant disease detection using image processing. IOSR J Electron Commun Eng 2012;2(6):31–4. ISSN: 2278-2834.

[5] Beucher S, Meyer F. The morphological approach to segmentation: the watershed transforms. In: Dougherty ER, editor. Mathematical morphology image processing, vol. 12. New York: Marcel Dekker; 1993. p. 433–81.

[6] Bhanu B, Peng J. Adaptive integrated image segmentation and object recognition. IEEE Trans Syst Man Cybern Part C 2000;30:427–41.

[7] Ghulam Mustafa Choudhary and Vikrant Gulati “Advance in Image Processing for Detection of Plant Diseases” International Journal of Advanced Research in Computer Science and Software Engineering 5(7), [ISSN: 2277 128X], pp. 1090-1093, July- 2015.

[8] S. M. Coakley, H. Scherm, and S. Chakraborty, “Climate change and plant disease management,” Annual Review of Phytopathology, vol. 37, no. 1, pp. 399–426, 1999.

[9] Camargo, A.; Smith, J., "An image-processing based algorithm to automatically identify plant disease visual symptoms", Biosyst. Eng. 2009, 102, 9–21.

[10] Munisami, T.; Ramsurn, M.; Kishnah, S.; Pudaruth, S. "Plant leaf recognition using shape features and colour histogram with K-nearest neighbour classifiers", Procedia Comput. Sci. 2015, 58, 740–747.

[11] Singh, J.; Kaur, H. "Plant disease detection based on region-based segmentation and KNN classifier", In Proceedings of the International Conference on ISMAC in Computational Vision and Bio-Engineering 2018, Palladam, India, 16–17 May 2018; pp. 1667–1675.

[12] Camargo, A.; Smith, J. "Image pattern classification for the identification of disease causing agents in plants", Comput. Electron. Agric. 2009, 66, 121–125.

[13] Chaudhary, A.; Kolhe, S.; Kamal, R. "An improved random forest classifier for multi-class classification", Inf. Process. Agric. 2016, 3, 215–222.

[14] Phadikar, S.; Sil, J.; Das, A.K. "Rice diseases classification using feature selection and rule generation techniques", Comput. Electron. Agric. 2013, 90, 76–85. [CrossRef]

[15] Islam M and Dinh A,Wahid K (2017) Detection of potato diseases using image segmentation and multiclass support vector machine. In: 2017 IEEE 30th canadian conference on electrical and computer engineering (CCECE)

[16] J. Boulent, S. Foucher, J. Théau and P. St-Charles, "Convolutional Neural Networks for the Automatic Identification of Plant Diseases", Frontiers in Plant Science, vol. 10, 2019. Available: 10.3389/fpls.2019.00941.

[17] S. Lee, H. Goëau, P. Bonnet and A. Joly, "New perspectives on plant disease characterization based on deep learning", Computers and Electronics in Agriculture, vol. 170, p. 105220, 2020. Available: 10.1016/j.compag.2020.105220.
[18] Xinda Liu; Weiqing Min; Shuhuan Mei; Lili Wang; Shuqiang Jiang, "Plant Disease Recognition: A Large-Scale Benchmark Dataset and a Visual Region and Loss Reweighting Approach" , IEEE Transactions on Image Processing ( Volume: 30), 2021.
[19] Lili Li; Shujuan Zhang; Bin Wang, "Plant Disease Detection and Classification by Deep Learning—A Review", IEEE Access ( Volume: 9), 2021
[20] Changjian Zhou; Sihan Zhou; Jinge Xing; Jia Song, "Tomato Leaf Disease Identification by Restructured Deep Residual Dense Network", IEEE Access ( Volume: 9), 2021.
[21] Punitha Kartikeyan and Gyanesh Shrivastava, "Hybrid Feature Approach for Plant Disease Detection and Classification using Machine Learning", IEEE World Conference on Applied Intelligence and Computing (AIC), 2022.
[22] PlantVillage. Available online: https://plantvillage.psu.edu/ (accessed on 29 September 2022 ).
[23] Pushpa BR, Shree Hari AV and Adarsh Ashok, "Diseased Leaf Segmentation from Complex Background Using Indices Based Histogram", 6th International Conference on Communication and Electronics Systems (ICCES), IEEE, 2021.

INTERNATIONAL JOURNAL OF INNOVATIVE TECHNOLOGY AND CREATIVE ENGINEERING (ISSN:2045-8711)
VOL.13 NO.08 AUG 2023

# GCD CONGRUITY SIGNCRYPTION FOR PROTECTED COMMUNICATION WITH BIG DATA

**P.Suresh Babu**
Associate Professor and Head, Department of Computer Science,
Bharathidasan College of Arts and Science, Erode,Tamil Nadu, India
**S.Sangeetha**
Research Scholar & Assistant Professor, Department of Computer Science,
Bharathidasan College of Arts and Science, Erode,Tamil Nadu, India

**Abstract - Big data is a collection of huge data employed to examine and extract information from large datasets. With the generation of a large volume of data, It faces severe security risks and challenges such as data leakage, malicious use, etc. Many researchers carried out their research for performing protected data communication. But, the data confidentiality level was not improved by using existing cryptographic methods. Therefore, a GCD congruity signcryption for protected communication with Big data (GCDPC) technique is introduced for protected data communication with higher data confidentiality and lesser communication overhead.**

**The GCDPC technique performs two processes, namely data classification and protected communication. Initially, a number of data are collected from the big dataset. After data classification, protected data transmission is performed using Multiplicative congruity signcryption technique. Multiplicative congruity signcryption includes three processes namely Multiplicative Congruity Key Generation, Signcryption and Unsigncryption. Experimental evaluation of the proposed GCDPC technique is carried out with respect to classification accuracy and communication overhead, with a different number of data. The discussed results indicates that the performance of GCDPC technique increases classification accuracy and minimum overheadthan the other state-of-the-art methods.**

**Keywords:** Protected data communication, Key Generation, Signcryption, Multiplicative congruity signcryption

## I. INTRODUCTION

Nowadays, security and privacy issues of big data systems are noteworthy and require precise observation. However, the large volume of big data is susceptible to attacks in modern communication systems.

Big data security events occur recurrently in recent years. Hence, Privacy- preserving techniques have been developed to guarantee the authorized individual to access the private information for enhancing the data security in communication systems.

A novel security scheme known as Lightweight Hybrid Scheme (LHS) was introduced in [1] based on Elliptic Curve Cryptography and provides the secure exchange mechanism. However, it failed to define a novel scheme for big data encryption and decryption that leads to a higher level of security and withstand attacks existence during the communication. The blockchain-Based Trusted Service Evaluation (BCSE) model was designed in [2] over Big Data. But, it failed to ensure the stability of the performance, especially in a large and increasing number of data samples.

In [3], different methods of data mining algorithms were considered to ensure security. However, the efficient machine learning method was not applied for minimizing the overhead of secure data transmission. A big data provenance model (BDPM) was developed in [4] to provide the entire data transformation through dissimilar components of a big data system. But the performance of data confidentiality and integrity was not analyzed. A trusted third-party-aided searchable and verifiable data protection approach was introduced in [5] to enhance the integrity of uploaded or downloaded data at any time. But it failed to use the secure data sharing scheme with big data applications.

## II. RELATED WORKS

Homomorphic encryption algorithms were designed in [6] for big data security analysis while preserving privacy. But the higher data confidentiality was not achieved. Attribute-Based Honey Encryption (ABHE) scheme was introduced in [7] for performing encryption-decryption on different sizes of files. This encryption scheme failed to use more robust approaches to ensure optimum security of big data.A novel blockchain technology was introduced in [8] for big data security solutions by integrating fragmentation, encryption, and access control. But the higher data confidentiality level was not achieved.

Blockchain technology was introduced in [9] to handle the healthcare system by providing the solution. The designed scheme guaranteed security with a smart contract. But, the overhead was not reduced. An effective Secure Information Propagation (SIP) approach was developed in [10] for E- health Networks. The designed security analysis was not feasible to find that illegitimate user.

### A. ORGANIZATION OF PAPER

The rest of the paper is organized into five different sections as follows. Section 2 provides a brief description of the proposed GCDPC technique with a neat architecture diagram. In section 3, the performance results of the proposed technique and existing methods are discussed with different metrics. At last, Section 4 concludes the paper.

## III. PROPOSED METHODOLOGY

Big Data is a vast term for implementation with the huge volume and complex datasets. While the data set is bigger in volume and the conventional methods are insufficient to extract the important and accurate data from large unstructured data. With the help of classification methods, unstructured data is turned into structured form as a result that any user accesses the required data easily. These classification techniques are applied over the big databases to offer data services from the huge volume of data sets. In addition, big data security is the major concern to perform the data point communication. A traditional security technique provides slow performance and is time-consuming in a Big Data context. Based on the above motivation, a novel technique called GCDPC is introduced.

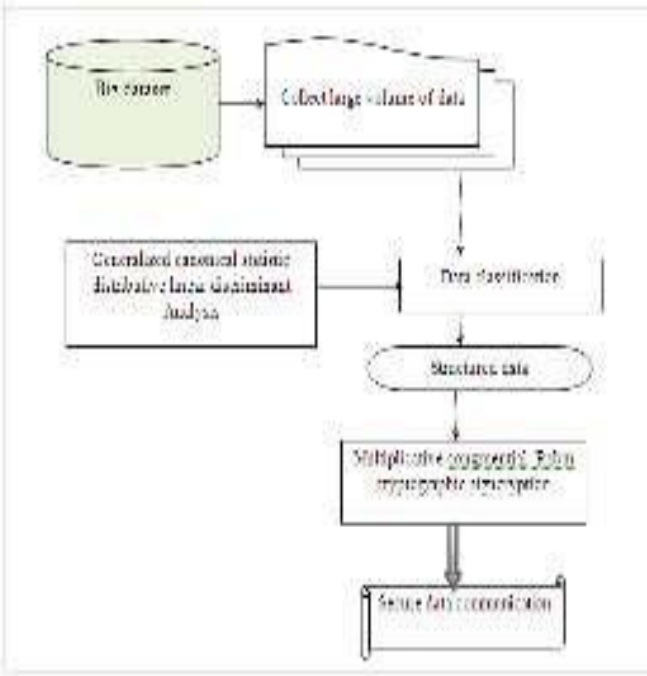


**Figure 1 Architecture of the Proposed GCDPC Technique**

The proposed GCDPC technique performs two major processes big data classification and secure communication. The classification process of the proposed GCDPC technique uses the machine learning technique called generalized canonical statistic distributive linear discriminant Analysis. Then the security of communication is achieved by applying Multiplicative congruity signcryption. With the use of the above-said two techniques, the architecture of the proposed GCDPC technique is constructed as given above.

Figure 1 demonstrates the structural representation of the proposed GCDPC technique that consists of two different main processes namely classification and secure data transmission with big data. Initially, the big dataset is considered. The proposed GCDPC technique first performs the classification of the data using Fisher's linear discriminant analysis. After classifying the data, secure data transmission from sender to receiver is said to be performed using Multiplicative congruity signcryption. These processes of the proposed GCDPC technique are briefly described in the following subsections.

### A. GENERALIZEDCANONICAL SDLD DATACLASSIFICATION

The first process of the proposed GCDPC technique is to perform the data classification using Generalized canonical statistic distributive linear discriminant analysis. The proposed linear discriminant analysis is a machine learning technique that helps to find the linear combination of data that distinguishes two or more classes using generalized canonical correlation and Hotelling t-squared statistic distribution. This helps to improve the data classification.

**Algorithm 1: Generalized canonical statistic distributivelinear discriminant data classification**

**Input:** Big dataset, Number of data$D_i = D_1, D_2, D_3 \dots D_n$

**Output:** Improve classification accuracy

**Begin**

1. **Define the number of classes** $C_j = C_1, C_2, \dots C_m$
2. **Calculate the mean of classes '**$m_c$**'**
3. **For each data** $D_i$ in dataset
4. **For each** mean **'**$m_c$**'**
5. Measure correlation '$G_c$'
6. **end for**
7. **end for**
8. Categorizes the data into a particular class
9. Measure variation within class'$Cov$(w)'
10. Measure variation between class '$Cov(b)$'
11. Obtain (classification results)

**End**

Algorithm 1 given above describes the step-by-step process of big data classification using a generalized canonical statistic distributive linear discriminant classifier. First, initializes the number of classes. Then the mean is estimated based on the number of data. Then the correlation between the mean of class and the data is measured. Based on the correlation measure, the data is classified correctly into a particular class. The result indicates that the variance within the class is minimized and maximizes the variance between the classes.

### B. MULTIPLICATIVE CONGRUITY CRYPTOGRAPHIC SIGNCRYPTION

After the data classification, the proposed GCDPC technique performs the secure data transmission from sender to receiver using a multiplicative congruity cryptographic signcryption scheme. The Multiplicative congruity Rabin cryptographic signcryption scheme consists of three major processes namely Key Generation, Signcryption, and Unsigncryption to increase the performance of secure data transmission.

**Multiplicative congruity Key Generation**

The proposed technique initially performs the key pair generation such as private and public keys. In a public key cryptographic, the multiplicative congruity Key Generation is used for the particular communication session. These keys are disabled after the session is completed. This helps to avoid unauthorized access hence it improves the confidentiality level. The private key is kept secret whereas the public key is distributed for further processing. By applying the multiplicative congruity linear pseudorandom number gene.

**Signcryption**

Once the pair of keys is generated, the proposed GCDPC technique simultaneously performs encryption digital signature generation using GCDPC technique.

**Unsigncryption**

Finally, the proposed technique performs the unsigncryption that comprises signature verification and decryption at thereceiver end.

## IV. PERFORMANCE RESULTS AND DISCUSSION

In this section, the performance of the GCDPC technique and existing methods LHS [1] and BCSE [2] are discussed with different metrics namely Classification accuracy and Communication overhead. The performance of two metrics is analyzed with the help of a table and graphical representation.

### A. IMPACTS OF CLASSIFICATION ACCURACY

The classification accuracy is defined as the ratio of a number of data that are correctly classified to the number of data. The formula for calculating the accuracy is given below,

$$Acc = \left[\frac{ncc}{n}\right] * 100 \quad (1)$$

From (1), $Acc$ represent a classification accuracy, '$ncc$' indicates the number of data correctly classified, $n$ denotes the total number of data taken as input. The classification accuracy is measured in terms of percentage (%).

**Signcryption**

Once the pair of keys is generated, the proposed GCDPC technique simultaneously performs encryption digital signature generation using GCDPC technique.

**Unsigncryption**

Finally, the proposed technique performs the unsigncryption that comprises signature verification and decryption at thereceiver end.

## V. PERFORMANCE RESULTS AND DISCUSSION

In this section, the performance of the GCDPC technique and existing methods LHS [1] and BCSE [2] are discussed with different metrics namely Classification accuracy and Communication overhead. The performance of two metrics is analyzed with the help of a table and graphical representation.

### B. IMPACTS OF CLASSIFICATION ACCURACY

The classification accuracy is defined as the ratio of a number of data that are correctly classified to the number of data. The formula for calculating the accuracy is given below,

$$Acc = [\frac{ncc}{n}] * 100 \quad (1)$$

From (1), $Acc$ represent a classification accuracy, '$ncc$' indicates the number of data correctly classified, $n$ denotes the total number of data taken as input. The classification accuracy is measured in terms of percentage (%).

**Table 1 Comparison of Classification Accuracy**

| Number of data | Classification accuracy (%) | | |
|---|---|---|---|
| | GCDPC | LHS | BCSE |
| 1600 | 93.75 | 88.12 | 86.25 |
| 3200 | 93.12 | 89.06 | 85.93 |
| 4800 | 94.16 | 90 | 86.45 |
| 6400 | 93.9 | 89.37 | 86.09 |
| 8000 | 94.75 | 91.87 | 88.12 |

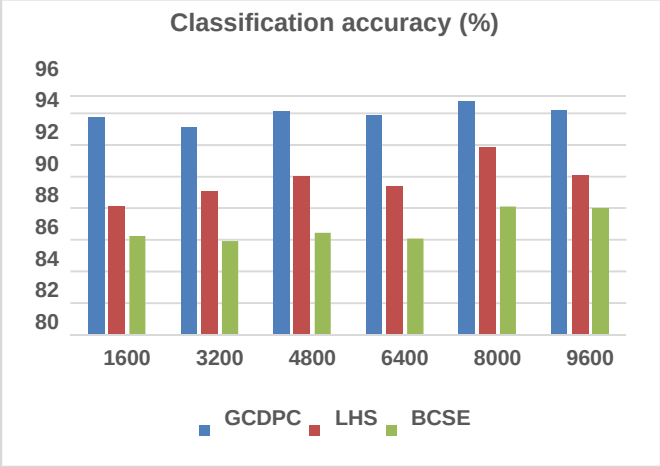



**Figure 2 Classification Accuracy Performance Comparison**

Figure 2 represents the graphical illustration of classification accuracy for five different numbers of data. As exposed in figure 2, numbers of data are represented on the '$x$ 'axis, and the accuracy of different methods is observed at the '$y$' axis. The graphical illustration indicates that the classification accuracy of the GCDPC technique is found to be improved than the other two existing methods.

### C. IMPACT OF COMMUNICATION OVERHEAD

The communication overhead is measured as the amount of time taken by the algorithm for secure data transmission from sender to receiver. The formula for calculating the communication overhead is expressed as given below,

$$CO = n * T(SDT) \quad (2)$$

Where $CO$ denotes a communication overhead, $n$ denotes the number of data, $T$ denotes a time for secure data transmission ($SDT$). It is measured in terms of milliseconds (ms).

**Table 2 Comparison of Communication Overhead**

| Number of data | communication overhead (ms) | | |
|---|---|---|---|
| | GCDPC | LHS | BCSE |
| 1600 | 27.2 | 31.2 | 35.2 |
| 3200 | 35.2 | 38.4 | 41.6 |
| 4800 | 39.84 | 43.2 | 45.6 |

technique is 93.75% and the accuracy of existing LHS [1] and BCSE [2] was found to be 88.12%, 86.25% respectively. Followed by, nine different performance results are observed for each method. Transmission was found to be '27.2$ms$'. However, the time consumption of existing LHS [1] and BCSE [2] was found to be 31.2$ms$' and 35.2$ms$. The observed results designate that the GCDPC technique decreases the communication overhead. After obtaining the six results, the overall time consumption of the GCDPC technique is reduced when compared to the existing results.

Figure 3 reveals the communication overhead involved in secure data transmission from the sender to the receiver. From Figure 3, it is obvious that the proposed method showed better performance over the other methods.

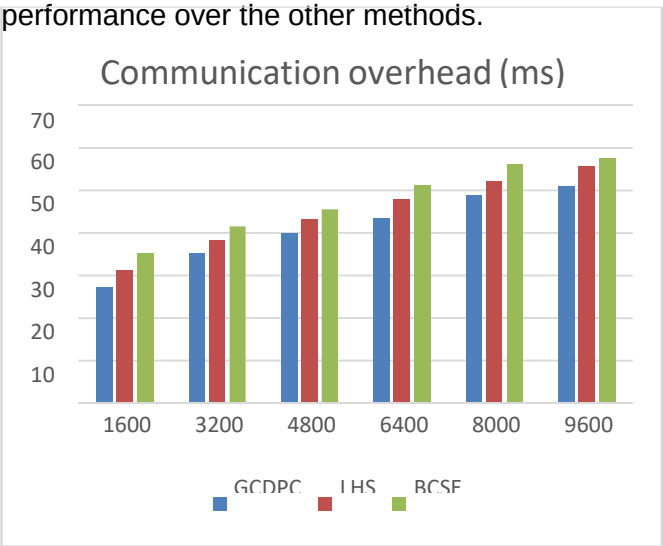



**Figure 3 Communication Overhead Performance Comparison**

## V. CONCLUSION

In this paper, a novel technique called GCDPC is implemented to deal with big data security. The technique is processed based on generalized canonical statistic distributive linear discriminant analysis and Multiplicative congruity Rabin cryptographic signcryption. First, the big data are taken as input, and perform the data classification to obtain the structured format to minimize the time consumption of the secure data communication. After the classification, the GCDPC performs the Multiplicative congruity Rabin cryptographic signcryption to improve the data confidentiality and integrity. In this way, secure big data transmission is performed.

The experimental assessments are carried out with respect to a number of data samples and compare the results of the proposed technique with two existing algorithms. The observed numerical results have confirmed that the proposed GCDPC technique provides improved performance results in terms of classification accuracy and communication overhead than the other cryptographic techniques.

## REFERENCES

[1] Mallepalli Prasanna Kumari and Tumma Srinivasa Rao, "A lightweight hybrid scheme for security of big data", Materials Today: Proceedings, Elsevier, March 2021, Pages 1-15. https://doi.org/10.1016/j.matpr.2021.03.151

[2] Fengyin Li, Xinying Yu, Rui Ge, Yanli Wang, Yang Cui, and Huiyu Zhou, "BCSE: Blockchain-Based Trusted Service Evaluation Model over Big Data", Big Data Mining and Analytics , Volume 5, Issue 1, 2022, Pages 1 — 14. DOI: 10.26599/BDMA.2020.9020028

[3] Veselska Olga, Ziubina Ruslana, Finenko Yuriy , Nikodem Joanna, "Big Data Analysis Methods Based on Machine Learning to Ensure Information Security", Procedia Computer Science, Volume 192, 2021, Pages 2633-2640. https://doi.org/10.1016/j.procs.2021.09.033

[4] Yuanzhao Gao, Xingyuan Chen, Xuehui Du, "A Big Data Provenance Model for Data Security Supervision Based on PROV-DM Model", IEEE Access, Volume 8, 2020, Pages 38742 38752.DOI:10.1109/ACCESS.2020.2975820

[5] Jian Shen, Member, Chen Wang, Anxi Wang, Sai Ji, and Yan Zhang, "A Searchable and Verifiable Data Protection Scheme for Scholarly Big Data", IEEE Transactions on Emerging Topics in Computing , Volume 9, Issue 1, 2021, Pages 216 — 225. DOI: 10.1109/TETC.2018.2830368

[6] Rafik Hamza, Alzubair Hassan, Awad Ali , Mohammed Bakri Bashir, Samar M. Alqhtani, Tawfeeg Mohmmed Tawfeeg and Adil Yousif, "Towards Secure Big Data Analysis via Fully Homomorphic Encryption Algorithms", Entropy, Volume 24, Issue 4, 2022, Pages 1-17. https://doi.org/10.3390/e24040519

[7] Gayatri Kapil1, Alka Agrawal1, Abdulaziz Attaallah2, Abdullah Algarni, Rajeev Kumar and Raees Ahmad Khan, "Attribute based honey encryption algorithm for securing big data: Hadoop distributed file system perspective", PeerJ Computer Science, Volume 6, 2020, Pages http://doi.org/10.7717/peerj-cs.259

[8] Hanan E. Alhazmi, Fathy E. Eassa; Suhelah M. Sandokji, “Towards Big Data Security Framework by Leveraging Fragmentation and Blockchain Technology”, IEEE Access, Volume 10, 2022, Pages 10768 — 10782. DOI:10.1109/ACCESS.2022.3144632

[9] Pratima Sharma, Malaya Dutta Borah and Suyel Namasudra, “Improving security of medical big data by using Blockchain technology”, Computers & Electrical Engineering, Elsevier, Volume 96, Part A, December 2021,Pages1-15. https://doi.org/10.1016/j.compeleceng.2021.107529

[10] Liping Zhang, Zhen Wei, Wei Ren, Xianghan Zheng, Kim- Kwang Raymond Choo, Neal N. Xiong, “SIP: An Efficient and Secure Information Propagation Scheme in E-Health Networks”, IEEE Transactions on Network Science and Engineering ,Volume 8, Issue 2, 2021,Pages 1502—1516. DOI: 10.1109/TNSE.2021.3063174.

# A STUDY ON DISTANCE IN GRAPH THEORY AND ITS APPLICATION

**D. Christia Jebakumari**[1]
Assistant Professor, Department of Mathematics, Sarah Tucker College (Autonomous),
Affiliated to Manonmaniam Sundaranar University, Tirunelveli,Tamill Nadu, India
**J. Vijaya Xavier Parthipan**[2]
Associate Professor, Department of Mathematics, St. John's College, Palayamkottai- 627002
Affiliated to Manonmaniam Sundaranar University, Tirunelveli,Tamill Nadu, India
**A. Revathi**[3]
Research Scholar(Reg.No:19211272092014), Department of Mathematics,
St. John's College, Palayamkottai- 627002
Affiliated to Manonmaniam Sundaranar University, Tirunelveli,Tamill Nadu, India

**Abstract- Graph theory is applicable in the section of computer science. It relates to the multiple fields of engineering, and geographical fragmentation. The concept of distance and time with a relational strategy to road networking is discussed with proper evaluation. The purpose of the paper is to discuss and understand the impact of distance in graph theory and to evaluate the impact of graph theory in the road mapping framework.**

## 1. INTRODUCTION

In the domains of computers, the field of engineering, and physical research, graphs is used to represent a variety of interactions and processes. These are employed in a wide range of basic and applied science fields, including information systems, the social sciences, and a number of others. In a variety of scientific domains, researchers use graph topologies to model the difficulties they face.

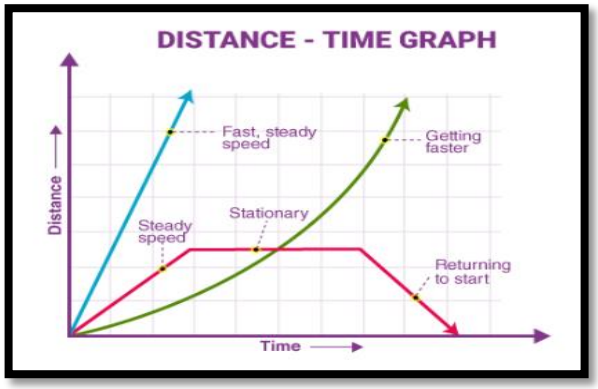



**Figure 1: Concept of Distance and Time graph**
(Source: Toppr, 2022)

Figure 1 states the concept of Distance and Time in Graph Theory with ease (Toppr, 2022). The slope line near the straight one depicts the speed of the elements. The uniformity in the motion is suggested by the straight line. The structure of a graph in which attributes are connected to the vertices and edges is referred to as a network of vertices. Image processing, optimizing, networking, recognition of patterns, and navigation often make use of distance in graphs.

The primary objectives of this research aim to make a representation of data that are essential for making a bond between dependent and independent variables.

**RO1** To understand the conceptual context of distance and time relation in the component's journey

**RO2** To evaluate the relation of graph theory in a road mapping scenario

**RO3** To create the framework of graphical networking with the bond of speed and distance

**RO4** To understand the area measurement in correspondence to distance traveling in a road network

The probable research questions of this topic include

**RQ1** What are the conceptual thought processes of the distance and time relation in graph theory?

**RQ2** What is the significance of graph theory in topographical road mapping structure?

**RQ3** How can speed and distance impact the graphical representation in graph theory?

**RQ4** What is the evaluation of understanding the area measurement in road networking?

## Discussion on the concepts of the relation of distance and time in graph theory

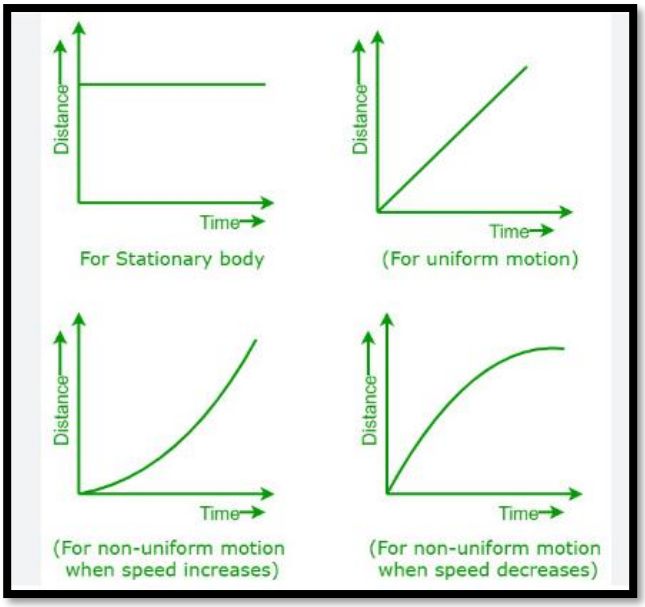



**Figure 2: Relational concept of Time and Distance**
(Source: Geeksforgeeks, 2022)

Figure 2 represents the concepts of time and distance with physical measurements (Geeksforgeeks, 2022). This figure dictates the time span through the horizontal line and the distance through the vertical line. As stated by Nosirov, Norov & Tashmetov (2022) this visual representation of how far a body has traveled over a predetermined period of time is known as a distance-time graph. The Y-axis is used to represent the connection between space and time, while the X-axis is employed to represent the passage of time. Understand the value of distance-time graphs first. As opined by Erb (2020) the formation and of these actor's bonding can be depicted through different clusters. First, the stationary speed of the component; second, the elemental body moving at a uniform speed; third, moving elements with accelerating speed; and moving with a decreased speed of individual bodily elements.

## Discussion on the significance of graph theory in topographical road mapping structure

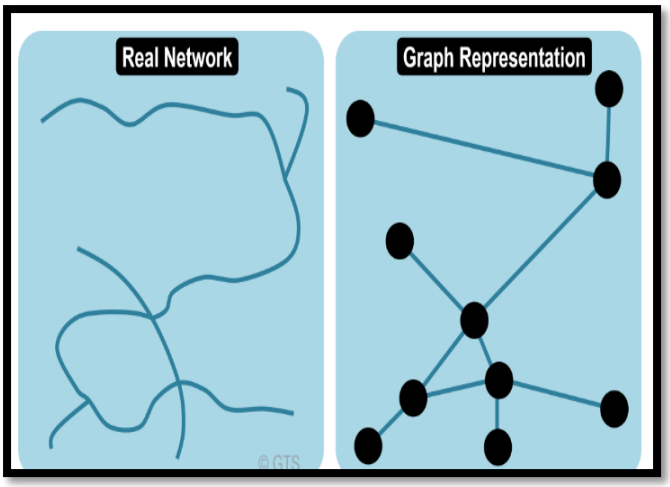



**Figure 3: Graphical representation of a road network**
(Source: Transportgeography, 2023)

Figure 3 represents the graphical structure of a road structure for real (Transport geography, 2023). The process of converting a planner graph into practical networking has to go through some significant processes. The primary point of networking is to make nodes at every juncture or intersection of the terminal. The later part of the process deals with joining and linking the nodes with straight segments of lines in a road network. As viewed by Guze (2019) the outcome of this abstraction is the network's actual structure, which appears in the above image. Depending on the level of detail, the real network can be challenging to understand when it involves connected frameworks. The easiest approach to understanding a network's interaction is through a graph representation.

| The attributes of road network | Activities of Nodes |
|---|---|
| Non-junctional node | It can represent the segments of shifting in 2-lane to 4-lane road mapping |
| Dummy node | A dummy node can be used for creative justification. It helps to compare the shape of the graphical project with the practical network. |

**Table 1: Road networking attributes and Node's activities**
(Source: Influenced by Saleem & Weng, 2022)

Table 1 discusses the geographical attributes of road networking along with node joints (Saleem & Weng, 2022). It is possible for each node's relative location to stay close to that of its counterpart in reality. Comprehending the geographical context of the network and the counterpart of road network mapping the attributes are important.

### Evaluation of the significance of area measurement in Topological graph theory

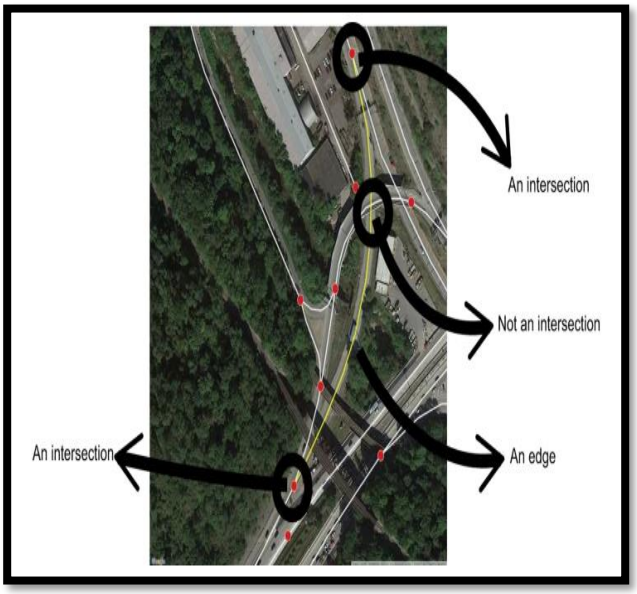



**Figure 4: Area measurement theory in Topological graph**
(Source: Anas, 2020)

Figure 4 depicts the area perspective of geographical locations in a topological view (Anas, 2020). It depicts that lead to the introduction of graph theory and the subsequent genre of the branch of geographical topology. As opined by Anna (2022) Topological Data Analytics (TDA) is a mixture of mathematical and geographical methodology that represents a useful computational framework and high dimensional data structure. Numerous spatially interconnected groups of constituents make up many geographical systems. As stated by Ma et al. (2020) therefore, by viewing these studies through the lens of graphs and networks, one can shed light on many intricate aspects of the elements that are built upon them, from analyzing and understanding human activities to studying the relationship between different urban regions.

## Methodology

The research survey follows the primary quantitative data collection. As opined by Saleem & Weng (2022) the significance of collecting fresh data is the first and foremost duty for analysis. The data is based on the responses of 80 participants around different geographic locations. The targeted participants of the project are the engineers of road networking frameworks. This methodology helps the research to get primary data for solving all topic-related objectives and questions. In this research Positivism philosophy and Deductive approach are followed for the methodology part. The whole process of analysis goes through the calculative methods of SPSS software. All the responses are analyzed in the SPSS software with proper steps and complex mathematical processes. As opined by Roy & Kesselman (2021) the hypothesis process is justified by the successful result as the outcome. All the collected data has helped a lot in the evaluation of the study.

## Findings

### Hypothesis testing

### Hypothesis 1

**H1.** Distance Graph theory helps to provide the natural motion of objects at the time

**H0.** Distance Graph theory does not help to provide the natural motion of objects at time

### Hypothesis 2

**H1** Distance graph theory can provide proper information regarding the velocity of particles

**H0** Distance graph theory cannot provide proper information regarding the velocity of particles

### Hypothesis 3:

**H1** Distance graph theory helps to analyze road mapping network

**H0** Distance graph theory does not help to analyze road mapping network

## Demographic data

### Gender

1. What is your Gender

| | | Frequency | Percent | Valid Percent | Cumulative Percent |
|---|---|---|---|---|---|
| Valid | | 10 | 11.1 | 11.1 | 11.1 |
| | Female | 36 | 40.0 | 40.0 | 51.1 |
| | Male | 31 | 34.4 | 34.4 | 85.6 |
| | Prefer not to say | 13 | 14.4 | 14.4 | 100.0 |
| | Total | 90 | 100.0 | 100.0 | |

**Table 2: Gender Analysis** (Source: SPSS)

Table 1 shows the classification of gender for the research analysis. This division shows gender division in cumulative and valid percentages.

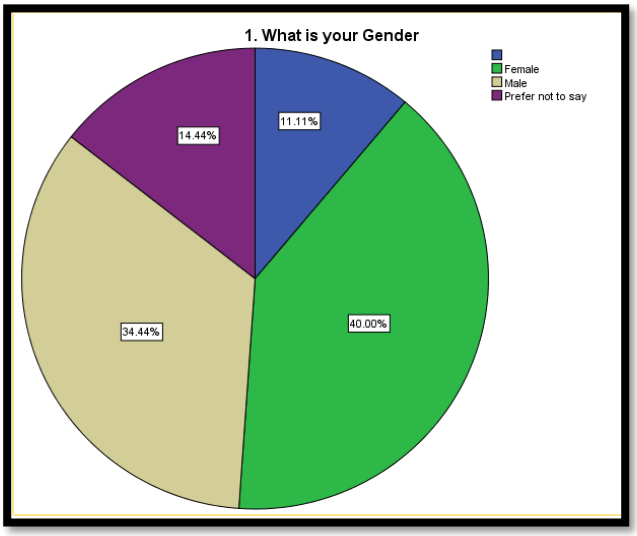



**Figure 5: Gender Analysis**
(Source: SPSS)

Figure 5 describes the sectional division of ages of the respondents of the survey. The female section holds 40% of the responses of all the participants. Whereas the male section covers 34.44% and 14.44 % of the respondents do not want to share their gender. As discussed by Yeh (2021), this pie chart describes the participation of individual gender tha help to maintain harmony in this analysis.

### Age

2. What is your age?

| | | Frequency | Percent | Valid Percent | Cumulative Percent |
|---|---|---|---|---|---|
| Valid | | 10 | 11.1 | 11.1 | 11.1 |
| | 10 to 20 | 18 | 20.0 | 20.0 | 31.1 |
| | 20 to 35 | 23 | 25.6 | 25.6 | 56.7 |
| | 35 to 50 | 21 | 23.3 | 23.3 | 80.0 |
| | Above 50 | 18 | 20.0 | 20.0 | 100.0 |
| | Total | 90 | 100.0 | 100.0 | |

**Table 3: Age analysis** (Source: SPSS)

Table 3 consists of the grouping of age of the participants with the percentile nature classifies the age group in a valid and cumulative format.

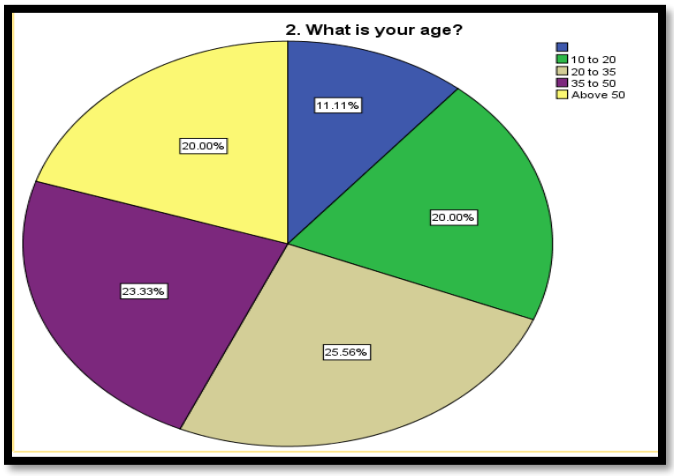



**Figure 6: Age Analysis**
(Source: SPSS)

Figure 6 discusses the age group of the respondents of the analysis. The age analysis is divided into four different stages of age starting from 10 to gradually ascending order of 50 above. People with the age group of 20-35 holds 25.56% of the overall respondents. The group of people between the ages of 35 to 50 presents 23.33% of the analysis. In addition the people of both the group of 10 to 20 and above 50 ages holds 20% each. As per the opinion of Peer (2021) the age analysis of each individual is divided into adolescence and adulthood. This approach assists researchers in keeping the focus on age differences in the context of the hypothesis.

## Monthly Income

3. What is your monthly Income?

| | Frequency | Percent | Valid Percent | Cumulative Percent |
|---|---|---|---|---|
| Valid | 10 | 11.1 | 11.1 | 11.1 |
| 20000 to 35000 | 23 | 25.6 | 25.6 | 36.7 |
| 35000 to 50000 | 39 | 43.3 | 43.3 | 80.0 |
| Above 50000 | 18 | 20.0 | 20.0 | 100.0 |
| Total | 90 | 100.0 | 100.0 | |

**Table 4: Analysis of monthly income**
(Source: SPSS)

Table 4 describes the monthly incomes of the participants of the research analysis. The segmentation shows the results in cumulative and valid percentages of hypothesis.

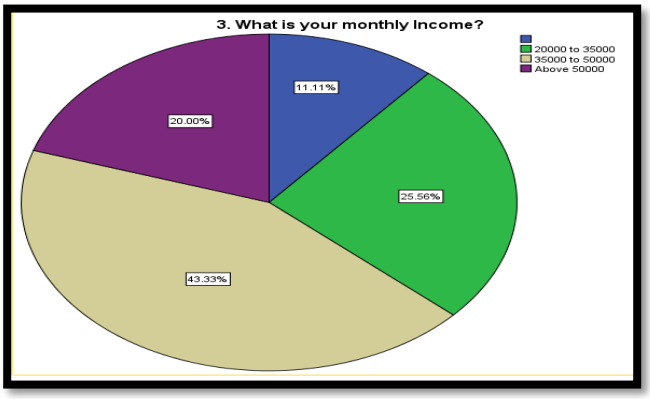



**Figure 7: Analysis of monthly income**
(Source: SPSS)

Figure 7 deals with the division of monthly income of all the respondents who participated in the response survey. This figure shows the sections of the monthly income of the respondents into 4 different parts. A maximum of 43.33% of people belong to the group with a monthly income of 35000 to 50000. 25.26% of the whole participants have a monthly income of 20000 to 35000. The significant group of people who have rs.50000 and above as their monthly income are 20% of the overall participants of this research survey. As commented by Miao (2019) monthly income depicts the class division in the current society as people with different earnings cannot have the same opinion on a single topic as their mental thinking capacity and social awareness.

## Descriptive Analysis
## Hypothesis 1

Model Summary[b]

| Model | R | R Square | Adjusted R Square | Std. Error of the Estimate | Change Statistics | | | | | Durbin-Watson |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | R Square Change | F Change | df1 | df2 | Sig. F Change | |
| 1 | .547[a] | .299 | .290 | 1.54581 | .299 | 33.235 | 1 | 78 | .000 | 2.707 |

ANOVA[a]

| Model | | Sum of Squares | df | Mean Square | F | Sig. |
|---|---|---|---|---|---|---|
| 1 | Regression | 79.416 | 1 | 79.416 | 33.235 | .000[b] |
| | Residual | 186.384 | 78 | 2.390 | | |
| | Total | 265.800 | 79 | | | |

Coefficients[a]

| Model | | Unstandardized Coefficients | | Standardized Coefficients | t | Sig. |
|---|---|---|---|---|---|---|
| | | B | Std. Error | Beta | | |
| 1 | (Constant) | 2.389 | .827 | | 2.889 | .005 |
| | IV1 | .450 | .078 | .547 | 5.765 | .000 |

**Table 5: Hypothesis 1**
(Source: SPSS)

Table 5 discusses the results of the first hypothesis analysis with the regression values. This hypothesis contains 3 different processes and expresses the data values in a model summary, ANOVA, and coefficients. The significant value of the first hypothesis analysis is 0.00 which means the variables of this hypothesis are correlated and share a strong bond among them.

## Hypothesis 2

Model Summary[b]

| Model | R | R Square | Adjusted R Square | Std. Error of the Estimate | Change Statistics | | | | | Durbin-Watson |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | R Square Change | F Change | df1 | df2 | Sig. F Change | |
| 1 | .158[a] | .025 | .013 | 1.82277 | .025 | 2.000 | 1 | 78 | .161 | 2.839 |

ANOVA[a]

| Model | | Sum of Squares | df | Mean Square | F | Sig. |
|---|---|---|---|---|---|---|
| 1 | Regression | 6.646 | 1 | 6.646 | 2.000 | .161[b] |
| | Residual | 259.154 | 78 | 3.322 | | |
| | Total | 265.800 | 79 | | | |

Coefficients[a]

| Model | | Unstandardized Coefficients | | Standardized Coefficients | t | Sig. |
|---|---|---|---|---|---|---|
| | | B | Std. Error | Beta | | |
| 1 | (Constant) | 5.630 | 1.025 | | 5.495 | .000 |
| | IV2 | .127 | .090 | .158 | 1.414 | .161 |

**Table 6: Hypothesis 2**
(Source: SPSS)

Table 6 states the successful results of the second hypothesis. The regression value of this result is 0.161 which is higher than the general parameter 0.05 in the hypothesis. It means the variables do not share a strong bond among themselves. As stated by Merschman (2020) the relationship of these variables could connect to a relationship if it is a relational analysis of the hypothesis.

### Hypothesis 3

**Model Summary[b]**

| Model | R | R Square | Adjusted R Square | Std. Error of the Estimate | Change Statistics | | | | | Durbin-Watson |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | R Square Change | F Change | df1 | df2 | Sig. F Change | |
| 1 | .373[a] | .139 | .128 | 1.71272 | .139 | 12.611 | 1 | 78 | .001 | 2.651 |

**ANOVA[a]**

| Model | | Sum of Squares | df | Mean Square | F | Sig. |
|---|---|---|---|---|---|---|
| 1 | Regression | 36.994 | 1 | 36.994 | 12.611 | .001[b] |
| | Residual | 228.806 | 78 | 2.933 | | |
| | Total | 265.800 | 79 | | | |

**Coefficients[a]**

| Model | | Unstandardized Coefficients | | Standardized Coefficients | t | Sig. |
|---|---|---|---|---|---|---|
| | | B | Std. Error | Beta | | |
| 1 | (Constant) | 4.325 | .791 | | 5.468 | .000 |
| | IV3 | .383 | .108 | .373 | 3.551 | .001 |

**Table 7: Hypothesis 3**
(Source: SPSS)

Table 7 analyzes the results of the third hypothesis. The significant value of regression is 0.01 which means the variables share a strong connection among themselves. As stated by Zhang, Cheng & Ren (2019) the statistical value of any hypothesis test goes through several mathematical calculations that have to be programmed manually and with accuracy to expect the proper results.

## Discussion

The discussion part of this research work consists of the overview evaluation of the relational impact of distance with graph theory. As commented by Virmani (2021) the mathematical educational process liked with the topological network system helps to create a practical networking system. The practical application of Distance segment in graph theory has gone far beyond than its traditional implication of limited usage.

The proposition of graph theory can be accessed by lining the concepts and ideas of different activities in different fields rather than engineering and demographics. This research study not only discusses the propaganda of the relevance of distance in graphical points but it also suggests the expansion of the usage of this particular . The significance of the Graph theory concept is used in almost every area of interest and for many reasons in daily life. As stated by Majeed & Rauf (2020) in the real world the theory is implied in connection with maps, social media, roads, and city planning. Apart from the road network, technological advancement helps to deliver real-time features with color classification maps with the congestion of roads. Graph theory can be additionally used to predict how diseases move across borders or across cities. In medicine and biology, graph theory is used to distinguish between pharmacological targets and to determine the quality of ambiguous capacity.

## Conclusion

In the concluding part, it can be stated that graph theory concepts are fundamental in SN analysis and modeling. In algorithmic execution, for example, graphs may be utilized to explain statements and their relationships, the execution order of the algorithm's statements, the control flow, the information flow, and the algorithm testing. The scope of graphs in many real-world issues can be exceedingly big, frequently encompassing over and services, happenings and operations, and trifles. It has been employed to organize the web according to interest instead of indexing only web pages with an infinite number of edges and nodes.

## References

1. Ahmed, H. (2019). Graph routing problem using Euler's theorem and its applications.Retrieved from: https://osf.io/gn8as/download (Retrieved on: 24th June, 2023)
2. Ajibade, C. A Survey of Graph Coloring Algorithm for Solving Constraint-Bound Problems.Retrieved from: https://cs.slu.edu/~chambers/spring20/advancedDS/homework/Cletus_Ajibade%20Proposal.pdf (Retrieved on: 24th June, 2023)

3. Anas AIT AOMAR (2020) https://towardsdatascience.com/notes-on-graph-theory-centrality-measurements-e37d2e49550a
4. Anna Jażdżewska, I. (2022). Use of graph theory to study connectivity and regionalisation of the Polish urban network. *Area*, *54*(2),290-303.Retrieved from: https://rgs-ibg.onlinelibrary.wiley.com/doi/pdf/10.1111/area.12774 (Retrieved on: 24th June, 2023)
5. Chung, M. G., Herzberger, A., Frank, K. A., & Liu, J. (2020). International tourism dynamics in a globalized world: A social network analysis approach. *Journal of Travel Research*, *59*(3), 387-403.Retrieved from: https://par.nsf.gov/servlets/purl/10194534 (Retrieved on: 24th June, 2023)
6. Erb, W. (2020). Shapes of uncertainty in spectral graph theory. *IEEE Transactions on Information Theory*, *67*(2), 1291-1307.Retrieved from: https://arxiv.org/pdf/1909.10865 (Retrieved on: 24th June, 2023)
7. Geeksforgeeks (2022), Retrieved from: https://www.geeksforgeeks.org/distance-time-graphs/ (Retrieved on: 24th June, 2023)
8. Guze, S. (2019). Graph theory approach to the vulnerability of transportation networks. *Algorithms*, *12*(12), 270.Retrieved from: https://www.mdpi.com/1999-4893/12/12/270/pdf (Retrieved on: 24th June, 2023)
9. Guze, S. (2019). Graph theory approach to the vulnerability of transportation networks. *Algorithms*, *12*(12), 270.Retrieved from: https://www.mdpi.com/1999-4893/12/12/270/pdf (Retrieved on: 24th June, 2023)
10. Kreuzer, D., Beaini, D., Hamilton, W., Létourneau, V., & Tossou, P. (2021). Rethinking graph transformers with spectral attention. *Advances in Neural Information Processing Systems*, *34*, 21618-21629.Retrieved from: https://proceedings.neurips.cc/paper/2021/file/b4fd1d2cb085390fbbadae65e07876a7-Paper.pdf (Retrieved on: 24th June, 2023)
11. Ma, Z., Wan, W., Song, L., Liu, C., Liu, H., & Wu, Y. (2022). An Approach of Path Optimization Algorithm for 3D Concrete Printing Based on Graph Theory. *Applied Sciences*, *12*(22), 11315.Retrieved from: https://www.mdpi.com/2076-3417/12/22/11315/pdf (Retrieved on: 24th June, 2023)
12. Majeed, A., & Rauf, I. (2020). Graph theory: A comprehensive survey about graph theory applications in computer science and social networks. *Inventions*, *5*(1), 10.Retrieved from: https://www.mdpi.com/2411-5134/5/1/10/pdf (Retrieved on: 24th June, 2023)
13. Majeed, A., & Rauf, I. (2020). Graph theory: A comprehensive survey about graph theory applications in computer science and social networks. *Inventions*, *5*(1), 10.Retrieved from: https://www.mdpi.com/2411-5134/5/1/10/pdf (Retrieved on: 24th June, 2023)
14. Merschman, E., Doustmohammadi, M., Salman, A. M., & Anderson, M. (2020). Postdisaster decision framework for bridge repair prioritization to improve road network resilience. *Transportation research record*, *2674*(3), 81-92.Retrieved from: https://www.researchgate.net/profile/Abdullahi-Salman/publication/339522165_Postdisaster_Decision_Framework_for_Bridge_Repair_Prioritization_to_Improve_Road_Network_Resilience/links/5e5fdc98299bf1bdb8540ac2/Postdisaster-Decision-Framework-for-Bridge-Repair-Prioritization-to-Improve-Road-Network-Resilience.pdf (Retrieved on: 24th June, 2023)
15. Miao, Z., Pan, L., Wang, Q., Chen, P., Yan, C., & Liu, L. (2019). Research on urban ecological network under the threat of road networks—a case study of Wuhan. *ISPRS International Journal of Geo-Information*, *8*(8), 342.Retrieved from: https://www.mdpi.com/2220-9964/8/8/342/pdf (Retrieved on: 24th June, 2023)
16. Nosirov, K., Norov, E., & Tashmetov, S. (2022). A Review of Shortest Path Problem in Graph Theory. *Eurasian Journal of Engineering and Technology*, *13*, 1-11.Retrieved from: https://www.geniusjournals.org/index.php/ejet/article/download/2755/2360 (Retrieved on: 24th June, 2023)

17. Peer, M., Brunec, I. K., Newcombe, N. S., & Epstein, R. A. (2021). Structuring knowledge with cognitive maps and cognitive graphs. *Trends in cognitive sciences*, *25*(1), 37-54.Retrieved from: https://www.sciencedirect.com/science/article/am/pii/S1364661320302503 (Retrieved on: 24th June, 2023)
18. Roy, A., & Kesselman, A. (2021). A Novel Approach to Topological Graph Theory with RK Diagrams and Gravitational Wave Analysis. *arXiv preprint arXiv:2201.06923*.Retrieved from: https://arxiv.org/pdf/2201.06923 (Retrieved on: 24th June, 2023)
19. Saleem, B., & Weng, Y. (2022). Explainable Graph Theory-Based Identification of Meter-Transformer Mapping. *arXiv preprint arXiv:2205.09874*.Retrieved from: https://arxiv.org/pdf/2205.09874 (Retrieved on: 24th June, 2023)
20. Toppr,(2022), Retrieved from https://www.toppr.com/guides/science/motion-and-time/measurement-of-speed-and-distance-time-graph/ (Retrieved on: 24th June, 2023) Transportgeography(2023), Retrieved fromhttps://transportgeography.org/contents/methods/graph-theory-definition-properties/graph-representation-real-network/ (Retrieved on: 24th June, 2023)
21. Virmani, N., Salve, U. R., Kumar, A., & Luthra, S. (2021). Analyzing roadblocks of Industry 4.0 adoption using graph theory and matrix approach. *IEEE Transactions on Engineering Management*.Retrieved from: https://repository.londonmet.ac.uk/6270/1/Updated-3-J.pdf (Retrieved on: 24th June, 2023)
22. Yeh, C. H., Jones, D. K., Liang, X., Descoteaux, M., & Connelly, A. (2021). Mapping structural connectivity using diffusion MRI: challenges and opportunities. *Journal of Magnetic Resonance Imaging*, *53*(6), 1666-1682.Retrieved from: https://onlinelibrary.wiley.com/doi/pdfdirect/10.1002/jmri.27188 (Retrieved on: 24th June, 2023)
23. Zhang, Y., Cheng, T., & Ren, Y. (2019). A graph deep learning method for short-term traffic forecasting on large road networks. *Computer-Aided Civil and Infrastructure Engineering*, *34*(10), 877-896.Retrieved from: https://discovery.ucl.ac.uk/id/eprint/10085240/7/Zhang_A%20graph%20deep%20learning%20method%20for%20short-term%20traffic%20forecasting%20on%20large%20road%20networks_AAM.pdf (Retrieved on: 24th June, 2023)

**Appendices**

**Survey Link:**

https://docs.google.com/forms/d/e/1FAIpQLSflmi0TCtX2EI3hTzjCzz_Td7GJJ94InZuTcbLSQ9A1zgu6xA/viewform?usp=sf_link

**Survey questions**

1. What is your Gender?
2. What is your age?
3. What is your monthly Income?

**DV. Importance distance in graph theory**

4. Graph theory can help to detect natural motions of particles
5. Distance in graph theory can locate the positional movements of elements

**IV1. Topological aspects of graph theory**

6. Image segmentation process can be cut through graph theory
7. Metric dimensional factors help to represent the road network in graph theory
8. Graphical presentation can enhance the topological theory more accessible

**IV2. Attributes in road networking**

9. The nodes in a graphical representation help to create virtual conception
10. The nodes in crossing creates the junction points for better road linking process
11. The attributes of graph can help to measure the practical networking design

**IV3. Key functions of road networking in graph theory**

12. Graphical theory helps to prioritize the ongoing project
13. Graphical theoretical designs can analyze the strategically necessity of road mapping

INTERNATIONAL JOURNAL OF INNOVATIVE TECHNOLOGY AND CREATIVE ENGINEERING (ISSN: 2045-8711)
VOL.13 NO.08 AUG 2023

# REVIEW ON DIFFERENT ENERGY OPTIMIZATION TECHNIQUES IN 5G/6G ASSISTED IOT NETWORK

**S.Sangeetha**
Research scholar
**Dr. L.Sudha**
Research Supervisor, Associate Professor, School Computer Science,
VET Institute of Arts and Science College, Erode,Tamilnadu,India

**Abstract- Energy efficiency is a considerable problem to be resolved in the progress of 5G/6G-IoT networks. Because these IoT applications are connected to larger number of devices and therefore utilizes more energy. As significance, designing an energy-efficient IoT application is supports to give long-term services. The IoT devices used in 5G/6G environment are energy limited. Hence, battery replacement and charging of IoT devices positioned in remote or hostile environments are expensive and hard to implement. Therefore, Energy harvesting methods have been introduced as better solutions to give constant energy to massive IoT organization. Energy harvesting methods gets energy from dissimilar sources i.e. wind, hydro, and solar energy and thereby facilitate autonomous power production. Though, owing to the uncertainties of dynamic environments, Energy harvesting techniques cannot promise continuous power supply to all the nodes in IoT system. In this scenario, implementing energy management techniques that provide better energy efficiency performance in 5G/6G-IoT environment is attained greater significance. In existing, there are many machine learning concepts were introduced for providing solution to energy efficient data transmission in IoT. Therefore, the key objective of this study is to present an investigation of diverse energy optimization techniques employed in conventional 5G/6G assisted IoT applications.**

**Keywords:** Energy Efficiency**,** Internet of Things (IoT), Devices, 5G, 6G

## 1. INTRODUCTION

The Internet of Things (IoT) is one of significant development application area which includes of several devices for communication through the Internet.

IoT-supported devices are organized in order to perform dissimilar operations as per the needs of the various applications i.e. environment monitoring, smart homes and hospitals, remote access control and monitoring, etc. As well, the devices utilized in an IoT environment are limited energy capacity. Hence, the devices are work together to utilize the network resources effectively. In a general IoT system, these devices are positioned either deterministically or arbitrarily to observe several constraints i.e. humidity, pressure, temperature, etc, from the environment. In a realistic IoT-organization, an enormous amount of diverse IoT-assisted devices are considered to cover a big geographical environment. Each device broadcasts its data to a sink node for processing or storage. Such networks face many confronts such as energy-efficient, cooperative communication among nodes, with the network's scalability problem.

### 1.1 IoT assisted 5G Network

The fifth generation (5G) cellular systems are constructed for reducing data traffic caused by the rising number of devices and their bandwidth-limited mobile applications. The growth of smart mobile devices rises in demand for a variety of applications in the 5G network. The data transmission on cellular networks is also a rising trend to communicate directly with other devices. In this scenario, energy-aware data communication is a key challenge in a 5G cellular network.

### 1.2 IoT enabled 6G Network

The growing wireless service constraints and user density directs evolution of 6G communication in the modern years. The benefit of 6G over the existing technologies is its high-level maintenance for heterogeneous applications and mobility support.

The design of 6G communications enhances the QoS constraints of the users by increasing data rates and controlling latency. This advantage is employed in different scenarios such as medical care, industry automation, smart homes and cities, intelligent transportation, etc. The incorporation of multi-level complex networks, information and communication technologies, and computational methods maintain through the cloud guarantees elastic and mobility-aware access to the 6G users.

In a 6G network, Device-to-Device (D2D) communication is referred the direct transmission of data between two devices. This assists the high-speed data transmission.

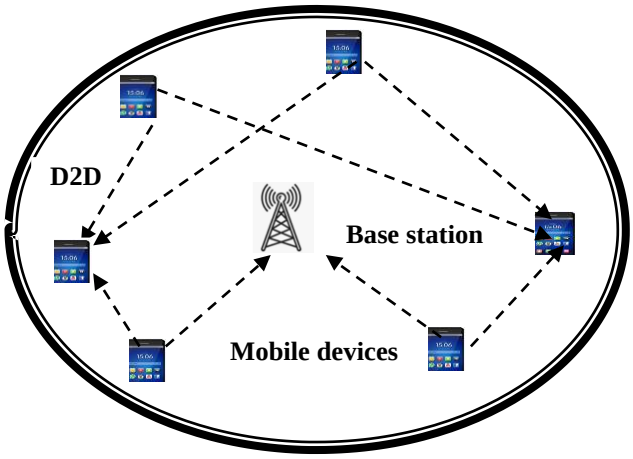



**Figure 1 IoT enabled 6G Network**

Figure 1 shows the D2D communication in the 6G network. In a 6G network, D2D communication guarantees a direct connection among the users for lessening the traffic. 6G makes the network connectivity among several devices. In a cellular network, D2D links want to allocate the same spectrum of resources as the cellular link. The applications of D2D include online gaming, multimedia images, audio, and video clip distribution, traffic offloading, and much more. In D2D communications, the nearby users are communicated without using any base stations. When considering short distance transmission, D2D communications can diminish energy utilization and progress the data delivery rate. When increasing the distance between the devices, achieving better energy efficiency is a key problem in the 6G network.

## 2. LITERATURE REVIEW

In the real environment of an IoT network with 5G support, the most vital problem is the adequate energy supply to fulfill the QoS requirements. Here, the challenge is to get better energy efficiency and providing the long-term services in network. Several techniques and algorithms have been implemented in the literature. Joint energy-efficient resource allocation (JEERA) algorithm was implemented in [1] with aiming at increasing the energy optimization performance in dynamic extensive 6G-IoT environment. The Lagrangian decomposition method was employed in JEERA to optimize power allocation and the KM algorithm to dynamically allocate resources to IoT users to obtain optimal solutions. This significantly reduces the computational complexity and makes the optimization process more scalable in dynamic large-scale IoT environment. The experimental outcomes demonstrated that the JEERA algorithm attained higher energy efficiency in realistic implementations in IoT networks.

An intelligent network solution was presented in [2] to get better energy efficiency results in a 6G-assisted larger IoT network. A cell-free massive multiple input multiple output (mMIMO) concepts were employed in this state-of-the-art work with the intention of boosting energy efficiency with optimal network resource management. A realistic power utilization model was also employed for the implemented network design that includes all the power constituents associated to information broadcasting and circuit power. The experimental outcomes proved that the conventional intelligent network solution provides higher energy efficiency and throughput.

SWIPT-Assisted Energy Efficiency Optimization concept was introduced in [3] for 5G/B5G Cooperative IoT system. In this work, several IoT-facilitated devices are organized to communicate with the source through relay nodes. The relay node broadcasts the information gathered by the source to the IoT devices. An Interference control model was intended in [4] to control the interference with the support of Lagrange optimization concept and thereby to increase the energy effectiveness and reliability in 5G systems. Here, the multi-objective optimization problem was solved in order to get better system reliability, throughput, and energy utilization. This interference control model has been analyzed and validated by considering different simulation conditions.

The acquired experimental results proved the effectiveness and the accuracy of interference control model was better in terms of network performance in IoT.

A power optimization model was constructed in [5] for 6G-supported big IoT association where it increases system performance and thus presents energy-preserving rate as compared to other works. The energy overhead due to the massive volume of linked devices was minimized via the optimal organization of power resources. The conventional power optimization model was analyzed for the maximum power distribution and the spectral efficiency for various network processes. Here, minimum distance scheduling and maximum channel gain scheduling algorithmic concepts were implemented to allocate users to a specific access points. The lower distance and higher channel gain among a user and an access point are taken as selection condition which minimizes the power resources. As a result, this power optimization model enhanced network efficiency.

An intelligent routing algorithm was planned in [6] for solving energy efficient management problem in 6G-powered wireless networks. The particle swarm optimization algorithm was employed in this work in order to increase the path discovery performance and thereby get better fault tolerance during the wireless network transmission when considering huge size of dynamic information. The experimental outcomes demonstrated that the state-of-the-art intelligent routing algorithm enhances the service life cycle and thereby guarantee the network transmission efficiency. Besides to that, this existing algorithm makes sure the energy utilization of the sensor and thus improves the service time via the node scheduling.

Deep Reinforcement Learning concept was designed in [7] for presenting an optimal computation resource distribution in IoT environment. Lyapunov optimization technique was utilized to get higher energy efficiency by considering broadcasting power, network stability, and latency. With the support of Markov decision algorithm, this model handles the transmission latency issues in the dynamic IoT application area.

Moreover, an adaptive method was introduced for continuous action-state spaces and thus lessens the completion time and total energy utilization of the IoT devices. The Lyapunov optimization and Deep Reinforcement Learning techniques obtained minimal computational complexity and better processing time.

A fuzzy logic and optimization-based resource management algorithmic concept was developed in [8] for huge IoT organization with the application of particle swarm optimization (PSO). The conventional Fuzzy logic allocates the MTC applications to access the channel. The available response slots are optimally allocated to the devices. The performance of optimization-based resource management algorithm was analyzed by considering the metrics such as path loss, energy and latency. The experimental results confirmed that PSO algorithm attained enhanced performance through finding the route with lesser path loss and minimal energy requirement.

A hybrid latency and power-aware approach was implemented in [9] with the goal of decreasing the latency with lesser overhead while giving a power-efficient solution for B5G-IoT-edge networks. This approach employed algorithm classifier tool (ACT) for choosing suitable optimization algorithms depends on the characteristics and necessities of B5G-IoT organization. Further, this approach minimizes latency and gets better overall network performance with the application of biogeography-based optimization (BBO) and grey wolf optimization (GWO) concepts.

An optimal mobile resource-sharing model was presented in [10] to resolve the constraint of 5G. on the contrary to state-of-the-art algorithms, the this optimal path planning and scheduling was implemented for mobile edge server includes the benefit of getting minimal delay and resource demand. This model employed path planning and optimal task scheduling with the objective of enhancing the network efficiency. From the simulation results, it is illustrated that there is a considerable improvement in resource employment and minimization in average response time.

An energy efficient data transmission rate allocation technique was implemented in [11] for 6G networks. This technique employed maximum likelihood estimation (MLE) concept in order to determine the true channel characteristics for a given set of observations.

The implemented allocation technique was robust against unknown channel statistics and also adjusts transmission rate to the channel characteristics which in turn lessens energy utilization and provides QoS. Both statistical investigation and simulation outcomes prove the efficiency of this technique with better energy efficiency and throughput.

## 4. CONCLUSION

As studied in the above literature, employment of 6G assisted IoT application environment results in high cost as compared 5G network system in terms of computational resources, energy necessity and efficient resource employment. Therefore, there is a need of optimal techniques to solve the energy management problems in huge IoT network. Several optimization, routing protocols and scheduling techniques were introduced for providing solution to energy efficiency problem in 5G/6G assisted IoT network. This paper provides a review of 5G/6G supported IoT technologies and optimal solutions to above problem by analyzing diverse conventional optimization, routing models and scheduling techniques with their advantages.

## REFERENCES

[1] Ansere, James Adu, Mohsin Kamal, Izaz Ahmad Khan, and Muhammad Naveed Aman (2023) “Dynamic Resource Optimization for Energy-Efficient 6G-IoT Ecosystems”, Sensors, Vol. 23, No. 10, pp. 1-20.

[2] Ankita Rana, Ashu Taneja, Nitin Saluja, Shalli Rani, Aman Singh, Fahd S. Alharithi and Sultan Mesfer Aldossary (2022), “Intelligent Network Solution for Improved Efficiency in 6G-Enabled Expanded IoT Network”, Electronics, ResearchGate, Vol. 11, pp. 1-15.

[3] Amjad, Maliha, Omer Chughtai, Muhammad Naeem, and Waleed Ejaz (2021), “SWIPT-Assisted Energy Efficiency Optimization in 5G/B5G Cooperative IoT Network”, Energies, ResearchGate, Vol. 14, No. 9, pp. 1-25.

[4] Radwa Ahmed Osman and Amira I. Zaki (2020), “Energy-Efficient and Reliable Internet of Things for 5G: A Framework for Interference Control”, Electronics, ResearchGate, Vol. 9, pp. 1-18,

[5] Ashu Taneja, Nitin Saluja, Neeti Taneja, Ali Alqahtani, M. A. Elmagzoub, Asadullah Shaikh, Deepika Koundal (2022), “Power Optimization Model for Energy Sustainability in 6G Wireless Networks”, Sustainability, ResearchGate, Vol. 14, No. 12, pp. 1-16.

[6] Tantong Zhang (2023), “An intelligent routing algorithm for energy prediction of 6G-powered wireless sensor networks”, Alexandria Engineering Journal, Elsevier, Vol. 76, pp. 35-49.

[7] James Adu Ansere, Eric Gyamfi, Yijiu Li, Hyundong Shin, Octavia A. Dobre, Trang Hoang, Trung Q. Duong (2023), Optimal Computation Resource Allocation in Energy-Efficient Edge IoT Systems with Deep Reinforcement Learning”, IEEE Transactions on Green Communications and Networking.

[8] Ajay Kumar Barapatre, Ritesh Sadiwala (2023), “Energy and resources management for Multiple Access in Massive IoT network”, Journal of Integrated Science and Technology, Vol. 11, No. 3, pp. 1-8.

[9] Ajay Kaushik and Hamed S. Al-Raweshidy (2022), “A Hybrid Latency- and Power-Aware Approach for Beyond Fifth-Generation Internet-of-Things Edge Systems”, IEEE Access, Vol. 10, pp. 1-16.

[10] Ajay K. Barapatre, Ritesh Sadiwala (2022), “Resource Allocation in Massive Internet of Things-Edge Network with Optimal Path Planning and Scheduling”, SAMRIDDHI: A Journal of Physical Sciences, Engineering and Technology, Vol.14, No 4, pp. 24-30, 2022

[11] A. Pourkabirian and M. H. Anisi (2021), “Robust Data Transmission Rate Allocation to Improve Energy Efficiency in 6G Networks”, IEEE Globecom Workshops (GC Wkshps), Madrid. Spain. pp. 1-6.

INTERNATIONAL JOURNAL OF INNOVATIVE TECHNOLOGY AND CREATIVE ENGINEERING (ISSN: 2045-8711)
VOL.13 NO.08 AUG 2023

# AGRICULTURAL INFORMATICS MONITORING SYSTEM USING IOT AND ARTIFICIAL INTELLIGENCE

**Dr. D. Padmapriya**
Assistant Professor, Department of Computer Technology
KG College of Arts and Science, Coimbatore, Tamilnadu, India

**Abstract- Agricultural activity on the farm requires effort for instance, planting, maintaining, and harvesting crops requires resources such as cash, time, energy, and labour. A group of researchers created an AI that can recognize plant ailments. To educate the AI to recognize crop illnesses and insect damage, this team employed a method known as transfer learning. Technology has altered farming practices over time and has had a wide range of affects on the agriculture industry. By 2050, when the population is anticipated to increase from 7.5 billion to 9.7 billion, only an additional 4% of the planet's surface will be cultivated, placing further pressure on the land. In many countries around the world, agriculture is the main source of employment. As a result, farmers will have to exert more effort while using fewer resources. According to the same report, food production would need to increase by 60% in order to feed an additional two billion people.**

**In today's world, agriculture must also keep up with the advancement of technology. IoT is crucial to smart agriculture. Internet of Things (IoT) sensors provide the necessary data regarding farm areas. The primary benefit of IoT is the ability to monitor agriculture using wireless sensor networks and gather data from numerous sensors that are placed in various locations and sent via wireless protocol. Node MCU is the IoT system that powers smart agriculture. It has a DC motor, moisture sensor, temperature sensor, and humidity sensor. This system begins to measure the level of moisture and humidity.**

**Keywords:** Internet of Things, Relay, Wi-Fi module ESP8266, Soil, Moisture and Temperature sensors.

## 1. INTRODUCTION

The support of human life is dependent on agriculture. Agriculture production increases in direct proportion to population growth. Agriculture output essentially depends on the seasonal conditions if there aren't enough water sources.

IoT-based smart agriculture systems are used to improve agricultural outcomes and solve difficulties.

Systems for monitoring agriculture on a global and regional scale are intended to offer current data on food production. A system is developed for agricultural field monitoring in IoT-based smart farming with the aid of sensors like light, humidity, temperature, soil moisture, etc. Farmers can keep an eye on the state of their fields from anywhere. Smart farming that is IoT-based is significantly more efficient than traditional farming. The ESP8266 Node MCU Module and DHT11 Sensor are used in the proposed IoT-based irrigation system. In addition to automatically irrigating the soil based on the amount of moisture present, it will also send data to the Thing Speak server to monitor the state of the land.

The development of automatic irrigation systems can take use of the most current improvements in sensors for agricultural irrigation systems as well as the development of WSN and IoT technologies. The system will identify the irrigation system parameters that are tracked for water quantity and quality, soil characteristics, weather, and fertiliser usage. It will also give an overview of the most frequently used nodes and wireless technologies used to implement WSN and IoT based smart irrigation systems.

## 2. LITERATURE SURVEY

Methods to monitor a crop field is explained in An IOT Based Crop-field Monitoring an Irrigation Automation System. A system is created utilizing sensors, and the irrigation system is automated based on a server's decision based on sensed data. The sensed data is transmitted wirelessly to a database on a web server. The moisture and temperature fields are reduced below the potential range if irrigation is automated.

With the use of a programmed that gives the user a web interface, the user can remotely monitor and control the system [1]. Farmers get the soil ready for sowing. During this process, large dirt clumps must be broken up, and trash like sticks, pebbles, and roots must be taken out. To further prepare the environment for crops, depending on the type of crop, add organic matter and fertilizers. When sowing seeds, it's crucial to think about the distance between each seed and the planting depth. At this point, climate variables including temperature, humidity, and rainfall are important [2].

Agricultural robotics, soil and crop monitoring, and predictive analytics are the three key areas of artificial intelligence (AI) in agriculture that are emerging as the sector increasingly transitions to digitalization. Farmers are increasingly using sensors and soil sampling, and the information they gather is stored on farm management systems for simpler processing and analysis. There are more possibilities for the use of AI in agriculture thanks to the availability of this data and related data.[3].

At the very end of the system, which can build a complete computer system from sensors to tools that view data from agricultural fields, the cloud computing devices are used. It offers a fresh approach to smart farming by incorporating wireless communication technologies with smart sensor systems and smart irrigation systems [4]. The installation cost of this system is low. Here, one can use a laptop, Smartphone, or computer to access and manage the agricultural system [5].

## 3. AGRICULTURE USING IOT

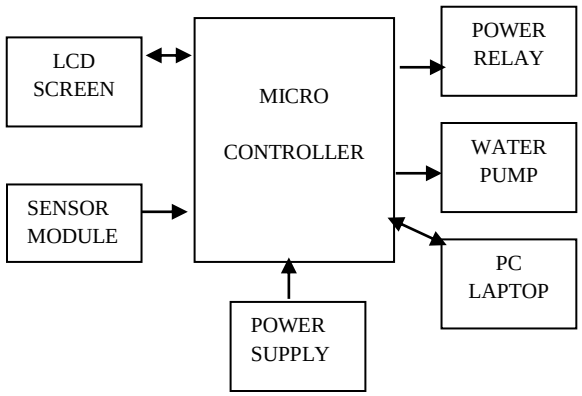



**Figure 1 depicts the proposed system's block diagram of Agriculture using IoT, which provides details on the necessary modules**.

**Required Modules and Hardware Requirements**

1. Soil moisture sensor

Both the irrigation industry and plant gardens depend heavily on the soil's moisture. In the same way that nutrients in the soil provide plants sustenance for growth. In order to adjust the plants' temperature, water must be provided to them. Utilizing a process similar to transpiration, water can be used to modify a plant's temperature. Additionally, plant root systems grow more effectively in damp soil. Extreme soil wetness might result in anaerobic conditions that can promote the growth of the plant and soil pathogens.[3]

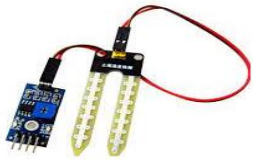


**Figure 2. Soil Moisture Sensor**

2. Temperature Sensor

A temperature and humidity sensor is an electronic device that measures, detects, and reports both ambient temperature and wetness at a cheap cost. the ratio of the maximum amount of moisture at a given air temperature to the overall amount of moisture. One of the most significant tools for measuring and monitoring temperature and humidity to a specific place, particularly in a data centre or server room, has been widely used in consumer, industrial, biomedical, and environmental applications. Temperature and humidity measurements are critical in most sectors since they contribute to the safety of all vital equipment that could have an impact on the entire operation.

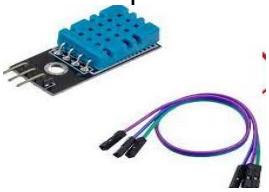


**Figure 3. Humidity and Temperature Sensor**

3. Power Relay

As an electrically operated switch, a relay is employed. It has a set of working contact terminals and a set of input terminals for one or more control signals. The switch may have several different types of contacts that can be used to make or break connections. In order to maintain the crop's moisture level, a relay is employed to activate the water pump.[4].

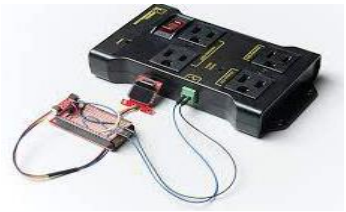


**Figure 3. Power Relay**

4. Water Pump

A cheap, compact submersible pump motor is the DC 3-6V Mini Micro Submersible Water Pump in Figure 5. Power between 2.5 and 6 volts is required for operation. With a relatively low current usage of just 220mA, it can pump up to 120 litres per hour. The motor outlet must only be connected to the tube pipe before being powered and submerged in water.

**Figure 4. Water Pump**

5. IoT (WI-FI module ESP8266)

The Node MCU (ESP8266) microcontroller depicted in Figure 6 has an integrated Wi-Fi module. This device has 30 total connections, 17 of which are GPIO (General Purpose Input/Output) pins that are connected to various sensors to receive data from the sensors and transmit output data to connected devices. The Node MCU contains 4MB of flash memory and 128KB of RAM for storing programmes and data. Through USB, the Node MCU receives the code and stores it. Every time the Node MCU receives input data from the sensors, it double-checks the information before storing it. Depending on the data received, the device sends a pulse to the relay module, which then serves as a switch to turn the pump on or off. The operating voltage and frequency of the Node MCU are 3 to 3.6 volts and 80 to 160 MHZ respectively. Between 46 (indoors) and 92 (outdoors) metres is the Wi-Fi module's range on the Node MCU.[5].

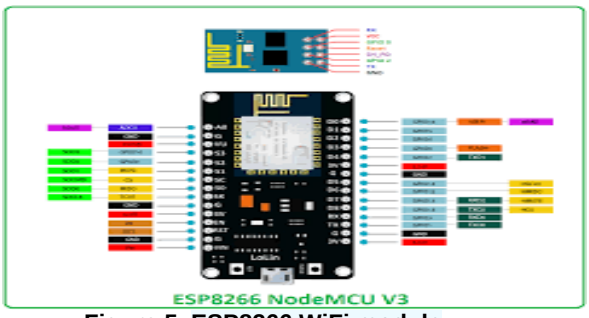



**Figure 5. ESP8266 WiFi module**

The ESP8266 WiFi module is highly affordable and user-friendly. It establishes a straightforward TCP/IP connection and is easily interfaced with microcontrollers through a serial port. The ESP-01, the first chip in this family, attracted a lot of market attention.

6. Power supply: 5V, 700mA Regulated power supply Software tools required

The general block diagram is shown below. It is extremely easy. The four primary sub-blocks are as follows, a Transformer, Circuit of the Rectifier, In the Filter, Regulatory body

1. Transformer ( Stepping Down)
2. Rectifier ( ac to dc conversion)
3. Filter ( Removing ripples from dc current)
4. Voltage Regulator ( To set regulated dc supply)

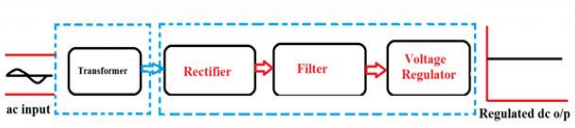



**Figure 6. Power Supply Block Diagram**

Voltage step-down: The voltage is first scaled down using a step-down transformer. The 220 ac voltage is changed to the lower ac voltage using a step down transformer. Now dispel a misunderstanding: Because we don't get shocked when we contact a transformer's output wire, the majority of people believe that it produces dc voltage. But this is just incorrect. The output of a step-down transformer is ac (alternating current).

Rectifiers are the pieces of machinery used to change alternating current into direct current. Rectification is the process of switching from alternating current (ac) to direct current (dc). These circuits are crucial to the design of the dc power supply.

7. Using Arduino

A cross-platform programme with functions built in C and C++ is the Arduino Integrated Development Environment (IDE). It is used to write programmes and dump them on Arduino compatible boards using third-party cores and other vendor development boards.

The IoT analytics platform Thing Speak is used to gather, visualise, and analyse real-time data streams in the cloud. It is available at the Thing speak website. Thing speak instantly visualises real-time data when it is sent from the devices and notifies the user when an alarm is sent. Operation of Thing Speak internally[6].

## 4. Working Principle

In various circumstances, the smart farm monitoring system is tested. The soil is tested using a soil moisture sensor under all climatic circumstances, and the results are properly interpreted. Readings of the moisture output are taken and updated under various weather situations. The wireless communication is made possible thanks to Wi-Fi.

The resistivity of the soil is the only factor that influences the values of soil moisture sensors. When the sensor first becomes wet, its value is 0. The motor pump receives the detected value through Node MCU and sends it to the microcontroller.

In this circumstance, OFF. On dry soil, the highest threshold value is 1023. The microprocessor activates the relay and turns on the motor when the measured value by the sensor reaches the threshold value. The motor pump automatically turns ON and OFF when plants receive a enough amount of water.

## 5. Features and Benefits

It is inexpensive to buy and simple to maintain. The necessary components are readily available. It is advantageous to check the status online using a laptop or smartphone. Even without a farmer present, the information is current. The farmer is aware of the crop's status thanks to the updated data that has been gathered. Several other sensors may also be incorporated in order to obtain more precise and effective crop details.

## 6. Precision Farming and Predictive Analytics

AI applications in agriculture have created tools and applications that assist farmers with accurate and controlled farming by giving them proper instructions on how to manage their water use, crop rotation, timely harvesting, and the type of crop to be grown, optimal planting techniques, pest attacks, and nutrition. AI-enabled technologies forecast weather conditions by analysing crop sustainability, evaluating farms for the presence of diseases or pests, and evaluating farms for poor plant nutrition using data such as temperature, precipitation, wind speed, and solar radiation in conjunction with images taken by satellites and drones. Farmers that lack internet can profit from AI right away with technologies as basic as an SMS-enabled phone and the Sowing App. A continuously AI-customized plan for a farmer's fields is currently available through AI applications for those with Wi-Fi connectivity. Farmers can fulfil the increased global need for food by using IoT and AI-driven solutions to enhance productivity and revenue while preserving priceless natural resources. The use of data to optimise yields down to individual plant rows will enable farmers to become agricultural technologists in the future. Robots that can easily carry out a variety of duties in farming fields are being created by AI businesses in the field of agricultural robotics. When compared to people, this kind of robot is trained to harvest crops more quickly and in greater quantities. These robots are taught to inspect the quality of crops, find weeds, and pick and pack harvests simultaneously. The difficulties faced by agricultural labour force labour can likewise be overcome by these robots. System that uses AI to find pests Pests are among the deadliest adversaries of farmers because they harm crops [9]. Artificial intelligence (AI) systems employ satellite photos and historical data to determine whether any insects have landed and, if so, what kind of insect has landed, such as a locust, grasshopper, or other species. Additionally, AI may send notifications to farmers' smartphones so they can take the appropriate precautions and employ the appropriate pest treatment. This aids farmers in their battle against pests.[7]

## 7. Result and Analysis using IoT Based Smart Agriculture Monitoring System

The primary goal of this initiative is to use contemporary technologies to necessary industries like agriculture. This solution makes farm monitoring simple by utilising IoT technologies in agriculture. The benefits, such as conserving manpower and water, that have already been discussed, are most needed in the current agricultural environment. As a result, smart irrigation is made while using sensor networks in agricultural fields. Utilising the cloud, IoT data is delivered to the client. Because of this, any changes in the crop can be easily seen, allowing for early study. The variables that are measured and watched, such as soil moisture, temperature, and humidity[8].

## 7.Conclusion

The technology can monitor and manage the irrigation system by forecasting the humidity and soil moisture levels. IoT helps improve time management, water management, crop monitoring, soil management, and pesticide and insecticide control in several areas of farming.

IoT will improve intelligent farming. The technology can monitor and manage the irrigation system by forecasting the humidity and soil moisture levels. IoT helps improve time management, water management, crop monitoring, soil management, and pesticide and insecticide control in several areas of farming. Additionally, this approach reduces the amount of labour required by humans, streamlines farming methods, and promotes smart farming. In addition to the benefits this system offers, smart farming can, with only a single touch and no work, aid in expanding the farmer market.[10]

## References

[1] M. A. Kekane, “Indian agriculture-status, importance and role in Indian economy”, International Journal of Agriculture and Food Science Technology, Vol. 4, No. 4, pp. 343-346, 2013

[2] B. F. Johnston, P. Kilby, Agriculture and Structural Transformation: Economic Strategies in Late-Developing Countries, Oxford University Press, 1975

[3] S. Kuznets, “Modern economic growth: Findings and reflections”, American Economic Association, Vol. 63, No. 3, pp. 247–258, 1973

[4] M. Syrquin, “Patterns on Structural Change”, in: Handbook of Development Economics, Vol. 1, Elsevier, 1988

[5] R. Dekle, G. Vandenbroucke, “A quantitative analysis of China’s structural transformation”, Journal of Economic Dynamics and Control, Vol. 36, No. 1, pp. 119-135, 2012

[6] M. Fan, J. Shen, L. Yuan, R. Jiang, X. Chen, W. J. Davies, F. Zhang, “Improving crop productivity and resource use efficiency to ensure food security and environmental quality in China”, Journal of Experimental Botany, Vol. 63, No. 1, pp. 13-24, 2012

[7] O. Oyakhilomen, R. G. Zibah, “Agricultural production and economic growth in Nigeria: Implication for rural poverty alleviation”, Quarterly Journal of International Agriculture, Vol. 53, No. 3, pp. 207-223, 2014

[8] T. O. Awokuse, “Does Agriculture Really Matter for Economic Growth in Developing Countries?”, The American Agricultural Economics Association Annual Meeting, Milwaukee, Newark, USA, July 28, 2009

[9] K. Jyostsna Vanaja, Aala Suresh et.al, “IOT based Agriculture System Using NodeMCU”, International Research Journal of Engineering and Technology, Vol.05.

[10] T. Rajesh, Y. Thrinayana and D. Srinivasulu “IoT based smart agriculture monitoring system”, International Research Journal of Engineering and Technology, Vol.07.

# THEIL-SEN LINEAR DATA TRANSMISSION IN IoT

**P. Suresh Babu**[1]
Associate Professor & Head, Department of Computer Science,
Bharathidasan College of Arts and Science, Erode, Tamilnadu, India
**S. Vadivelu**[2]
Research Scholar, Department of Computer Science,
Bharathidasan College of Arts and Science, Erode, Tamilnadu, India

**Abstract - The Wireless Sensor Network is a self-organizing system of multiple-hop sensor nodes that communicate with one another wirelessly. The sensor node collects data while keeping an eye on the surrounding environment and transmits and receives data from/to many nodes. Energy is the most crucial factor to improve the total network longevity during data transmission. While massive data transfer poses a serious challenge to WSN, sensor nodes typically have minimal energy consumption.**

**A cutting-edge method of machine learning called IoT aware Energy Indexed To extend the network lifetime in WSN, Theil-Sen Linear Regressive Time Instantaneous Data transfer (IoT-ETLR) is introduced. IoT-ETLR performs Theil-Sen Linear Regression Analysis to examine the remaining energy of the sensor nodes based on Camargo's index. Machine learning technique called Theil-Sen regression analysis is used to estimate the relationship between one or more variables. The more energy-efficient sensor nodes are found using a linear regression analysis. Following that, the Time difference of the arrival technique is used to determine the route way between the source and sink nodes via nearby higher energy sensor nodes. Finally, the route path is built, and data transmission is completed effectively. Energy consumption, packet delivery ratio, packet loss rate, and end-to-end delay are all experimentally evaluated in relation to various numbers of sensor nodes and data packets. The observed results show that our suggested IoT-ETLR technique outperforms state-of-the-art works in terms of energy-aware data transfer with a greater delivery ratio and less latency.**

Keywords: IoT, WSN, Energy Efficient Data Transmission, Camargo's index, Theil-Sen Linear Regression, Time difference of arrival method

## I. Introduction

A WSN comprises spatially scattered sensor nodes meant to monitor and collect the variety of physical and environmental conditions with the help of the Internet of Things (IoT).

The different types of sensor nodes are generally energy disposed of in nature which directs to the formulation of novel techniques to limit any redundant energy dissipation. The source nodes in the wireless network consume a lot of energy in communicating data. A comparative study of different routing techniques has been developed to enhance network lifetime. Energy Soaring-based Routing Algorithm (ESRA) was designed in [1] for IoT Applications to monitor the environment and enhance the network lifetime as well as minimize the delay. However, higher data delivery was not achieved between the source and sink node. A Q-learning-based data aggregation-aware energy-efficient routing (Q-DAEER) technique was designed in [2] to find the best path for enhancing the lifetime and minimizing energy consumption of the network. However, the designed Q-DAEER technique failed to consider the delay-aware routing.

Power-Efficient Gathering in Sensor Information Systems (PEGASIS) was introduced in [3] to transmit packets for energy-balanced WSNs. But the performance analysis of packet delivery and loss rate was not carried out. Region-Based Mobile (RBM) Routing Protocol was designed in [4] for Wireless Sensor Networks. The region-based routing consumes more time for effectively transmitting the data packets. An Original Queen Honey Bee Migration (QHBM) technique was introduced in [5] for solving efficient mobile routing. However, higher data delivery was not achieved. An energy-efficient deep belief network (DBN) based routing method was introduced in [6] to achieve improved data transmission via the selected path. The selection of the shortest path was not performed for efficient data transmission to enhance network lifetime and energy efficiency.An Improved Stable Election Protocol (I-SEP) was introduced in [7] for IoT-based Environmental Monitoring. However, the algorithm was not applied for a mobile network where sensor nodes progress from one point to another with a stable speed.

An efficient data transmission model was introduced in [8] for next-hop node selection for data transmission throughout the network. But it failed to maximize the data routing balance among the IoT nodes.

## 2. Related Works

Three different algorithms were developed in [9] to improve the performance of a sensor network using nature-inspired computational methods. However, it failed to minimize the time complexities of optimizing the solution. A dynamic multi-hop energy-efficient routing protocol (DMEERP) was designed in [10] for WSN. However, it failed to balance the energy of data transmission. A Dynamic Routing Algorithm Based On Energy-Efficient Relay Selection (DRAWERS) was introduced in [11]. However, the complexity of DRA-EERS was increased while increasing the network size. An Energy-Efficient Cooperative Routing method was designed in [12] for Heterogeneous WSN. However, the performance analysis of delivery and loss rate was not focused. Distance Aware Residual Energy- Efficient Stable Election Protocol (DARE-SEP) was designed in [15] to provide an optimal transmission from sensor nodes to the other node. But it failed to maximize the network lifetime and protect the total energy of the network.

*B. Organization of Paper*

The rest of the paper is arranged into five different sections as follows: Here, Section 2 describes the proposed IoT-ETLR technique in detail. In Section 3, the performance of proposed and existing methods is compared with different performance metrics. Finally, Section 4 concludes the paper.

## 3. Proposed Methodology

In recent years, huge attention and dramatic development have been exposed for the Internet of Things (IoT) based constrained Wireless sensor network (WSN) to attain efficient resource utilization and better data transmission. IoT requires a better communication network for data transmission between heterogeneous devices and an optimally organized energy-efficient WSN. An IoT node is a small device and the delivered power is generally provided by batteries. Therefore, energy is a major constraint in the designing process of a WSN. In this paper, a novel technique called IoT-ETLR is introduced to attain a better network lifetime by performing the energy-efficient routing in WSN.

A. System Model

The system model of the proposed IoT-ETLR architecture is discussed in this subsection. In WSN, the number of IoT aware sensor nodes $S_i \in SN_1, SN_2, SN_3 \dots SN_n$ are randomly positioned in a squared grid structure '$m * m$' regarding certain transmission range '$T_R$' to collect the environmental conditions from the environment.

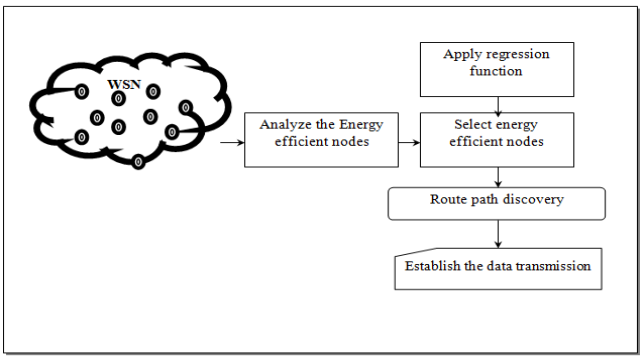



**Figure 1 Architecture of the Proposed IoT-ETLR Technique**

The collected data is transferred into the sink node in terms of data packets '$Dp_j = Dp_1, Dp_2, \dots, Dp_m$ with the help of energy-efficient intermediate neighboring nodes '$In_1, In_2, \dots, In_n$ to expand the network lifetime.

Fig.1. shows the architecture diagram of the proposed IoT-ETLR technique used for improving the data transmission with minimum delay and better network lifetime. The IoT aware WSN comprises sensor nodes $S_i \in SN_1, SN_2, SN_3 \dots SN_n$ for data transmission from source to sink node. The proposed IoT-ETLR technique consists of three different processes such as energy-efficient node selection, route discovery, and route maintenance. First, the Theil-Sen linear regression is applied for finding the energy-efficient sensor nodes. Second route path discovery is performed between sources to sink node based on time of flight method. Finally, the data transmission is performed along the route path.

B.Camargo's Indexive Theil-Sen Linear Regression

The first process of the proposed IoT-ETLR technique is to find the energy-efficient sensor nodes among the numerous sensor nodes. Energy efficiency is the most essential parameter in the design process of any routing technique in WSN. Therefore, energy-efficient nodes are identified through the Theil-Sen linear regression.

Regression analysis is a set of statistical processes for measuring the relationships between the dependent variable (i.e. Sensor nodes) and one or more independent variables (i.e. energy). The most common form of regression analysis is most closely finding the energy-efficient nodes according to a specific mathematical criterion. Theil-Sen linear regression uses Camargo's index for finding the energy-efficient nodes.

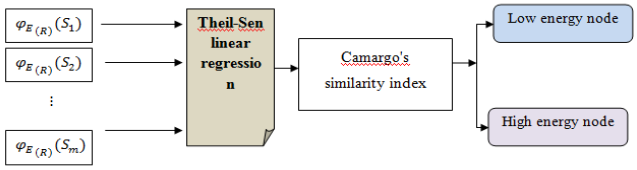



**Figure 2 Camargo's indexive Theil-Sen linear regression**

Fig.2. illustrates Camargo's indexive Theil-Sen linear regression that classifies the sensor nodes into high or low energy sensor nodes. First, the residual energy of the sensor nodes is given to the regression function. Then Camargo's similarity index is applied to measure the residual energy of node and median The higher energy nodes are identified when the $CI$ index is greater than the 0.5. Otherwise, the nodes are identified as lesser energy nodes. In this way, all the higher energy sensor nodes are identified for efficient data delivery from source to sink node.

### C. Time Difference of Arrival Method Based Route Path Discovery

The second process of the proposed IoT-ETLR technique is to find the Route path between source and sink node through the higher energy sensor nodes. In the route discovery phase, a source node in the WSN tries to discover a path to a specific destination i.e. sink node. The discovered path is used by the source node as the pathway for all communication until the discovered path becomes invalid. Therefore, the proposed IoT-ETLR technique uses the Time difference of the arrival method to find the neighboring sensor nodes through the Route request and reply message distribution.The Time difference of the arrival method is defined as the difference between the reply arrival time and request sending time at the source node. The node which has a minimum time difference is selected as a neighboring node. In this way, the route path from source to sink node is established via neighboring sensor nodes.

### D. Data Packet Transmission

Finally, the proposed IoT-ETLR technique performs the data packet transmission from source to sink node. During the transmission of data, there is a possibility of route failure. In this case, route maintenance is performed between the nodes. In this proposed technique, each source and other nodes have a routing table that establishes various paths in the route discovery phase and they are stored in their route collections. The sensor nodes store all paths in the routing table. In case of route failure in the main route, a node selects the alternative route for efficient data transmission and minimizes the delay. In this way, efficient data packets transmission is performed from source to destination. In IoT-ETLR, first, the high-energy efficient nodes are identified using Theil-Sen linear regression for measuring Camargo's similarity index. After that, the Time difference of the arrival method is applied to find the multiple route paths between the sensor nodes based on route request and route reply. The route path discovery between the source and sink node via neighboring higher energy sensor nodes is established based on the Time difference of the arrival method. Finally, the route path gets constructed and data transmission is carried out in an efficient manner. In case of any route failure, the route maintenances are carried out to minimize the delay of data transmission.

## 4. PERFORMANCE ANALYSIS

Simulation analysis of IoT-ETLR technique and existing methods ESRA[1], Q-DAEER [2] are discussed with different performance metrics such as packet delivery ratio and end to end delay.

***Packet delivery ratio***: It is formulated as the ratio of the number of data packets that are successfully delivered at the sink node to the total number of data packets sent. The formula for calculating the packet delivery ratio is given below,

$$DpD = \left(\frac{DpSD}{nDp}\right) * 100 \quad (1)$$

Where, $DpD$ signifies a packet delivery ratio, $DpSD$ indicates the number of packets successfully delivered, $nDp$ represents the number of data packets. The packet delivery ratio is measured in percentage (%).

**Table 1 Packet Delivery Ratio Versus Number of Data Packets**

| Number of data packets | Packet delivery ratio (%) | | |
|---|---|---|---|
| | IoT-ETLR | ESRA | Q-DAEER |
| **100** | 91 | 89 | 87 |
| **200** | 90 | 88 | 85.5 |
| **300** | 91.66 | 89.33 | 85 |
| **400** | 91 | 87 | 84 |
| **500** | 90.2 | 86.6 | 84.2 |

Performance analysis of packet delivery ratio versus the number of data packets is reported in table 4 and figure 4. The above results indicate that the performance of the IoT-ETLR technique is comparatively higher than the conventional methods. Let us consider the 100 data packets and the number of data packets successfully delivered to the sink node is 91 data packets. Therefore, the percentage of packet delivery ratio is 91%. By applying the existing methods ESRA[1], Q-DAEER [2], the 89 and 87 data packets are successfully delivered at the sink node, and observed delivery ratios are 77% and 87%. Likewise, nine remaining results are observed for all three methods. Totally, five results are obtained and compared.

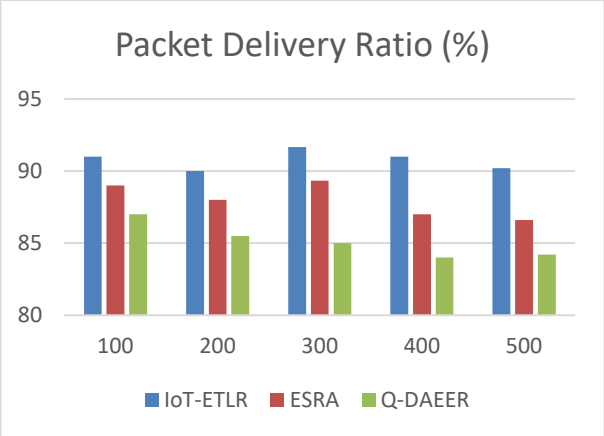



**Figure 4 Graphical Analysis of Packet Delivery Ratio**

This improvement of the proposed IoT-ETLR technique is to apply the Time difference of the arrival method. The multiple route paths are established between the source and sink node by distributing the route request and reply messages. Based on the message distribution, the time difference between the request sending and reply arrival is measured to identify the neighboring node and construct the route path.

This process helps to successfully deliver the data packets to the sink node.

***End-to-end delay***: It is measured as the difference between the actual arrival time of the data packets at the sink node and the expected arrival time. It is formulated as given below,

$$DEL = [T_{AC}] - [T_{EX}] \quad (2)$$

Where $DEL$' indicates an end to end delay, $T_{AC}$ indicates an actual arrival time $T_{EX}$ designates an expected arrival time. The delay is measured in terms of milliseconds (ms).

**Table 2 End to End Delay Versus Number of Data Packets**

| Sensor nodes (Numbers) | end to end delay (ms) | | |
|---|---|---|---|
| | IoT-ETLR | ESRA | Q-DAEER |
| **50** | 14.3 | 15.8 | 18.2 |
| **100** | 15.4 | 17.3 | 20.1 |
| **150** | 16.7 | 18.3 | 21.2 |
| **200** | 18.3 | 20.1 | 23.4 |
| **250** | 21.5 | 23 | 25.3 |
| **300** | 24 | 26.2 | 28 |

Table 6 given above illustrates the performance results of end-to-end delay of data packet transmission versus a number of sensor nodes. The observed simulation results indicate that the proposed IoT-ETLR technique provides better performance in terms of minimizing the delay than the conventional routing methods. For example, the number of nodes considered is 50 in the first iteration. The end-to-end delay of the IoT-ETLR technique is $14.3ms$ whereas the performance of delays using conventional methods ESRA[1], Q-DAEER [2] are $15.8ms$, and $18.2ms$ respectively. The overall observed statistical analysis demonstrates that the proposed IoT-ETLR technique provides better routing performance by minimizing the delay than the conventional methods.

Figure 5 presents the graphical results of an end-to-end delay along with a number of sensor nodes. The end-to-end delay of all the methods gets increased while increasing the number of sensor nodes. But comparatively, the proposed IoT-ETLR technique provides a lesser delay in order to accurately perform the data transmission. The IoT-ETLR technique is to perform route maintenance.

During the data transmission, any link between the nodes is broken, another alternative link path is efficiently chosen and improves the data delivery between the nodes resulting it minimizing the end-to-end delay.

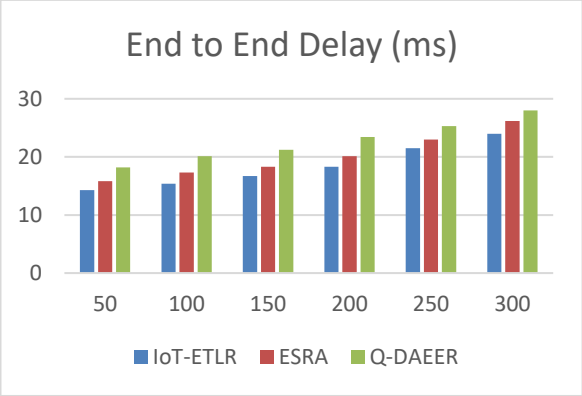



**Figure 5 Graphical Analysis of End to End Delay**

## 5. CONCLUSION

In this paper, a novel energy-efficient technique called the IoT-ETLR technique is introduced to enhance the network lifetime. The proposed IoT-ETLR technique performs three major processes. First, the Theil-Sen Linear Regression is applied to analyze the residual energy of the sensor nodes and find the higher energy nodes based on Camargo's similarity index. After that, the multiple route paths between the source and sink nodes are established by applying a Time difference of arrival method. Finally, the route maintenances are carried out to improve data transmission and minimize the delay. The simulation analysis is carried out on certain performance factors such as packet delivery ratio and less end-to-end delay. The observed result shows that the proposed IoT-ETLR technique provides better results has a higher delivery ratio and energy consumption than the conventional methods.

## REFERENCES

[1] N. Samarji and M. Salamah, “ESRA: Energy soaring-based routing algorithm for IoT applications in software-defined wireless sensor networks”, Egyptian Informatics Journal, 2021, Pages 1-10. https://doi.org/10.1016/j.eij.2021.12.004

[2] Wan-Kyu Yun; Sang-Jo Yoo, “Q-Learning-Based Data-Aggregation-Aware Energy-Efficient Routing Protocol for Wireless Sensor Networks”, IEEE Access, Volume 9, 2021, Pages 10737 – 10750. DOI: 10.1109/ACCESS.2021.3051360

[3] Kun Wang, Chih-Min Yu, Li-Chun Wang, “DORA: A Destination-Oriented Routing Algorithm for Energy-Balanced Wireless Sensor Networks”, IEEE Internet of Things Journal Volume 8, Issue 3, 2021, Pages 2080 – 2081. DOI: 10.1109/JIOT.2020.3025039

[4] Muhammad Fahad Mukhtar, Muhammad Shiraz, Qaisar Shaheen, Kamran Ahsan, Rizwan Akhtar, and Wang Changda, “RBM: Region-Based Mobile Routing Protocol for Wireless Sensor Networks”, Wireless Communications and Mobile Computing, Hindawi, Volume 2021, February 2021, Pages 1-11. https://doi.org/10.1155/2021/6628226

[5] Apripriharta, Wang Zhi Hao, Muladi, Gwo-Jiun Horng, Gwo-Jia Jong, “A New Bio-Inspired for Cooperative Data Transmission of IoT”, IEEE Access, Volume 8, 2020, Pages 161884 – 161893. DOI: 10.1109/ACCESS.2020.3021507.

[6] Greeshma Arya, Ashish Bagwari, Durg Singh Chauhan, “Performance Analysis of Deep Learning-Based Routing Protocol for an Efficient Data Transmission in 5G WSN Communication”, IEEE Access ,Volume 10, 2022, Pages 9340 – 9356. DOI: 10.1109/ACCESS.2022.3142082

[7] Trupti Mayee Behera, Sushanta Kumar Mohapatra, Umesh Chandra Samal, Mohammad. S. Khan, Mahmoud Daneshmand, and Amir H. Gandomi, “I-SEP: An Improved Routing Protocol for Heterogeneous WSN for IoT-Based Environmental Monitoring”, IEEE Internet of Things Journal , Volume 7, Issue 1, 2020, Pages 710 – 717. DOI: 10.1109/JIOT.2019.2940988

[8] Alireza Izaddoost and Matthew Siewierski, “Energy Efficient Data Transmission in IoT Platforms”, Procedia Computer Science, Elsevier, Volume 175, 2020, Pages 387-394. https://doi.org/10.1016/j.procs.2020.07.055

[9] Rajendra Prajapat, Ram Narayan Yadav, Rajiv Misra, “Energy-Efficient k-Hop Clustering in Cognitive Radio Sensor Network for Internet of Things”, IEEE Internet of Things Journal, Volume 8, Issue 17, 2021, Pages 13593 – 13607. DOI: 10.1109/JIOT.2021.3065691

[10] Zhihao Peng , Mehdi Sajedi Jabloo , Yahya Dorostkar Navaei , Morteza Hosseini , Rozita Jamili Oskouei , Poria Pirozmand , and Seyedsaeid Mirkamali, “An Improved Energy-Aware Routing Protocol Using Multiobjective Particular Swarm and Mobile Computing, Hindawi, Volume 2021, May 2021, Pages 1-16. https://doi.org/10.1155/2021/6677961

INTERNATIONAL JOURNAL OF INNOVATIVE TECHNOLOGY AND CREATIVE ENGINEERING (ISSN: 2045-8711)
VOL.13 NO.08 AUG 2023

# AN IOT - BASED APPROACH FOR CONTROLLING ELECTRICAL APPLIANCES IN THE SMART ROOM USING HUMAN MOVEMENT

**G. Yashodha**
Assistant Professor, Department of Computer Technology
KG College of Arts and Science, Coimbatore,Tamilnadu,India

**Abstract - The demand for more integration density, higher bandwidth, and lower power are always rising. Due to the rapid rise in the use of resources used on a daily basis during the past few years, there has been an increase in the demand for energy. Energy is produced from fossil fuels, which eventually run out. We therefore have a pressing need for an energy-saving system. The answer to completing the duties required for the creation of this system is the Internet of Things (IoT). In this study, an Internet of Things (IoT)-based system that allows different devices to be connected over the Internet in order to save energy is presented. The "IoT Based Electricity Conservation System" prototype's main goal is to incorporate the aforementioned features utilising a PIR sensor. If the PIR detects any motion, it sends a signal to Node MCU which in turn triggers the relay to turn on the lights and turn off if no motion is detected.**

Keywords - PIR sensor, NodeMCU, Energy conservation, Automation, Internet of Things, Cost-effective, Time conserving, Energy efficient.

## I. INTRODUCTION

Devices with various functionality can connect to the Internet via IoT (Internet of Things). In order to leverage the IoT concept in this project, two distinct characteristics must be integrated. If there are people nearby, turn on or off the lights as part of energy conservation. In [1], a PIR sensor is utilised to identify human presence. An IR sensor called a PIR measures changes in the temperature radiated by adjacent objects to look for motion. The PIR sensor outputs a high signal on its output pin when motion is detected, and NodeMCU receives this signal as an input. A microcontroller called NodeMCU is equipped with serial communication protocols, analogue and digital ports, and Wi-Fi capabilities. NodeMCU can read the PIR sensor's logic signal, and then it can send a signal to the relay [1].

The associated devices, which could be a bulb, a tube light, a fan, or any other electrical device, are then given the volt via the relay. Therefore, Node MCU [2] is used to regulate the light intensity. The portions that make up the remaining parts of the paper are listed below. The paper's second portion, portion II, provides a summary of related papers and problem statements. Process flow, a conceptual model including a flowchart, connection, and implementation, is presented in Section III. Section IV includes the results, discussion, and conclusion and wraps up the work by outlining potential future research avenues.

## II. RELATED WORKS

This section includes a list of some of the earlier papers connected to this one. IoT technology have been acknowledged as playing a significant part in energy conservation across a number of industries, including manufacturing, transportation, health, and smart cities. According to R. Yasodharan et al., their IoT project will enable the in-class staff to manage the environment via an Android app. The design of the entire system is mostly based on the Arduino Mega 2560. The Blynk software or Blynk Android application is used in the development of the Android application. By simply turning on certain features in the Android application on our smartphone[5], we can monitor the condition of sensors linked to the Arduino board and operate the modules.Isanka Diddeniya, Niroshan Gunawardana, Kaveendra Maduwantha, Kaveenga Koswattage, Mahadurage Viduni Randima, and Vasanthapriyan proposed an Internet of Things (IoT)-based device controlling system in their paper that can be used to operate any electrical device with the least amount of user interaction.

The "IoT Based Energy Efficient Smart Classroom" prototype that has been suggested is put into use to cut down on electricity waste in a lecture hall at the Sabaragamuwa University of Sri Lanka. By detecting the presence of people in a particular location, the system can manage the operation of electrical devices (such as turning them ON/OFF) [6]. The year 2021 saw the introduction of a proposal by Mohd Wafi Nasrudin, Nur Asyikin Nordin, Iszaidy Ismail, Mohd Ilman Jais, Amir Nazren Abdul Rahim, and Wan Azani Mustafa that energy efficiency measures, such as shutting off lights and electrical equipment when not in use, may be used to reduce the amount of electricity used. In order to avoid wasting electricity in the classroom, this work proposed the smart classroom for electricity-saving with an integrated IoT System. The primary goal of this study is to use an IoT application and sensor system to control the lighting and fan systems. The Blynk application software is utilised to show the classroom's status[7].

## III. PROBLEM IDENTIFICATION

Since 2000, India has used more energy, with 80% of the country's energy coming from coal, oil, and solid biomass. Greater energy consumption is a result of population growth, economic progress, and technological improvements. These resources diminish with time because the majority of them are non-renewable resources. If all resources were exhausted, it would be difficult for living organisms to survive. We should therefore start preserving the available resources for later demands. We must also begin practising time management. Time management skills enable us to work more productively and quickly towards our goals.

UPKEEP OF ELECTRICITY

A PIR sensor is used to focus on the first module, "Conservation of Electricity." PIR sensors are used to detect motion and typically determine when a person enters or exits their detection range. They are not only lightweight, affordable, and low-power, but they are also simple to use and never need replacing. PIR, passive infrared, and pyroelectric sensors are other names for IR motion sensors. Utilising all these benefits, the presence of students in the classroom is detected, and lights are turned on and off as necessary.

## IV. ARCHITECTURAL FRAMEWORK

The advantages of an IoT-based system in today's society, notably for conserving electricity, will be demonstrated via a classroom electricity-saving system that uses less electricity than manually turning off appliances. A NodeMCU microcontroller, a PIR (passive infrared sensor), and a relay control for energy conservation will be used to implement this research. Figure1 shows the architectural framework of conservation of electricity.

## V. UPKEEP OF ELECTRICITY

1) Connections: The NodeMCU ESP8266 is linked to the PIR sensor's output, input, and ground pins as well as the relay module's output, input, and ground pins.

2) Implementation: We used a 6-by-6-foot classroom with two fluorescent tube lights. Based on the Fresnel zone, the sensors were placed above the fan and the lamp so that the clearance requirements were met. This allows the sensors to detect every motion in the classroom and immediately switch on or off the fan or lamp. The sensor will be able to recognise body heat between 0°C and 50°C. An Internet of Things-based automated system for electrical equipment in smart rooms that is based on human movement The average range that is comfortable for the system to detect the presence of users or an empty classroom is 1811 °C, according to Section A of the research report published in the European Chemical Bulletin 2023, 12(Special Issue 5), 1807-1818.

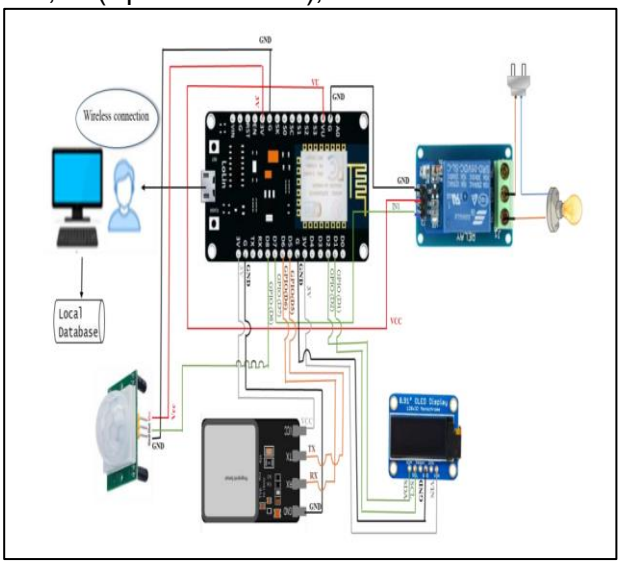



**Figure 1 : Architectural Framework**

The microcontroller will receive the data when the PIR sensor has recorded it. The functions thet have been programmed onto the microcontroller was listed in Table 1.

**Table: 1 Steps to be processed by the microcontroller**

| Steps | Instructions |
|---|---|
| Step: 1 | Data is collected through the infrared radiation emitted by hot bodies. |
| Step: 2 | An electrical signal is generated when the sensor detects infrared radiation. An internal PIR sensor has two halves, one of which is positive and the other negative. |
| Step: 3 | Thus, one half detects a hot body's motion and the other half creates a signal. |
| Step: 4 | Output signals are generated by the difference between the two signals. |

This sensor's main components are bifurcated Fresnel lenses, which can be used to detect infrared radiation that is created when a hot body moves over a wide region or a smaller area. The following software was employed to carry out the design: 1) The Windows 10 operating system ii) The Integrated Development Environment (IDE) for Arduino. Shortly mentioned is the Pseudo Code, which uses a PIR sensor to conserve energy. The idea of using pseudocode in this work is to clarify some ideas, such as how it is simple to comprehend even for non-programmers and how it can be rapidly and easily transformed into a real programming language because of its similarities to one. Even if there are syntactic mistakes, the intended meaning is usually still clear.

3) Hardware Specification: i) Specification of NodeMCU In addition to its ability to act as a standalone system, ESP32 can be used as a slave device to a host MCU, reducing communication stack overhead on the main application processor. Through its SPI / SDIO as well as I2C / UART interfaces, the ESP32 can interface with other systems for Wi-Fi and Bluetooth functionality. There is 128 KB of RAM and 4MB of Flash memory on the NodeMCU for storing data and programs. The device's high processing power along with its Wi-Fi and Bluetooth capabilities make it ideal for Internet of Things projects

| DESCRIPTION | SPECIFICATION |
|---|---|
| Microcontroller | ESP-8266 32-bit |
| Node MCU Model | Clone LoLin |
| Node MCU Size | 58mm x 32mm |
| Pin Spacing | 1.1" (27.94mm) |
| Clock Speed | 80 MHz |
| USB to Serial | CH340G |
| USB Connector | Micro USB |
| Operating Voltage | 3.3V |
| Input Voltage | 4.5V-10V |
| Flash Memory/ SRAM | 4MB/64KB |
| Digital I/O Pins | 11 |
| Analog In Pins | 1 |
| ADC Range | 0-3.3V |
| UART/ SPI/ I2C | 1/1/1 |
| Wi-Fi Built-In | 802.11 b/g/n |

Specification of PIR Sensor

| DESCRIPTION | SPECIFICATION |
|---|---|
| Model | PIR HC-SR501 |
| Operating Voltage(VDC) | 4.5~ 20 |
| Average Current Consumption(mA) | 0.06 |
| Distance Measuring Range(CM) | 300 ~ 700 |
| Output Type | (High/ Low-level Signal)3.3V TTL output |

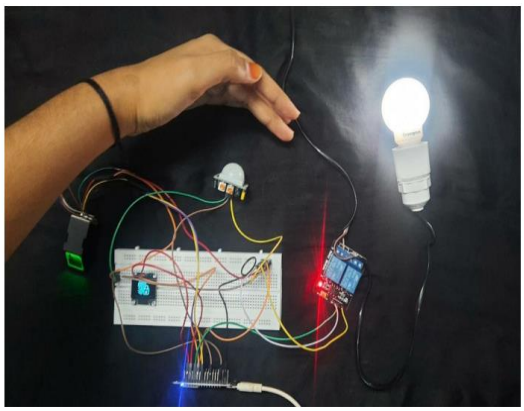


**Figure 2: Energy Conservation System**

Based on the specified objective, energy conservation has been accomplished cost effectively with the help of a PIR sensor as shown in Fig.2.

This figure demonstrates energy conservation through motion detection. When the user enters the room, the motion detection load is operated. Therefore, the user need not consider turning on/off the appliances.

## VI. CONCLUSION

This study describes an Internet of Things (IoT)-based energy-saving system that turns off classroom lights when students are not in the room using passive infrared radio sensors. The electricity-saving device was installed in an area with a 6-foot radius and linked to the lighting fixture that illuminates that specific space. Additionally, this system can be improved by adding a sensor that monitors light intensity and temperature in order to turn on fans when it becomes too hot and turn on lights only when it is dark outside. We will be able to conserve energy even more than usual as a result.

## REFERENCES

[1] Cynthia Twumasi, Koffi Agbeblewu Dotche, Francois Sekyere, Energy saving system using a PIR sensor for classroom monitoring, IEEE, 2017

[2] Talal H Noor, El-Sayed Atlam, Abdulqader M. Almars, Ayman Noor, Amer S.Malki, IoT Based Energy Conservation Smart Classroom System, Intelligent Automation and Soft Computing, Volume 35 Issue 3, 2023

[3] M. A. Meor Said, M.H. Misran, M. A Othman, M M Ismail, Biometric Attendance, International Symposium on Technology Management and Emerging Technologies (ISMET 2014), May 24, 2014

[4] Aswinth Raj, IoT-Based Biometric Attendance System using Arduino and Thingsboard, May 2019

[5] Yasodharan R, Bennaiah D, Hari Krishnan V, Karthick S, Prince Roy, IoT Based Classroom Automation Using Arduino, International Journal of Trend in Scientific Research and Development, Volume 2 Issue 2, February 2018

[6] Isanka Diddeniya, Niroshan Gunawardana, Kaveendra Maduwantha, Kaveenga Koswattage, Mahadurage Viduni Randima, Vasanthapriyan, IoT based Energy efficient smart classroom, International Journal of Multidisciplinary Engineering Science Studies, Volume 6 Issue 12, December 2020

[7] Mohd Wafi Nasrudin, Nur Asyikin Noordin, Iszaidy Ismail, Mohd Ilman Jais, Amir Nazren Abdul Rahim, Wan Azani Mustafa, Smart Classroom for Electricity saving with Integrated IoT system, Journal of Physics: Conference Series, Volume 2107, October 2021

[8] Tripti Jain, Urvashi Tomar, Umang Arora, Swati Jain, IoT-based Biometric Attendance System, International Journal of Electrical Engineering and Technology, Volume 11 Issue 2, May 2020

[9] D Sathiyaraj, Jeevanantham M, Dayalan S, Hariharan M, Suraj Punj, Classroom Automation Based IoT, Journal for Multidisciplinary Research Volume 11 Issue 4, 2021

[10] Shyamli u. Kamane, Akshay S. Bhoyar, Shubham V. Kusange, Swati M. Gajbe, Stuti S.Patil, Smart Classroom Automation System using Arduino and IoT, International Journal of Research and Analytical Reviews, Volume 7 Issue 1, March 2020

[11] Vishal Kolambkar, Tejas Patil, Sagar Anute, Prof. Rasika Shintre, Smart Classroom Automation Based on IoT, International Research Journal of Engineering and Technology Volume 7 Issue 3, March 2020

[12] S.A.I.P. Diddeniya, H.D.C.N. Gunawardana, K.Maduwantha, K.R.Koswattage, IoT Based Energy Efficient Smart Classroom, Journal of Multidisciplinary Engineering Science Studies, Volume 6 Issue 12, December 2020

[13] Kaustav Mani Pathak, Priyanka Jain, Swati Yadav, Veena Mittal, Prateek Jain, An IoT-based Mechanism for Automatic Classroom Electricity Saving, IEEE, October 2019

[14] Ganesh Kakade, Prathamesh Shivpuje, Sanket Awale, Shubham Rawale, IoT Based Classroom Automation System, International Research Journal of Engineering and Technology (IRJET), Volume 6 Issue 12, December 2019

[15] Sebastian Chennattu, Aditya Kelkar, Aaron Anthony, Sushma Nagdeote, Portable Biometric Attendance System using IoT, IEEE, November 2019

[16] Sri Madhu, Kavya Kanagotagi, Devansh, IoT-based Automatic Attendance Management System, International Conference on Current Trends in Computer, Electrical, electronics and Communication (ICCTCEEC-2017)

[17] K C Arun, Mubashir Ali, Ayesha Siddique, Muhammad A Asim, Zaka Ullah, Energy Efficient Classroom Automation using IoT, June 2017

[18] Mohammad Kamrul Hasan, Musse Mohamud Ahmed, Bishwajeet Pandey, Hardik Gohel, Shayla Islam, Izzul Fitrie Khalid, IoT-Based Smart Electricity Monitoring Control System Using Usage Data, Wireless Communications and Mobile Computing, Volume 2021, October 2021

[19] Prabesh Paudel, Sangkyoon Kim, Soonyoung Park, A Context-Aware Architecture for Energy Saving in Smart Classroom Environments, IEEE International Conference On Consumer Electronics, 2019

[20] Fawaz Alassery ,A Smart Classroom of Wireless Sensor Networks for Students Time Attendance System, IEEE Integrated STEM Education Conference (ISEC),2019.

# DIGITAL VIDEO SCRAMBLING AND DESCRAMBLING USING DWT IN CHROMINANCE CHANNEL

**V. Mounash**
III B.Sc.IT, Department of Information Technology,
Sri Krishna Adithya College of Arts & Science, Coimbatore, Tamilnadu, India
**I.Gobi**
Assistant Professor, Department of Information Technology,
Sri Krishna Adithya College of Arts & Science, Coimbatore, Tamilnadu, India

**Abstract-Video scramblers are commonly employed to prevent unauthorized access to video data. Several video scrambling systems rely on methods of directly distorting the visual image data (in the spatial domain) such that, without de-scrambling, the video appears unintelligible to a viewer. These scrambling techniques are not efficient for transmitting digital video signals because they, in general, will significantly change the statistical property of the original video signal, thus making it very difficult to compress.**

**Our objective is to create the existing video file unintelligible to the Viewer using the Block cipher algorithm such as DES for intruders and hackers difficult to break the information. The algorithm is used to make the video file in the compressed format, which makes efficient to travel across the Network without any traffic problems and with higher security. The video file is converted into individual images at first in order to get the values of each frames and make the values interchange that makes the information unintelligible. Then the rearranged frames are converted again to video file it cannot be accessed except the client without the Descrambling process. In the Client side the Scrambled video is converted back to get the original Information as they have sent.**

**Keywords:** DES, IPR, DVD, MANETs, AVI, BMP

## I. INTRODUCTION

Use of digital media has exploded in the past few years, primarily due to several distinct advantages that digital media can offer over analog media. These advantages include higher quality, easier editing, and perfect copying and easier and more efficient transmission over information network.

The wide dissemination of digital media also creates some potential Problems.Due to the popularity of Internet commerce and digital library applications, the intellectual property right (IPR) protection is becoming increasingly important. Content providers will be reluctant to provide their valuable contents if they are not assured that their contents are securely protected. Some good examples are the deployments of the digital versatile disk (DVD) market and the online music market. The IPR management and protection issue is currently being addressed in the emerging MPEG-4 standards for moving pictures compression. Several technologies have been developed for IPR protection. One is conditional access through encryption. The digital media will be scrambled before it is distributed. Only authorized users who have the proper key for decryption can access the clear content. The other one is digital watermarking that securely embeds hidden message into the multimedia data to identify the owner, or the buyer of a digital media. These two techniques are complementary to each other. We focus on conditional access through encryption in this project. Digital images/video are often communicated or distributed over non-private channels, such as satellite links, cable television networks, wireless networks, and the Internet. Conditional access systems for private digital image/video transmission or storage are a necessity for many applications, for example, pay-TV, confidential videoconferences, confidential facsimile transmissions, and medical image transmission and storage in a database.

In general, complex cryptography techniques make cracking of the system difficult, but are also expensive to implement. Since digital video transmission system usually includes a compression module that aims to reduce the transmitted bit rate, the cryptography techniques have to be carefully designed to avoid potential adverse impact on the compression efficiency, and on the functionalities that the compression format provides.

**FEATURES**

Video scramblers are commonly employed to prevent unauthorized access to video data. Several video scrambling systems rely on methods of directly distorting the visual image data (in the spatial domain) such that, without de-scrambling, the video appears unintelligible to a viewer. These scrambling techniques are not efficient for transmitting digital video signals because they, in general, will significantly change the statistical property of the original video signal, thus making it very difficult to compress.A more secure MPEG encryption algorithm is described in that encodes half of the data using DES and has a 47% gain in terms of number of XOR operations over DES. Although it may achieve real-time encryption/decryption for low-resolution, low bit rate MPEG sequence, it will have real-time implementation problem for high resolution, and high bit rate sequences.The above mentioned systems of scrambling the compressed bitstream may also be vulnerable to possible plain text attacks that take advantage of the syntax of the compression algorithm such as the known synchronization word or End of Block symbol that are often used to limit error propagation. To selectively encrypt some segments of the compressed data such as Intra-coded blocks often incurs additional header overhead to locate such segments. In addition, in some applications, a scrambled compressed video has to be transcoded, e.g., from a high bit rate to a low bit rate, due to change of the nature of the transmission channels. The above mentioned classical approach is subject to security threat for transcoding at intermediate routers of the transmission channel because the scrambling key has to be available at the transcoding points in order to decrypt, decompress, recompress and re-encrypt. This imposes a big challenge to the key management system and the intermediate routers have to be decryption/encryption capable. In some other applications, it is desirable that the encryption allows the content to be somewhat transparent instead of totally indiscernible. For example, a roadcaster of pay-TV may want to promote a contract with nonpaying watcher, by allowing them to access some of the video information. Classical approaches give no transparency.

**OBJECTIVE**

Our objective is to create the existing video file unintelligible to the Viewer using the Block cipher algorithm such as DES for intruders and hackers difficult to break the information. The algorithm is used to make the video file in the compressed format which makes efficient to travel across the Network without any traffic problems and also with higher security. The video file is converted into individual images at first in order to get the values of each frames and make the values interchange that makes the information unintelligible. Then the rearranged frames are converted again to video file it cannot be accessed except the client without the Descrambling process. In the Client side the Scrambled video is converted back to get the original Information as they have sent.A fundamental issue arising in mobile ad hoc networks (MANETs) is the selection of the optimal path between any two nodes. A method that has been advocated to improve routing efficiency is to select the most stable path so as to reduce the latency and the overhead due to route reconstruction

## II. LITERATURE STUDY

System analysis will be performed to determine if it is flexible to design information based on policies and plans of organization and on user requirements and to eliminate the weakness of present system. This chapter discusses the existing system, proposed system and highlights of the system requirements. Traditional cryptographic

algorithms / systems for data security are often not fast enough to process the vast amount of data generated by the multimedia applications to meet the real-time constraints.General cryptographic algorithms are not suitable for digital video data. Video scramblers are commonly employed to prevent unauthorized access to video data. These video scrambling systems rely on methods of directly distorting the visual image data and appears unintelligible to a viewer. These scrambling techniques are not efficient transmitting digital video signals Digital video transmission system usually includes a compression module to reduce transmitted bit rate. Scrambling digital images, the images are first subject to compression, and then the compressed image data is treated as ordinary data and is encrypted/decrypted using traditional cryptographic algorithms such as Data Encryption Standard (DES). Due to the high data rate of video (even compressed video), they usually add a large amount of processing overhead to meet the real-time video delivery requirement.

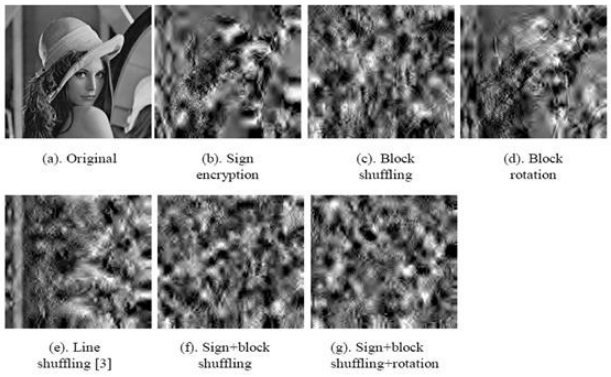
**Drawbacks**

- General cryptographic algorithms are not suitable for digital video data.
- Traditional cryptographic algorithms/systems for data security are often not fast enough to process the vast amount of data.
- Are not efficient transmitting digital video signals, digital video transmission system usually includes a compression module to reduce transmitted bit rate.

## III. DEVELOPMENT DIGITAL VIDEO SCRAMBLING AND DESCRAMBLING USING DWT IN CHROMINANCE CHANNEL

The proposed system includes both scrambling and compression framework in which digital video data is efficiently scrambled in the frequency domain without affecting the compression efficiency significantly. For the wavelet transform based scrambling system described above, the user can select to leave the low-resolution sub bands unscrambled in order to provide some level of transparency. For example, one potential application is to allow free access to low-resolution digital TV signal while requiring a key to watch high definition TV programs.

In fact, similar transparency is also possible for the DCT based system to be described where the DC and low frequency components can be left unscrambled.




Scrambling in the frequency domain may only involve changing the spatial positions of individual frequency coefficients.

This proposed system first Opens AVI Video and Extracts frames into BMP images. Then for each image the following steps repeated

- Read Image data
- Apply Block Rotation
- Apply Line Shuffling
- Apply Sign Encryption
- Finally build AVI video from the scrambled images.

**Benefits**

- The proposed framework has the nice feature scrambling process is very simple and efficient. Provides different levels of security, has very limited adverse impact on the compression efficiency and no adverse impact on the error resiliency.
- Allows more flexible selective encryption, transcodability/scalability, transparency and some other useful features.
- Effectiveness of scrambling motion vector information, which does not seem to have been properly, addressed in prior selective encryption schemes.
- Scrambling in the frequency domain may only involve changing the spatial positions of individual frequency coefficients

## MODULE DESIGN

### List of Modules

1. Wavelet Based Systems
2. Selective bit scrambling
3. Block shuffling

### Wavelet Based Systems

We assume the input video frames (original or residual error after motion compensation) are transformed using the wavelet filter banks.It shows 16 subbands that represents a five level wavelet decomposition of an input frame obtained by separable filtering along the vertical and the horizontal directions. Each subband represents selected spatial frequency information of the input video frame. The statistics of the coefficient distribution generally differ from subband to subband. In addition, because the coefficients of the subbands are arranged in the spatial arrangement of the original image, neighboring coefficient correlation exists that can be exploited by a bitstream coder. The goal here is to provide a coefficient scrambling/shuffling method that does not significantly destroy these statistical properties.

### Selective bit scrambling

The first basic approach scrambles selected bits in the transform coefficients to encrypt an image.Each bit of a coefficient can be viewed as one of three types. Significance bits for a coefficient are the most significant bit with a value of 1, and any preceding bits with a value of 0. These bits limit the magnitude of the coefficient to a known range. Refinement bits are the remaining magnitude bits, used to refine the coefficient within the known range. The sign bit determines whether the known range is positive or negative. It is recognized that the efficiency of a bitstream coder such as the bit plane coders proposed in differs, depending on the bit type being coded. Most transforms create a large number of coefficients having small magnitude, and tend to group small magnitude coefficients together. Thus the significance bits have relatively low entropy, and are therefore highly compressible. On the other hand, most transforms produce coefficients with sign bits that have an approximately equal probability of being a 1 or a 0, and that are highly uncorrelated with the sign bits of neighboring coefficients. Refinement bits also tend to have approximately equal probabilities of 1 or 0, and are highly uncorrelated with neighboring refinement bits. Because of their high entropy (and limited predictability), the sign bits and refinement bits are not highly compressible.

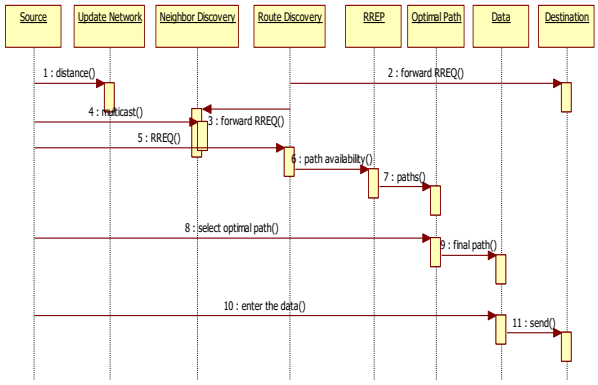
### Block shuffling

To increase the level of security, block shuffling is proposed. We divide each subband into a number of blocks of equal size. The size of the block can vary for different subbands. Within each subband, blocks of coefficients will be shuffled according to a shuffling table generated using a key. The shuffling table generally will be different for different subbands, and can vary from frame to frame. Since the scrambling performed by block shuffler is block-based, it retains most of the local 2-D statistics of the subband signal. Only coefficients around the block boundary may be slightly affected. Therefore, the negative impact on subsequent statistical coding is very small (e.g., less than 5% bit rate increase while the visual effect of the shuffling on a decompressed encrypted image is dramatic.

### Block rotation

To further improve security with little impact on statistical coding, each block of coefficients can be rotated to form an encrypted blocks. The encrypted block is selected from a set of eight blocks that are rotated versions of the original block. The key controls the selection process. Block rotation retains most of the local 2-D statistics of the subband signal. It could be considered as a special case of shuffling coefficients within a block taken from certain subband.

### Sequence Diagram



## IV. RESULT AND DISCUSSION

The implementation phase focuses how the engineer attempts to develop the system. It also deals with how data are to be structured, how procedural details are to be implemented, how interfaces are characterized, how the design will be translated into programming and hoe the testing will be performed. The methods applied during the development phase will vary but three specific technical tasks should always occur.

- The software design
- Code generation
- Software testing

The system group has changed with responsibility to develop a new system to meet requirements and design and development of new information system. The source of these study facts is variety of users at all level throughout the organization.

**Stage of Development of a System**

- Feasibility assessment
- Requirement analysis
- External assessment
- Architectural design
- Detailed design
- Coding
- Debugging
- Maintenance

**Feasibility Assessment**

In Feasibility this stage problem was defined. Criteria for choosing solution were developed, proposed possible solution, estimated costs and benefits of the system and recommended the course of action to be taken.

**Requirement Analysis**

During requirement analysis high-level requirement like the capabilities of the system must provide in order to solve a problem. Function requirements, performance requirements for the hardware specified during the initial planning were elaborated and made more specific in order to characterize features and the proposed system will incorporate.

**External Design**

External design of any software development involves conceiving, planning out and specifying the externally observable characteristic of the software product.

These characteristics include user displays, report formats, external data source and data links and the functional characteristics.

**Internal Design Architectural and Detailed Design**

Internal design involved conceiving, planning out and specifying the internal structure and processing details in order to record the design decisions and to be able to indicate why certain alternations were chosen in preference to others. These phases also include elaboration of the test plans and provide blue prints of implementation, testing and maintenance activities. The product of internal design is architectural structure specification.

The work products of internal design are architectural structure specification, the details of the algorithm, data structure and test plan. In architectural design the conceptual view is refined.

**Detailed Design**

Detailed design involved specifying the algorithmic details concerned with data representation, interconnections among data structures and packaging of the software product. This phase emphasizes more on semantic issues and less synthetic details.

**Coding**

This phase involves actual programming, i.e, transacting detailed design into source code using appropriate programming language.

**Debugging**

This stage was related with removing errors from programs and making them completely error free.

**Maintenance**

During this stage the systems are loaded and put into use. They also get modified accordingly to the requirements of the user. These modifications included making enhancements to system and removing problems.

## V. CONCLUSION AND FUTURE ENHANCEMENT

The unique feature of multimedia data is that the data rate is high while the information value is usually lower than ordinary electronic information. This feature justifies the employment of lightweight cryptographic algorithms to reduce the computational overhead while still maintaining reasonable level of security. In this project that by jointly considering encryption and compression, efficient compression-friendly digital video scrambling techniques can be designed to facilitate simple implementation and to allow for video transcodability/scalability, transparency, and other useful encryption-domain signal processing. The proposed scrambling techniques appear to achieve a very good compromise between several desirable properties such as speed, security, file size and transcodability, therefore is very suitable for network video applications

### Scope for Future Enhancement

In future we plan to improve the performance significantly by developing algorithms for:

- 100% quality at both the ends – client terminals will also receive the video with full quality.
- Further Video compression upto half the original size of video without losing its quality.

## REFERENCES

[1] Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2016–202, Cisco, 2017.

[2] TERA Consultants, “Building a digital economy. The importance of saving jobs in EuropesCreative Economy Paris: International Chamber of commerce, 2010.

[3] N. Nikolaidis, A. Tefas, I. Pitas, „Chaotic sequence for digital watermarking“, Advances in Nonlinear Signal and Image Processing, Ed. S. Marshall and G. L. Sicuranza, Hindawi, 2006.

[4] L. Agilandeeswari, K. Ganesan, “A robust color video watermarking scheme based on hybrid embedding techniques,”, Multimed Tools Appl, 75: 8745-8780, 2017. https://doi.org/10.1007/s11042-015- 2789-9, 2016.

[5] K. Surachat, “A Robust Digital Watermarking using Discrete Wavelet Transformation in Chrominance Channel” Journal of Automation and Engineering, vol. 1, no. 2, pp. 99102, Jan. 2013.

[6] Md. Asikuzzaman, Md. J. Alam, A. J. Lambert, M. R. Pickering, “A Robust Digital Watermarking using Discrete Wavelet Transformation in Chrominance Channel”, IEEE Trans. Inf. And Forensics and Security, vol. 9, no. 9, pp. 1502-1517, 2014.

[7] P. Singh, B. Raman, “A Secured Robust Watermarking Scheme Based on Majority Voting concept for Rightful Ownership assertion”, Multimed Tools Appl, 76: 21497, 2017. https://doi.org/10.1007/s11042-016-4006-x.

INTERNATIONAL JOURNAL OF INNOVATIVE TECHNOLOGY AND CREATIVE ENGINEERING(ISSN:2045-8711)
VOL.13 NO.08 AUG
2023

# ENSEMBLE DEEP LEARNING STRATEGIES FOR CHEST X-RAY IMAGES TO IMPROVE COVID-19 CASE IDENTIFICATION

**S Deepankumar**[1]
Research Scholar, Department of Computer Science,
Dr.SNS Raja Lakshmi College of Arts and Science,
Bharathiar University, Coimbatore, Tamil Nadu, India
**Dr R Maruthaveni**[2]
Assistant Professor, Department of Computer Science,
Dr.SNS Raja Lakshmi College of Arts and Science,
Bharathiar University, Coimbatore, Tamil Nadu, India

**Abstract: People's lives continue to be drastically affected by COVID-19 all around the world. It is essential to quickly and cheaply screen the affected patients in order to fight this disease. Radiological testing, with a chest X-ray being the most accessible and affordable alternative, is one of the most effective ways to get there. Using chest X-ray pictures, we have suggested a Deep Convolutional Neural Network-based approach that can identify COVID-19 +ve patients. In the proposed work, many cutting-edge CNN models—including DenseNet201, Resnet50V2, and Inceptionv3—have been used. They everyone received individualized training to develop their own judgment. To forecast a class value, the models are then integrated using a novel weighted average ensembling strategy.Using publicly accessible chest X-ray pictures of COVID +ve and -ve cases, we evaluated the effectiveness of the treatment. To create training, test, and validation sets, 538 photos of COVID +ve patients and 468 images of COVID -ve patients were used. With a classification accuracy of 91.62%, the suggested method outperformed both the benchmark algorithm and the most recent CNN models. We created a GUI-based program for general use. Any medical professional can use this application to quickly identify COVID +ve patients using chest X-ray pictures by running it on any computer.**

Keywords : COVIF,X-ray,PCR

## 1. Introduction

The largest metropolitan area of the Chinese province of Hubei, Wuhan, reports the first official case of the coronavirus, as certified by WHO.

Millions of verified cases have been reported worldwide, and it has already claimed thousands of lives. A pandemic outbreak has disastrous effects on the health and wellbeing of the enti]re world's population. This resulted in the development of the coronavirus illness 2019 (acronym for Severe Acute Respiratory Syndrome coronavirus, or SARS-CoV).Reverse transcription polymerase chain reaction (RT-PCR) genetic tests are currently the mostcommon ones used to diagnose COVID-19 These assessments are quite precise. It is possible to identify and measure viruses in patient samples, even if they are very little. However, it is important to remember that the PCR test is exceedingly difficult, time-consuming, and expensive. Consequently, not all healthcare facilities have the capacity to carry it out. Given these restrictions, radiography scanning can be used as a stand-in method to identify the condition. Chest radiography images can be examined to look for, or rule out, various diseases. the new coronavirus's symptoms. According to studies This family of viruses exhibits a substantial radiographic appearance. As a result, it may be argued that categorization using radiographic images, such as a chest X-ray (CXR), is not only significantly quicker and less expensive than a PCR test, but also more accurate. Additionally, chest X-rays may be obtained in practically every clinic and are more affordable than other radiological examinations like CT scans.The only issue with CXR-based detection of COVID-19 patients, particularly in remote places, is that qualified physicians might not always be accessible.

Additionally, many professionals lack prior experience with COVID-19 positive patient CXRs, making the radiological symptoms associated with COVID-19 new and unexpected.

## 2. Related Works

Medical diagnosis has seen widespread use of computer vision. It is helpful in medical specialties like dermatology that call for visual examinations. To determine whether a skin irregularity is a possible early sign of skin cancer, computer vision is employed as a diagnostic tool. Additionally, it is used to find problems with the body, particularly with the tissues, blood vessels, joints, etc. It aids in the early diagnosis of conditions like diabetic retinopathy in ophthalmology, preventing blindness. In addition to medicines, it has demonstrated excellent success in procedures. Computer vision applications use a variety of medical imaging formats, including ultrasound, chest X-rays, magnetic resonance imaging (MRI), PET scans, and computed tomography (CT) scans. Medical imaging, according to studies, aid in the analysis of viruses in the lungs. Deep learning-based algorithms have been established in several works to diagnose pneumonia, other classes of thoracic disorders , skin cancer, hemorrhage categorization, etc. from medical images. With very modest architecture, some of these research have shown encouraging results.In a study COVID-19 individuals were identified using CT scan pictures and a convolutional neural network (CNN) model. There are numerous further studies being done to determine whether a CT scan can identify the COVID-19 virus in human lungs [11,20–23]. Throughout Yan's work Using CXR pictures, COVID-19 patients can be identified[4,5,7,12,26–32]. Transfer learning has been utilized in one such study by Makris et al. [4] to classify normal, pneumonia, and COVID-19 patients using CXR pictures. In a different study by Mangal et al. [6],normal people, patients with viral and bacterial pneumonia, and COVID-19 patients were separated using the DenseNet with ChexNet design. Xception and ResNet50V2 were concatenated by Rahimzadeh. In addition to utilizing individual state-of-the-art deep learning models, Wang work from has created a unique architecture known as the COVID Net architecture for the categorization of COVID-19 patients, healthy participants, and pneumonia patients.

With a classification accuracy of 94%, this custom network exceeds laboratory tests. It was created utilizing the lightweight projection-expansion-projection-extension (PEPX) construction pattern. It is evident that the majority of studies investigating COVID-19 identification in CXR pictures have used specific deep learning models, such as DenseNet, ResNet, Xception, etc. To increase the models' classificational power, no effort has attempted to integrate them. Numerous studies on ensemble learning with deep neural networks demonstrate the effectiveness of ensemble learning. Numerous studies on ensemble learning with deep neural networks demonstrate that these methods are more accurate predictors than single models and also aid in avoiding over fitting. As a method for ensembling, a weighted average of the output probabilities has been introduced It is discovered that it is superior to the unweighted average.The relative effectiveness of various ensemble approaches with Convolutional Neural Networks, such as unweighted average, majority voting, Bayes Optimal Classifier, and Super Learner, has been examined in a different study In this study, a novel approach to classify COVID-19 +ve patients from CXR pictures using three cutting-edge CNN models

## 3. Ensemble Learning

Ensemble approaches outperformed a single model, according to reports in the literature. In addition to enhancing classification performance, it can also lessen the danger of overfitting. Predictions are created by ensemble learning using estimates from other classifiers or models. To improve the performance in this case, we aggregated the top five models from each strategy. We used majority voting-based ensemble learning since it is an easy method to come to a conclusion from many predictions produced by different models. Hard voting and soft voting are the two distinct majority voting techniques.Ensemble learning with both hard and soft majority voting techniques. Hard voting selects the final forecast with the highest total of model votes. According to reports in the literature, ensemble approaches outperformed single models in terms of performance.When two classes receive an equal number of votes, we decided to give the final label to the class with the smallest index. Soft voting adds up the models' estimated probabilities for each class, and the class label with the highest probability is used as the chosen class label.

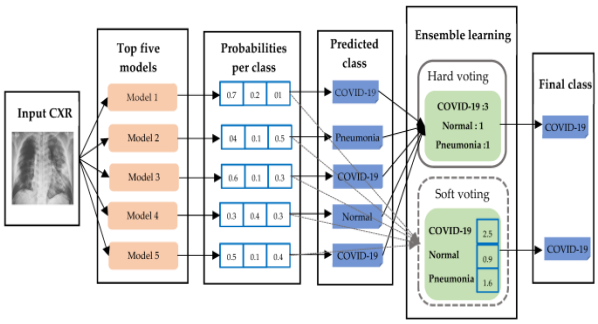



**Fig.1 Ensemble learning using majority hard voting and soft voting**

**Data Set Generation**

We have gathered the photographs for this study from various free online sources These publicly available, open-source datasets largely from Europe include CXR images of COVID-19 positive patients, patients with pneumonia and other illnesses, and healthy participants. This data consists of CXR scans taken by different patients, only the frontal photos are taken into account, and the lateral images are ignored. This is so that the area of interest, which is the lungs, may be inspected more effectively from the front than from the side. The COVID-19, Pneumonia, and Normal labels on the initial set of CXR pictures are still in use today. To simplify things for this project, we have divided the photographs into COVID-19 POSITIVE (also known as class 0) and COVID-19 NEGATIVE.

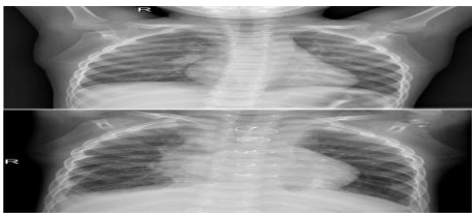


**Fig.2 CXR images of COVID-19 positive subjects**

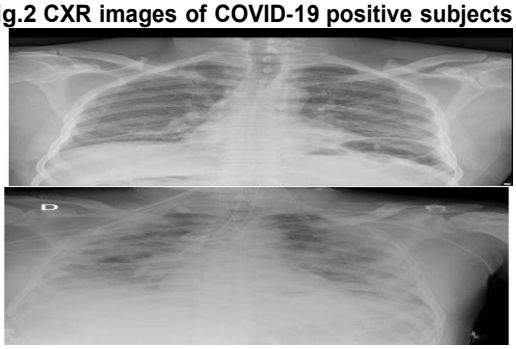


**Fig.3 CXR images of COVID-19negative subjects**

Similar to that, there are 117 photos in the validation data, 57 of which are for class 0 and 60 for class 1. We used 5-fold cross validation while dividing the image collection to ensure optimal training.

The images for the training and test datasets were chosen by splitting the combined collection of photos (i.e., training and test images combined) into 5 folds and selecting a different fold each time as the test set. A fixed set of images was chosen as the validation data. Five iterations, one with each fold of test data and associated training data, have been executed.

## 4. Tools and Performance Metrics

| MODEL | NORMAL | PNEUMONIA | COVID-19 |
|---|---|---|---|
| MODEL 1 | 96.3 | 89.6 | 79 |
| MODEL 2 | 96.2 | 91.5 | 94 |
| MODEL 3 | 95.4 | 94.6 | 93.3 |
| MODEL 4 | 96.4 | 93.1 | 95 |
| MODEL 5 | 94.8 | 92.7 | 95 |

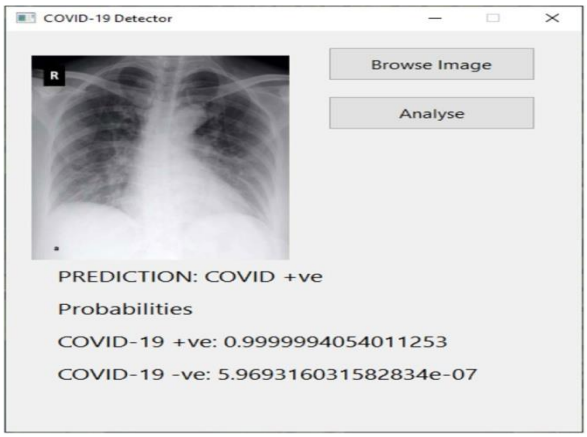
Tensor Flow 2.2.0, Python 3.7, and Google Colab GPU (Tesla K80 12GB GDDR5 VRAM) have all been used. TensorFlow 2.2.0's deep learning library is utilized for the implementation of CNN, while the Google Colab platform is used for training and testing. Classification accuracy, sensitivity, and F1-score are the measures used to assess how well the proposed approach performs.

Classification accuracy = TP+TN/TP+TN+FP+FN

Sensitivity=TP/TP+FN

F1Score=2xSensitivityxPrecision / Sensitivity+Precision

Where TP denotes True Positive, FP denotes False Positive, FN is False Negative, and TN denotes True Negative. In a confusion matrix, the COVID-19 +ve situations that the model correctly classifies are referred to as True Positives, while the COVID -ve cases that the model mistakenly classifies are referred to as False Positives. Similarly, COVID -ve subjects correctly identified as COVID +ve are referred to as False Negative, and COVID +ve subjects wrongly classified as COVID -ve are referred to as True Negative.In this study, we suggest a brand-new ensemble technique based on weighted aver- ages. We have contrasted the outcomes of the proposed algorithm with those of a recent study on COVID-19 detection from X-ray images that used a concatenation f benchmark CNNs (henceforth referred to as "Concatenated network") in addition to the outcomes from the individual networks used in the ensemble. This provides a clear idea of how the proposed algorithm performs when compared to recent work in the same field.




## 5. Conclusion

Based on the suggested remedy, a straightforward desktop tool for identifying COVID-19 positive and negative cases has been created. Any medical staff member can now browse a chest X-ray image by inputting it to the application. This will therefore identify the COVID +ve and COVID -ve situations as well as their probability, as illustrated in Fig. 11. It is compatible with operating systems like Windows, Mac OS, and Linux. Any COVID-19 testing facility or other healthcare facility can utilize this interface for the quick detection of the illness. To stop the disease from spreading and maintain control over it, quick and accurate diagnosis of COVID +ve patients is required. To find the COVID +ve, this investigation was conducted. +ve patients in an easy and affordable manner from Chest X-Ray pictures. Three cutting-edge deep learning models have been adopted and combined in the work that is being proposed in this study. The classification accuracy of the suggested model is 91.62%. The fact that it produces a sensitivity of almost 95% for COVID +ve cases—i.e., that out of 100 COVID +ve patients, more than 95 can be accurately diagnosed by our suggested model is even more significant. It help the doctors to detect the affected patients with the help of computer-aided analysis.The medical field will benefit greatly from this, in our opinion.

## REFERENCES

[1] S. Chen, E. Dobriban, and J. H. Lee, “Invariance reduces variance: Understanding data augmentation in deep learning and beyond,” 2019, arXiv:1907.10905.

[2] Y.-Y. Cheng, P. P. Chan, and Z.-W. Qiu, “Random forest based ensemble system for short term load forecasting,” in Proc. Int. Conf. Mach. Learn. Cybernet., 2012, pp. 52–56.

[3] M. E. Chowdhury *et al.*, “Can AI help in screening Viral and COVID-19 pneumonia?,” 2020, *arXiv:2003.13145*.

[4] F. Divina, A. Gilson, F. Goméz-Vela, M. García Torres, and J. F. Tor- res, “Stacking ensemble learning for short-term electricity consumption forecasting,” *Energies*, vol. 11, no. 4, 2018, Art. no. 949.

[5] F. A. Faria, J. A. Dos Santos, A. Rocha, and R. d. S. Torres, “A framework for selection and fusion of pattern classifiers in multimedia recognition,” *Pattern Recognit. Lett.*, vol. 39, pp. 52–64, 2014.

[6] M. Farooq and A. Hafeez, “COVID-ResNet: A deep learning framework for screening of COVID19 from radiographs,” 2020, *arXiv:2003.14395*.

[7] W.-J. Guan *et al.*, “Clinical characteristics of coronavirus disease 2019 in China,” *New England J. Med.*, vol. 382, no. 18, pp. 1708–1720, 2020.

[8] C. Huang *et al.*, “Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China,” *Lancet*, vol. 395, no. 10223, pp. 497–506, 2020.

[9] G. Huang, Y. Li, G. Pleiss, Z. Liu, J. E. Hopcroft, and K. Q. Weinberger, “Snapshot ensembles: Train 1, get M for free,” 2017, *arXiv:1704.00109*.

[10] A. Jacobi, M. Chung, A. Bernheim, and C. Eber, “Portable chest x-ray in coronavirus disease-19 (COVID-19): A pictorial review,” *Clin. Imag.*, vol. 64, pp. 35–42, 2020.

[11] D. Jakubovitz, R. Giryes, and M. R. Rodrigues, “Generalization error in deep learning,” in *Compressed Sensing and Its Applications*, Springer, 2019, pp. 153–193.

[12] A. Kumar, P. K. Gupta, and A. Srivastava, “A review of modern technolo- gies for tackling COVID-19 pandemic,” *Diabetes Metabolic Syndrome Clin. Res. Rev.*, vol. 14, no. 4, pp. 569–573, 2020.

[13] Ilyas M, Rehman H, Nat-Ali A (2020) Detection of COVID-19 from chest X-ray images using artificial intelligence: an early review. arXiv:2004.05436

[14] Mahmud T, Rahman MA, Fattah SA (2020) CovXNet: a multidila- tion convolutional neural network for automatic COVID-19 and other pneumonia detection from chest X-ray images with trans- ferable multi-receptive feature optimization. Comput Biol Med 122:103869

[15] Ozturk T, Talo M, Yildirim EA, Baloglu UB, Yildirim O, Acha- rya UR (2020) Automated detection of COVID-19 cases using deep neural networks with X-ray images. Comput Biol Med 121:103792

[16] L. J.Muhammad, E. A.Algehyne, S. S.Usman, A.Ahmad, C.Chakraborty, and I. A. Mohammed, “Supervised machine learningmodels for prediction of COVID-19 infection usingepidemiologydataset,” SN computer science, vol. 2, pp. 1-13, 2021.

[17] S.Mohan, A.Abugabah, S.Kumar Singh, A.KashifBashir, and L. Sanzogni, “An approach to forecast impact of Covid-19 usingsupervised machine learning model,” Software: Practice and Experience, vol. 52, no. 4, pp. 824-840,2022.

[18] S.Chakraborty, B.Murali, andA. K.Mitra, “An efficient deeplearning model to detect COVID-19 usingchest X-ray images,” International Journal of EnvironmentalResearch and Public Health, vol. 19, no. 4, 2013.

[19] E.Irmak, “A noveldeepconvolutional neural network model for COVID-19 diseasedetection,”In 2020 Medical Technologies Congress (TIPTEKNO)pp. 1-4,2020. IEEE.

[20] A.Akhtar, S.Akhtar, B.Bakhtawar, A. A.Kashif, N.Aziz, and M. S.Javeid, “COVID-19 detectionfrom CBC using machine learning techniques,” International Journal of Technology, Innovation and Management (IJTIM), vol. 1, no. 2, pp. 65-78,2021.

[21] D.Assaf, Y. A.Gutman, Y.Neuman, G.Segal, S.Amit, S.Gefen-Halevi, and A.Tirosh, “Utilization of machine-learningmodels to accuratelypredict the risk for critical COVID-19,” Internal and emergency medicine, vol. 15, pp. 1435-1443,2020

[22] M.AlazabSoui, N.Mansouri, R.Alhamad, M.Kessentini, andK.Ghedira, “NSGA-II as featureselection technique and AdaBoost classifier for COVIDpredictionusingpatient’ssymptoms,” Nonl ineardynamics, vol. 106, no. 2, pp. 1453-1475,2021.

[23] F. M.Salman, S. S.Abu-Naser, E.Alajrami, B. S.Abu-Nasser and B. A. Alashqar, “Covid-19 detectionusingartificial intelligence,2020.

# LEARNING SYSTEMS AND THEIR CHALLENGES IN CURRENT TECHNOLOGIES

**P.R.Pameela Rani**
Assistant Professor
Department of computer Technology
KG College of Arts and Science, Coimbatore, Tamilnadu, India

**Abstract-Instructors are continuously endeavoring to customize learning for understudies. Innovation can assist them with arriving at new levels with admittance to continuous understudy information, longitudinal data, content, applications, and that's just the beginning. Innovation can assist instructors with establishing mixed learning conditions and influence computerized devices for developmental and summative evaluations, bringing new models for learning and educating to homerooms.**

**Innovation in schooling and the right gadgets in understudies' grasp sets them up with the profession and specialized abilities they should find actual success today and in the upcoming labor force. Significant growth opportunities in STEAM can rouse imagination, assist understudies with applying importance to their learning, and set them up for future vocation open doors and occupations that haven't even been made at this point. Explicit abilities in coding, programming, actual processing, and computational reasoning have become normal necessities in the labor force. However making, understudies can acquire these abilities and sharpen their critical thinking and decisive reasoning abilities for the 21st 100 years. Advancing by doing with creator outlooks and conditions can be extremely captivating when planned and incorporated with the right innovation.**

## How important is Technology in Education

The COVID-19 pandemic is rapidly showing why online schooling ought to be an imperative piece of educating and learning. By coordinating innovation into existing educational plans, instead of utilizing it exclusively as an emergency the board instrument, educators can saddle web based advancing as a strong instructive device.

The viable utilization of computerized learning apparatuses in home rooms can increment under study commitment, assist educators with further developing their illustration designs, and work with customized learning. It additionally assists understudies with building fundamental 21st-century abilities.

Virtual homerooms, video, increased reality (AR), robots, and other innovation devices might not just make at any point class more exuberant, they can likewise establish more comprehensive learning conditions that cultivate joint effort and curiosity and empower instructors to gather information on understudy execution.

In any case, it's critical to take note of that innovation is a device utilized in schooling and not an end in itself. The commitment of instructive innovation lies in how teachers manage it and the way things are utilized to best help their understudies' necessities.

## Educational Technology Challenges

Built In reports that 92% of educators figure out the effect of innovation in training. As per Project Tomorrow, 59% of center school understudies say computerized instructive devices have assisted them with their grades and grades. These devices have become so well known that the instructive innovation market is projected to grow to $342 billion by 2025, as indicated by the World Economic Forum.

In any case, instructive innovation has its difficulties, especially with regards to execution and use. For instance, regardless of developing revenue in the utilization of AR, computerized reasoning, and other arising innovation, under 10% of schools report having these apparatuses in their homerooms, as per Project Tomorrow. Extra worries incorporate unnecessary screen time, the viability of educators utilizing the innovation, and stresses over innovation value.

Conspicuously ascending from the COVID-19 emergency is the issue of content. Teachers should have the option to create and say something regarding on the web instructive substance, particularly to urge understudies to think about a subject according to alternate points of view. The earnest moves initiated during this emergency didn't give adequate chance to this. Access is an additional worry — for instance, only one out of every odd school locale has assets to furnish understudies with a PC and web network can be problematic in homes.

Also, while certain understudies flourish in web-based schooling settings, others slack for different elements, including support assets. For instance, an understudy who previously battled in up close and personal conditions might battle much more in the ongoing circumstance. These understudies might have depended on assets that they never again have in their homes.

In any case, most understudies normally show trust in utilizing on the web schooling when they have the assets, as review have proposed. Nonetheless, online schooling might present difficulties for educators, particularly where it has not been the standard. In spite of the difficulties and concerns, it's critical to take note of the advantages of innovation in schooling, including expanded joint effort and correspondence, worked on nature of training, and drawing in illustrations that assist with starting creative mind and a quest for information in understudies.

## Benefits of Technology in Education

Instructors need to further develop understudy execution, and innovation can assist them with achieving this point. To moderate the difficulties, executives ought to assist educators with acquiring the capabilities expected to improve learning for understudies through innovation. Furthermore, innovation in the homeroom ought to make educators' positions simpler without adding additional chance to their day.

Innovation furnishes understudies with simple to-get to data, sped up learning, and fun chances to rehearse what they realize. It empowers understudies to investigate new subjects and extend how they might interpret troublesome ideas, especially in STEM. Using innovation inside and outside the homeroom, understudies can acquire 21st-century specialized abilities fundamental for future occupations.

In any case, youngsters learn all the more successfully with heading. The World Economic Forum reports that while innovation can help youthful understudies learn and obtain information through play, for instance, proof recommends that gaining is additional viable through direction from a grown-up, like an instructor.

Pioneers and executives ought to assess where their personnel are regarding how they might interpret online spaces. From illustrations got the hang of during this troublesome time, they can carry out arrangements now for what's to come. For instance, executives could allow instructors possibly 14 days to consider cautiously about how to show courses not already on the web. Not withstanding an investigation of arrangements, adaptability during these difficult times is of central significance.

The following are instances of how significant innovation is in training and the advantages it offers to understudies and educators.

## Customized Learning Opportunities

Innovation permits all day, every day admittance to instructive assets. Classes can occur totally online through the utilization of a PC or cell phone. Mixture adaptations of gaining consolidate the utilization of innovation from anyplace with normal in-person homeroom meetings. In the two situations, the utilization of innovation to tailor learning plans for every understudy is conceivable. Educators can make illustrations in light of understudy interests and qualities. An additional advantage is that understudies can learn at their own speed. At the point when they need to survey class material to get a superior comprehension of fundamental ideas, understudies can survey recordings in the illustration plan. The information produced through these web-based exercises empower instructors to see which understudies battled with specific subjects and deal extra help and backing.

## Interest Driven by Engaging Content

Through connecting with and instructive substance, educators can ignite curiosity in kids and lift their interest, which examination says has connections to scholastic achievement. Interest assists understudies with getting a superior comprehension of math and understanding ideas. Making connecting with content can include the utilization of AR, recordings, or webcasts.

For instance, while submitting tasks, understudies can incorporate recordings or cooperate with understudies from across the globe.

### Separating/Breaking down The Impact Of Educational Technology

Late years have seen an extraordinary development in the mechanical world. With an ever increasing number of individuals getting into advanced media, it is fundamental for educators to utilize the most recent devices that anyone could hope to find in their work to draw in understudies. To draw in understudies in learning, one should be imaginative and novel thoughts ought to be presented so understudies become amped up for what they are realizing. The utilization of instructive innovation has become fundamental for educators due to its significance in the present schooling industry.

The mechanical advances made throughout the years have influenced each field of life. As innovation has become more open, it has entered into all aspects of our lives and, thus, we presently expect each gadget that we use to be all around as innovatively progressed as could be expected. Schools have likewise embraced innovation as a vital piece of their schooling interaction and, accordingly, they presently offer software engineering classes, where understudies figure out how to code PCs and use them for their work or activities.

The utilization of instructive innovation has consequently become significant for the present students since it offers them a chance to learn at a lot quicker pace than they would in the event that they were not utilizing such devices and projects.

### The Benefits

- It Offers A Wider Choice Of Materials That Can Be Accessed Easily.
- It Helps Improve Learners' Communication Skills And Performance In School And Workplace Settings.
- It Provides a Fun and Engaging Learning Experience For Students.
- It Allows Learners to Access the Internet from Anywhere at Any Time.
- It Helps Learners Learn New Skills and Acquire New Knowledge.
- It Allows Students To Improve Themselves Both Mentally And Physically.
- It Helps Learners Stay Up To Date with New Technological Advancements.

### Students and teachers can access information at any time!

This is conceivably the clearest advantage of innovation. At the point when old-teachers were understudies, they needed to go through hours in the library searching for the data they required. Today, innovation combination makes everything unique and easier. Understudies can undoubtedly get to papers, logical articles, studies, and some other sort of satisfied on the web. They can compose better, further scholastic papers since they can uphold their contentions with more proof. At the point when you give a talk the understudies don't have any idea, they can find less complex directions and data with a solitary Google search.

### Educational Technology Trends In 2020-2021

Large Data, Machine Learning, and the Internet of Things (IoT) were the greatest instructive innovation patterns of 2019. Notwithstanding, distance learning has turned into the one pattern that rules them all. The COVID-19 pandemic has radically altered the manner in which we educate and learn. Understudies currently need to become accustomed to separate learning through computerized stages because of social removing. Despite the fact that a few schools are returning, this pattern might go on until 2021. The most recent EdTech patterns in 2020, and further into 2021 are being reformed with serious areas of strength for an on network, flexibility, and understudy focused learning. We should investigate the most recent top 10 patterns in instructive innovation.

### Top Current Educational Technology Trends

#### 1. eLearning

Distance learning turned into the main 2020 instructive innovation pattern for the time being a result of the quick spread of COVID-19 and school terminations. This prompted a rising interest for online instructive stages. eLearning is schooling or preparing conveyed electronically. It very well may be slide-based web-based exercises, or it can likewise be an internet based course that assists a business with preparing workers in important abilities.

With eLearning, instructive substance is conveyed to students through PCs, workstations, tablets, or cell phones. Saving time as well as opening numerous entryways for intelligent learning. As opposed to being in a latent encounter, students can pick what they need to realize rapidly and effectively, any place.

They likewise learn through collaborating straightforwardly with on-screen data through, for example, hauling content starting with one spot then onto the next. In addition, the dynamic situations in eLearning likewise urge students to pursue their own decisions on what they will realize straightaway.

In eLearning, students simply absorb information through perusing or survey content, it alters how training is conveyed. Likewise, numerous eLearning courses incorporate movement, webcasts, and recordings that make a multimodal and functional opportunity for growth.

The last point is, despite the fact that eLearning has been around for quite a while, it is remaining green and consistently creating. Instructors are utilizing the benefits of innovation to make learning more powerful. That is the reason increasingly on the web and mixed learning courses are delivered these days.

Assortment is the extraordinary element of web based learning stages. You can show your understudies progressively (coordinated) by means of live stream or gathering gatherings utilizing Zoom or Microsoft Teams, or you can utilize recorded (non concurrent) systems with a large number of media and computerized capabilities accessible to improve illustrations. A decent internet learning stage can likewise be joined with a Learning Management System (LMS) so you can monitor your understudies' learning results.

**2.AI-based Personalised Analysis of Individual**

Artificial Intelligence (AI) is modern day buzzword in the world today. Not solely science area however all sectors together with finance, manufacturing and healthcare are going to be revolutionized due to implementation of AI. Education zone is additionally in all likelihood to get extra thoughts and implementations the usage of AI in the upcoming years.

Feedback from instructors (human) has some boundaries in phrases of capacity to know pupil overall performance whilst taking into issues historic performance. AI can assist presenting special evaluation of scholar overall performance in the examinations.

AI can be developed to analyse a candidate's responses whilst evaluating psychometric tests. AI can take algorithm-based choices which would show very beneficial to analyse a candidate's responses and in flip recognize the candidate's mind-set and behaviour patterns. It can additionally assist to make predictions about the candidate's behaviour in future. Thus, customized gaining knowledge of and evaluation of the studying is viable the use of AI pushed assessments of the students.

**3.Gamification**

The want for immersive and interactive studying techniques in educating has led to instructors and educators lending their assignments and tasks in the structure of games. While now not being a new prospect in the subject of education, as it has been in the industry's motto for over 20 years, today's generational mastering relies upon on it with full force. While most human beings assume that gamification refers to the system of integrating solely video games into the instructing process, it is so a whole lot wider in its attain than it may appear at face value.

Introducing gamification into one's instructing agenda may additionally encompass making the mastering journey for an man or woman as clear as viable through supplementing him guidelines noted absolutely and boundaries made especially to indulge in the innovative behaviour of the student.

Excluding stereotypical modules from the minds of creators as properly as dad and mom and instructors is a must, as barring that the have faith in gamification by using the college students will be restricted ensuing in lesser consequences than expected.

## CONCLUSION

The timing has by no means been higher for the use of science to allow and enhance getting to know at all levels, in all places, and for humans of all backgrounds. From the modernization of E-rate to the proliferation and adoption of brazenly licensed instructional resources,thekey portions vital to understand satisfactory the transformations made viable by means of science in schooling are in place.

Educators,policymakers,administrators,and trainer practise and expert improvement packages now need to embed these equipment and sources into their practices. Working in collaboration with families, researchers, cultural institutions, and all different stakeholders, these companies can

Dispose of inefficiencies, attain past the partitions of ordinary classrooms, and shape sturdy partnerships to guide everywhere, all-the-time learning.

Although the presence of science does no longer make sure fairness and accessibility in learning, it has the electricity to decrease limitations to each in methods in the past impossible. No count number their perceived competencies or geographic locations, all rookies can get admission to resources, experiences, planning tools, and data that can set them on a direction to obtaining know-how inconceivable a technology ago.

All of this can work to increase the knowledge, skills, and capabilities of educators. Tools and statistics structures can be built-in seamlessly to supply data on scholar getting to know growth past the static and dated rankings of regular assessments. Learning dashboards and collaboration and verbal exchange equipment can assist join instructors and households with immediate ease. This all is made extra possibly with the training of robust imaginative and prescient and management at all tiers from teacher-leaders to school, district, and nation administrators. For these roles, too, technological know-how approves larger communication, aid sharing, and extended exercise so that the imaginative and prescient is owned via all and committed to assisting each person in the machine enhance getting to know for students.

## REFERENCES

[1] Indonesian government, Law No. 20 of 2003 on National Education System, 2003.
[2] A. Januszewski and M. Molenda, Educational technology: A definition with commentary. Routledge, 2013.
[3] N. Tshuma, “The vulnerable insider: navigating power, positionality and being in educational technology research,” Learn. Media Technol., vol. 46, no. 2, pp. 218-229, 2021, doi: doi.org/10.1080/17439884.2021.1867572.
[4] M. Bond and S. Bedenlier, “Facilitating Student Engagement through Educational Technology: Towards a Conceptual Framework,” J. Interact. Media Educ., vol. 2019, no. 1, pp. 1-14, 2019, doi: https://doi.org/10.5334/jime.528.
[5] S. Lidolf and D. Pasco, “Educational Technology Professional Development in Higher Education: A Systematic Literature Review of Empirical Research,” in Frontiers in Education, vol. 5, pp. 1-20. 2020, doi: https://doi.org/10.3389/feduc.2020.00035.
[6] A. Januszewski, Educational technology: The development of a concept. Libraries Unlimited, 2001.
[7] D. Ariani, “Actualization of the Educational Technology Profession in Indonesia (in Indonesia),” Indonesian Journal of Curriculum and Educational Technology Studies, vol. 5, no. 1, pp. 1-9, 2017.
[8] M. B. Miles, A. M. Huberman, and J. Saldana, Qualitative Data Analysis: A Methods Sourcebook. SAGE Publications, 2013.
[9] Sugiyono, Educational Research Methods (Quantitative, Qualitative, Combination, R&D and Educational Research) (in Indonesia). Bandung: Alfabeta, 2019.
[10] M. A. Almaiah, A. Al-Khasawneh, and A. Althunibat, “Exploring the critical challenges and factors influencing the E-learning system usage during COVID-19 pandemic,” Educ. Inf. Technol., vol. 25, pp. 5261-5280, 2020, doi: 10.1007/s10639-020-10219-y.
[11] Association for Education Communication and Technology, Meaning of educational technology. Washington DC: AECT definition and terminology document, 2004.
[12] M. E. Mahmud, Educational Technology: Basic Concepts and Applications (in Indonesia). Samarinda: Mulawarman University Press, 2020.

# A REVIEW OF MACHINE LEARNING AND ITS APPLICATIONS

**A. Lavanya**
Assistant Professor, Department of Computer Science
KG College of Arts And Science, Coimbatore, Tamilnadu, India

**Abstract: Application Machine learning has been quite helpful in finding things that can lead to crucial decisions as data has become more and more abundant over the past few decades. Artificial intelligence is a field that enables machines to learn from experience and examples, much like people do, and discover intriguing patterns without being taught. The algorithm is provided data, and it creates a model using that data. It can forecast new values using this model. Finding something new and unknown can lead to the discovery of a wide range of new possibilities. Health, finance, retail, travel, media, image processing and computer vision, natural language processing, automated trading, automotive, aerospace, manufacturing, and many other industries can all benefit from machine learning. In this study, the fundamentals of machine learning, its methods, and its applications in diverse industries are reviewed.**

Keyword: Machine, Vision, Finance

## 1.INTRODUCTION

The area of computer science known as machine learning enables computers to learn without explicit programming. In accordance with [1] "A computer programme is said to learn from experience E with some class of tasks T and performance measure P if its performance at tasks in T, as measured by P, improves with experience E." — Tom Mitchell. Instead of using programming to carry out tasks, machine learning involves the creation of algorithms and procedures. [2] Machines learn from prior instances and historical trends, and based on their prior knowledge, a model that can be used to forecast future values can be developed.

Machine learning can be used to help find answers to problems based on the study of data when the data and questions are too big to be resolved organically. It can facilitate quicker searches for crucial items. Because machines learn more quickly and can even outperform humans in some fields, complex issues are easily handled. As a result, its demand is always increasing.

Machine learning is becoming more important as big data and cloud computing do since it uses computational power to solve various problems. It has numerous applications. It can aid in the development of new drugs and enable clinicians to make accurate diagnoses that enable early illness prognosis.

Customers are targeted on social media, where they are divided into groups based on factors like location, gender, and age to analyse their shopping patterns. It is simple to spot online fraud. It is highly helpful in areas such as automated trading, natural language processing, automotive, aerospace, and face and speech recognition.

## 2. TYPES OF LEARNING

As described below, there are three different types of learning:

Supervised learning is a sort of learning where input and output variables are present. A function can be derived using an algorithm from input to output. When we have information on the outcome we need to forecast, we use it. Data is separated into training and test data. It examines the training data to generate an inferred function that can be used to map test data for classification or prediction. Let's say we need to determine whether a new fruit will be an apple, grapes, or banana from a basket full of fruits like bananas, apples, and grapes. Simply by analysing the characteristics of existing fruits, we may predict the characteristics of the new fruit. As another illustration, imagine that there are two categories of emails spammed and nonspammed and that we must categorise new emails accordingly. We can categorise incoming emails based on instances of older, already-classified emails. Classification and regression issues are two categories of supervised tasks. Regression predicts a true value, whereas classification predicts a label. Regression is used for replies that are a real number, such as the miles per gallon of a specific car, while classification method is used for nominal, not ordinal, response values. [3,4]

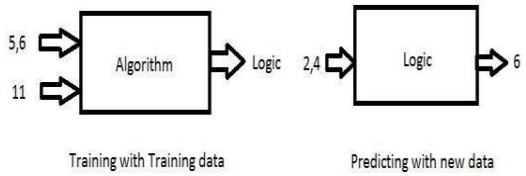



**Fig.1 Example of Supervised Learning [5]**

Here, the three most popular supervised learning algorithms are discussed.

**i) Naive Bayes**

It is a classification method. It is founded on conditional probability's Bayes theorem. The idea behind how it operates is that each feature is independent of the others. This classifier is straightforward, effective, and performs well despite this drawback. We are able to categorise the fresh samples using the training data and its attributes. For instance, it is considered that characteristics such as hair colour, height, eye colour, and other attributes can be used to categorise a person as male or female. Eq. 1.1 displays the Bayes theorem.

$$P(X/Y)=(P(Y/X)*P(X))/P(Y) \quad 1.1$$

P(X/Y) = Posterior Probability
P(Y/X) = Conditional Probability
P(X) = Prior Probability
P(Y) = Probability of Y which is fixed

**ii) Decision Tree:**

Decision tree classifiers are successfully applied in a wide range of fields, including speech recognition, character identification, remote sensing, medical diagnostics, and expert systems, to mention a few. The divide and conquer strategy is used. [6] There is a root node, branches that are internal nodes, and leaf nodes in this tree-like arrangement. The property that reflects the most information is referred to as the root node. A leaf node indicates the class label, and branches indicate the outcome of the test performed on the inside nodes. Regression trees and classification trees are two different types of trees where the class label is given real values and discrete values, respectively. There are various decision tree algorithms, including CART, C4.5, ID3, etc. The techniques for identifying the feature that separates the training data most effectively include information gain and gini index.

Decision trees have the following benefits and drawbacks:

**Advantages**

- It employs a white box model, handles category and numerical data
- It requires no prior data preparation,
- It is incredibly simple to grasp.

**Disadvantages**

- Complexity: Large trees with lots of data have a tendency to become complex.
- Instability: Redrawing the entire tree may result from changing data or variables.
- Cost: Increasing complexity may result in higher costs.

**iii)Support Vector Machine**

It is a method of supervised learning. Both classification and regression use it. It considers input factors and provides a hyperplane as an output to categorise fresh samples. For instance, we must divide the squares and circles in the figure below. SVM merely drew a line.

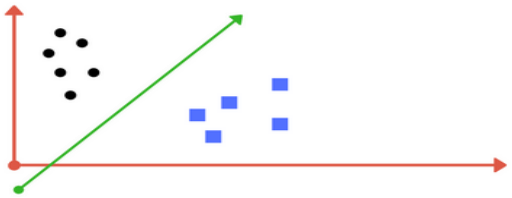


**Fig 2: Shows a hyperplane to separatecircle and squares**

SVM has the following advantages and drawbacks:

**Advantages**

- Even with a high number of properties, it still functions well.
- As SVM is kernel based, it can model complicated, real-world issues and scales rather well to high dimensional data.
- SVM has a lower overfitting risk.

**Disadvantages:**

- Picking a useful kernel function is a difficult process.
- Training takes a lengthy time when the dataset is large.[7]

**B. Unsupervised learning**

Unsupervised learning is a sort of learning where the only variables are the inputs. Algorithms are used to discover the intriguing patterns from unlabeled data. They pick up the features on their own, and as new data comes in, they organise it according to the features they've picked up from the old data. As an illustration, we may state that we only have questions and no appropriate solutions. For clustering and association issues, it is mostly used. [8].

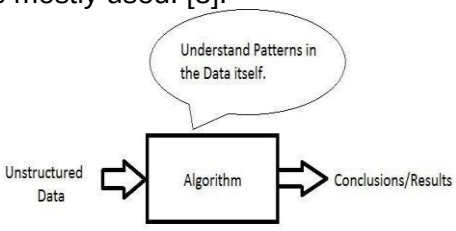



**Fig. 3 Unsupervised Learning Model**

The following algorithms are used for unsupervised learning:

**i)Clustering:**

This process entails breaking up the data into various subgroups. Data items are grouped into one or more subgroups known as clusters when they share the same characteristics or are comparable to one another. For instance, a store owner can segment the clientele based on how they often make purchases. K Means Clustering and Hierarchical Clustering are two different methods of clustering. Market segmentation, social network research, medical imaging, and other fields can all benefit from clustering.

**ii)Association:**

These methods are used to discover connections among the database's data pieces. They are helpful in market behaviour analysis, where we can forecast which products will be purchased in pairs by identifying their relationships. Antecedent and consequent are the two halves of an association rule, where antecedent is a data item from the transaction and consequent is a data item discovered in conjunction with antecedent. Rules are made up of support and confidence components. Support refers to how frequently things occur, and confidence refers to how frequently if/then assertions are verified as true. For instance, a customer is more likely to purchase butter if he also purchases bread.

**B.Reinforcement Learning:**

In this kind of learning, a computer is placed in a setting where it continuously honed its skills via trial and error.[10] To maximise performance and reduce risk, every conceivable state is frequently identified. The agent must receive simple reward feedback in order to learn its actions.

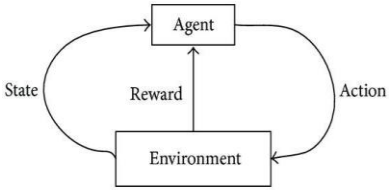



**Fig.4 Model of Reinforcement Learning[11]**

## 3. Applications

Machine learning techniques find use in wide number of applications like health, finance, social media, travel, email spam and malware filtering, online customer support, search engine result refining, product recommendations and lot more.

Some of which are explained in detail below:

**i)Health:**

- ✓ Today lots of patients are suffering from various diseases which prove to be fatal. Doctors can look for symptoms and can even predict how long they will live.
- ✓ New drugs can be discovered and built easily.

**ii) Finance:**

- ✓ Online fraud is rising steadily. Algorithms that use machine learning can quickly identify these scams. PayPal, for instance, can quickly determine whether a transaction is legitimate or not.
- ✓ Ask for advice on the investment that will yield the highest returns.
- ✓ Character recognition is one application for it in banking.[13]

**(iii)Social media**

- ✓ Face recognition: Facebook recognises traits that are similar and recommends tags with those names. Facebook suggests a good name before we choose our friend's name for the tag.
- ✓ People whose profiles we browse or search for are suggested as Facebook friends.
- ✓ Targeted marketing: Based on customer analysis and an understanding of their purchasing patterns, clusters are developed.

**iv) Travel:**

- ✓ GPS can help us get to our location and even foresee traffic.
- ✓ UBER is able to calculate the cost of our ride.
- ✓ When we reserve airfare for our trip, we receive hotel recommendations.
- ✓ TripAdvisor displays reviews and commentary from numerous users, which can aid individuals in finding their own preferences.

**v) Filtering for viruses and spam in email:**

- ✓ Similar patterns are examined over numerous lines of code, and malware that could otherwise jeopardise data security is found.
- ✓ Regular updates are made to spam filters.

**vii) Refining of Search Engine Results:**

- ✓ When we use a search engine like Google to find results, the backend maintains note of whether we looked at the first page and got the results or the second or third page, etc.

**vii) Online support for customers:**

- ✓ If a live expert is not available, chatbots are available, which give us with answers by searching for answers from the websites.
- ✓ Some websites offer an online engagement system to handle their clients' problems right away.

**viii) Product Suggestions**

- ✓ When we conduct a website search for a product, we are continually given recommendations for products of a similar type, with the same price, from various brands, etc.

**xi)Processing of Natural Language x**

- ✓ It aids in text classification, allowing us to assign a class label to a specific passage of text.
- ✓ Speech Recognition: Voice can be identified from an audio signal or voice clip.
- ✓ Because there are photos, we can anticipate what they will be captioned.

**x) Automative Sector**

- ✓ Drivers are even helped with speed limits and given access to an automatic braking system;
- ✓ The mood of the driver may be predicted;
- ✓ Many traffic incidents can be avoided owing to driver drowsiness; etc.[14,15]

## 4.Conclusion

The fundamentals of machine learning, its three subtypes—supervised, unsupervised, and reinforcement learning—as well as the most popular supervised and unsupervised learning algorithms—have been covered in this review paper. Also covered in detail here are the applications of machine learning in a number of industries, including automated trading, the automotive, aerospace, and manufacturing industries, as well as the domains of health, finance, retail, travel, and media.

## References

[1] S. Diksha, K. Neeraj. A Review on Machine Learning Algorithms, Tasks and Applications, International Journal of Advanced Research in Computer Engineering & Technology (IJARCET) 6(2017) 1548-1551.

[2] H. Sandhu, Tejenderkaur MACHINE LEARNING AND NATURAL LANGUAGE PROCESSING – A REVIEW, International Journal of Advanced Research in Computer Science 9(2018) 582-584.

[3] Anish. T, Yogesh. K. Machine Learning: An artificial intelligence methodology, International Journal Of EngineeringAnd Computer Science, 2(2013) 3400-3404

[4] I.Muhammad, Z.Yan. SUPERVISED MACHINE LEARNING APPROACHES A SURVEY, School of Information Sciences and Technology, Southwest Jiaotong University,China, 942-951

[5] Different types of machine learning [Online]https://medium.com/deep-math-machine-learning-ai/different-types-of- machine-learning-and-their-types- 34760b9128a2

[6] S, Umadevi, K.S., Jeen Marseline A Survey on Data Mining Classification, Algorithms International Conference on Signal Processing and Communication (ICSPC'17),(2017) 264-268

[7] https://statinfer.com/204-6-8-svm-advantages-disadvantages-applications, lastaccessed on 10 Jan

[8] A. Dey A Brief Review of Machine Learning and its Application, International Journal of Computer Science and Information Technologies, 7(2017) 1174-1179

[9] C.C, Aggarwal, S.Y, Philip. Data mining techniques for associations, clustering and classification, Proceedings of the Third Pacific-Asia Conference on Methodologies for knowledge Discovery and Data Mining, (1999) 13–23.

[10] Shalev-Shwartz, S.Ben-David Understanding Machine Learning: From Theory to Algorithms, Cambridge University Press (2014)

[11] https://towardsdatascience.com/types-of-machine-learning-algorithms-you-should- know-953a08248861, last accessed on 14Jan

[12] https://medium.com/app-affairs/9 applications-of-machine-learning-from-day-to- day-life-112a47a429d0, last accessed on 14Jan

[13] Bharti, A., Srivastava, S., Singh, J.,Pandey, S., Alankar, 2018. Machine Learning: Handwritten Character Recognition, International Journal of Innovations & Advancement in Computer Science vol 7, pp. 320-324

[14] https://en.wikipedia.org/wiki/Outline_of_machine_learning, last accessed on 15 Jan

[15] M, Xue, C,Zhu. A Study and Application on Machine Learning of Artificial Intellligence, International Joint Conference on Artificial Intelligence 2009) 272-274.

INTERNATIONAL JOURNAL OF INNOVATIVE TECHNOLOGY AND CREATIVE ENGINEERING (ISSN:2045-8711)
VOL.13 NO.8 AUG 2023

# A SURVEY ON APPLICATIONS OF HEALTHCARE BY USING AI TECHNIQUES

**Dr Vanaja K[1]**
Assistant Professor, Department of IT and AI&ML,
Adithya College of Arts and Science, Coimbatore, Tamil Nadu, India
**Alin Glency E[2]**
Student, Department of IT and AI&ML,
Adithya College of Arts and Science, Coimbatore, Tamil Nadu, India
**Keerthana S[3]**
Student, Department of IT and AI&ML,
Adithya College of Arts and Science, Coimbatore, Tamil Nadu, India

**Abstract-The Internet of Things (IoT) and artificial intelligence (AI) are the driving forces behind the digital transformation of modern healthcare. Security issues must be taken into account from the beginning of any digital transformation's design. Due to the sensitivity of healthcare data, any breach jeopardizes patient privacy. Even more so with IoT networks where the attached devices are open to intrusion. Cyber attacks have the potential to have fatal repercussions. This paper gives a thorough overview and analysis of current research on using AI as a cybersecurity tool to secure IoT networks used in healthcare settings. It thoroughly reviews relevant works and finds a specific topic within the field that offers scholars an ideal learning opportunity.**

Keywords—Internet of Things, Artificial Intelligence, Security, Privacy, Healthcare

## 1. INTRODUCTION

Technologies are being developed to simplify our lives, and they are evolving quickly everyday. This technical innovation has been extremely beneficial to the healthcare business. It facilitates the completion of any protracted and laborious task and enablesmedical professionals like doctors to safely use this incredible technology. Before technology was brought to the healthcare sector, staff members had to do duties manually and patients had to go through lengthy examination wait times. The only ways patients communicated with the doctors were through hospital visits, phone calls, andtexts.

The patient's condition could not be continually monitored in order to provide an early and precise diagnosis. AI and IoT in healthcare have the potential to significantly alter medical research, illness diagnosis, andpatient treatment.

IoT is a network of real-world objects that can communicate and share data. Internet of Medical Things (IoMT) refers to the exchange of data via linked IoT devices in healthcare. In order to handle complicated medical data, do analysis, reasoning, pattern identification, carry out specified tasks, and solve issues without any direct human involvement or input, AI in healthcare employs various software and algorithms to approximate human intelligence.

AI is a group of several technologies. Wearable's, biosensors or body sensors, and medical equipment are all examples of IoT devices that are employed in the healthcare industry. The medical information gathered may take the form of EHR, claims, patient registries, health surveys, or information from clinical trials. AI is the most efficient method for processing all huge data and doing accurate real-time analysis.

Including lowering doctor-patient visits, tracking chronic illness patients in real-time, helping patients recover, and providing healthcare to remote places.

With every new technology, security is a concern, and this is no exception. Medical data breaches involving IoT devices are a matter of life and death.

A patient may become powerless if their pacemaker, life support system, or oxygen delivery system is connected to the internet and is therefore open to many types of assaults. The security and privacy of the gadget, the patient, and the hospital may all be at danger when using connected IoT devices in healthcare.

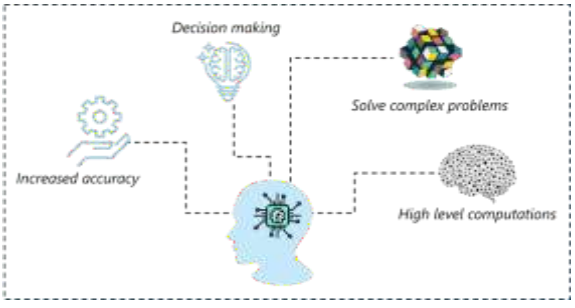
## 2. AI SYSTEMS RELEVANT TO HEALTHCARE

In fact, many different technologies make up artificial intelligence. Although the majority of these technologies are immediately applicable to the healthcare industry, the precise procedures and jobs they assist differ greatly. Below are definitions and descriptions of some specific AI technologies that are very significant to healthcare.

### 2.1 Machine learning – Neural Networks and Deep Learning

Precision medicine, which determines which treatment procedures are likely to be effective on a patient based on a variety of patient traits and the context of the therapy is the most popular use of classical machine learning in the healthcare industry. The vast majority of machine learning and precision medicine applications ask for supervised learning, which calls for a training dataset for which the end variable (such the beginning of illness) is known.

The neural network is a more advanced type of machine learning. This technology, which has been around since the 1960s and has been widely used in medical research for several decades, is used for categorization applications like predicting whether a patient will contract a specific disease. It approaches issues in terms of variables' weights, or "features," that link inputs and outcomes. It has been compared to how neurons interpret signals, however the comparison to how the brain works is not very strong.




Deep learning, or neural network models with many levels of features or variables that predict outcomes, is one of the most difficult types of machine learning. These models could contain thousands of hidden elements that today's graphics processing units and cloud systems can more quickly process. Deep learning is frequently used in the healthcare industry to identify possibly malignant tumours in radiological pictures.

Deep learning is increasingly being applied to radiomics, or the human eye can perceive the detection of clinically relevant features in imaging data beyond what. The most prevalent applications of deep learning and radiomics in picture analysis are in cancer. Compared to the previous generation of automated techniques for image analysis, known as computer-aided detection or CAD, their combination seems to offer improved accuracy in diagnosis.

Deep learning is a type of natural language processing (NLP), which is covered in more detail below, and is increasingly utilised for voice recognition. Each feature in a deep learning model often has minimal significance to a human observer, in contrast to older types of statistical analysis. As a result, it could be exceedingly challenging or perhaps impossible to comprehend the explanation of the model's results.

### 2.2 Rule-based expert systems

In the 1980s and succeeding decades, expert systems built on databases of "if-then" rules dominated the field of artificial intelligence. Over the last couple of decades, they were used extensively in the healthcare industry for "clinical decision support" purposes, and they are being used extensively today. Today, a lot of suppliers of electronic health records (EHRs) provide a set of guidelines with their systems.

A set of rules in a certain knowledge domain must be built by human experts and knowledge engineers for expert systems. They are simple to grasp and operate well up to a point. However, they frequently fail when there are a lot of rules (typically above a few thousand) and when the rules start to clash with one another. Additionally, it might be challenging and time-consuming to update the rules if the knowledge domain does. More techniques based on data and machine learning algorithms are gradually replacing them in the healthcare industry.

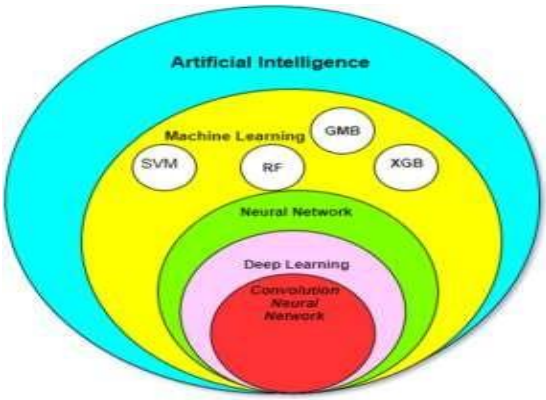
### 2.3 Natural Language Processing

Since the 1950s, AI researchers have been trying to understand how human language works. NLP is a discipline that comprises language-related applications including speech recognition, text analysis, translation, and others. The two main methods are statistical NLP and semantic NLP. StatisticalNLP is based on machine learning, specifically deep learning neural networks, and has helped to improve recognition accuracy recently. To learn it, you need a sizable "corpus" or body of language.

The generation, comprehension, and categorization of clinical documentation and published research are the primary uses of NLP in the field of healthcare. NLP systems are able to perform conversational AI, create reports (for example, on radiological tests), evaluate unstructured clinical notes on patients, and record patient interactions.

### 2.4 Robotic Process Automation

This technology executes organized digital administration tasks—those involving information systems—as if they were being carried out by a human user who was following a set of instructions or guidelines. They are less costly, simpler to programme, and more transparent than other types of AI. Robotic process automation (RPA) mostly use server-based software rather than actual robots. To behave as a semi-intelligent user of the information systems, it depends on a mix of workflow, business rules, and 'presentation layer' integration. They are employed in the healthcare industry for routine duties including billing, prior authorization, and patient record updates. They may be used to extract data from, for instance, faxed images in order to feed it into transactional systems when paired with other technologies like image recognition.




Although we have only discussed these technologies individually, they are progressively being merged and integrated. For example, robots are acquiring AI-based "brains," and RPA and image recognition arebeing coupled. Perhaps these technologieswill become so intertwined in the future that composite solutions will become more possible or practical.

## 3. RESULTS & DISCUSSION

The significance of artificial intelligence in the healthcare industry was discussed in this article. We looked into administrative, patient monitoring, and medical diagnosisapplications in the healthcare industry. We also examined and evaluated the most prevalent illnesses for which AI has been applied. To handle unstructured data (images, EP data), a good AI system must have both a machine learning (ML) and a natural language processing (NLP) component. The system's sophisticated algorithms must first be trained using healthcare data before they can help professionals with sickness diagnosis and treatment. There is a wealth of medical research that demonstrates the benefits of AI. It is well known that AI uses algorithms to find traits in a lot of medical data.

The subsequent analysis of this data yields crucial insights that support clinical training. Decisions are subsequently made using the data that was collected from running the algorithms. The absence of prejudice in AI is one way in which it differs from human behavior. This indicates that the choice was made with the patient's best interests or clinical actions in mind.

AI is also capable of self-correction, which helps it become better at forecasting and being accurate. This is crucial for finding biomarkers in an individual's molecular structure. A potential for a wrong diagnosis would exist if the system lacked self-correcting capabilities. The approach has been shown to be helpful in preserving information for the healthcare sector. It has achieved this by giving current details on patients' medical backgrounds. This is made feasible by automation; every time a patient enters a healthcare institution; their information is documented and kept, making it simple for doctors to trace the data.

### Future of AI in Healthcare

Securing AI's acceptability in routine clinical practice is the most difficult task for the industry. These machine learning and artificial intelligence algorithms have been useful in several crucial medical applications. The market for AI systems is anticipated to reach $6 billion by 2021, according to Frost & Sullivan.

## 4. CONCLUSION

In terms of healthcare, artificial intelligence has shown to be incredibly helpful to the American population. Although first deployments are expensive, it is very conceivable to see the advantages outweigh the costs in the long run and become affordable for everyone. Therefore, we may conclude that the jobs listed below are those for which AI is most frequently used in healthcare: 1) Assistance with diagnosis; 2) management of healthcare organizations. The main obstacles to using AI in healthcare, however, are as follows: 1) the demand for certain business architecture; 2) the general attitude towards AI; 3) the demand for information security and privacy; and 4) the demand for high-quality data. Services that are very reliable and of great quality.

## REFERENCES

1. Jack Jr, C. R., Knopman, D. S., Jagust, W. J., Petersen, R. C., Weiner, M. W., Aisen, P. S & Lesnick, T.G. (2013). Tracking pathophysiological processes in Alzheimer's disease: an updated hypothetical model of dynamic biomarkers. The Lancet Neurology, 12(2), 207-216.
2. Kubat, M. (2015). Artificial neural networks. In An Introduction to Machine Learning (pp. 91-111). Springer, Cham.
3. Efron B, Tibshirani RJ. An Introduction to the Bootstrap. New York: Chapman and Hall;1994.https://doi.org/10.1201/9780429246593
4. Oresko, J. J., Jin, Z., Cheng, J., Huang, S., Sun, Y., Duschl, H., & Cheng, A. C. (2010). A wearable smartphone-based platform for real-time cardiovascular disease detection via electrocardiogram processing. IEEE Transactions on Information Technology in Biomedicine, 14(3), 734-740.
5. Zang, Y., Zhang, F., Di, C. A., & Zhu, D. (2015). Advances of flexible pressure sensors toward artificial intelligence and health care applications. Materials Horizons, 2(2), 140- 156.
6. Tripathy, A. K., Carvalho, R., Pawaskar, K., Yadav, S., & Yadav, V. (2015, February).Mobile based healthcare management using artificial intelligence. In 2015 International Conference on Technologies for SustainableDevelopment (ICTSD) (pp. 1-6). IEEE.
7. James G, Witten D, Hastie T, et al. An introduction to Statistical Learning with applications in R. First Edition: Springer,2013.
8. Agah, A. (2013). Introduction to medical applications of artificial intelligence. Medical Applications of Artificial Intelligence, 19-26.
9. Dilsizian, S. E., & Siegel, E. L. (2014). Artificial intelligence in medicine and cardiac imaging:harnessing big data and advanced computing to provide personalized medical diagnosis and treatment. Current cardiology reports,16(1), 441.
10. Kreps, G. L., & Neuhauser, L. (2013). Artificial intelligence and immediacy: designing health communication to personally engage consumers and providers. Patient education and counseling, 92(2), 205-210.
11. Luxton, D. D. (2014). Artificial intelligence in psychological practice: Current and future applications and implications. Professional Psychology: Research and Practice, 45(5), 332.
12. Noël, L., Appel, H., Huang, B., & Hefley, W. (2013). Overall Health and Health Care Utilization Among Latino American Men in the United States. American Journal of Men'sHealth,7(1), 6–17. https://doi.org/10.1177/1557988312452752
13. Riccardi, G. (2014, November). Towards healthcare personal agents. In Proceedings ofthe 2014 Workshop on Roadmapping the Future of Multimodal Interaction Research including Business Opportunities and Challenges (pp. 53-56).
14. Salih, A., & Abraham, A. (2013). A review of ambient intelligence assisted healthcare monitoring. International Journal of Computer Information Systems and Industrial Management (IJCISIM), 5, 741-750.
15. Riccardi, G. (2014, November). Towards healthcare personal agents. In Proceedings ofthe 2014 Workshop on Roadmapping the Future of Multimodal Interaction Research including Business Opportunities and Challenges (pp. 53-56).

# SMART ID CARD FOR EMPLOYED WOMEN SAFETY BASED ON IOT SYSTEM

**C. Saranya**
Assistant Professor, Department of Computer Science
Kangeyam Institute of Commerce, EBET Knowledge Park,
Nathakadaiyur, Kangeyam, Tamil Nadu, India
csaranchandru.ya@gmail.com

Abstract: **At Work premises and in public venues women must be protected from assault in many parts of the world. A smart ID card could make female employers safer and be one solution to this problem. Because of the built-in sensors and connectivity components, this ID card will be able to sense and respond to its surroundings based on the Internet of Things (IoT). If the sensors detect abnormal activities, such as a sudden movement or a loud noise, heart rate rise, GPS location will send a signal to the system. Following the completion of the data analysis, the system takes any necessary actions, such as sending alerts to security personnel or notifying on-call personnel. The proposed system will provide the user with access to an emergency panic button. The system's GPS component will track the user's location and send alerts to the appropriate parties. There will also be a cloud-based monitoring system with the proposed approach. This allows the employer or loved ones of the worker to track the worker's whereabouts and activities in real time. The system will analyze the data using machine learning techniques, and the results will reveal the worker's behavior and level of safety. A smart ID card linked tothe Internet of Things (IoT) could help women work in safer environments. The proposed technology would enable real-time monitoring, predictive alarms, and data analysis. The precise data greatly help the relatives and cops to recover them from unpleasant situations.**

Keywords— IoT, Smart ID Card, Woman Safety, ESP32 Controller, IoT Interface, GPS, BMS

## I. INTRODUCTION

People have been concerned about the well-being of working women since the beginning of time. Even though there are laws and policies in place to protect women, sexual harassment and assault still happen. So, it is very important to come up with new ways to keep working women safe and give them confidence.

In this context, a "smart ID card" is a type of ID card that uses the Internet of Things (IoT) to transmit data to a central database. It is a one-of- a-kind ID card with numerous sensors and other technological components. Women today are aware of men in all aspects of their lives, but this awareness comes at a cost: they must deal with antagonism, aggression, and sexual assault in both public and private spaces. They can't just get up, put on whatever they want, and walk out the door to work together.

Women must deal with a stigma that undermines their beliefs, goals, and sense of what is possible. Because of these factors, it is clear that women's safety in this country is precarious. This study looks at the effectiveness of a special insurance system designed to ensure that women never feel helpless in such social situations. It is possible to build a strong system that can assist women in danger. In this experiment, the Raspberry Pi is combined with a temperature sensor, GSM, GPS, and camera module. We outline three methods for contacting the appropriate people in our post. Using voice, temperature, and heart-rate data that we uploaded to the Raspberry Pi, we created a small device that a lady can activate by pressing a button. This was done to ensure women's safety [1].

The Internet of Things (IoT) is a global network of computers, mobile devices, and other sensors that can communicate with one another via the internet. Each device is given a UID, or unique identifier, so that it can communicate and share data with other devices. The importance of computerizing these locations has grown as a result of numerous business structures and social hierarchies. The application sends an SMS with the user's location to a list of contacts using the Global Positioning System (GPS).

On a separate website, a photograph of the event has been posted. The software will also contact someone on its contact list. It can also be used in situations such as chain stealing, accidents, personal problems, and molesters. Anyone who witnessed the incident is also eligible to apply. Users must provide personal information such as an email address and phone numbers for emergency contacts. Individuals can create their own contact lists, which can include family members and close friends. The smart ID card was created to make businesses, schools, and other locations safer for customers, employees, and students. This is made possible through the use of the card.

The smart ID card could have sensors like a global positioning system (GPS), an accelerometer, and temperature and humidity sensors that record in real time where the cardholder is, what he or she is doing, and what the weather is like. These sensors' collected data could be transmitted to a central database using IoT technology. As a result, the system can now monitor and analyze data in real time. Because the smart ID is adaptable, it can be customized to meet the needs of a variety of businesses and institutions. A smart ID card, for example, can be used at work to track not only an employee's presence but also the time they spend on the task. This ensures that workers are properly compensated for their time and effort.

Schools can use the smart ID card to locate their students and protect them from potential risks on campus.

## II. LITERATURE SURVEY

A smart ID card can track a person's mental and physical health. Medical personnel can save lives by keeping track of a patient's vital signs with a smart ID card. This is extremely useful in times of crisis. CSR, or continuous automatic speech recognition, can be used for a variety of tasks, including command and control, dictation, transcription of recorded speech, searching audio files, and interactive spoken conversations. A large number of statistical models that accurately describe the vast array of phonemes in the target language form the foundation of any voice recognition system.

The hidden Markov model (HMM) is an excellent starting point for speech detection systems due to its temporal structure and representation as a sequence of spectral vectors that encompass the audible frequency range.

This is because words and phrases in human speech can be broken down into smaller components. Because HMM is based on complex mathematics, it can serve as a theoretical foundation for a wide range of disciplines. When used correctly, the HMM model performs admirably for a wide range of critical tasks. Almost every operation has a measurable output or signal that others can observe. You can use both one-time signals and continuous signals. Both of these are possible. Signal sources can be stationary, meaning their statistical properties do not change over time, or non-stationary, meaning they do (i.e., the signal properties vary over time). The signals could have been tampered with during transmission or by other means. Transmitted voice signals are used to create word sequences in the speech recognition process. Many technologies, including dynamic programming and neural networks, have been used to develop an effective and precise method for voice recognition. In this post, we'll look at the connection between speech recognition and HMM as well as how HMM can be used to solve common voice recognition problems. The major objective of this essay is to provide a strategy for women in India to keep themselves secure. This device is intended to make it easy and quick for abused women to seek help. When the owner of this item feels threatened, they will press the "panic" button. Individuals who have agreed to receive cell phone alerts will receive a text message with the danger's latitude and longitude. After locating the woman using the coordinates provided, the appropriate assistance can be dispatched to her. The concept's creators hoped it would reduce the frequency of violent crimes against women in India [2].

| Ref | Findings |
|---|---|
| [3] | The purpose of this study is to look into the potential benefits of putting Internet of Things (IoT)-based wearable gadgets on women. To assist working women in perilous situations, the authors propose an Internet of Things (IoT)-connected wearable smart device with a panic button, GPS, and Wi-Fi. |

| | |
|---|---|
| [4] | According to this study, a smart card system with an integrated microprocessor, GPS, and GSM module could help women stay safe. In the event of an emergency, it allows working women to notify their parents or the police of their location. |
| [5] | Based on the Internet of Things, this study suggests a smart safety device for women. The authors create a wearable device with a microprocessor, Wi-Fi module, and GPS that allows women to send vital signals to their guardians or authorities. |
| [6] | This article discusses how to ensure women's safety by combining machine learning and the Internet of Things. Using IoT concepts, the authors create a wearable device with GPS and GSM modules to alert authorities or carers when a person has fallen. |
| [7] | This project makes use of cloud computing and the Internet of Things (IoT) to develop a method for improving female commuter safety. The authors design a wearable device with a microprocessor, GPS, GSM, and Wi-Fi module to alert authorities or guardians of a woman's location in an emergency. |
| [8] | According to the results of this survey, women should use an IoT-based smart security system. In an emergency, the authors create a wearable device that uses a microcontroller, GPS, and Wi-Fi to allow women to contact their guardians or the appropriate individuals. This technology can be used to track women's movements and locations. |
| [9] | The findings suggest that the Internet of Things (IoT) be used to implement a smart public safety and security solution for women. The authors create a wearable device for women that uses a microcontroller, GPS, and GSM to send emergency notifications to their guardians or the appropriate authorities. This technology can be used to track women's movements and locations. |
| [10] | This study looks at how IoT-compatible equipment could be used to make cities safer for women. The authors hope to help women in distress by creating a high-tech wearable device equipped with a panic button, GPS, and Wi-Fi. The authors recommend machine learning for improved performance. |
| [11] | This article discusses a smart safety device based on IoT for working women. The authors create a wearable device for women that uses a microcontroller, GPS, and GSM to send emergency notificationsto their guardians or the appropriate authorities. This technology can be used to track women's movements and locations. |
| [12] | The study's authors believe that an Internet of Things (IoT) smart wearable could be used to monitor women's health. The authors design a wearable device with a microcontroller, various sensors, and Wi- Fi to send emergency signals to the women's guardians or the appropriate authorities. The device can also monitor other women's health parameters. |
| [13] | The authors of this paper propose using big data and the internet of things to increase women's safety. In an emergency, the authors create a wearable device that uses a microcontroller, GPS, and Wi-Fi to allow women to contact their guardians or the appropriate individuals. The technology could be used to track women's movements and activities. This information could then be used to generate future forecasts. |
| [14] | The Internet of Things and wearable technologies are combined in this study to create a method for ensuring women's safety. The authors develop a wearable device that uses a microprocessor, numerous sensors, and Wi-Fi to send distress signals to the women's guardians or the appropriate authorities. Other women's health parameters can also be monitored by the device. |
| [15] | To keep women safe, the author suggests using cloud computing, big data analytics, and the internet of things. In an emergency, the authors create a wearable device that uses a microcontroller, GPS, and Wi- Fi to allow women to contact their guardians or the appropriate individuals. This method can also be used to determine women's interests and destinations. |

**Table I. Literature Review Comparison**

## III. PROPOSED METHODOLOGY

The proposed system constructed around ESP 32 microcontroller enabled within build Wi-fi connectivity. The mobile phone or local Wi-Fi hot spot will ensure the connectivity of entire system to cloud. The camera, GPS, Press button will generate the data continuously and send them to cloud based monitoring system which analysis the data and make certain decision based on that alerts

messages, location coordinates are sent to care takers of women. The vibrator is employed to intimate the women about the status and action taken by monitoring system through generation of vibration in defined pattern. The system also records the last GPS location where the connectivity lost, which greatly useful in recovery operation.

The mobile application installed within care takers mobile provides snaps from camera deployed in smart ID card. The text messages also received from cloud monitoring system which includes heart rate, GPS location continuously. The ESP 32 microcontroller is the one of low power device employed in the proposed system that ensures the long battery life and system stand by duration.

## IV. BATTERY MANAGEMENT SYSTEM

The system is powered by Lithium-Ion (Li-ion) battery pack for which Battery Management System (BMS) is mandatory Therefore the advanced battery management algorithm is adopted for charging and discharging each Li-ion cell efficiently. The BMS also sent the battery charge level to ESP 32 that generates notification message about low battery and the same is intimated through care taker mobile application.

**Figure 1. Architecture of the Smart ID card for woman**

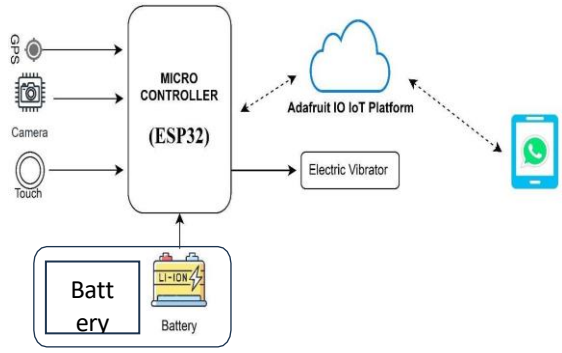
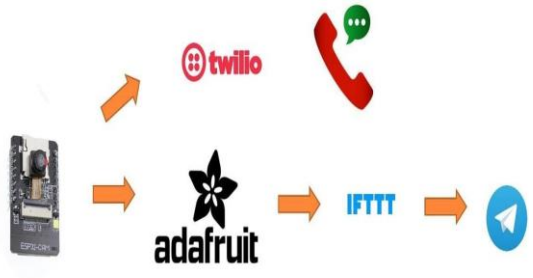



**Figure 2. IoT interface for the Smart Identity Card**




## V. SYSTEM WORKING

The proposed system sends the GPS location data to cloud monitoring system in periodic interval. The women who have wear the smart ID card suspects any abnormal situation and activities of neighbor person she can activate the press button to wake up the system to active mode. Now the camera is turned on to capture the person present in-front and the snap shot is send to cloud monitoring system. The cloud monitoring system generates alert message and sent the snap shot, alert message and GPS coordinates to care taker mobile application. Once the intimation is received the care taker shoot acknowledge the message with in certain duration of time (in seconds) by clicking acknowledge icon. Else the same content will be sent to police station that situated near to the location of victim presence. Suppose the care taker acknowledged the intimation then he or she is proved with option of informing the police for the same. The cloud monitoring system may generate the continuous update from smart ID card along with some suggestion messages based on situation such as location of victim, care taker and police station.

The cloud monitoring system also compare the attacker photo with data base to identify the person details, which plays vital role in securing the victim from the attacker. The GPS satellite connectivity may be poor at indoor locations therefore location data of victim is extracted from internet connectivity through Wi-Fi communication protocol. Therefore, the reliability proposed system is highly ensured to save the victims from the unpleasant situations.


**Figure 4. Model Prototype**

## VI. RESULTS AND DISCUSSIONS

An identity card attached with a camera and vibrator provides safety to women. The proposed system will take a photo of the attacker and send it to police and the woman's parents via WhatsApp. The vibrator will distract the attacker and help the woman escape the situation. A mobile app is developed to monitor the safety of women. Figure 1 shows the architecture of the Smart ID card for woman. The MQTT protocol is used to send the data to the cloud and from the cloud to the parent's mobile, phones via a WhatsApp group with a photo of the attacked and his current location. Figure 2 shows the IoT interface for the Smart Identity Card. The woman will activate the camera by touching the pin, and immediately the photo of the attacker and current location will be sent to the IoT cloud platform, and from there an alert call will be sent to the police and parents through the Twilio API, with the photo and location shared with the WhatsApp group and police personnel. Figure 3 shows the work flow of designed Smart Identity Card. Figure 4 show the prototype model. Woman ID card fitted with camera and processor send to cloud and from it will be hooked to the whatsapp with location. Figure 5 show the result of prototype where the whatsapp group members will receive the photo of attacker along with location and message. It will help the membersto act quickly and rush to spot

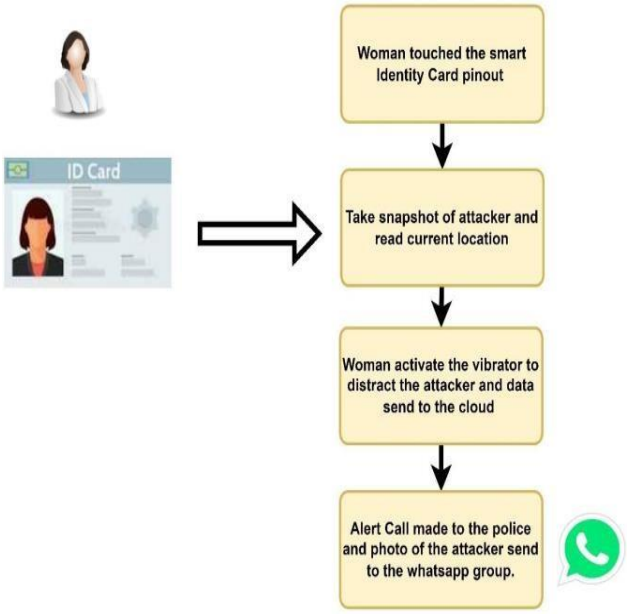



**Figure 3. Work Flow of designed Smart Identity Card**


**Figure 5. Result Snapshot**

Superior hardware is required for Internet of Things-based smart ID cards to function. Sensors, microcontrollers, wireless communication modules, batteries, and other components make up the hardware. Card hardware may improve and become more affordable in the near future. After the hardware is finished, the software for the microcontroller can be written. Over the network, the application must be able to communicate with the card's sensors and other components. It may be necessary in the future to develop more reliable card software. The user interface for smart ID cards should be simple enough for working women to use. The cardholder's current location and emergency contacts should be displayed in an interface element. Cards' interfaces may become more intuitive and user-friendly in the near future. The smart ID card should work with the current safety and rescue infrastructure. Women's employment may be protected in the future if systems withuser-friendly interfaces are developed. Before issuing smart ID cards, make sure they can be used, that they perform their intended functions, and that they are secure. More testing and validation may be required to ensure the card's functionality and dependability in the real world.

## VII. CONCLUSION

Providing Internet of Things-enabled smart ID cards to working women could be one way to improve their safety. This smart ID card contains a number of sensors and communication devices that enable women to send an SOS signal in an emergency. Authorities can track the woman's movements in real time and respond quickly to any potential problems by linking her smart ID card to a central monitoring system. This would not only keep women physically safe, but it would

also boost their self-esteem and enable them to succeed in male- dominated industries. It's critical to remember that the IoT-powered smart ID card will only work if it's installed and linked to an existing security system. Women should also learn how to use their smart ID cards in an emergency.

## REFERENCES


1. G. Lulla, A. Kumar, G. Pole, and G. Deshmukh, “IoT based Smart Security and Surveillance System,” 2021 International Conference on Emerging Smart Computing and Informatics(ESCI), Mar.2021, doi: https://doi.org/10.1109/esci50559.2021.939684 3.
2. A. Sarker, A. D. Gupta, and P. Das, “IOT Based Safe and Cost Efficient Smart Vehicle Management System,” IEEE Xplore, May 01, 2021.
3. G. C. Harikiran, K. Menasinkai, and S. Shirol, “Smart security solution for women based on Internet Of Things(IOT),” IEEE Xplore, Mar. 01,2016.
4. G. Kalyana Abenanth, K. Harish, V. Sachin, A. Rushyendra, and N. Mohankumar, “Enhancing the Security for Smart Card-Based Embedded Systems,” Computer Networks and Inventive Communication Technologies, pp. 673–686, Sep. 2021, doi: https://doi.org/10.1007/978-981-16-3728-5_50.
5. K. Venkatesh, S. Parthiban, P. S. Kumar, and C.N. S. Vinoth Kumar, “IoT based Unified approach for Women safety alert using GSM,” IEEE Xplore, Feb. 01, 2021.
6. H. Qiu, Q. Zheng, G. Memmi, J. Lu, M. Qiu, and B. Thuraisingham, “Deep Residual Learning based Enhanced JPEG Compression in the Internet of Things,” IEEE Transactions on Industrial Informatics, pp. 1–1, 2020, doi: https://doi.org/10.1109/tii.2020.2994743.
7. M. Acharya and P. S N, “Enhancing data security on a self-defensive Jacket Along with Microphone using IoT Technology and Cloud Computing,” IEEE Xplore, Oct. 01, 2022.
8. W. Akram, M. Jain, and C. S. Hemalatha, “Design of a Smart Safety Device for Women using IoT,” Procedia Computer Science, vol. 165, pp. 656– 662, 2019, doi: https://doi.org/10.1016/j.procs.2020.01.060.
9. D. Sunehra, V. S. Sreshta, V. Shashank, and B. U.Kumar Goud, “Raspberry Pi Based Smart Wearable Device for Women Safety using GPS and GSM Technology,” IEEE Xplore, Nov. 01, 2020.
10. K. Sharma, D. Anand, M. Sabharwal, P. K. Tiwari, O. Cheikhrouhou, and T. Frikha, “A Disaster Management Framework Using Internet of Things-Based Interconnected Devices,” Mathematical Problems in Engineering, vol. 2021, pp. 1–21, May 2021, doi: https://doi.org/10.1155/2021/9916440.
11. N. Prakash, E. Udayakumar, N. Kumareshan, and R. Gowrishankar, “GSM-Based Design and Implementation of Women Safety Device Using Internet of Things,” Intelligence in Big Data Technologies—Beyond the Hype, pp. 169–176. Jul. 2020, doi: https://doi.org/10.1007/978-981-15-5285- 4_16.
12. T. Yoldemir, “Internet Of Things and women’s health,” Climacteric, vol. 23, no. 5, pp. 423–425, Sep. 2020, doi: https://doi.org/10.1080/13697137.2020.1811563.2
13. G. Aceto, V. Persico, and A. Pescapé, “Industry4.0 and Health: Internet of Things, Big Data, and Cloud Computing for Healthcare 4.0,” Journal of Industrial Information Integration, vol. 18, p. 100129, Jun.2020, doi: https://doi.org/10.1016/j.jii.2020.100129.
14. X. Li, Y. Lu, X. Fu, and Y. Qi, “Building the Internet of Things platform for smart maternal healthcare services with wearable devices and cloud computing,” Future Generation Computer Systems, vol. 118, pp. 282–296, May 2021, doi: https://doi.org/10.1016/j.future.2021.01.016.
15. P. D. Singh, G. Dhiman, and R. Sharma, “Internet of Things for sustaining a smart and secure healthcare system,” Sustainable Computing: Informatics and Systems, vol. 33, p. 100622, Jan. 2022, doi: https://doi.org/10.1016/j.suscom.2021.1006 22.
16. C. S. Tang, “Innovative Technology and Operations for Alleviating Poverty through Women’s Economic Empowerment,” Production and Operations Management, Dec. 2020, doi: https://doi.org/10.1111/poms.13349.

# UNION, INTERSECTION AND CARTESIAN PRODUCT OF INTUITIONISTIC MULTI L – FUZZY SUBGROUPS WITH EXAMPLES

**C.GOWRISHANKAR**[1]
Research Scholar, Department of Mathematics,
SRMV College of Arts and Science, Coimbatore, Tamilnadu, India
**M.AKILESH**[2]
Department of Mathematics,
SRMV College of Arts and Science, Coimbatore, Tamilnadu, India *

**ABSTRACT**

**In this paper, we define the algebraic structures of Union, Intersection and Cartesian Product of an Intuitionistic Multi L – Fuzzy Subgroups and some related theorems with its examples are investigated.**

**Keywords:** Intuitionistic Multi L-Fuzzy Set (IMLFS), Intuitionistic Multi L–Fuzzy Subgroup (IMLFSG), Union, Intersection and Cartesian Product of Intuitionistic Multi L–Fuzzy Subgroups.

## INTRODUCTION

L.A.Zadeh introduced the notion of fuzzy sets in 1965. S.Sabu and T.V.Ramakrishnan proposed the theory of multi fuzzy sets. The idea of intuitionistic fuzzy set was given by K.T. Atanassov. After that in 2011, P.K.Sharma initiated the concept of intuitionistic fuzzy groups. The algebraic structures of intuitionistic multi fuzzy subgroup was introduced by R.Muthuraj and S.Balamurugan in 2014.

In this paper, we define Union, Intersection and Cartesian Product of intuitionistic multi L-fuzzy subgroups with suitable examples are discussed.

## PRELIMINARIES

In this section, we have to list some basic definitions and concepts of intuitionistic multi L-fuzzy subgroups.

2.1 Definition

Let X be a non-empty set. A fuzzy set $\mu$ of X is defined by $\mu: X \to [0,1]$.

### 2.2 Definition

Let $(G,.)$ be a group. A fuzzy subset $\mu$ of G is said to be L-fuzzy subgroup (LFSG) of G, if the following conditions are satisfied:

(i) $\mu(xy) \geq min\{\mu(x), \mu(y)\}, \ \forall\, x, y \in G.$

(ii) $\mu(x^{-1}) = \mu(x), \forall\, x \in G.$

### 2.3 Definition

Let X be a fixed non-empty set. A Multi L-fuzzy subset (MLFS) $\mu$ in X is defined as a set of ordered sequences. $\mu = \{(x, \mu_1(x), \mu_2(x), \ldots, \mu_i(x), \ldots.) : x \epsilon X\}$, where $\mu_i: X \to [0,1]$, for all $i$. Also note that, for all $i, \mu_i(x)$ is a decreasingly ordered sequence of elements.

ie., $\mu_1(x) \geq \mu_2(x) \geq \ldots\ldots \geq \mu_i(x) \geq \ldots\ldots$ , for all $x \epsilon X$.

**B**

Let X be a fixed non-empty set. An Intuitionistic L-fuzzy subset (ILFS) $\mu$ of X is an object of the form

$\mu = \{(x, \mu_A(x), \gamma_A(x)) : x \epsilon X\}$, where $\mu_A : X \to [0,1]$ and $\gamma_A : X \to [0,1]$.

Define the degree of membership and degree of non-membership of the element $x \epsilon X$ respectively with

$0 \leq \mu_A(x) + \gamma_A(x) \leq 1$, for all $x \epsilon X$.

REMARK

(a) When $\mu_A(x) + \gamma_A(x) = 1$, ie.,when $\gamma_A(x) = 1 - \mu_A(x) = \mu_A^c(x)$. Then $\mu$ is called as L-fuzzy set.

(b) We use an inscription $\mu = (\mu_A, \gamma_A)$ to denote the Intuitionistic L-fuzzy subset (ILFS) $\mu$ of X.

### 2.5 Definition

Let X be non-empty set. Let $\mu = \{(x, \mu_A(x), \gamma_A(x)) : x \epsilon X\}$ in X is defined as a set of ordered sequences.

ie., $\mu = \left\{\left(x, \left(\mu_{A_1}(x), \mu_{A_2}(x), \ldots, \mu_{A_i}(x), \ldots\right), \left(\gamma_{A_1}(x), \gamma_{A_2}(x), \ldots, \gamma_{A_i}(x), \ldots\right)\right) : x \epsilon X\right\}$.

Where $\mu_{A_i} : X \to [0,1]$, $\gamma_{A_i} : X \to [0,1]$ and $0 \leq \mu_{A_i}(x) + \gamma_{A_i}(x) \leq 1$ for all $i$.

Here, $\mu_1(x) \geq \mu_2(x) \geq \ldots \geq \mu_i(x) \geq \ldots$ , for all $x \epsilon X$ are decreasingly ordered sequence. Then the set $\mu$ is said to be an Intuitionistic Multi L-fuzzy subset (IMLFS) of X.

REMARK

Since we arrange the membership sequence in decreasing order, the corresponding non-membership sequence may not be in decreasing or increasing order.

### 2.6 Definition

The Intuitionistic Multi L-fuzzy subset $\mu = \{(x, \mu_A(x), \gamma_A(x)) : x \epsilon X\}$ of a group G is said to be Intuitionistic Multi L-fuzzy subgroup of G (IMLFSG) if it satisfies the following: For all $x, y \epsilon G$,

(i) $\mu_A(xy) \geq min\{\mu_A(x), \mu_A(y)\}$ and $\gamma_A(xy) \leq max\{\gamma_A(x), \gamma_A(y)\}$,

(ii) $\mu_A(x^{-1}) = \mu_A(x)$ and $\gamma_A(x^{-1}) = \gamma_A(x)$.

Or Equivalently$\mu$ is IMLFSG of G iff

$$\mu_A(xy^{-1}) \geq min\{\mu_A(x), \mu_A(y)\}$$

and $\gamma_A(xy^{-1}) \leq max\{\gamma_A(x), \gamma_A(y)\}$.

## 3. CHARACTERIZATIONS OF INTUITIONISTIC MULTI L – FUZZY SUBGROUPS

In this fragment, we discuss about the Union, Intersection and Cartesian Product of Intuitionistic multi L-fuzzy subgroups with given some of its examples.

3.1 Definition

Let $A = \{(x, \mu_A(x), \gamma_A(x)) : x \epsilon X\}$ and $B = \{(x, \mu_B(x), \gamma_B(x)) : x \epsilon X\}$ be any two ILMFSG's of X, then

**1. Union:** $A \cup B = \{(x, (\mu_A \cup \mu_B)(x), (\gamma_A \cup \gamma_B)(x)) : x \epsilon X\}$

Where, $(\mu_A \cup \mu_B)(x) = max\left(\mu_A(x), \mu_B(x)\right)$ and $(\gamma_A \cup \gamma_B)(x) = min\left(\gamma_A(x), \gamma_B(x)\right)$

**2. Intersection:** $A \cap B = \{(x, (\mu_A \cap \mu_B)(x), (\gamma_A \cap \gamma_B)(x)) : x \epsilon X\}$

Where, $(\mu_A \cap \mu_B)(x) = min\left(\mu_A(x), \mu_B(x)\right)$ and $(\gamma_A \cap \gamma_B)(x) = max\left(\gamma_A(x), \gamma_B(x)\right)$

**3. Cartesian Product:** $AxB = \{(\mu_A(x)\mu_B(x), \gamma_A(x)\gamma_B(x)): x\epsilon X\}$

**3.2 Theorem**

Let $H_1$ and $H_2$ be any two intuitionistic multi L- fuzzy subgroups of a group G. Then $H_1 \cap H_2$ is also an intuitionistic multi L- fuzzy subgroup of G.

**Proof**

Let, $\mu_A = \{(x, \mu_A(x), \gamma_A(x)): x\epsilon G\}$ in $H_1$ and

$\mu_B = \{(y, \mu_B(y), \gamma_B(y)): y\epsilon G\}$ in $H_2$ be any two IMLFSG's of G.

Then $H_1 \cap H_2 \neq \boldsymbol{\phi}$, since at least the identity element e is common to both $H_1$ and $H_2$.

i.e., $e \in H_1 \cap H_2$.

In order to prove that $H_1 \cap H_2$ is an IMLFSG.

It is sufficient to prove that $x \in H_1 \cap H_2$, $y \in H_1 \cap H_2 \Rightarrow xy^{-1} \in H_1 \cap H_2$.

i.e.,to prove (i) $\mu_A(x) \in H_1 \cap H_2, \mu_B(y) \in H_1 \cap H_2 \Rightarrow \mu_{A\cap B}(xy^{-1}) \in H_1 \cap H_2$ and

(ii) $\gamma_A(x) \in H_1 \cap H_2, \gamma_B(y) \in H_1 \cap H_2 \Rightarrow \gamma_{A\cap B}(xy^{-1}) \in H_1 \cap H_2$.

**To prove (i)**

Now, $\mu_A(x) \in H_1 \cap H_2 \Rightarrow \mu_A(x) \in H_1$ and $\mu_A(x) \in H_2$ and

$\mu_B(y) \in H_1 \cap H_2 \Rightarrow \mu_B(y) \in H_1$ and $\mu_B(y) \in H_2$.

But our hypothesis $H_1, H_2$ are IMLFSG's of G.

Then by definition of IMLFSG,

$\mu_A(x) \in H_1, \mu_B(y) \in H_1 \Rightarrow \mu_{A\cap B}(xy^{-1}) \in H_1$
…… (1)

and $\mu_A(x) \in H_2, \mu_B(y) \in H_2 \Rightarrow \mu_{A\cap B}(xy^{-1}) \in H_2$
…… (2)

From (1) & (2), we get

$\mu_{A\cap B}(xy^{-1}) \in H_1$ and $\mu_{A\cap B}(xy^{-1}) \in H_2 \Rightarrow \mu_{A\cap B}(xy^{-1}) \in H_1 \cap H_2$.

**To prove (ii)**

Now, $\gamma_A(x) \in H_1 \cap H_2 \Rightarrow \gamma_A(x) \in H_1$ and $\gamma_A(x) \in H_2$ and

$\gamma_B(y) \in H_1 \cap H_2 \Rightarrow \gamma_B(y) \in H_1$ and $\gamma_B(y) \in H_2$.

But our assumption that $H_1, H_2$ are IMLFSG's of G.

Then by definition of IMLFSG,

$\gamma_A(x) \in H_1, \gamma_B(y) \in H_1 \Rightarrow \gamma_{A\cap B}(xy^{-1}) \in H_1$
…… (3)

and $\gamma_A(x) \in H_2, \gamma_B(y) \in H_2 \Rightarrow \gamma_{A\cap B}(xy^{-1}) \in H_2$
…….. (4)

From (3) & (4), we get

$\gamma_{A\cap B}(xy^{-1}) \in H_1$ and $\gamma_{A\cap B}(xy^{-1}) \in H_2 \Rightarrow \gamma_{A\cap B}(xy^{-1}) \in H_1 \cap H_2$.

Hence, $H_1 \cap H_2$ is an IMLFSG of G.

Hence the proof.

**3.3 Theorem**

If $\{H_i / i \in N\}$ be a family of Intuitionistic Multi L-Fuzzy Subgroups of a group G. Where $H_i = \{< x, \mu_{H_i}(x), \gamma_{H_i}(x) > / x\epsilon G\}$, then $\bigcap_i H_i$ is also an IMLFSG of G.

**Proof**

Let $G$ be a group and let $\{H_i/i\epsilon N\}$ be any family of IMLFSG's of $G$. Here N is an index set and is such that

$\forall\, i \in N, H_i$ is an IMLFSG of G.

Let, $H = \bigcap_{i\in N} H_i = \bigcap\{(x, \mu_{H_i}(x), \gamma_{H_i}(x)) \in G/x \in H_i, \forall\, i \in N\}$ be the intersection of this family of IMLFSG's
of G.

Obviously $H \neq \boldsymbol{\phi}$, since at least the identity element $e \in H_i$, $\forall\, i \in N$.

Now let $x, y$ be any two elements of $H$.

Then, $x \in \bigcap_{i\in N} H_i \Rightarrow x \in \mu_{H_i}\, and\, x \in \gamma_{H_i}$, $\forall\, i \in N$
…… (1)

And $y \in \bigcap_{i\in N} H_i \Rightarrow y \in \mu_{H_i}\, and\, y \in \gamma_{H_i}$, $\forall\, i \in N$
…… (2)

But $\forall\, i \in N$, $H_i$ is an IMLFSG of G.

From (1) & (2), we get

$x \in \mu_{H_i} \bigcap \gamma_{H_i},\ y \in \mu_{H_i} \bigcap \gamma_{H_i} \Rightarrow xy^{-1} \in \mu_{H_i} \bigcap \gamma_{H_i}, \forall\, i \in N.$

Consequently, $xy^{-1} \in \bigcap_{i\in N} H_i$ .

Thus we have shown that $x, y \in \bigcap_{i\in N} H_i \Rightarrow xy^{-1} \in \bigcap_{i\in N} H_i$ .

Therefore, $H = \bigcap_{i\in N} H_i$ is an IMLFSG of G.

EXAMPLE 1

The intersection of two IMLFSG's is an IMLFSG.

**Proof**

Let G be the additive group of integers.

Then $\mu_A = \{0, 0.2, 0.4, 0.6, 0.8, \ldots\ldots\ldots\}$ and

$\mu_B = \{0, 0.3, 0.6, 0.9, 0.12, \ldots\ldots\ldots\}$ are both IMLFSG's of G.

Therefore, $\mu_{A\cap B} = \{0, 0.6, 0.12, 0.18, \ldots\ldots\ldots\}$ is an IMLFSG of G.

3.4 Theorem

The union of two IMLFSG's is an IMLFSG if and only if one is contained in the other.

**Proof**

Let $\mu_A = \{(x, \mu_A(x), \gamma_A(x))/x \in G\}$ in $H_1$ and

$\mu_B = \{(y, \mu_B(y), \gamma_B(y))/y \in G\}$ in $H_2$ be any two IMLFSG's of G.

(i) Suppose $H_1$ and $H_2$ are two IMLFSG's of a group G.
Let, $H_1 \subseteq H_2$ or $H_2 \subseteq H_1$.

Then $H_1 \cup H_2 = H_2$ if $H_1 \subseteq H_2$

and $H_1 \cup H_2 = H_1$ if $H_2 \subseteq H_1$.

But $H_1, H_2$ are IMLFSG's and therefore $H_1 \cup H_2$ is also an IMLFSG.

(ii) Conversely, suppose $H_1 \cup H_2$ is an IMLFSG.
To prove that $H_1 \subseteq H_2$ or $H_2 \subseteq H_1$.

Let us assume that $H_1$ is not a intuitionistic multi L-fuzzy subset of $H_2$ and $H_2$ is also not a intuitionistic multi L-fuzzy subset of $H_1$.

Now, $H_1$ is not a intuitionistic $\Rightarrow \exists\ \mu_A(x) \in H_1$ or $\gamma_A(x) \in H_1$ and $\mu_A(x) \notin H_2$ or $\gamma_A(x) \notin H_2$.
multi L- fuzzy subset of $H_2$.

$\Rightarrow \mu_A(x) \in H_1 \cup H_2$ and $\gamma_A(x) \in H_1 \cup H_2$
…….…….. (1)

And also $H_2$ is not a intuitionistic $\Rightarrow$ $\exists$ $\mu_B(y) \in H_2$ or $\gamma_B(y) \in H_2$ and $\mu_B(y) \notin H_1$ or $\gamma_B(y) \notin H_1$.

multi L-fuzzy subset of $H_1$.

$\Rightarrow$ $\mu_B(y) \in H_1 \cup H_2$ and $\gamma_B(y) \in H_1 \cup H_2$ …….…….. (2)

From (1) & (2), we have $\mu_A \in H_1 \cup H_2$ and $\mu_B \in H_1 \cup H_2$.

Since $H_1 \cup H_2$ is an IMLFSG, therefore $\mu_{A\cup B}(xy) = \mu_{A\cup B}(z)$ (say) is also an element of $H_1 \cup H_2$.

But, $\mu_{A\cup B}(xy) = \mu_{A\cup B}(z) \in H_1 \cup H_2$

$\Rightarrow$ $\mu_{A\cup B}(xy) = \mu_{A\cup B}(z) \in H_1$ *or* $H_2$.

Suppose, $\mu_{A\cup B}(xy) = \mu_{A\cup B}(z) \in H_1$.

Then, $\mu_{A\cup B}(y) = \mu_{A\cup B}(x^{-1}z) \in H_1$ [$\because H_1$ is an IMLFSG, then $\mu_A(x) \in H_1 \Rightarrow \mu_A(x^{-1}) \in H_1$]

But from (2), we have $\mu_B(y) \notin H_1$.

Thus we get a contradiction.

Again suppose $\mu_{A\cup B}(xy) = \mu_{A\cup B}(z) \in H_2$.

Then, $\mu_{A\cup B}(x) = \mu_{A\cup B}(y^{-1}z) \in H_2$
[$\because H_2$ is an IMLFSG, then $\mu_B(y) \in H_2 \Rightarrow \mu_B(y^{-1}) \in H_2$]

But from (1), we have $\mu_A(x) \notin H_2$.

Thus here also we get a contradiction.

Hence either $H_1 \subseteq H_2$ or $H_2 \subseteq H_1$.

**3.5 Theorem**

Let $\mu_A$ and $\mu_B$ be any two IMLFSG's of a group $G$. Then $\mu_{A\cup B}$ need not be an IMLFSG of G.

**Proof**

By above theorem 3.6, follows that union of two IMLFSG'S of a group G need not be an IMLFSG of G and hence the proof is clear.

**Remark**

In general, $\cup_i H_i$ is need not be an IMLFSG of G.

## Example 2

The union of two IMLFSG's is not necessarily an IMLFSG.

**Proof**

Let G be the additive group of integers.

Then $\mu_A = \{0,0.2,0.4,0.6,0.8,. \ldots \ldots .\}$ and

$\mu_B = \{0,0.3,0.6,0.9,0.12,. \ldots \ldots .\}$ are both IMLFSG's of G.

(i) We have, $\mu_{A\cup B} = \{0,0.2,0.3,0.4,0.6,. \ldots \ldots .\}$

Obviously $\mu_{A\cup B}$ is not closed with respect to addition as,

$0.2 \in \mu_{A\cup B}$, $0.3 \in \mu_{A\cup B}$ but $0.2 + 0.3 = 0.5 \notin \mu_{A\cup B}$.

Therefore, $\mu_{A\cup B}$ is not an IMLFSG of G.

(ii) If we take the IMLFSG, $\mu_C = \{0,0.4,0.8,. \ldots \ldots .\}$ of G.

Then $\mu_{A\cup C} = \{0,0.2,0.4,0.6,0.8,. \ldots \ldots .\}$

$= \mu_A$ and $\mu_A$ is an IMLFSG of G. [$\because \mu_C \subseteq \mu_A$]

So we shall prove in one of the following examples that the union of two IMLFSG's is an IMLFSG iff one is contained in the other.

**3.6 Definition**

Let $\mu_A$ and $\mu_B$ be any two IMLFSG's of a groups $G_1$ and $G_2$ respectively. Then the Cartesian product of $\mu_A$ and $\mu_B$ is denoted by $\mu_{A\times B}$ of $G_1 \times G_2$ is defined as follows:

$$\mu_{A\times B} = \{\langle (x,y), \mu_{A\times B}(x,y), \gamma_{A\times B}(x,y)\rangle / (x,y) \in G_1 \times G_2\}$$

Where, $\mu_{A\times B}(x,y) = min\{\mu_A(x), \mu_B(y)\}$ and

$$\gamma_{A\times B}(x,y) = max\{\gamma_A(x), \gamma_B(y)\}.$$

### 3.7 Definition

Let $\mu_{A_1}, \mu_{A_2}, \mu_{A_3}, \dots\dots, \mu_{A_n}$ be any n-IMLFSG's of a groups $G_1, G_2, G_3, \dots\dots, G_n$ respectively. Then the Cartesian product of n- IMLFSG's of $\mu_{A_1}, \mu_{A_2}, \mu_{A_3}, \dots\dots, \mu_{A_n}$ is defined as follows:

$$(\mu_{A_1\times A_2\times A_3\times \dots\dots\times A_n})(x_1, x_2, x_3, \dots\dots, x_n) = min\{\mu_{A_1}(x_1), \mu_{A_2}(x_2), \mu_{A_3}(x_3), \dots\dots\dots, \mu_{A_n}(x_n)\}$$ and

$$(\gamma_{A_1\times A_2\times A_3\times \dots\dots\times A_n})(x_1, x_2, x_3, \dots\dots, x_n) = max\{\gamma_{A_1}(x_1), \gamma_{A_2}(x_2), \gamma_{A_3}(x_3), \dots\dots\dots, \gamma_{A_n}(x_n)\},$$ for all $x_i \in G_i$.

### 3.8 Theorem

Let $(\mu_A, \gamma_A)$ and $(\mu_B, \gamma_B)$ be two intuitionistic multi L- fuzzy subsets of the groups G and H respectively. Suppose that $e$ and $e'$ are the neutral elements of G and H respectively. If $\mu_{A\times B}$ and $\gamma_{A\times B}$ be the IMLFSG of GxH, then at least one of the following two statements must hold:

(i) $\mu_B(e') \geq \mu_A(x), \forall\, x \in G$ $and$ $\gamma_B(e') \leq \gamma_A(x), \forall\, x \in G.$

(ii) $\mu_A(e) \geq \mu_B(y), \forall\, y \in H$ $and$ $\gamma_A(e) \leq \gamma_B(y), \forall\, y \in H.$

**Proof**

Let $\mu_{A\times B}$ and $\gamma_{A\times B}$ be the intuitionistic multi L- fuzzy subgroup of GxH.

Let, $\mu_{A\times B} = \{\langle (x,y), \mu_{A\times B}(x,y), \gamma_{A\times B}(x,y)\rangle / (x,y) \in G \times H\}$

By contraposition, suppose that none of the statements above (i) and (ii) holds.

Then we can find $a \in G$ and $b \in H$ such that

(1) $\mu_A(a) > \mu_B(e')$ and $\mu_B(b) > \mu_A(e)$.

We have, $\mu_{A\times B}(a,b) = \min\{\mu_A(a), \mu_B(b)\}$

$\geq \min\{\mu_B(e'), \mu_A(e)\}$

$= \min\{\mu_A(e), \mu_B(e')\}$

$= \mu_{A\times B}(e,e')$

Thus $\mu_{A\times B}$ is not an IMLFSG of GxH.

Hence, either (i) or (ii) is true.

(2) Similarly, $\gamma_A(a) < \gamma_B(e')$ and $\gamma_B(b) < \gamma_A(e)$.

We have, $\gamma_{A\times B}(a,b) = \max\{\gamma_A(a), \gamma_B(b)\}$

$\leq \max\{\gamma_B(e'), \gamma_A(e)\}$

$= \max\{\gamma_A(e), \gamma_B(e')\}$

$= \gamma_{A\times B}(e,e')$

Thus again $\gamma_{A\times B}$ is not an IMLFSG of GxH.

Hence, either (i) or (ii) is true.

### 3.9 Theorem

Let $(\mu_A, \gamma_A)$ and $(\mu_B, \gamma_B)$ be two intuitionistic multi L- fuzzy subsets of the groups G and H respectively, such that $\mu_A(x) \leq \mu_B(e')$ and $\gamma_A(x) \geq \gamma_B(e')\ \forall\, x \in H$, $e'$ being neutral element of H. If $\mu_{A\times B}$ and $\gamma_{A\times B}$ be the IMLFSG's of $G \times H$, then $\mu_A$ and $\gamma_A$ be an IMLFSG's of G.

**Proof**

(i) If $\mu_{A\times B}$ is an IMLFSG of $G\times H$ and $x,y\in G$ then $(x,e'),(y,e')\in G\times H$.

Now, using the property $\mu_A(x)\le\mu_B(e')\ \forall\ x\in G$.

We get $\mu_A(xy)=\min\{\mu_A(xy),\mu_B(e'e')\}$

$=\mu_{A\times B}(x,e')(y,e')$

$\ge\min\{\mu_{A\times B}(x,e'),\mu_{A\times B}(y,e')\}$ [$\because$ *by definition* 3.6]

$=\min\{\min(\mu_A(x),\mu_B(e')),\min(\mu_A(y),\mu_B(e'))\}$

$=\min\{\mu_A(x),\mu_A(y)\}$

i.e., $\mu_A(xy)\ge\min\{\mu_A(x),\mu_A(y)\}$

Also, $\mu_A(x^{-1})=\min\{\mu_A(x^{-1}),\mu_B(e^{-1})\}$

$=\mu_{A\times B}(x^{-1},e^{-1})=\mu_{A\times B}(x,e')^{-1}$

$=\mu_{A\times B}(x,e')$ [$\because\ \mu_{A\times B}$ is an IMLFSG]

$=\min\{\mu_A(x),\mu_B(e')\}$

i.e., $\mu_A(x^{-1})=\mu_A(x)$

Hence, $\mu_A$ is an IMLFSG of G.

(ii) Similarly, if $\gamma_{A\times B}$ is an IMLFSG of $G\times H$ and $x,y\in G$ then $(x,e'),(y,e')\in G\times H$.

Now, using the property $\gamma_A(x)\ge\gamma_B(e')\ \forall\ x\in G$.

We get $\gamma_A(xy)=\max\{\gamma_A(xy),\gamma_B(e'e')\}$

$=\gamma_{A\times B}(x,e')(y,e')$

$\le\max\{\gamma_{A\times B}(x,e'),\gamma_{A\times B}(y,e')\}$ [$\because$ *by definition* 3.6]

$=\max\{\max(\gamma_A(x),\gamma_B(e')),\max(\gamma_A(y),\gamma_B(e'))\}$

$=\max\{\gamma_A(x),\gamma_A(y)\}$

i.e., $\gamma_A(xy)\le\max\{\gamma_A(x),\gamma_A(y)\}$

Also, $\gamma_A(x^{-1})=max\{\gamma_A(x^{-1}),\gamma_B(e^{-1})\}$

$=\gamma_{A\times B}(x^{-1},e^{-1})=\gamma_{A\times B}(x,e')^{-1}$

$=\gamma_{A\times B}(x,e')$ [$\because\ \gamma_{A\times B}$ is an IMLFSG]

$=\max\{\gamma_A(x),\gamma_B(e')\}$

i.e., $\gamma_A(x^{-1})=\gamma_A(x)$

Hence, $\gamma_A$ is an IMLFSG of G.

**3.10 Theorem**

Let $(\mu_A,\gamma_A)$ and $(\mu_B,\gamma_B)$ be two intuitionistic multi L- fuzzy subsets of the groups G and H respectively, such that $\mu_A(e)\ge\mu_B(x)$ and $\gamma_A(e)\le\gamma_B(x)\ \forall\ x\in H$, $e$ being neutral element of G. If $\mu_{A\times B}$ and $\gamma_{A\times B}$ be the IMLFSG's of $G\times H$, then prove that $\mu_B$ and $\gamma_B$ be an IMLFSG's of H.

**Proof**

(i) If $\mu_{A\times B}$ is an IMLFSG of $G\times H$ and $x,y\in G$ then $(x,e),(y,e)\in G\times H$.

Now, using the property $\mu_A(e)\ge\mu_B(y)\ \forall\ y\in H$.

We get $\mu_B(xy)=\min\{\mu_B(xy),\mu_A(ee)\}$

$= \mu_{B\times A}((x,e)(y,e))$

$\geq \min\{\mu_{B\times A}(x,e), \mu_{B\times A}(y,e)\}$
[$\because$ *by definition* 3.6]

$= \min\{\min(\mu_B(x), \mu_A(e)), \min(\mu_B(y), \mu_A(e))\}$

$= \min\{\mu_B(x), \mu_B(y)\}$

i.e., $\mu_B(xy) \geq \min\{\mu_B(x), \mu_B(y)\}$

Also, $\mu_B(x^{-1}) = \min\{\mu_B(x^{-1}), \mu_A(e^{-1})\}$

$= \mu_{B\times A}(x^{-1}, e^{-1}) = \mu_{B\times A}(x,e)^{-1}$

$= \mu_{B\times A}(x,e)$
[$\because \mu_{B\times A}$ is an IMLFSG]

$= \min\{\mu_B(x), \mu_A(e)\}$

$\mu_B(x^{-1}) = \mu_B(x)$

Hence, $\mu_B$ is an IMLFSG of H.

(ii) Similarly, if $\gamma_{A\times B}$ is an IMLFSG of $G \times H$ and $x, y \in G$ then $(x,e), (y,e) \in G \times H$.

Now, using the property $\gamma_A(e) \leq \gamma_B(y)$ $\forall\, y \in H$.

We get $\gamma_B(xy) = \max\{\gamma_B(xy), \gamma_A(ee)\}$

$= \gamma_{B\times A}((x,e)(y,e))$

$\leq \max\{\gamma_{B\times A}(x,e), \gamma_{B\times A}(y,e)\}$
[$\because$ *by definition* 3.6]

$= \max\{\max(\gamma_B(x), \gamma_A(e)), \max(\gamma_B(y), \gamma_A(e))\}$

$= \max\{\gamma_B(x), \gamma_B(y)\}$

i.e., $\gamma_B(xy) \leq \max\{\gamma_B(x), \gamma_B(y)\}$

Also, $\gamma_B(x^{-1}) = \max\{\gamma_B(x^{-1}), \gamma_A(e^{-1})\}$

$= \gamma_{B\times A}(x^{-1}, e^{-1}) = \gamma_{B\times A}(x,e)^{-1}$

$= \gamma_{B\times A}(x,e)$
[$\because \gamma_{B\times A}$ is an IMLFSG]

$= \max\{\gamma_B(x), \gamma_A(e)\}$

$\gamma_B(x^{-1}) = \gamma_B(x)$

Hence, $\gamma_B$ is an IMLFSG of H.

### 3.11 Theorem

Let $(\mu_A, \gamma_A)$ and $(\mu_B, \gamma_B)$ be two intuitionistic multi L- fuzzy subgroups of G and H respectively, then $\mu_{A\times B}$ and $\gamma_{A\times B}$ is an IMLFSG of $G \times H$.

### Proof

Let $(a,b), (c,d), (a^{-1}, b^{-1}) \in G \times H$.

(i) Then, $\mu_{A\times B}\big((a,b)(c,d)\big) = \mu_{A\times B}(ac, bd)$

$= \min\{\mu_A(ac), \mu_B(bd)\}$

$\geq \min\{\min(\mu_A(a), \mu_A(c)), \min(\mu_B(b), \mu_B(d))\}$
[$\because$ *by definition* 3.6]

$= \min\{\min(\mu_A(a), \mu_B(b)), \min(\mu_A(c), \mu_B(d))\}$

$= \min\{\mu_{A\times B}(a,b), \mu_{A\times B}(c,d)\}$

i.e., $\mu_{A\times B}\big((a,b)(c,d)\big) \geq \min\{\mu_{A\times B}(a,b), \mu_{A\times B}(c,d)\}$

Also, $\mu_{A\times B}((a,b)) = \min\{\mu_A(a), \mu_B(b)\}$

$= \min\{\mu_A(a^{-1}), \mu_B(b^{-1})\}$
[$\because \mu_A$ & $\mu_B$ are IMLFSG of G]

$= \mu_{A\times B}(a^{-1}, b^{-1}) = \mu_{A\times B}(a,b)^{-1}$

i.e., $\mu_{A\times B}((a,b)) = \mu_{A\times B}(a,b)^{-1}$

Hence, $\mu_{A\times B}$ is an IMLFSG of $G \times H$.

(i) Also, $\gamma_{A\times B}\big((a,b)(c,d)\big) = \gamma_{A\times B}(ac,bd)$

$= max\{\gamma_A(ac), \gamma_B(bd)\}$

$\leq \max\{\max(\gamma_A(a),\gamma_A(c)), \max(\gamma_B(b),\gamma_B(d))$ [$\because$ $by\ definition$ 3.6]

$= \max\{\max(\gamma_A(a),\gamma_B(b)), \max(\gamma_A(c),\gamma_B(d))\}$

$= \max\{\gamma_{A\times B}(a,b), \gamma_{A\times B}(c,d)\}$

i.e., $\gamma_{A\times B}\big((a,b)(c,d)\big) \leq \max\{\gamma_{A\times B}(a,b), \gamma_{A\times B}(c,d)\}$

Also, $\gamma_{A\times B}((a,b)) = max\{\gamma_A(a), \gamma_B(b)\}$

$= \max\{\gamma_A(a^{-1}), \gamma_B(b^{-1})\}$
[$\because$ $\gamma_A$ & $\gamma_B$ are IMLFSG of G]

$= \gamma_{A\times B}(a^{-1}, b^{-1}) = \gamma_{A\times B}(a,b)^{-1}$

i.e., $\gamma_{A\times B}((a,b)) = \gamma_{A\times B}(a,b)^{-1}$

Hence, $\gamma_{A\times B}$ is an IMLFSG of $G \times H$.

Hence the proof.

### 3.12 Theorem

Let $\mu_{A_1}, \mu_{A_2}, \ldots\ldots, \mu_{A_n}$ and $\gamma_{A_1}, \gamma_{A_2}, \ldots\ldots, \gamma_{A_n}$ be the IMLFSG's of $G_1, G_2, \ldots\ldots, G_n$ respectively. Then $\mu_{A_1\times A_2\times\ldots\ldots\times A_n}$ and $\gamma_{A_1\times A_2\times\ldots\ldots\times A_n}$ is an IMLFSG of $G_1 \times G_2 \times \ldots\ldots \times G_n$.

### Proof

Let $\mu_{A_1}, \mu_{A_2}, \ldots\ldots, \mu_{A_n}$ and $\gamma_{A_1}, \gamma_{A_2}, \ldots\ldots, \gamma_{A_n}$ be an IMLFSG's of $G_1, G_2, \ldots\ldots, G_n$ respectively.

(i) To prove that $\mu_{A_1\times A_2\times\ldots\ldots\times A_n}$ is an IMLFSG of $G_1 \times G_2 \times \ldots\ldots \times G_n$.

Then,
$(\mu_{A_1\times A_2\times\ldots\ldots\times A_n})\big((x_1,x_2,\ldots\ldots,x_n)(y_1,y_2,\ldots\ldots,y_n)\big)$
$= (\mu_{A_1\times A_2\times\ldots\ldots\times A_n})(x_1y_1, x_2y_2, \ldots\ldots, x_ny_n)$

$= \min\{\mu_{A_1}(x_1y_1), \mu_{A_2}(x_2y_2), \ldots\ldots\ldots, \mu_{A_n}(x_ny_n)\}$
[$\because$ $by\ definition$ 3.7]

$\geq \min\{\min\{\mu_{A_1}(x_1), \mu_{A_1}(y_1)\}, \min\{\mu_{A_2}(x_2), \mu_{A_2}(y_2)\},$

$\ldots\ldots\ldots, \min\{\mu_{A_n}(x_n), \mu_{A_n}(y_n)\}\}$

$\geq \min\{\min\{\mu_{A_1}(x_1), \mu_{A_2}(x_2), \ldots\ldots, \mu_{A_n}(x_n)\},$

$\min\{\mu_{A_1}(y_1), \mu_{A_2}(y_2), \ldots\ldots, \mu_{A_n}(y_n)\}\}$

$= \min\{(\mu_{A_1\times A_2\times\ldots\ldots\times A_n})(x_1,x_2,\ldots\ldots,x_n),$

$(\mu_{A_1\times A_2\times\ldots\ldots\times A_n})(y_1,y_2,\ldots\ldots,y_n)\}$

i.e.,
$(\mu_{A_1\times A_2\times\ldots\ldots\times A_n})\big((x_1,x_2,\ldots\ldots,x_n)(y_1,y_2,\ldots\ldots,y_n)\big)$
$\geq \min\{(\mu_{A_1\times A_2\times\ldots\ldots\times A_n})(x_1,x_2,\ldots\ldots,x_n),$

$(\mu_{A_1\times A_2\times\ldots\ldots\times A_n})(y_1,y_2,\ldots\ldots,y_n)\}$

Also,
$(\mu_{A_1\times A_2\times\ldots\ldots\times A_n})((x_1,x_2,\ldots\ldots,x_n)^{-1}) =$
$(\mu_{A_1\times A_2\times\ldots\ldots\times A_n})(x_1^{-1}, x_2^{-1}, \ldots\ldots, x_n^{-1})$

$= \min\{\mu_{A_1}(x_1^{-1}), \mu_{A_2}(x_2^{-1}), \ldots\ldots\ldots, \mu_{A_n}(x_n^{-1})\}$

$= \min\{\mu_{A_1}(x_1), \mu_{A_2}(x_2), \dots\dots\dots, \mu_{A_n}(x_n)\}$

[$\because \mu_{A_i}$ is an IMLFSG of G]

i.e.,

$(\mu_{A_1\times A_2\times\dots\dots\dots\times A_n})((x_1, x_2, \dots\dots\dots, x_n)^{-1}) = (\mu_{A_1\times A_2\times\dots\dots\dots\times A_n})(x_1, x_2, \dots\dots\dots, x_n)$

Hence, $\mu_{A_1\times A_2\times\dots\dots\dots\times A_n}$ is an IMLFSG of $G_1 \times G_2 \times \dots\dots\dots\times G_n$.

(ii) Similarly, to prove that $\gamma_{A_1\times A_2\times\dots\dots\dots\times A_n}$ is an IMLFSG of $G_1 \times G_2 \times \dots\dots\dots\times G_n$.

Then,

$(\gamma_{A_1\times A_2\times\dots\dots\dots\times A_n})\big((x_1, x_2, \dots\dots\dots, x_n)(y_1, y_2, \dots\dots\dots, y_n)\big)$
$= (\gamma_{A_1\times A_2\times\dots\dots\dots\times A_n})(x_1y_1, x_2y_2, \dots\dots\dots, x_ny_n)$

$=$

$\max\{\gamma_{A_1}(x_1y_1), \gamma_{A_2}(x_2y_2), \dots\dots\dots\dots, \gamma_{A_n}(x_ny_n)\}$

[$\because$ *by definition* 3.7]

$\leq$

$\max\{\max\{\gamma_{A_1}(x_1), \gamma_{A_1}(y_1)\}, \max\{\gamma_{A_2}(x_2), \gamma_{A_2}(y_2)\}, \dots\dots\dots\dots, \max\{\gamma_{A_n}(x_n), \gamma_{A_n}(y_n)\}\}$

$\leq$

$\max\{\max\{\gamma_{A_1}(x_1), \gamma_{A_2}(x_2), \dots\dots\dots, \gamma_{A_n}(x_n)\}, \max\{\gamma_{A_1}(y_1), \gamma_{A_2}(y_2), \dots\dots\dots, \gamma_{A_n}(y_n)\}\}$

$=$

$\max\{(\gamma_{A_1\times A_2\times\dots\dots\dots\times A_n})(x_1, x_2, \dots\dots\dots, x_n), (\gamma_{A_1\times A_2\times\dots\dots\dots\times A_n})(y_1, y_2, \dots\dots\dots, y_n)\}$

i.e.,

$(\gamma_{A_1\times A_2\times\dots\dots\dots\times A_n})\big((x_1, x_2, \dots\dots\dots, x_n)(y_1, y_2, \dots\dots\dots, y_n)\big)$

$\leq$

$\max\{(\gamma_{A_1\times A_2\times\dots\dots\dots\times A_n})(x_1, x_2, \dots\dots\dots, x_n), (\gamma_{A_1\times A_2\times\dots\dots\dots\times A_n})(y_1, y_2, \dots\dots\dots, y_n)\}$

Also,

$(\gamma_{A_1\times A_2\times\dots\dots\dots\times A_n})((x_1, x_2, \dots\dots\dots, x_n)^{-1}) = (\gamma_{A_1\times A_2\times\dots\dots\dots\times A_n})(x_1^{-1}, x_2^{-1}, \dots\dots\dots, x_n^{-1})$

$= \max\{\gamma_{A_1}(x_1^{-1}), \gamma_{A_2}(x_2^{-1}), \dots\dots\dots\dots, \gamma_{A_n}(x_n^{-1})\}$

$= \max\{\gamma_{A_1}(x_1), \gamma_{A_2}(x_2), \dots\dots\dots\dots, \gamma_{A_n}(x_n)\}$

[$\because \gamma_{A_i}$ is an IMLFSG of G]

i.e.,

$(\gamma_{A_1\times A_2\times\dots\dots\dots\times A_n})((x_1, x_2, \dots\dots\dots, x_n)^{-1}) = (\gamma_{A_1\times A_2\times\dots\dots\dots\times A_n})(x_1, x_2, \dots\dots\dots, x_n)$

Hence, $\gamma_{A_1\times A_2\times\dots\dots\dots\times A_n}$ is an IMLFSG of $G_1 \times G_2 \times \dots\dots\dots\times G_n$.

Hence the proof.

## Example 3

The Cartesian product of two IMLFSG's is an IMLFSG.

**Proof**

Let $\mu_A = \{0, 0.2, 0.4, 0.6, 0.8\}$ and

$\mu_B = \{0, 0.3, 0.6, 0.9\}$ are both IMLFSG's of G.

We have, $\mu_{AXB} = \{(0,0), (0,0.3), (0,0.6), (0,0.9), (0.2,0), (0.2,0.3), (0.2,0.6), (0.2,0.9), (0.4,0), (0.4,0.3), (0.4,0.6), (0.4,0.9), (0.6,0), (0.6,0.3), (0.6,0.6), (0.6,0.9), (0.8,0), (0.8,0.3), (0.8,0.6), (0.8,0.9)\}$

Therefore, $\mu_{AXB}$ is an IMLFSG of G.

## CONCLUSION

In this paper, we have to studied the development of the theory of Intuitionistic Multi L-Fuzzy Subgroups using Intuitionistic Multi L-Fuzzy Subsets. Also in that paper dealt with the basic concepts and theorems based on Union, Intersection and Cartesian Products of Intuitionistic Multi L-Fuzzy Subgroups. In this concepts extend basically starting with two non-empty Intuitionistic Multi L-Fuzzy Subgroups in to arbitrary n – Intuitionistic Multi L-Fuzzy Subgroups and some of their examples were discussed.

## REFERENCES

[1] Akilesh.M and Senthil Kumar.A, "Anti product of multi anti L-fuzzy subgroup", International Journal of Pure and Applied Mathematics, ISSN:1314-3395 Vol.117, No. 20 (2017) , pp.51-62.

[2] Anthony.J.M and Sherwood.H, "A Characterization of fuzzy subgroups", Fuzzy sets and system, Vol.7, 297-305, (1982).

[3] Atanassov.K.T, "Intuitionistic fuzzy sets", Fuzzy Sets and Systems, vol. 20, pp. 87-96, 1986.

[4] Azriel Rosenfeld, 1971, "Fuzzy Groups", Journal of mathematical analysis and applications, 35, pp. 512- 517.

[5] Balasubramanian.K.R and Revathy.R, "Product of $(\lambda,\mu)$- Multi Fuzzy Subgroups of a Group", Annals of R.S.C.B., Vol.25, Issue 6, (2021), ISSN:1583-6258, pp. 2448-2460.

[6] Balamurugan.S, "Characterization of Intuitionistic Multi Fuzzy Subgroup", International Journal of Engineering, Science and Mathematics, Vol.10, Issue 09 (2021), ISSN:2320- 0294, pp. 1-20.

[7] Ejegwa.P.A., "On Fuzzy multigroups and Fuzzy Submultigroups,"Journal of Mathematics, J. Fuzzy Math. 26(3), 2018, 641-654, pp.1-13.

[8] Fathi.M and Abdul Razak Salleh, "Intuitionistic fuzzy groups," Asian Journal of Algebra,vol. 2(1), pp1-10, 2009.

[9] Muthuraj.R and Balamurugan.S, " A Study on Intuitionistic Multi-Fuzzy Subgroups", International Journal of Applications of Fuzzy sets and Artificial Intelligence, Vol.4, 153-172 (2014).

[10] Palaniappan.N , Naganathan.S and Arjunan.K, "A Study on Intuitionistic L-Fuzzy Subgroups", Applied mathematical Sciences, Vol.3, no.53 (2009), 2619-2624.

[11] Prabir Bhattacharya, "Fuzzy subgroups", Some characterizations, J. Math. Anal. Appl. 128 (1981), 241-252.

[12] Rosenfeld.A, "Fuzzy groups", J. Math. Anal. Appl, vol. 35, pp. 512-517, 1971.

[13] Sabu.S and Ramakrishnan .T.V, "Multi-fuzzy sets", International Mathematical Forum, 50, pp. 2471-2476.

[14] Sabu.S, Ramakrishnan.T.V, "Multi Fuzzy Sub Groups", Int.J. Contemp. Math. Sciences, Vol.6, No.8, pp.365-72 (2011).

[15] Sharma.P.K, "Intuitionistic Fuzzy Representations of Intuitionistic Fuzzy Groups", Asian Journal of Fuzzy and Applied Mathematics (ISSN;2321-564X) ( Vol. 03, 2015, pp.81-94.

[16] Sinoj.T.K. and Sunil.J.J., "Intuitionistic Fuzzy Multi Sets", International Journal of Engineering Sciences and Innovative Technology, Vol.2, No.6, 2013, pp.1-24.

[17] Souriar Sebastian, Babu Sundar.S, "Commutative L-Fuzzy Subgroups Fuzzy sets and Systems", 68 (1994), pp.115-121.

[18] Yuying Li, Xuzhu Wang and Liqiong Yang, "A Study of $(\lambda,\mu)$-Fuzzy Subgroups", Journal of Applied Mathematics Volume 2013.

[19] Zadeh.L.A, "Fuzzy sets", Information and control, 8, (1965), pp. 338-353.

[20] Zhan.J and Tan.Z, "Intuitionistic M-fuzzy groups," Soochow journal of mathematics, vol. 30(1), pp. 85-90, 2004.

# ANALYSIS OF PREY PREDATOR MODEL USING GRAPH DIFFERENTIAL EQUATIONS

**S.Vidhya**
Assistant Professor in Mathematics, Kongu Arts and Science College, Erode, Tamilnadu, India

**Abstract- In this paper, we introduce new concepts like a pseudo simple graph, the study of graph differential equations or equivalently matrix differential equations. We formulate a matrix differential equation for the famous prey predator model and Criteria are obtained to guarantee the existence of a solution and an iterative technique for convergence to the solution of a matrix differential equation is developed.**

Keywords: Simple Graph, Pseudo simple graph, graph differential equation, matrix differential equation and prey predator model.

## I. INTRODUCTION

Any natural or a man made system involves interconnections between its constituents, thus forming a network, which can be expressed by a graph [1, 2, 3]. Graphs arise naturally when trying to model organizational structures in social sciences. It has been noted that a graph which is static in nature is not suitable for social phenomena whose changes with time are natural. This led to the introduction of a dynamic graph and a Graph Differential Equation (GDE) in [2]. The introduced concepts were successfully utilized to study stability of complex dynamic systems through its associated adjacency matrix [2]. In [3] we have utilized the concepts defined in [2] including a graph linear space and its associated matrix linear space. Using the notion of a dynamic graph and the graph differential equations we observed that the study of GDEs falls into the realm of differential equations in abstract spaces. This study, through highly mathematical, would be of little use for practical problems. On the other hand, if we consider the associated matrix differential equation (MDE) then the approach appeared more reasonable and practical for the study of GDEs. Hence in [3], we considered a weighted directed simple graph as the basic element and developed the theory. We have obtained existence and uniqueness of solutions of a GDE through its associated MDE using the monotone iterative technique. In [4] through we have developed significant results, the basic concept involved was weighted directed simple graph. Since a simple graph has no loops, this fact when translated into differential equations frame work states that there is no way to accommodate the rate of change of anedge eii and its relation with other edges including the edge eii . This is a drawback that had to be handled to model physical phenomena using graph differential equations, which called for a new concept that we plan to introduce in this paper. Further, since there exists an isomorphism between graphs and their adjacency matrices, we successfully exploited it and defined the product of two graphs. A good example, will go a long way in support of the theory, we have considered the prey predator problem and developed the corresponding matrix differential equation and showed how the nonlinearity is preserved in this set up. The rest of the paper is as follows. In section two, we introduce the concepts of pseudo simple graph and product of two graphs and have obtained a result that can be of practical importance in this set up. In section three we obtained the matrix differential equation for prey predator problem and extended it to three species and further generalized it. In section four we conclude our work.

## II. MAIN RESULTS

In this section, we begin with the concept of a pseudo simple and later introduce the product of two graphs.

Definition 1 (Pseudo simple graph). A simple graph having loops is called as a pseudo simple graph. Parallel to the definitions and theory developed in [3] we proceed to state the results in this set up. We avoid the details for fear of repetition.

Let v1, v2, . . . , vN be N vertices, N fixed. Let DN be the set of all weighted directed pseudo simple graphs D = (V, E). Then (DN ,+, .) is a

linear space with the definitions given in [3] and [2].

Let the set of all corresponding adjacency matrices be EN . Then (EN ,+, .) is a matrix linear space where '+' denotes matrix addition and '.' indicates scalar multiplication. With this basic structure defined, the comparison theorems, existence and uniqueness results of solutions of MDE and the corresponding GDE follow as in [3].

Taking cue from matrix multiplication we define the product of two graphs as follows. Product of graphs: Let G1 and G2 be two graphs with edges (ei j) N×N and (di j)N×N respectively. Then the product of the two graphs G1 and G2 is the graph G in which the weight gi j of the edge from vj tovi is the dot product of the vectors one having the weights of the edges inwards to vi and the other having weights of the edges outwards from vj .We now proceed to develop a result on the nature of solutions of a graph differential equation.

Let

D ' = g(t, D) (1)

be a graph differential equation.

Now if possible suppose g(t, D) can be written as a product of two graphs CD where C is a graph having constant weights. Then the GDE (1) can be written in the form

D ' =CD

Dt0 = D0 (2)

where C is a graph called a coefficient graph and D0 is the initial graph.

Let

E ' =AE

Et0=E ' (3)

be the corresponding IVP of the MDE where E0 is the adjacency matrix corresponding to the initial graph D0. Then we have the following result relating to the solutions of MDE and hence to that of GDEs.

## III. MODELING OF THE PREY-PREDATOR PROBLEM

In this section, we formulate a matrix differential equation for the famous prey predator model and later extend it to three species and N-species. Let x denote the prey population and y denote the predator population, Then the rate of change of prey and that of predator gives rise to a system of nonlinear differential equations given by

$$\frac{dx}{dt} = ax - bxy \quad (4)$$

$$\frac{dy}{dt} = dyx - cy \quad (5)$$

$\frac{dx}{dt}$ and $\frac{dy}{dt}$ represent the growth rates of the two populations over time.

$t$ represents time

$a$, $b$, $c$, $d$ are just positive real parameters describing the interaction of the two species:

$a$ − Growth rate

$b$ − Searching efficiency or attack rate

$c$ − Predator mortality rate

$d$ − Growth rate of predator or predator's ability at turning food into offspring.

It is well known that the above differential equations are linearized and solved as a linear system of differential equations.

We now express the above system as a Graph Differential Equation and consider the corresponding Matrix Differential Equation. We show that the nonlinearity is preserved in this set up.

Let the vertex v1 denote the prey and v2 denote the predator. Set e11 = x as population of the prey and e22 = y as the population of the predator. It can be seen that e12 is the edge going outward from v1 and is incident on v2. This means that e12 denotes the interaction between prey and predator. Actually, e21 gives the status of prey finding the predator. Similarly e21 denotes the edge outward from v2 and incident on v1. In terms of our model, this edge indicates the status of prey that fall preyto predators.

Now the graph of the prey predator model is of the form

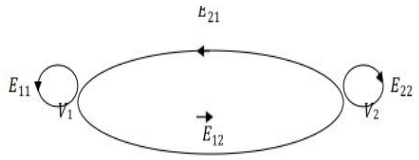



Figure 1: The Graph of the prey-predator problem

And its adjacency matrix is given by

$$\begin{bmatrix} e_{11} & e_{12} \\ e_{21} & e_{22} \end{bmatrix}$$

The equation (4) and (5) reduce the form

$$e'_{11} = ae_{11} - be_{21} \quad (6)$$

$$e'_{22} = de_{12} - ce_{22} \quad (7)$$

Our aim is to obtain a Matrix Differential Equations of the form

$$\begin{bmatrix} e'_{11} & e'_{12} \\ e'_{21} & e'_{22} \end{bmatrix} = A \begin{bmatrix} e_{11} & e_{12} \\ e_{21} & e_{22} \end{bmatrix}$$

Where $A_{2\times 2}$ is the coefficient matrix.

It can be easily seen that

$$\begin{bmatrix} e'_{11} \\ e'_{22} \end{bmatrix} = \begin{bmatrix} a & -b \\ d & -c \end{bmatrix} \begin{bmatrix} e_{11} \\ e_{22} \end{bmatrix}$$

And hence we propose to choose

$$A = \begin{bmatrix} a & -b \\ d & -c \end{bmatrix}$$

And obtain matrix differential equation of the form

$$\begin{bmatrix} e'_{11} & e'_{12} \\ e'_{21} & e'_{22} \end{bmatrix} = \begin{bmatrix} a & -b \\ d & -c \end{bmatrix} \begin{bmatrix} e_{11} & e_{12} \\ e_{21} & e_{22} \end{bmatrix} \quad (8)$$

The system (8) yields the equation (6),(7) and the following two differential equations given by

$$e'_{12} = ae_{12} - be_{22} \quad (9)$$

$$e'_{21} = de_{11} - ce_{21} \quad (10)$$

The equation (9) describe the rate of change of predator finding prey and it is negatively proportional to the predator finding prey and positively proportional to the predator population. The equation (10) gives the rate of change prey coming in way of predator and this is negatively proportional to prey available and positively proportional to prey falling to predator. Hence it can be seen that all the four equations given by (6), (7), (9) and (10) are consistent with the standard prey predator problem.

The beauty in this set up is that the nonlinearity is preserved and effectively used. The system obtained reduces to a Matrix linear differential equation and the solution is immediately given by

$$\begin{bmatrix} e_{11}(t) & e_{12}(t) \\ e_{21}(t) & e_{22}(t) \end{bmatrix} = e^{A(t-t_0)}E_0$$

Where $E0$ is the given matrix of initial conditions at t =$t0$, see [1]. Observe that $e^{A(t-t0)}$ is a matrix. If A is diagonalizable then $e^{A(t-t0)}$ can be replaced by the diagonal matrix $e^{H(t-t0)}$ , where H = diag[$\lambda 1$, $\lambda 2$] where $\lambda 1, and\ \lambda 2$ are the eigen values of A and the matrix has the form

$$e^{H(t-t_0)} = \begin{bmatrix} e^{\lambda_1(t-t_0)} & 0 \\ 0 & e^{\lambda_2(t-t_0)} \end{bmatrix}$$

Thus it has been effectively shown that a physical phenomena can be described through a graph and using the standard models we can preserve the nonlinearity and obtain more information using its associated matrix differential equation.

## V. CONCLUSION

In this paper we have introduced the notions of a pseudo simple graph and the product of two graphs we have given sufficient conditions under which a solution of a Graph Differential Equation has the same nature as its graph of initial conditions. Further, we have obtained a matrix differential equation for a prey predator problem and explicitly gave its solutions preserving the nonlinearity. From the model, it is clear that the nonlinearity in the prey predator problem is preserved by using a graph differential equation.

## REFERENCES

[1] S.G. Deo, V.Lakshmikantham, and V.Raghavendra. Textbook of Ordinary Differential Equations, 2 nd Edition. Tata McGraw Hill Education Private Limited, New Delhi.1998.

[2] D.D. Siljak. Dynamic Graphs, Nonlinear Analysis, Hybrid Systems 2, 544-567, 2008.

[3] J.Vasundhara Devi, R.V.G. Ravi Kumar, and N.G. Babu. On graph differential equations and its associated matrix differential equations. Malaya Journal of Mathematik 1 (1) 82- 90.2012

[4] J.Vasundhara Devi, R.V.G. Ravi Kumar. Modeling the Prey Predator Problem by a Graph Differential Equation. European Journal of Pure and Applied Mathematics vol.7, No.1,2014, 37-4.

# ADVANCED FEATURE FUSION FOR MULTIMODAL BIOMETRIC USING HYBRID APPROACH TEMPLATE

**M Dharshna Moorthy[1]**
III B.Sc.IT, Department of Information Technology,
Sri Krishna Adithya College of Arts & Science,
Coimbatore, Tamil Nadu, India
**Dr A Vinoth[2]**
Assistant Professor, Department of Information Technology,
Sri Krishna Adithya College of Arts & Science,
Coimbatore, Tamil Nadu, India

**Abstract**

**Multimodal biometrics, which combines two or more biometric modalities/traits in a single identification system are becoming popular. Fingerprint and Iris are the most used and reliable biometrics having high performance but fail under certain conditions. We address this problem by combining fingerprint and iris biometrics to overcome limitation of unimodal fingerprint and iris biometric system. Besides enhancing accuracy, this framework also address non-universality, noisy input data, intraclass variation and spoof attacks. In this work we present a transformation based approach of combining fingerprint and iris at confidence level. The significance of this approach is that, it does not require any estimation, probabilistic interpretation or a large number of training score. The features are extracted from individual fingerprint and iris biometric modalities by efficient algorithm. These features are first matched with their corresponding templates to compute the corresponding scores at confidence level. Match scores obtained from these traits are transformed using different transformation techniques and combined by different fusion rule to generate a fused match score. Practical investigation demonstrate improved accuracy of this system over its unimodal counterpart.. The experimental result demonstrates significant improvement in performance. This, framework is easier to implement and requires less memory.**

Keywords: VANET, TDG, UAV, NDN, RSU, QIP

## 1. INTRODUCTION

The need for reliable user authentication techniques has increased in the wake of heightened concerns about security and rapid advancements in networking, communication and mobility. A wide variety of systems require reliable personal recognition schemes either to confirm or determine the identity of an individual requesting their facilities. The aim of such systems is to guarantee that the rendered services are accessed only by genuine user, and not by anyone else. Examples of such applications include secure access to organizations, cell phones, laptops, computer systems and different electronic gadgets. In lack of robust personal recognition arrangements, these systems are exposed to the tricks of an impostor.

Traditional authentication techniques based on passwords and tokens are limited in their ability to address following issues:

Client attack - person guessing, robbing or stealing tokens or password

Eavesdropping - password shared by many individuals

Trojan horse attack - Stealing password by installing fake log-in screen.

Denial of service - Using incorrect password to disable the system.

Non-Repudiation - individual claiming that password was lost or misplaced. A warranty that an individual who used or accesses a certain facility cannot later deny using it.

Negative recognition - A process by which a system determines that a certain individual is indeed registered in the system although the individual might later deny it.

## II. LITERATURE STUDY

Pre classification integration or fusion before matching can be divided into two categories sensor level and feature level. Sensor and feature integration is used to merge information or raw data of two multiple sensors or multiple samples of single biometric using single sensor. The benefit of these level integrations feature attaching and reduction techniques, ones the feature sets are joined, to reduce the size of the feature which subsequently ease the matching process.

Past work shows that sensor integration and feature level integration are mainly carried on fingerprint and face traits. Complimentary evidences is primarily extracted from different sensor.

DRAWBACK OF EXISTING SYSTEM

However fusion at image level and feature level is difficult to succeed because of following reasons:

(a) Feature set of multiple traits may be incompatible, as in our case (e. g. texture feature of iris and minutiae set of fingerprint

(b) Combining feature vector may lead to large dimension resultant feature vector resulting in 'curse of dimensionality" which is well known phenomenon in pattern recognition.

(c) Relation between feature set of multiple biometrics is difficult to find and

(d) Complex matching algorithm is required to operate on this combined feature vector.

Researcher in some application have demonstrated that integration at these level do not provide good result compared to score level. Notwithstanding this fact, when features from multiple biometrics not related or independent as in fingerprint and iris, in some circumstances feature combination may allow dependences to be exploited fully. Hence an attempt is made to fuse fingerprint and iris at feature level than merely using score level integration.

## III. DEVELOPMENT ON USER AUTHENTICATION BASED ON THE KEYSTROKE DYNAMICS

Fingerprint and Iris are considered to be most reliable biometrics having high performance accuracy. The two individual classifiers perform better individually but fail under certain conditions. In case of fingerprint recognition poor quality fingerprint image may create problem. The enhancement subsystem extracts the ridges present but the loss due to cuts and scars present on the fingerprint image may create problem in extraction of minutiae points. Fingerprints of a small section of the population may be not suitable for automatic identification because of aging, environmental and or occupational reasons (e.g., workers may have a number of cuts and bruises on their fingerprints and that may also change with time) and genetic disorders. A problem can also arise at the time of iris image acquisition where the user has to be co-operative while giving iris image.

The International Biometric Group (IBG) evaluated the performance of specific iris recognition software (Iridian KnoWho OEM SDK) on iris images obtained using three different iris cameras (LG Iris Access 3000, OKI Electronics IRISPASS-WG and Panasonic BM-ET300) from 1,224 subjects. It was reported that between 1.6% and 7% of the participants could not be successfully enrolled in the system based on the camera that was used (International Biometric Group, 2005). Thus, there is a failure to enroll (FTE) rate associated with using a single biometric trait.

Iris systems have a very low False Accept Rate (FAR) compared to other biometric traits, but the False Reject Rate (FRR) of these systems can be rather high (International Biometric Group, 2005). Secondly fingerprints and iris traits are also susceptible to spoof attacks. Fingerprint can be lifted off from car doors and iris data can be extracted from high resolution photographs. Possibility of generating digital artifacts of biometric characteristics in order to circumvent a biometric system has also been demonstrated. Spoof attacks, when successful, can severely undermine the security afforded by a biometric system.

Some ad hoc techniques have been tried using fusion at different levels using a combination of uncorrelated modalities (e.g. fingerprint and face, two fingers of a person, etc.) and combination of correlated modalities (e.g. different fingerprint matchers, different face matcher etc.). Still there are problem associated with fusion techniques, and which fusion technique is the best is yet to be decided. We believe that fusion in multibiometric is still a largely unexplored area, and problems related to fusion, normalization and correlation among different biometric algorithms needs to be addressed. Although there has been research on combining different biometrics modalities for a variety of purposes, however, biometric fusion techniques have been developed only for fingerprint and face,face and iris, audio and face. Little work has been focused on the combination of finger print and iris, the two characteristics that can reach the best recognition performance for high security applications.

Here we combine fingerprint and iris biometric in which the information is integrated at confidence level as well as data or image level.

BENEFITS OF PROPOSED SYSTEM:

In our research work we present a novel approach of combining image level information of fingerprint and iris. The raw information or data is fused at image level or feature level. A unique feature vector is created from the textural information of fused image of fingerprint and iris. The proposed framework consists of fingerprint feature extraction module, iris feature extraction module, fusion module and matching module.

Fingerprint feature extraction module comprises of fingerprint acquisition, enhancement and resizing of fingerprint image and discrete wavelet transformation (DWT).

In iris module first the iris region is segmented from eye image and then it is normalized before applying DWT. Haar wavelet transform is been used. This breaks the fingerprint and iris image into four subsamples LL, LH, HL, HH, where L and H means high pass and low pass filter. LL of both fingerprint and iris are fused together and a unique feature vector is extracted from this fused image.

Classification of feature vector is carried out using Hamming distance. The block diagram of proposed framework is as shown in Fig.1

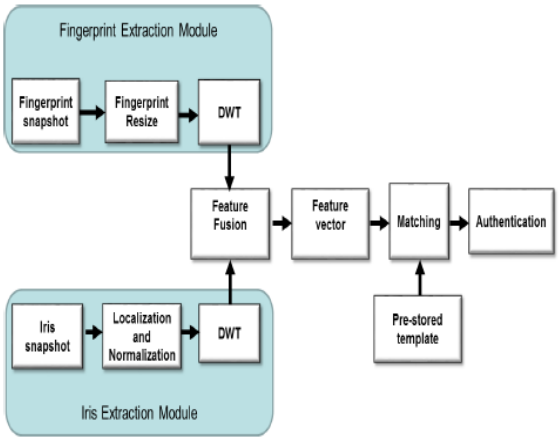



**Fig.1 Block Diagram of the proposed Framework**

## MODULE DESIGN

### Fingerprint Feature Extraction Module

There are several standard fingerprint databases available. The standard database acquired and used for our work was the FVC database. FVC is International fingerprint Verification Competition which was organized in 2000, 2002, 2004 and 2006. This FVC database consists of four databases which were acquired using three different sensors and the SFinGE synthetic generator, which was used for these competition.

The fingerprint image is then pre-processed. This step mainly consists of image enhancement and resizing the image for further operations. This is done as explained in previous chapter.

Wavelets are used to decompose the image data into components that appear at different resolutions. Wavelets have the benefit over traditional Fourier transform. A number of wavelet filters, also called a bank of wavelets, is applied to the image, one for each resolution with each wavelet a scaled version of some basis function. The output of applying the wavelets is then encoded in order to provide a compact and discriminating representation of the finger pattern. Discrete Wavelet Transform (DWT) is carried out on enhanced fingerprint image.

The Haar wavelet is one of the simplest wavelet transforms which can transform huge data sets to considerably smaller representations. Here we use Haar wavelet transform. De-composing images with wavelet transform yields a multi-resolution from a detailed image to approximation images in each level. The sub-images (quadrants) within the image indicated as LH, HL, and HH represents horizontal, vertical, and diagonal orientation, respectively. The sub-image LL representing an approximation image that is further used as an input to feature fusion block.

**Iris Feature Extraction Module**

We have collected CASIA [Chinese Academy of Sciences Institute of Automation] Iris Image databases. The CASIA database is the only standard databases publicly available and used for evaluation of the iris biometric modality. It has been widely used for research and evaluation. The main stages iris feature extraction module are iris localization, normalization and discrete wavelet transform. Localization and normalization are implemented. Next discrete wavelet transform is applied to normalized iris image. Haar wavelet is used to generate four bands of coefficients. This breaks the iris image into four subsamples LL, LH, HL, HH, where L and H means high pass and low pass filter. LL sub image is used for further processing.

**Feature Fusion Module**

In multimodal biometric system employing feature fusion usually features are extracted from single trait using different sensors of same modality. These features are not independent, and hence it is reasonable to concatenate the two vectors into single new vector. The new vector has high dimensionality and represents identity in different space. Further feature reduction techniques are used to reduce the size of feature vector. Feature fusion is used when feature set are homogenous.

Feature set of iris and fingerprint are incompatible, non-homogenous and connection between them is not known hence a new distinct method of fusion is been proposed. Fingerprint and iris features are extracted as explained earlier.

The outputs of feature extraction modules are sub images (LL band) of original fingerprint and iris images is the input to fusion module. The DWT image of each fingerprint LL subsample and iris LL subsample are fused together. The fused subsample is converted into one dimensional vector which is treated as a single template vector.

**Matching Module**

When a live fingerprint and iris is presented for comparison, the fingerprint and iris pattern is processed, textural information of fingerprint and iris are extracted and encoded into single template vector. The feature template vector derived from this process is compared with previously stored feature vector. This process is called matching. Matching appraises the level of match between the newly acquired vector and the candidate's data base entry. Based on this level of match, final decision is taken whether acquired data does or doesn't come from the same fingerprint and iris as does the database entry. Hamming distance is the distance measure used which is given as

Where, Xi and Yi are query and database codes and B number of bits.

**INPUT DESIGN**

Input design is the process of converting user oriented input to computer based format. In the System design phases, the expanded data flow diagramed identified local data flows, data stores, sources and destination. A system flowchart specifies master files, transaction files computer programs, input data are collected and organized into group of similar data. Once identified, appropriate input media are selected for processing. The input screen is design for easy use of the users. The next boxes are laid out to enable the users logically follow the data input. Combo boxes allow the user to selected various option in the search form.

The goal of designing input data is to make data entry as easy, logical and free from error as possible. In entering data, the operations included are. The allocated space for each field. Filed sequences that matches sources document. The format in which data fields are entered.

**NORMALIZATION**

The first stage is normalization is to reduce the data to its first normal form, by removing repeating items showing them as separate records but including in them the key fields of the original record.

The next stage of reduction to the second normal form is to check that the record, which one is first normal form, all the items in each record are entirely dependent on the key of the record. If a data item is not dependent on the key of the record, but on the other data item, then it is removed with its key to form another record. This is done until each record contains data items, which are entirely dependent on the key of their record.

The final stage of the analysis, the reduction of third normal form involves examining each record, which one is in second normal form to see whether any items are mutually dependent. If there are any item there are removed to a separate record leaving one of the items behind in the original record and using that as the key in the newly created record.

**BUSINESS MODELING**

The information flow among business function is modeled in a way that answers the following questions: what information drives the business process? What information is generated? What generate it? Where does the information go? Who process it?

**DATA MODELING**

The information flow defined as a process of the business modeling is refined into a set of data objects that are needed to support the business. The characteristics (called attributes) of each object are identified and relationships between these objects are defined

**PROCESS MODELING**

The data objects defined in the data-modeling phase are transformed to achieve the information flow necessary to implement a business function. Processing description is created for addition, modifying, deleting, or retrieving a data object.

## IV. RESULT AND DISCUSSION

Integration of fingerprint and iris, the two most unique, universal, invariant and accurate traits is carried out in this research work. The results demonstrate that combining fingerprint and Iris, which are uncorrelated or negatively correlated traits results in increase in matching performance than combining positively correlated sources (e.g. different fingerprint matchers, different iris matcher etc.). It also confirms that combining strong biometric can provide improved performance accuracy contradictory to the belief that combining strong biometric degrades the performance.

This multimodal biometrics authentication framework of fingerprint and iris overcome few of the limitation of its unimodal counterpart. The issue of non-universality is addressed, as multiple modality of iris and fingerprint can ensure sufficient population coverage compared to limited coverage in unimodal systems. Subjects unsuitable for one modality can use the other modality. (e.g., manual workers having large number of cuts and bruises on their fingerprints or whose fingerprint keep changing and individual unable to give a iris image). It becomes extremely difficult for an intruder to spoof fingerprint and iris traits simultaneously of a legitimately registered individual. The framework effectively addresses noisy data problems and facilitate in indexing and filtering large biometric databases. Continuous monitoring is possible when certain source become unreliable due to sensor replacement, software malfunction or deliberate manipulation by user. It also reduces the overall failure to enroll rate and enhance system security.

In future different density based integration methods and classifier based method can be tried for integrating fingerprint and iris biometrics.

## V. CONCLUSION AND FUTURE ENHANCEMENT

In this work, we have proposed a comprehensive DAVN architecture to integrate drones with ground vehicular networks to efficiently improve system performance. We

have discussed the challenges and open research issues in DAVN. The effectiveness of DAVN has been verified. In this paper, we proposed a routing protocol that uses infrastructure drones for boosting VANET communications to achieve a minimum vehicle-to-drone packet delivery delay. This work proposed a closed-form expression for the probability distribution of the vehicle-to-drone packet delivery delay on a two-way highway. In addition, based on that closed-form expression, we calculated the minimum drone density (maximum separation distance between two adjacent drones) that stochastically limits the worst case of the vehicle to drone packet delivery delay.

Moreover, we proposed a drones-active service (DAS) that is added to the location service in a VANET to dynamically and periodically obtain the required number of active drones based on the current highway connectivity state by obtaining the maximum distance between each two adjacent drones while satisfying a probabilistic constraint for vehicle-to-drone packet delivery delay. The simulation results show the accuracy of our analysis and reflect the relation between the drone density, vehicular density and speed, other VANET parameters, and the vehicle-to-drone packet delivery delay.

## REFERENCES

[1] A. K. Jain, A. Ross, and S. Prabhakar, “An introduction to biometric recognition,” IEEE Trans. on Circuits and Systems for Video Technology, 2004, vol. 14, pp. 4–20.

[2] A. K. Jain and A. Ross, “Multibiometric Systems.”, Communications of the ACM, Special Issue on Multimodal Interfaces, 2004, vol. 47(1), pp. 34–40,

[3] A. K. Jain and A. Ross, “Multimodal biometric an overview”, 12th European Signal Processing Conference (EUSIPCO), (Vienna, Austria), 2004, pp. 1221-1224..

[4] Marcialis, G. L. and Roll. F. “Fingerprint Verification by Fusion of Optical and Capacitive Sensors.”, Pattern Recognition Letters, 2004,vol. 25(11), pp.1315-1322,

[5] Chang, K. I., Bowyer, K. W., and Flynn, P J. “An Evaluation of Multimodal 2D+3D Face Biometrics.” IEEE Transactions on Pattern Analysis and Machine Intelligence, 2005, 27(4):619-624.

[6] K. I. Chang, K. W. Bowyer, and P. J. Flynn, “Face recognition using 2D and 3D facial data,”, Proc. Of Workshop on Multimodal User Authentication, [1] A. K. Jain, A. Ross, and S. Prabhakar, “An introduction to biometric recognition,” IEEE Trans. on Circuits and Systems (Santa Barbara, CA), 2003, pp. 25–32.

[7] Lu, X. and Jain, A. K. “Integrating Range and Texture Information for 3D Face Recognition”, IEEE Computer Society Workshop on Application of Computer Vision, (WACV), Breckenridge, USA. 2005, pp 156-163.

[8] Chen, X., Flynn, P. J., and Bowyer, K. W. “IR and Visible Light Face Recognition. Computer Vision and Image Understanding”, 2005, vol. 99(3):332-358.

[9] Chang, K. I., Bowyer, K. W, Flynn, P. J., and Chen, X. “Multibiometrics Using Facial Appearance, Shape and Temperature.”, Sixth IEEE International Conference on Automatic Face and Gesture Recognition, Seoul, Korea, 2004, pp 43-48.

[10] Ross, A. and Govindarajan, R. “Feature Level Fusion Using Hand and Face Biometrics”, Proceedings of SPIE Conference on Biometric Technology for Human Identification II, Orlando, USA, 2005, vol. 5779, pp 196-204.

[11] Wilson, C, Hicklin, A. R., Bone, M., Korves, H., Grother, R, Ulery, B., Micheals, R., Zoepfl, M., Otto, S., and Watson, C. “Fingerprint Vendor Technology Evaluation 2003: Summary of Results and Analysis Report.” NIST Technical Report NISTIR 7123, National Institute of Standards and Technology. (2004).

[12] Garris, M. D., Watson, C. I., and Wilson, C. L. “Matching Performance for the US-Visit IDENT System Using Flat Fingerprints.” Technical Report 7110, National Institute of Standards and Technology (NIST). NIST Internal Report 7110. (2004).

[13] Prabhakar, S. and Jain, A. K. “Decision-level Fusion in Fingerprint Verification.”, Pattern Recognition, 2002, vol. 35(4), pp 861-874.

[14] Jain, A. K. and Ross, A. “Fingerprint Mosaicking”, IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Orlando, USA, 2002, vol. 4, pp 4064 -4067.

[15] Jain, A. K., Prabhakar, S., and Chen, S., “Combining Multiple Matchers for a High Security Fingerprint Verification System”, Pattern Recognition Letters, 1999, vol. 20, pp 1371-1379.

# USER AUTHENTICATION BASED ON THE KEYSTROKE DYNAMICS

**P Sunitha**[1]
III B.Sc.IT, Department of Information Technology,
Sri Krishna Adithya College of Arts & Science,
Coimbatore, Tamil Nadu, India
**Dr S Ganeshmoorthy**[2]
Assistant Professor, Department of Information Technology,
Sri Krishna Adithya College of Arts & Science,
Coimbatore, Tamil Nadu, India

Abstract:

**Authentication is frequently referred as the most critical part of a computer system security. Users commonly identify themselves using a combination of username and password, but sometimes this is not enough. Concerning web-based services, attacks like phishing or social engineering can easily result in identity theft. In addition, the widespread use of single sign-on services can seriously increase the consequences of such attacks. In these circumstances strong authentication is mandatory. Strong authentication is often implemented using additional authentication steps or specialized hardware modules, which is not suitable for web-based systems, However, biometrics can used to overcome these limitations. More specifically, behavioural biometrics based on keyboard typing patterns can provide an extra security layer on top of conventional authentication methods, with no additional cost and no impact to the user experience. This work aims to evaluate the feasibility of the implementation of strong authentication on the web using keystroke dynamics. This is carried out through the creation of a application prototype using python environment.**

Keyword : User, Keystroke, cryptography

## I. INTRODUCTION

Our dependence on computers and digital platforms has been observed to be overwhelmingly increased to simplify our lives. The use of such automated information systems together has resulted in improved performance of the available networking services in the form of reliability and computational costs.

With such efficient utilization the advances in technology have generated a collective interest in global access to such online platforms. However, at the same time there is a rise in the threats regarding to the security of computers. Usage of advanced methodologies to safeguard the system from attacks and frauds come under the topmost concerning priority of many research scientists. Hence there is a need to generate foolproof measures to prevent such unauthorized access is being worked upon. One such preventive method to give access to individuals by detecting their unique and behavioral pattern is an individual's typing rhythm. This unique typing rhythm tends to become a natural choice for security of computers and is commonly known as Keystroke Dynamics. It is observed that when a person types; the placement of his fingers, applied pressure on the keys and the regularly typed strings appears to be consistent for a specific individual. Hence this concept is used to differentiate between an intruder and a legit user as they will be typing on the keyboard anyway. Therefore, making such kind of typing rhythms easily accessible to track computer activities.

## II. LITERATURE STUDY

Keystroke dynamics is observed to be an emerging field of interest in terms of security that is responsible to validate the authenticity of the system based on users typing rhythm. This authentication process has been overpowering other fields of technology due to the following reasons:

- Easy implementation as only the typing data is required
- Zero hardware requirement
- Low computational cost
- Does not require special permission from the user

Various statistical models have been proposed to implement the studies of keystroke dynamics [1-5]. Classifiers such as [6-10] have been built using machine learning approaches. Apart from this hybrid models are also taken into consideration [11-13] to build the model with accuracy. These models can distinguish and establish a certain pattern of typing frequency and are also capable to access multitudes of data. The argument of keystroke dynamics is observed to differ in certain studies with regards to neuro-physiological behavior of an individual.

This behavior exhibits a user's specific typing pattern [14]. In [15] the concept of keystroke was further classified using MLP and clustering algorithms. Alternatively, [16] designed a system that incorporates the features of applied pressure and typing latency time difference to create and organize a typing pattern for a user. To authenticate individual users, they made use of Artificial Neural Network as classifiers. With an average training time of 0.9094 seconds, the classification rate was 100%. Using the RBFN method, the same concept of pressure and time latency was established in [17]. [18] Made use of MLP with Radial Basis Function and achieved an authentication of 97%.

The contribution in [19] achieved an accuracy of 97.5%. [20] Performed an experiment with a 17-digit password, where each individual typed in the password several times in each session and features such as the size of the finger and timing frames were taken into consideration to reduce the error rate.

Similarly [21] made use of the time intervals between keystrokes utilizing the concepts of MLP and fundamentals of principal component analysis (PCA). This experimentation resulted in an accuracy of 80%. [22] Examined the dataset of keystrokes without the touchscreen features where the data was collected from Android devices and classification algorithms such as SVM, Naïve Bayes were used.

This implementation resulted in a 10% percent increase than a normal authentication process.

**Existing System**

Distortion in keystroke dynamics is not a well researched field, and we did not find many results when doing literature search. However, it measured two typing samples of keystroke dynamics data and used two different measures to compare the samples. However, one method that could be used for detecting distorted timing information is Benford's Law and ZIPF's Law. Benford's Law, or the first-digit law, is an observation in a set of numerical data where the first digit, or leading digit, is more likely to be small. In a balanced distribution of numbers between 1 and 9 there would be exactly 11% for each number to be the leading digit. However, if Benford's Law is obeyed then the change of the leading bits to be small increases.

**DISADVANTAGES OF EXISTING SYSTEM**

- However, the results showed that only latency values from keystroke dynamics timing information followed the law.
- While duration values did not follow the law. More resources used for re-keying because it is done for each join or leave operations.

## III. DEVELOPMENT ON USER AUTHENTICATION BASED ON THE KEYSTROKE DYNAMICS

In order to authenticate a user using keystroke dynamics we need to create a reference average value for each user that will represent, as accurately as possible, their specific typing behaviour. This template varies a lot depending on whether static or continuous authentication is used. In authentication we want to create a template that reflects the typing rhythm that the genuine user uses in order to type the password.

This average value is created based on enrolment samples, where the user would get requested to type their password a number of times. The features, such as duration and latency, are then extracted and the average

typing rhythm is calculated and stored as a reference average. When a user tries to authenticate, the system will check their typing rhythm, which is referred to as a probe, against the reference template and then either reject or accept the user based on a criterion. This criteria for decision making are decided by a threshold which is created by a minimum and maximum average time in milliseconds.

**Advantages:**

- The proposed Keystroke dynamics is based on the assumption that each user can be authenticated because of their unique typing manner. This is because keystroke dynamics performs on a millisecond's precision level meaning it is impossible to accurately recreate the way another user types.
- This is true even for a user who is typing their own password, as they would not be able to type exactly the same way they did last time. Even though the user might type one of the keys or key pairs the same, there are still other keys that they could type in a different way. It is because of this reason keystroke dynamics works as an authentication method.

## Module Design

### Keystroke Timing:

From the KeyDown and KeyUp time of each keystroke we can calculate the duration and latency of a key. The duration of a key is how long the key was held down, this can also be referred to as dwell or hold time. While the latency of a key is the time between releasing one key and pressing another key, and this can sometimes be referred to as flight time. We differentiate between 4 different latencies, given as:

- **pp-latency:** The timing it takes to press down one key and the next key.
- **rr-latency:** The timing it takes to release one key and the next key.
- **rp-latency:** The timing it takes to release one key and press the next key.
- **pr-latency:** The timing it takes to press down one key and release the next key.

In order to get the pp-, rr- and pr-latency we have to use the timing information from duration and rp-latency.We can calculate pp-latency as *latpp* = *durA*+*latrp*, rr-latency as *latr r* = *latrp* + *durB* and pr-latency as *latpr* = *durA* + *latrp* + *durB* where *durA* and *durB* represents the duration of two different keys. The following figure shows the timing values we can extract if a user types the keys *A* and *B*.

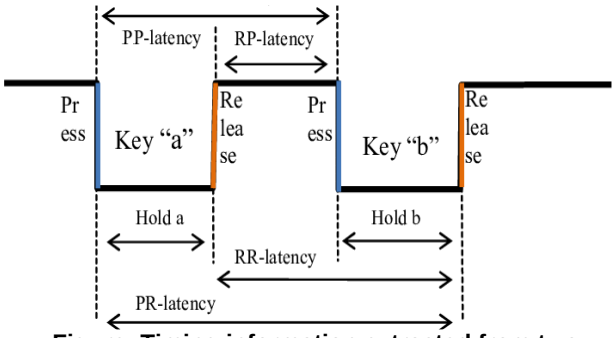



**Figure: Timing information extracted from two keystrokes**

From these latencies, only the rr-latency and rp-latency can be negative. For example, for the rr-latency we can press the shift key, followed by pressing the C key, and then release the C key before releasing the shift key. The same can be said with rp-latency as we can press the next key before releasing the previous key, for example we can press the C key before releasing the shift key.

By using pp-latency, rp-latency and duration for our timing values. The naming of pp-latency refers to press-press-latency, however, we will refer to this as KeyDown-KeyDown latency (DD). While rplatency refers to press-release-latency, which will be called KeyUp-KeyDown latency (UD).

```
timeMAX = 0#slowest
timeMIN = 0#fastest
for #yourForLoop
    if sys.stdin.readline() > timeMAX:
        timeMAX = sys.stdin.readline()
    if sys.stdin.readline() < timeMIN:
        timeMIN = sys.stdin.readline()
```

### Matching Module

In this module, extracted timing features are compared against those stored in the

database, using a matching algorithm from statistical approaches. Ultimately, this process results in a matching score, that represents a similarity measure between the extracted features and the ones previously stored in the database.

**Decision Module**

Given the matching score and a decision threshold, this module is responsible for either accepting or rejecting the claimed user identity, based on the matching score and a predefined threshold.

**Input Design**

The input design is the link between the information system and the user. It comprises the developing specification and procedures for data preparation and those steps are necessary to put transaction data in to a usable form for processing can be achieved by inspecting the computer to read data from a written or printed document or it can occur by having people keying the data directly into the system. The design of input focuses on controlling the amount of input required, controlling the errors, avoiding delay, avoiding extra steps and keeping the process simple. The input is designed in such a way so that it provides security and ease of use with retaining the privacy. Input Design considered the following things:'

- What data should be given as input?
- How the data should be arranged or coded?
- The dialog to guide the operating personnel in providing input.
- Methods for preparing input validations and steps to follow when error occur.

**Objectives**

1.Input Design is the process of converting a user-oriented description of the input into a computer-based system. This design is important to avoid errors in the data input process and show the correct direction to the management for getting correct information from the computerized system.

2. It is achieved by creating user-friendly screens for the data entry to handle large volume of data. The goal of designing input is to make data entry easier and to be free from errors. The data entry screen is designed in such a way that all the data manipulates can be performed. It also provides record viewing facilities.

3.When the data is entered it will check for its validity. Data can be entered with the help of screens. Appropriate messages are provided as when needed so that the user will not be in maize of instant. Thus the objective of input design is to create an input layout that is easy to follow.

**Output Design**

A quality output is one, which meets the requirements of the end user and presents the information clearly. In any system results of processing are communicated to the users and to other system through outputs. In output design it is determined how the information is to be displaced for immediate need and also the hard copy output. It is the most important and direct source information to the user. Efficient and intelligent output design improves the system's relationship to help user decision-making.

1. Designing computer output should proceed in an organized, well thought out manner; the right output must be developed while ensuring that each output element is designed so that people will find the system can use easily and effectively. When analysis design computer output, they should Identify the specific output that is needed to meet the requirements.

2.Select methods for presenting information.

3.Create document, report, or other formats that contain information produced by the system.

The output form of an information system should accomplish one or more of the following objectives.

- Convey information about past activities, current status or projections of the
- Future.
- Signal important events, opportunities, problems, or warnings.
- Trigger an action.
- Confirm an action.

**Database Design**

Databases are normally implemented by using a package called a Data Base Management System (DBMS). Each particular DBMS has somewhat unique characteristics,

and so such, general techniques for the design of database are limited. One of the most useful methods of analyzing the data required by the system for the data dictionary has developed from research into relational database, particularly the work of E.F.Codd. This method of analyzing data is called "Normalization". Unnormalized data are converted into normalized data by three stages. Each stage has a procedure to follow.

**Normalization**

The first stage is normalization is to reduce the data to its first normal form, by removing repeating items showing them as separate records but including in them the key fields of the original record.

The next stage of reduction to the second normal form is to check that the record, which one is first normal form, all the items in each record are entirely dependent on the key of the record. If a data item is not dependent on the key of the record, but on the other data item, then it is removed with its key to form another record. This is done until each record contains data items, which are entirely dependent on the key of their record.

The final stage of the analysis, the reduction of third normal form involves examining each record, which one is in second normal form to see whether any items are mutually dependent. If there are any item there are removed to a separate record leaving one of the items behind in the original record and using that as the key in the newly created record.

**Business Modeling**

The information flow among business function is modeled in a way that answers the following questions: what information drives the business process? What information is generated? What generate it? Where does the information go? Who process it?

**Data Modeling**

The information flow defined as a process of the business modeling is refined into a set of data objects that are needed to support the business. The characteristics (called attributes) of each object are identified and relationships between these objects are defined.

**Process Modeling**

The data objects defined in the data-modeling phase are transformed to achieve the information flow necessary to implement a business function. Processing description is created for addition, modifying, deleting, or retrieving a data object.

## SYSTEM TESTING

Software testing is a critical element if software quality assurance represents the ultimate reviews of specification, design and coding. Testing is vital of the system. Errors can be injected at any stage during development. During testing, the program is executed with correctness. A series of testing are performed for the proposed systems before the system is delivered to the user.

**Unit Testing**

In the unit testing the testing is performed on each module and this module is known as module testing. This testing was carried out during programming state itself. In this testing all the modules working satisfactorily as regard to the expected output from the module.Unit testing is a method by which individual units of source code are tested to determine if they are fit for use. A unit is the smallest testable part of an application. In procedural programming a unit may be an individual function or procedure. Unit tests are created by programmers or occasionally by white box testers.

Unit test cases embody characteristics that are critical to the success of the unit. These characteristics can indicate appropriate/inappropriate use of a unit as well as negative behaviors that are to be trapped by the unit. A unit test case, in and of itself, documents these critical characteristics, although many software development environments do not rely solely upon code to document the product in development. Unit testing provides a sort of living documentation of the system.

**Acceptance Testing**

Acceptance testing is black-box testing performed on a system (e.g. software, lots of manufactured mechanical parts, or batches of chemical products) prior to its delivery. It is also known as functional testing, black-box testing, release acceptance, QA testing, application testing, confidence testing, final testing, validation testing, or factory acceptance testing.

Acceptance testing generally involves running a suite of tests on the completed system. Each individual test, known as a case, exercises a particular operating condition of the user's environment or feature of the system, and will result in a pass or fail, or Boolean, outcome. There is generally no degree of success or failure. The test environment is usually designed to be identical, or as close as possible, to the anticipated user's environment, including extremes of such. These test cases must each be accompanied by test case input data or a formal description of the operational activities (or both) to be performed—intended to thoroughly exercise the specific case—and a formal description of the expected results.

**Types of Acceptance Testing**

Typical types of acceptance testing include the following

**User Acceptance Testing**

This may include factory acceptance testing, i.e. the testing done by factory users before the factory is moved to its own site, after which site acceptance testing may be performed by the users at the site.

**Operational Acceptance Testing**

Also known as operational readiness testing, this refers to the checking done to a system to ensure that processes and procedures are in place to allow the system to be used and maintained.

**Contract and Regulation Acceptance Testing**

In contract acceptance testing, a system is tested against acceptance criteria as documented in a contract, before the system is accepted. In regulation acceptance testing, a system is tested to ensure it meets governmental, legal and safety standards.

**Alpha and beta testing**

Alpha testing takes place at developers' sites, and involves testing of the operational system by internal staff, before it is released to external customers. Beta testing takes place at customers' sites, and involves testing by a group of customers who use the system at their own locations and provide feedback, before the system is released to other customers. The latter is often called “field testing”.

**Integration Testing**

One module can have adverse effect on another such functions when combined may not produce the desired results. Integration testing is a systematic technique for constructing the program structure and conducting test to uncover errors associated with interface. All the modules are combined in this testing step. The entire program is tested as the whole. The errors uncovered are corrected for the next testing step.

**Black Box Testing**

The black box approach is attesting method in which test data are delivered from the functional requirement without regard to the final program structure. Because only functionality of the software is concerned.

In black box testing, only the functionality is determined by observing the outputs to the corresponding input. In this testing various input images are exercised and the output images are compared as required by the content retriever.

**White Box Testing**

White box testing are the software predicates on close examination of procedure details. It provides test cases that exercise specific test for conditions and loops. White box testing was carried out in the order to guarantee that

- All independent parts within a module exercised at least once.
- All logical decision on this true and false side was exercised

**Validation Testing**

Computer input procedures are designed to detect errors in the data at the lower level of detail which is beyond the capability of the control procedures. The validation succeeds when the software functions in the manner that can be reasonably expected by the customer.

## IV. IMPLEMENTATION

The implementation phase focuses how the engineer attempts to develop the system. It also deals with how data are to be structured, how procedural details are to be implemented, how interfaces are characterized, how the design will be translated into programming and hoe the testing will be performed. The methods applied during the development phase will vary but three specific technical tasks should always occur.

- The software design
- Code generation
- Software testing

The system group has changed with responsibility to develop a new system to meet requirements and design and development of new information system. The source of these study facts is variety of users at all level throughout the organization.

**Stage of Development of a System**

- Feasibility assessment
- Requirement analysis
- External assessment
- Architectural design
- Detailed design
- Coding
- Debugging
- Maintenance

**Feasibility Assessment**

In Feasibility this stage problem was defined. Criteria for choosing solution were developed, proposed possible solution, estimated costs and benefits of the system and recommended the course of action to be taken.

**Requirement Analysis**

During requirement analysis high-level requirement like the capabilities of the system must provide in order to solve a problem. Function requirements, performance requirements for the hardware specified during the initial planning were elaborated and made more specific in order to characterize features and the proposed system will incorporate.

**External Design**

External design of any software development involves conceiving, planning out and specifying the externally observable characteristic of the software product. These characteristics include user displays, report formats, external data source and data links and the functional characteristics.

**Internal Design Architectural and Detailed Design**

Internal design involved conceiving, planning out and specifying the internal structure and processing details in order to record the design decisions and to be able to indicate why certain alternations were chosen in preference to others. These phases also include elaboration of the test plans and provide blue prints of implementation, testing and maintenance activities. The product of internal design is architectural structure specification.

The work products of internal design are architectural structure specification, the details of the algorithm, data structure and test plan. In architectural design the conceptual view is refined.

**Detailed Design**

Detailed design involved specifying the algorithmic details concerned with data representation, interconnections among data structures and packaging of the software product. This phase emphasizes more on semantic issues and less synthetic details.

**Coding**

This phase involves actual programming, i.e, transacting detailed design into source code using appropriate programming language.

**Debugging**

This stage was related with removing errors from programs and making them completely error free.

**Maintenance**

During this stage the systems are loaded and put into use. They also get modified accordingly to the requirements of the user. These modifications included making enhancements to system and removing problems.

## V. CONCLUSION AND FUTURE ENHANCEMENT

With the conclusion of this work, and regarding the original research goals, keystroke dynamics can be considered a viable choice to implement strong authentication on the web. We've shown that this technology can be integrated on top of a tradition authentication procedure, taking advantage of user input to seamlessly capture and classify the associated typing behaviour.

Keystroke dynamics presented some notable advantages over other alternatives. No specialized hardware was needed. Biometric data acquisition was completely transparent and multifactor authentication was effectively achieved with a single conscious action from the user. In addition, even simple matching algorithms, as the ones used in this study, can yield reasonable accuracy on authentication.

Although the proposed goals of this work were successfully met, some points could have been improved. An assessment of the system usability should have been done by the users that participated in the study. Furthermore, the target population was greatly biased, mainly computer science students, which are probably better typists than the average individual. Although more complex, applying this study to a more diverse population would have been a closer representation of a real environment.

**SCOPE FOR FUTURE ENHANCEMENT**

For instance, it would be interesting to develop a complementary study, focused on the evaluation of more sophisticated algorithms, in order to further improve matching accuracy.

## VI. BIBLIOGRAPHY

[1] J. V. Monaco and C. C. Tappert, "The partially observable hidden Markov model and its application to keystroke dynamics," Pattern Recognition, 2018.

[2] S. Roy, U. Roy, and D. D. Sinha, "Security enhancement of knowledge-based user authentication through keystroke dynamics," in Proc. MATEC Web Conf., 2016, vol. 57.

[3] K. S. Killourhy. (2012). A scientific understanding of keystroke dynamics. [Online]. Available: http://reports-archive.adm.cs.cmu.edu/anon/2012/CMU-CS-12-100.pd f

[4] A. Messerman, T. Mustafić, S. A. Camtepe, and S. Albayrak, "Continuous and non-intrusive identity verification in real-time environments based on free-text keystroke dynamics," in Proc. Int. Jt. Conf. Biometrics, 2011.

[5] P. S. Teh, B. J. Andrew Teoh, T. S. Ong, and H. F. Neo, "Statistical fusion approach on keystroke dynamics," in Proc. Int. Conf. Signal Image Technol. Internet Based Syst., January 2007, pp. 918–923.

[6] J. Ho and D. K. Kang, "Mini-batch bagging and attribute ranking for accurate user authentication in keystroke dynamics," Pattern Recognit., vol. 70, pp. 139–151, 2017.

[7] S. Maheshwary and V. Pudi, "Mining keystroke timing pattern for user authentication," Lect. Notes Comput. Sci., vol. 10312, pp. 213–227, 2017.

[8] A. Darabseh and A. S. Namin, "On accuracy of classification-based keystroke dynamics for continuous user authentication," in Proc. 2015 Int. Conf. Cyberworlds, 2016, pp. 321–324.

[9] P. H. Pisani, A. C. Lorena, and A. C. P. L. F. de Carvalho, "Adaptive Positive Selection for Keystroke Dynamics," J. Intell. Robot. Syst. Theory Appl., vol. 80, pp. 277–293, 2015.

[10] R. Giot, M. El-Abed, and C. Rosenberger, "Keystroke dynamics with low constraints SVM based passphrase enrollment," in Proc. IEEE 3rd Int. Conf. Biometrics Theory, Appl. Syst., 2009.

[11] H. Mohabeer and S. K. M. Soyjaudah, "Application of predictive coding in neuroevolution," Int. J. Comput. Appl., vol. 114, no. 2, pp. 41–47, 2015.

[12] J. Nisha and R. P. Kumar, "User authentication based on keystroke dynamics analysis," Int. J. Eng. Res. Appl., vol. 4, no. 3, pp. 345–349, 2014.

[13] H. B. K. Bharadi, P. S. Shah, and A. Ambardekar, "Keystroke dynamic analysis using relative entropy & timing sequence

euclidian distance," in Proc. Int. Conf. Work. Emerg. Trends Technol., 2011, p. 220

[14] R. Joyce, G. Gupta, Identity authorization based on keystroke latencies, Commun. ACM 33 (2) (1990) 168–176.

[15] L. K. Maisuria , C. S. Ong and W. K. Lai, " A comparison of artificial neural network and cluster analysis for typing biometrics authentication", International Joint Conference on Neural Network, IJCNN'99, vol.5, pp 3295-3299, 1999.

[16] H. Ali, Wahyudi and Momoh J.E Salami, "Intelligent Keystroke Pressure-Based Typing Biometrics Authentication System by Combining ANN and ANFIS-Based Classifiers", International Colloquium on Signal Processing & Its Applications (CSPA), pp198, 2009.

[17] A. Sulong, Wahyudi and M.U Siddiqi, "Intelligent Keystroke Pressure-Based Typing Biometrics Authentication System by Using Radial Basis Function Network", International Colloquium on Signal Processing & Its Applications (CSPA), pp151, 2009.

[18] N. Capuano, M.Marsella, S.Miranda and S. Salerno, "User Authentication with Neural networks", Univerity of Salerno Italy. http://www.capuano.biz/Papers/EANN_99.pdf

[19] M.S. Obaidat and D.T Macchairolo, "A multilayer neural network system for computer access security", IEEE transactions on Systems, Machine and Cybernetics, vol 24(5), 1994.

[20] R. Gaines, W. Lisowski, S. Press, N. Shapiro, Authentication by keystroke timing: some prelimary results. Rand Rep. R-2560-NSF, Rand Corporation, 1980.

[21] J. Leggett, G. Williams, Verifying identity via keystroke characteristics, Int. J. Man-Mach. Stud. 28 (1) (1988) 67–76.

[22] Trojahn, M., Arndt, F., & Ortmeier, F. (2013). Authentication with keystroke dynamics on Touchscreen Keypads—effect of different N-Graph combinations. In MOBILITY 2013, The third international conference on mobile services, resources, and users (pp. 114–119).

[23] Harun, N., Dlay, S.S., & Woo, W.L. (2010). Performance of keystroke biometrics authentication system using multilayer perceptron neural network (mlp nn). In 7th International symposium on communication systems networks and digital signal processing (CSNDSP' 2010) (pp. 711–714). IEEE, Newcastle upon Tyne.

[24] Antal, M., Szabo, L.Z., & Laszlo, I. (2014). Keystroke dynamics on android platform. In INTER-ENG 2014, 8th international conference inerdisciplinarity in engineering (pp. 114–119). Tirgu Mures: Elsevier

# FUZZY-BASED INTRUSION DETECTION AND PREVENTION OF ATTACKS IN MANETS REQUIRE SUPPORT

**Dr P Prabakaran[1]**
Department of Computer Applications, School of Computing Sciences,
Vels Institute of Science Technology and Advanced Studies, Chennai, Tamil Nadu, India.
**Dr R Shalini[2]**
Department of BCA & IT, School of Computing Sciences,
Vels Institute of Science Technology and Advanced Studies, Chennai, Tamil Nadu India.
**Dr S Nagasundaram[3]**
Department of Computer Applications, School of Computing Sciences,
Vels Institute of Science Technology and Advanced Studies, Chennai, Tamil Nadu, India.

Abstract— ***The topology is chiefly inter-reliant on transmission powers of nodes and locations of mobile nodes. MANETs possess applicability in numerous domains such as military applications, corporate houses or meeting halls, dangerous mission programmer for assistance matters in case of disaster actions. Because of dynamic network topology and absence of centralized organization, the network security turns into a most significant issue in MANETs. In the numerous attacks, the DoS attack, being a multi-layer occurrence shows a foremost character in distracting the network. Generalized Intrusion Detection and Prevention (GIDP) mechanism has been anticipated, which employs the combination of anomaly-based and knowledge-based ID for safeguarding the MANETs from diversity of attacks. Fuzzy rules are then employed for isolating the misbehaving nodes to prevent intrusion.***

Keywords: QoS, MAC, IDT, GIDP and SVM

## I. INTRODUCTION

MANETs are a group of dynamic, self-governing, wireless devices which combine the communication networks, necessitating any backing of endless infrastructure. Based on the position and transmission range, each node in MANETs functions as a router and tends to move randomly and dynamically linked together for forming a network. The topology is chiefly inter-reliant on transmission powers of nodes and locations of mobile nodes. MANETs possess applicability in numerous domains such as military applications, corporate houses or meeting halls, dangerous mission programmer for assistance matters in case of disaster actions like large-scale accidents like war or terrorist attacks, natural disasters and all. The disadvantages pertaining towards quality of service (QoS) in MANETs are the bandwidth restrictions, vibrant and non-predictive topology, and restricted processing and smallest storage at the mobile nodes. MANETs familiarize numerous security issues because of their open communication medium, node movement, absence of centralized security services, and absence of previous security associations. In high-security MANETs, user verification is dangerous in averting illegal users from retrieving or revising network resources.

## II. ATTACKS IN MANETS

The wireless nature and characteristic features of MANETs make them susceptible towards wider diversity of attacks. The attacks on MANETs are categorized into passive and active attacks, depending on whether the normal procedures of networks are disturbed or not. The malicious node(s) are attacked in MANETs using dissimilar means, like sending fake messages numerous times, fake routing data and publicizing fake links to interrupt routing processes.

In MANETs, each and every node functions as both routers and normal nodes. Because of dynamic network topology and absence of centralized organization, the network security turns into a most significant issue in MANETs. In the numerous attacks mentioned above, the DoS

attack, being a multi-layer occurrence, shows a foremost character in distracting the network. DoS could be categorized basically into two main types: routing layer attack and MAC layer attack.

## III. INTRUSION DETECTION SYSTEMS

As MANETs possess distributed architecture and varying topology, traditional centralized monitoring techniques are no longer practicable. In these cases, there becomes necessary developing an intrusion detection system (IDS) (Shakshuki *et al.* 2013). IDS procedures improve the reliability, thereby prolonging the lifetime. IDSs endeavour in detecting and mitigating an attack after it is launched. Numerous monitoring-based intrusion detection techniques (IDTs) have been anticipated. Several nodes monitor transmission actions of other nodes and investigate the packet contents for detecting and mitigating the active attackers (Anantvalee *et al.* 2006). Monitoring-based intrusion detection is not prospective to be precise for ad hoc networks because of their fluctuating noise levels and fluctuating signal propagation features in dissimilar directions. IDT employs additional procedures like trust values for nodes before consideration of the nodes to be suspicious. Even with such supplementary contrivances, monitoring neighbour‘s transmissions is the important procedure that activates the detection process for several IDTs.

## IV. PROBLEM IDENTIFICATION

In Nadeem and Howarth (2011), a generalized intrusion detection and anticipation procedure has been proposed by Adnan Nadeem, where the Adaptive Intrusion Detection and Prevention (AIDP) is employed in detecting the DoS attacks. Subsequently, Generalized Intrusion Detection and Prevention (GIDP) mechanism has been anticipated, which employs the combination of anomaly-based and knowledge-based ID for safeguarding the MANETs from diversity of attacks. It possesses the aptitude in detecting fresh intrusive activities which degrade the Network Performance (NP). Also, they have analyzed numerous attacks and their impacts on NP, and have compared them with the impact on NP of GIDP‘s intrusion response, which is to segregate the intruder from the network. GIDP oversees the network and gathers audit information pertaining to intrusion detection during the network‘s entire lifetime.

The foremost benefit of this scheme is that it protects the MANETs from numerous attacks with reasonable processing overhead. With response to intrusion, GIDP separates the intruding nodes from the network. But, a few nodes in the network are more dangerous than others because of their location in the network topology. As the critical nodes often vary with time, separating the significant routing nodes may distress the number of routes in the network when compared to other nodes. Additionally, re-routing reasons to noteworthy routing disruptions that degrade the NP significantly. Therefore, it could be concluded that the intrusion response of totally isolating the critical nodes could result in net degradation of NP. Furthermore, the anomaly-based detection procedure does not include any typical training procedures like neural networks, support vector machines (SVMs), etc. For testing purposes, they have employed a rule-based approach that cannot produce precise results.

Therefore, in order to evade these issues and to advance the NP, a design of SVM and Fuzzy-based Intrusion Detection and Prevention (SFIDP) technique has been proposed for MANETs attacks.

## V. SVM AND FUZZY-BASED IDS

In attacker model, the wireless channel is considered as apprehensive so that the communicating nodes cannot be trusted. This is because MANETs are frequently deployed in unattended and aggressive environments. Therefore, these nodes are vulnerable to security attacks. The attackers could be categorized as passive attackers that could eavesdrop the messages without troubling the communication, and active attackers that inoculate few packets to these networks together with eavesdropping. These attackers could distract the network communication and burden the traffic.

In MANETs, the nodes are organized in the hostile environments. After organizing the nodes, clusters are formed. The CMs pick the CH based on stability index. The stability index could be defined as the ratio of the time difference of the last signal and the first signal that is received from each neighbours to the difference between the strength of those signals. The node with supreme stability index has been selected as the CH. The CMs send and receive the information through the CH. A node can operate as CH for a pre- defined time. After the expiration of the specified time, the CH

selection process is modified. It is normally assumed that the nodes that are considered as malicious do not partake in the CH selection procedure.

The Support Vector Machine (SVM) could be referred to as supervised learning models with associated learning algorithms. This methodology analyses the data and recognizes the patterns that ate typically utilized for data classification and regression analysis. The following steps entail the corresponding process involved.

Let (*ai*,*bi*)be the training data set where *i* =1, 2,…., *n*

$$nb \in \{-1, 1\}, a \in \{R^d\}$$

*i* *i*

The SVM is considered to be the supervised learning models with related learning mechanisms.

$$q.a - v$$

Where q is the normal vector, and v is the offset.

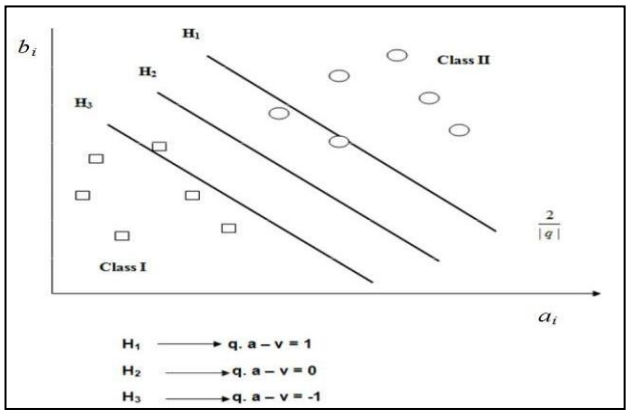
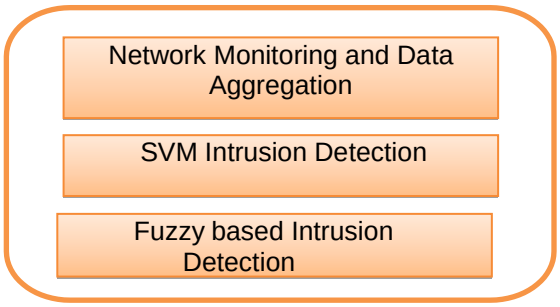



**Figure 5 Maximum margin hyper planes for SVM**

## VI. SFIDP ARCHITECTURE

It includes network monitoring and aggregation of data to perform intrusion detection all through the network. Following network deployment, the SVM is utilized to construct a classifier to discriminate misbehaving nodes from well-behaved nodes. Subsequently, fuzzy rules are used for decision.

**Figure 6. SFIDP architecture**




A clustered MANETs organization has been considered. On the basis of stability index of the deployed nodes, the nodes with supreme stability index are selected as CHs, and the other nodes become CMs. Similarly, secure communication is established among CH and CMs.

During data aggregation, CH collects the Network Characteristics (NC) and NP information and stores in its route cache. The NC information comprises route request message (R_REQ), route request source sequence number ($R\_REQ_{ss}$), route request destination sequence number ($R\_REQ_{ds}$), route reply message (R_REP), route reply destination sequence ($R\_REP_{ds}$), route error message (R_ERR) and time to live value (TTL). The NP information comprises routing protocol overhead (RO), packet delivery ratio (PDR), number of control packets dropped (CPD) and throughput ($\eta$).

## VII. CONCLUSION

The design of an SFIDP against MANETs attacks is presented in this paper. Primarily, the nodes with supreme stability index are selected as CHs and other nodes turn into CMs. Secure communication has been established between CH and CMs. Subsequently, SVM is utilized to discriminate misbehaving nodes from well-behaved nodes. Fuzzy rules are then employed for isolating the misbehaving nodes to prevent intrusion. From the simulation results, it could be seen that the proposed technique improves the NP by 24% on PDR, 76% on packet drop, 17% on miss detection ratio and 21% on false positives while varying the attackers.

## REFERENCES

[1] Akshat Jain, Shekher Singh Sengar 2013, ‗Colluding Black Holes Detection in MANET', International Journal of Engineering Research and Technology, vol. 2.

[2] Anjum, Farooq, and Petros Mouchtaris. Security for wireless ad hoc networks. John Wiley & Sons, 2007.

[3] Axelsson, S. ‗Intrusion Detection Systems: A Taxonomy and Survey', Technical Report No 99-15, Department of Computer Engineering, Chalmers University of Technology, Sweden, 2000.

[4] Bing Wu, Jianmin Chen, Jie Wu, Mihaela Cardei, ‗A Survey on Attacks and Countermeasures in Mobile Ad Hoc Networks', Wireless/Mobile Network Security, Springer, 2006.

[5] Bu, Shengrong, et al. ‗Distributed combined authentication and intrusion detection with data fusion in high-security mobile ad hoc networks.' IEEE transactions on vehicular technology 60.3 (2011).

[6] Chi-Ho Tsang, , Sam K wong, and Hanli Wang. Genetic-fuzzy rule mining approach and evaluation of feature selection techniques for anomaly intrusion detection.' Pattern Recognition 40.9.

[7] Da Silva, APR, Martins, MHT, Rocha, BPS, Loureiro, AAF, Ruiz, LB & Wong, HC, Decentralized Intrusion Detection in Wireless Sensor Networks', InProceedings of the 1st ACM International Workshop on Quality of Service and Security in Wireless and Mobile Networks

[8] Dijiang Huang & Deep Medhi 2008, ‗A secure group key management scheme for hierarchical mobile ad hoc networks', Ad Hoc Networks, vol. 6.

[9] Ertaul, L & Ibrahim, D 2009, ‗Evaluation of Secure Routing Protocols in Mobile Ad Hoc Networks (MANETs)', International Conference on Security and Management.

[10] Garcia, Mario A, and Tung Trinh. ‗Detecting Simulated Attacks in Computer Networks Using Resilient Propagation Artificial Neural Networks.' Polibits 51 (2015).

[11] Gimer Cervera, Michel Barbeau, Joaquin Garcia-Alfaro & Evangelos Kranakis, 2012, Preventing the cluster formation attack against the hierarchical OLSR protocol', Foundations and Practice of Security- Springer Berlin Heidelberg.

[12] Haimabati Dey & Raja Datta 2012, Transmission-efficient group-key generation in large dynamic MANET environments', Third International Conference on Emerging Applications of Information Technology (EAIT), 2012.

# EMPLOYING A UAV UNIT FOR THE DATA DISSEMINATION PROTOCOL IN VANET WITH THE NAMED DATA ARCHITECTURE

**P Harish Pandi[1]**
III B.Sc.IT, Department of Information Technology,
Sri Krishna Adithya College of Arts & Science,
Coimbatore, Tamil Nadu, India.
**Dr K Santhiya[2]**
Assistant Professor, Department of Information Technology,
Sri Krishna Adithya College of Arts & Science,
Coimbatore, Tamil Nadu, India

**Abstract**
**The demand for more integration density, higher bandwidth, and lower power are always rising. Due to the rapid rise in the use of resources used on a daily basis during the past few years, there has been an increase in the demand for energy. Energy is produced from fossil fuels, which eventually run out. We therefore have a pressing need for an energy-saving system. The answer to completing the duties required for the creation of this system is the Internet of Things (IoT). In this study, an Internet of Things (IoT)-based system that allows different devices to be connected over the Internet in order to save energy is presented. As a result, a peer's reputation motivates it to cooperate and desist from malicious activities. The cryptographic protocol is coupled with self-certification and cryptographic mechanisms for identity management and countering Sybil attack. It illustrates the security and the efficiency of the system analytically andby means of simulations in a completely decentralized Gnutella-like P2P network. The model also evaluates the effect of control strategies like node quarantine on stifling the spread of malware. The model is then extended to consider the impact of P2P networks on the malware spread in networks of smart cell phones.**

Keywords: TTL, P2P, UAV, NDN, RSU, QIP

## I. INTRODUCTION

The use of peer-to-peer (P2P) networks as a vehicle to spread malware offers some important advantages over worms that spread byscanning for vulnerable hosts. This is primarily due to the methodology employed by the peers to search for content. For instance, in decentralized P2P architectures such as Gnutella where search is done by flooding the network, a peer forwards the query to it's immediate neighbors and the process is repeated until a specified threshold time-to-live, TTL, is reached. Here TTL is the threshold representing the number of overlay links that a search query travels. A relevant example here is the Mandragore worm that affected Gnutella users. Having infected a host in the network, the wormcloaks itself for other Gnutella users. Every time a Gnutella user searches for media files in the infected computer, the virus always appears as an answer to request, leading the user to believe that it is the file the user searched.

## II. LITERATURE STUDY

System analysis will be performed to determine if it is flexible to design information based on policies and plans of organization and on user requirements and to eliminate theweakness of present system. This chapter discusses the existing system, proposed system andhighlights of the system requirements.

Reputation Models (EXISTING MODELS)

Resnick et al defines the reputation system as "a system that collects, distributes, and aggregates feedback about consumer's past behavior."

Resnick et al. explain the problem of pseudo spoofing. Pseudo spoofing is “the use of multiple pseudonyms in a system by the same real-life entity”.

**Drawbacks of Existing System**

- The disadvantage is that any entity can discard a handle or a pseudonym with which a bad reputation is associated and join the system as a new user, under a new pseudonym. This can possibly nullify the usefulness of a reputation system, which assigns reputations to handles.
- The authors also advocate that the newcomers should pay their dues in order to mitigate the effect of pseudo spoofing. In other words, the newcomers should not only use the services of the system but should also contribute to the system as per the system guidelines.

## III. DEVELOPMENT A UAV UNIT FOR EMPLOYING A UAV UNIT FOR THE DATA DISSEMINATION PROTOCOL IN VANET WITH THE NAMED DATA ARCHITECTURE

A more ambitious approach to protect the P2P network without using any central component, and thereby harnessing the full benefits of the P2P network. The reputations of the peers are used to determine whether a peer is a malicious peer or a good peer. Once detected, the malicious peers are ostracized from the network as the good peers do not perform any transactions with the malicious peers. Expulsion of malicious peers from the network significantly reduces the volume of malicious activities.

All peers in the P2P network are identified by identity certificates (aka identity). The reputation of a given peer is attached to its identity. The identity certificates are generated using self-certification, and all peers maintain their own (and hence trusted) certificate authority which issues the identity certificate(s) to the peer. Each peer owns the reputation information pertaining to all its past transactions with other peers in the network, and stores it locally.

A two-party cryptographic protocol not only protects the reputation information from its owner, but also facilitates secure exchange ofreputation information between the two peers participating in a transaction.

### 2.2.2 BENEFITS OF PROPOSED SYSTEM

- This infrastructure reduces the percentage of malicious transactions in the network.
- It also generates significantly less network traffic as compared to other reputation-based security solutions for P2P networks.

### MODULE DESIGN

### INPUT DESIGN

The input design is the link between the information system and the user. It comprises the developing specification and procedures for data preparation and those steps are necessary to put transaction data in to a usable form for processing can be achieved by inspecting the computer to read data from a written or printed document or it can occur by having people keying the data directly into the system. The design of input focuses on controlling the amount of input required, controlling the errors, avoiding delay, avoiding extra steps and keeping the process simple. The input is designed in such a way so that it provides security and ease of use with retaining the privacy. Input Design considered the followingthings:’

- What data should be given as input?
- How the data should be arranged or coded?
- The dialog to guide the operatingpersonnel in providing input.
- Methods for preparing input validations and steps to follow when error occur.

### OBJECTIVES

1. Input Design is the process of converting a user-oriented description of the input into a computer-based system. This design is important to avoid errors in the data input process and show the correct direction to the management for getting correct information from the computerized system.
2.

3. It is achieved by creating user-friendly screens for the data entry to handle large volume of data. The goal of designing input is to make data entry easier and to be free from errors. The data entry screen is designed in such a way that all the data manipulates can be performed. It also provides record viewing facilities.

4. When the data is entered it will check for its validity. Data can be entered with the help of screens. Appropriate messages are provided as when needed so that the user will not be in maize of instant. Thus the objective of input design is to create an input layout that is easy to follow.

**OUTPUT DESIGN**

A quality output is one, which meets the requirements of the end user and presents the information clearly. In any system results of processing are communicated to the users and to other system through outputs. In output design it is determined how the information is to be displaced for immediate need and also the hard copy output. It is the most important and direct source information to the user. Efficient and intelligent output design improves the system's relationship to help user decision-making.

1. Designing computer output should proceed in an organized, well thought out manner; the right output must be developed while ensuring that each output element is designed so that people will find the system can use easily and effectively. When analysis design computer output, they should Identify the specific output that is needed to meet the requirements.

2. Select methods for presenting information.

3. Create document, report, or other formats that contain information produced by the system. The output form of an information system should accomplish one or more of the following objectives.

- Convey information about past activities, current status or projections of the
- Future.
- Signal important events, opportunities, problems, or warnings.
- Trigger an action.
- Confirm an action.

## IV. RESULT AND DISCUSSION

The implementation phase focuses how the engineer attempts to develop the system. It also deals with how data are to be structured, how procedural details are to be implemented, how interfaces are characterized, how the design will be translated into programming and hoe the testing will be performed. The methods applied during the development phase will vary but three specific technical tasks should always occur

- The software design
- Code generation
- Software testing

The system group has changed with responsibility to develop a new system to meet requirements and design and development of new information system. The source of these study facts is variety of users at all level throughout the organization.

**Stage of Development of a System**

- Feasibility assessment
- Requirement analysis
- External assessment
- Architectural design
- Detailed design
- Coding
- Debugging
- Maintenance

**Feasibility Assessment**

In Feasibility this stage problem was defined. Criteria for choosing solution were developed, proposed possible solution, estimated costs and benefits of the system and recommended the course of action to be taken. Requirement Analysis During requirement analysis high-level requirement like the capabilities of the system must provide in order to solve a problem. Function requirements, performance requirements for the hardware specified during the initial planning were elaborated and made more specific in order to characterize features and the proposed system will incorporate.

**External Design**

External design of any software development involves conceiving, planning out and specifying the externally observable characteristic of the software product. These characteristics include user displays, report

formats, external data source and data links and the functional characteristics. Internal Design Architectural and Detailed Design Internal design involved conceiving, planning out and specifying the internal structure and processing details in order to record the design decisions and to be able to indicate why certain alternations were chosen in preference to others. These phases also include elaboration of the test plans and provide blue prints of implementation, testing and maintenance activities. The product of internal design is architectural structure specification. The work products of internal design are architectural structure specification, the details of the algorithm, data structure and test plan. In architectural design the conceptual view is refined.

**Detailed Design**

Detailed design involved specifying the algorithmic details concerned with data representation, interconnections among data structures and packaging of the software product. This phase emphasizes more on semantic issues and less synthetic details.

**Coding**

This phase involves actual programming, i.e, transacting detailed design into source code using appropriate programminglanguage.

**Debugging**

This stage was related with removing errors from programs and making them completely
error free.

**Maintenance**

During this stage the systems are loaded and put into use. They also get modified accordingly to the requirements of the user. These modifications included making enhancements to system and removing problems.

## V. CONCLUSION AND FUTURE ENHANCEMENT

This project presents self-certification, an identity management mechanism, reputation model, and a cryptographic protocol that facilitates generation of global reputation data ina P2P network for cyber analysis, in order to expedite detection of rogues. A reputationsystem for peer-to-peer networks can be thwarted by a consortium of malicious nodes. Such a group can maliciously raise the reputation of one or more members of the group. There is no known method to protect a reputation system against liar farms and the absence of a third trusted partymakes the problem of liar farms even more difficult. In this system we developed an analytic model to understand the dynamics of malware spread in P2P networks for cyber analysis. The need for an analytic framework incorporating user characteristics (e.g., offline to online transitional behavior) and communication patterns (e.g., the average neighborhood size) was put forth by quantifying their influence on the basic reproduction ratio.

**SCOPE FOR FUTURE ENHANCEMENT**

This system can be extended to encapsulate the reputations of both the provider and the requester. In addition, instead of generic number values, the reputation values can be modified in accordance with the context of the reputation.

## REFERENCES

[1] Clip2, “The Gnutella Protocol Specification v0.4,” http://www.clip2.com/GnutellaProtocol04.pdf, Mar. 2001.30

[2] E. Damiani, D. di Vimercati, S. Paraboschi, P. Samarati, and F. Violante, “A Reputation-Based Approach for Choosing Reliable Resources in Peer-to- Peer Networks,” Proc. ACM Conf Computer and Comm. Security (CCS), pp. 207- 216, Nov. 2002.

[3] X. Yang and G. de Veciana, “Service Capacity in Peer-to-Peer Networks,” Proc. IEEE INFOCOM ’04, pp. 1-11, Mar. 2004.

[4] D. Qiu and R. Srikant, “Modeling and Performance Analysis of BitTorrent- Like Peer-to-Peer Networks,” Proc. ACM SIGCOMM, Aug. 2004.

[5] J. Mundinger, R. Weber, and G.Weiss, “Optimal Scheduling of Peer-to-Peer File Dissemination,” J. Scheduling, vol. 11, pp. 105-120, 2007.

[6] A. Bose and K. Shin, “On Capturing Malware Dynamics in Mobile Power- Law Networks,” Proc. ACM Int’l Conf. Security and Privacy in Comm. Networks (SecureComm), pp. 1-10, Sept.2008.

[7] L. Zhou, L. Zhang, F. McSherry, N. Immorlica, M. Costa, and S. Chien, “A First Look at Peer-to-Peer Worms: Threats and Defenses,” Int’l Workshop Peer-To-Peer Systems, Feb. 2005.

[8] F. Wang, Y. Dong, J. Song, and J. Gu, “On the Performance of Passive Worms over Unstructured P2P Networks,” Proc. Int’l Conf. Intelligent Networks and Intelligent Systems (ICINIS), pp. 164-167, Nov. 2009.

[9] R. Thommes and M. Coates, “Epidemiological Models of Peer-to-Peer Viruses and Pollution,” Proc. IEEE INFOCOM ’06, Apr. 2006.

[10] J. Schafer and K. Malinka, “Security in Peer-to-Peer Networks: Empiric Model of File Diffusion in BitTorrent,” Proc. IEEE Int’l Conf. Internet Monitoring and Protection (ICIMP ’09), pp. 39-44, May 2009.

[11] J. Luo, B. Xiao, G. Liu, Q. Xiao, and S. Zhou, “Modeling and Analysis of Self-Stopping BT Worms Using Dynamic Hit List in P2P Networks,” Proc. IEEE Int’l Symp. Parallel and Distributed Processing (IPDPS ’09), May 2009.

[12] W. Yu, S. Chellappan, X. Wang, and D. Xuan, “Peer-to-Peer System-Based Active Worm Attacks:Modeling, Analysis and Defense,” Computer Comm., vol. 31, no. 17, pp. 4005- 4017,Nov. 2008.

# SHARING SECRET WITH MULTI PARTY USING EFFICIENT VERIFIABLE THRESHOLD ALGORITHM

**R Abhilash[1]**
III B.Sc.IT, Department of Information Technology,
Sri Krishna Adithya College of Arts & Science,
Coimbatore, Tamil Nadu, India
**Dr B P Sreejith Vignesh[2]**
Head of the Department, Department of Information Technology,
Sri Krishna Adithya College of Arts & Science,
Coimbatore, Tamil Nadu, India

**Abstract - One of the most significant cryptographic primitives used for data outsourcing is secret sharing. One of the well-known secret sharing strategies in cryptography is the threshold-based scheme. Multiple secret sharing schemes can significantly increase the efficiency of secret sharing while a single secret sharing scheme has low efficiency. Numerous secrets are shared among participants in a single sharing process using our efficient verified threshold multi-secret sharing system, which has numerous steps. A few weak subsets of participants can collectively uncover these secrets in this way. Because it is verifiable, each member may confirm their individual share.**

**Our proposed scheme can share secrets to multiple participants and each has its own threshold access structure. Additionally, each participant only keeps one share, yet during the recovery process they are able to rebuild all secrets. The quantity of the participant's share is equal to that of each secret in this multi-stage-use secret sharing method, and the amount of information released meets the optimal lower bound. This multi-secret sharing system is more computationally secure and practical than the earlier verified (t; n)-threshold scheme.**

Keyword : threshold, RSU, Vanet

## I. INTRODUCTION

Intelligent Transportation Systems (ITS) are a set of innovative technologies which are expected to enable a wide range of road traffic management and safety services. As a key part of ITS, Vehicular Ad hoc Networks (VANETs) are expected to play a major role in improving road safety and making commuting a pleasant experience. VANETs rely on inter-vehicular communication (IVC) for establishing an efficient exchange of information among vehicles, roadside units (RSU), and possibly nearby pedestrians. VANET routing protocols allow vehicles to organize themselves in mobile communication networks, therefore allowing ITS to deploy various services without relying on fixed communication infrastructures.

However, due to the high mobility of vehicles and the dynamically changing topology of such networks, routing data packets between vehicles turns out to be a challenging task. Therefore, one of the main challenges of ITS in successfully deploying road and traffic services is to ensure that control data and application data are routed and delivered with acceptable reliability. That is, one of the major focuses of the research community concerned with VANET applications is to develop routing mechanisms adapted to these specific VANET scenarios.

Several methods have been developed in the literature for robust routing 20 protocols which address the issues of high mobility in VANET and of the unpredictable movements of the vehicles caused by the frequent change in the network topology. In an urban environment, vehicles usually are not uniformly distributed around the city resulting in network fragmentation that reduces the efficiency and reliability of communication between vehicles. In addition, the existence of obstacles such as buildings and landmarks can cause disruption to the radio signal resulting in frequent communication failures between vehicles, even when within acceptable ranges of each other.

Most of the proposed routing techniques designed for urban environments disregard the possible obstacles that can adversely impact the quality of data routing. In urban environments, there are many types of mobile vehicles such as trains, cars, airplanes and even pedestrians that can be considered as part of a VANET.

All of these entities can serve as a relay to support routing in VANETs and forwarding data packets. However, these relays are localized in predefined areas and sometimes cannot participate in relaying data between vehicles. During the last few years, the number of Unmanned Aerial Vehicles (UAVs) that fly over urban areas has increased significantly. In the near future, UAVs are expected to be deployed to support a plethora of applications. Equipped with communication and navigation accessories, we will not wait too long before a need to organize UAVs in an ad hoc network of flying vehicles communicating with each other and exchanging data. As a result, it is only natural to expect a network of UAVs extending VANETs in an integrated way and cooperating with vehicles on the ground. UAVs are mobile flying vehicles whose movements are generally defined beforehand.

In this work, we consider the cooperation of an ad hoc network of flying UAVs with existing VANETs through heterogeneous communications. The purpose of this research is to take advantage of the UAVs to re-establish communication links in case of disconnections due to obstacles. UAVs are used here for both delivering data packets and re-linking disconnected road segments. A secret sharing scheme is a method of distributing a secret amongst a set of participants by giving each participant a share in such a way that only certain specified subsets of participants can reconstruct the secret from a pooling of their shares.

Secret sharing schemes are highly versatile cryptographic primitives and, as a result, have been employed in a vast range of different applications including protection of cryptographic keys, access control, key recovery mechanisms, electronic voting, distributed certificate authorities, online auctions and secure multiparty computation. They are also objects of inherent mathematical interest and have also been researched as such.

The access structure of a secret sharing scheme normally partitions the set of all subsets of participants into authorized sets who are able to recover the secret and unauthorized sets who cannot. (Some schemes feature a third class of subsets who are neither authorized or unauthorized.) The two fundamental properties of a secret sharing scheme are thus:

1. Privacy: Unauthorized subsets of participants should be prevented from learning the secret.
2. Recoverability: Authorized subsets of participants should be able to recover the secret by pooling their shares.

Secret sharing schemes also involve two functionalities that are, in many cases, carried out by a dedicated entity. The dealer is normally responsible for generating system parameters, generating the secret, creating initial shares and sending initial shares to participants. The combiner is responsible for pooling shares and reconstructing the secret. The dealer is normally a fully trusted third party, while the combiner is often left

unspecified (but can be a third party or even one of the participants).

## II. LITERATURE STUDY

The earliest secret sharing scheme is the threshold-based secret sharing scheme, which was proposed by Bailey and Shamir, respectively. In general, the system of secret sharing assumes that both dealer and participants are honest, but this is unrealistic in real life. Therefore, these systems cannot effectively stop the dealer from cheating the participants (the dealer distributes fake shares to some participants) and stop the participants from deceiving the other participants (some participants distribute false shares when reconstructing to the secret).

In order to solve these problems, several verifiable secret sharing schemes have been proposed. In these schemes, a verification algorithm is added and participants contest whether the dealer distributes the false shares to restore the secret and moreover each participant can test whether other participants have provided a valid share. Several effective multi-secret sharing schemes have been proposed based on the Shamir threshold scheme.

Drawbacks of Existing System

- The shares are selected by the dealer and distributed to the participants through a secure channel.
- Since a secure channel needs to be established between the participant and the dealer, the scheme has relatively high requirements of the system.
- In a secret sharing scheme, dishonest participants and the dealer are likely to cheat the other participants during execution.

## III. DEVELOPMENT ON SHARING SECRET WITH MULTI PARTY USING EFFICIENT VERIFIABLE THRESHOLD ALGORITHM

To alleviate these concerns, we improve our scheme with verifiability properties, such as variable and publicly verifiable secret sharing. Verifiability stops the dealer from sharing wrong shares and hence public verifiability forces participants to submit their sub-shares correctly.

In the dynamic secret sharing schemes, the main concern is the updates of secret and share, as well as the addition of new individuals or the deletion of the participant, which does not involve changes of the threshold value tin the multi-secret sharing scheme, we propose a multi-stage secret sharing scheme, in which each secret corresponds to an independent threshold. When reconstructing multiple secrets with different threshold values, the secret holder needs to disclose part of the information in order to keep the number of shares saved by the participants as low as possible.

In our proposed scheme, the participants' share is generated independently and randomly which has nothing to do with a single secret. All the secrets are also generated independently and randomly. For each secret, the secret holder publishes its own independent public information in advance, and the corresponding threshold value of each secret is determined by its public information.

### 2.2.2 BENEFITS OF PROPOSED SYSTEM

- Our protocol provides a 2FA security
- Our protocol supports fine-grained attribute-based access which provides a great flexibility for the system to set different access policies according to different scenarios. At the same time, the privacy of the user is also preserved.

### MODULE DESIGN

### DESCRIPTION OF MODULES

TRUSTEE

It is responsible for generating all system parameters and initializes the security device.

AUTHORITY

It is responsible to generate a user secret key for each user according to their attributes.

USER

It is the player that makes authentication with the cloud server. Each user has a secret key issued by the attribute-issuing authority and a security device initialized by the trustee.

CLOUD SERVICE PROVIDER

It provides services to anonymous authorized users. It interacts with the user during the authentication process

## IV. RESULT AND DISCUSSION

The implementation phase focuses on how the engineer attempts to develop the system. It also deals with how data are to be structured, how procedural details are to be implemented, how interfaces are characterized, how the design will be translated into programming and how the testing will be performed. The methods applied during the development phase will vary but three specific technical tasks should always occur.

- The software designs
- Code generation
- Software testing

The system group has changed with responsibility to develop a new system to meet requirements and design and development of a new information system. The source of these study facts is a variety of users at all levels throughout the organization. Stage of Development of a System

- Feasibility assessment
- Requirement analysis
- External assessment
- Architectural design
- Detailed design
- Coding
- Debugging
- Maintenance

Feasibility Assessment

In Feasibility this stage problem was defined. Criteria for choosing a solution were developed, proposed possible solutions, estimated costs and benefits of the system and recommended the course of action to be taken.

Requirement Analysis

During requirement analysis high-level requirements like the capabilities of the system must be provided in order to solve a problem. Function requirements, performance requirements for the hardware specified during the initial planning were elaborated and made more specific in order to characterize features the proposed system will incorporate.

External Design

External design of any software development involves conceiving, planning out and specifying the externally observable characteristic of the software product. These characteristics include user displays, report formats, external data source and data links and the functional characteristics.

Internal Design Architectural and Detailed Design

Internal design involved conceiving, planning out and specifying the internal structure and processing details in order to record the design decisions and to be able to indicate why certain alterations were chosen in preference to others. These phases also include elaboration of the test plans and provide blueprints of implementation, testing and maintenance activities. The product of internal design is architectural structure specification. The work products of internal design are architectural structure specification, the details of the algorithm, data structure and test plan. In architectural design the conceptual view is refined.

Detailed Design

Detailed design involved specifying the algorithmic details concerned with data representation, interconnections among data structures and packaging of the software product. This phase emphasizes more on semantic issues and less synthetic details.

Coding

This phase involves actual programming, i.e., transacting detailed design into source code using appropriate programming language.

Debugging

This stage was related to removing errors from programs and making them completely error free.

Maintenance

During this stage the systems are loaded and put into use. They also get modified according to the requirements of the user. These modifications included making enhancements to the system and removing problems.

## V. CONCLUSION AND FUTURE ENHANCEMENT

This project will explore a new verifiable threshold algorithm multi-secret sharing scheme and outsource the process of the secret reconstruction to the cloud service provider. Our proposed scheme enjoys four advantages such as the multiple secret sharing, the privacy of the shared secrets, the efficient secret reconstruction, and the efficient verification of the hare and the returned result. The last two properties are thanking to the outsourcing computation of the secret reconstruction and the share verification. Furthermore, our proposed system supports the participants to recover the desired return as well as to identify the hacker.

Scope for Future Enhancement

We can also expect further developments in the formalization of models for such schemes, as only robust secret sharing schemes have been set in a framework compatible with much of the recent theoretical formalization of other types of cryptographic primitive. There would thus seem to remain some room in this area for further application of interesting mathematical techniques to provide secret sharing schemes with the capability of coping with sophisticated adversarial behavior.

## REFERENCES

1. M. P. Jhawar, A. Venkatachalam, and R. Safavid-Naini, ``Parlier-based publicly verifiable (non-interactive) secret sharing,'' Des., Codes Cryptor., vol. 73, no. 2, pp. 529-546, 2014.
2. Y. Tian, C. Peng, and J. Ma, ``Publicly verifiable secret sharing schemes using bilinear pairings,'' Int. J. Newt. Secure., vol. 14, no. 3, pp. 142-148, 2012.
3. T.-Y. Wu and Y.-M. Tseng, ``Publicly verifiable multi-secret sharing scheme from bilinear pairings,'' IET Inf. Secure., vol. 7, no. 3, pp. 239-246, Sep. 2013.
4. S. Mashhad, ``Secure publicly verifiable and proactive secret sharing schemes with general access structure,'' Inf. Sci., vol. 378, pp. 99-108, Feb. 2017.
5. M. H. DeKorte and R. Gha semi, ``A lightweight public verifiable multi secret sharing scheme using short integer solution,'' Wireless Pers. Common., vol. 91, no. 3, pp. 1459-1469, 2016.
6. Q. Peng and Y. Tian, ``Publicly verifiable secret sharing scheme and its application with almost optimal information rate,'' Secure. Common. Newt., vol. 9, no. 18, pp. 6227-6238, 2016.
7. Q. Peng and Y. Tian, ``A publicly verifiable secret sharing scheme based on multilinear Diffie Hellman assumption,'' Int. J. Newt. Secure., vol. 18, no. 6, pp. 1192-1200, 2016.
8. E. Zhang, J. Peng, and M. Li, ``Outsourcing secret sharing scheme based on homomorphism encryption,'' IET Inf. Secure., vol. 12, no. 1, pp. 9499, 2018.
9. H. Zhang, J. Yu, C. Tian, P. Zhao, G. Xu, and J. Lin, ``Cloud storage for electronic health records based on secret sharing with variable reconstruction outsourcing,'' IEEE Access, vol. 6, pp. 40713-40722, 2018.

# MAINTAINABILITY METRIC MODEL FOR COMPONENT-BASED SOFTWARE EMPLOYING SOFT COMPUTING METHODOLOGIES

**Praveen Kumar**
#Research Scholar, Department of Computer Science & Application,
Dr. A.P.J. Abdul Kalam University, Indore, MP, India
**Dr. Kailash Patidar**
* Department of Computer Science & Application,
Dr. A.P.J. Abdul Kalam University, Indore, MP, India

**Abstract- Due to their interdependencies and complexity, component-based software systems are difficult to maintain. Using soft computing techniques, this research investigates the creation of a maintainability metric model for component- based software. The difficulties in assessing and enhancing the maintainability of component-based software may be mitigated by the use of soft computing methods like fuzzy logic, evolutionary algorithms, and neural networks. In the present study, two assessment models for software maintainability are provided to aid software designers in creating such programmes. Fuzzy logic is used for the first 16 DTCM maintainability metric implementations. DRCE maintainability model employs Fuzzy logic, neural networks, and the ANFIS (Neuro-Fuzzy) approach to enhance the efficacy. Values from the model and those from other researchers are compared and contrasted. The model shows an increase of about 6% to 10%, depending on the inputs.**

Keywords: Metric, Component-Based System (CBS), Soft computing

## I. INTRODUCTION

These days, embedded systems are employed in a wide variety of contexts, including but not limited to satellites, factories, cars, hospitals, and other medical devices (Guan and Offutt 2015). As a result, system engineers are facing increasing difficulty in developing this system due to the many quality standards that must be met, including those for the system's dependability, availability, maintainability, security, safety, independence, and portability (Garousi et al., 2018). The success of the project and the happiness of the customers depend on meeting these quality criteria. Therefore, there was a rising need to forecast and assess these quality parameters for early issue detection.

Measures of quality attributes are often made using the tried-and-true techniques of the past, such as Markov models, Queueing networks, Fault trees, Stochastic Petri Nets, and monolithic models (Cai et al., 2017). However, a component- based approach is required in designing the sophisticated embedded system (Chatzipetrou et al., 2019) due to rising system scalability and complexity.The capacity to detect and fix bugs within a predetermined time frame and set of conditions is known as maintainability. This is an example of corrective maintenance including resource change. Fixing software flaws and implementing new fixes in the area of Performance may significantly boost the program's productivity.

In software engineering, maintainability refers to how easily bugs can be fixed, which has become more important as more and more software relies on it. Adaptive upkeep: Many businesses make use of correction services. Insufficient practises in remaining useful modified or changing environment are a common cause of software issues after delivery.With the fast advancement of component technology, the discipline of software engineering has adopted a new methodology known as Component-Based Software (CBS), which places an emphasis on the aggregation of componentsinto large software systems. This method has several benefits, including increased efficiency, higher quality, more reusability, lower maintenance costs, and faster time to market. Individual component reliability calculations and an analysis of the connectivity approach may be used to forecast system reliability.

Failure predicting approaches that statistically assess system reliability are at the heart of CBS reliability forecasting. Dimensions, Difficulty, System Customers, Operating Preventative upkeep: After software has been delivered, it is repaired in such a way that any latent problems in the programme's Facilities, Maintenance Team, and Documentation

are found and fixed (Vale et al. 2016)Component-based development has become more popular in the field of software engineering because it improves reusability, maintainability, and productivity. However, there is still a difficulty in making sure component-based software systems can be maintained. The use of soft computing approaches to this problem has shown encouraging results. Using soft computing techniques, this research hopes to investigate the idea of a maintainability metric model for component-based software.

## COMPONENT-BASED SOFTWARE DEVELOPMENT

Component-based software development (CBSD) is an approach that emphasizes the assembly of software systems from reusable software components. This methodology allows for modular design, reusability, and faster development cycles. However, the maintenance of component-based software systems poses unique challenges due to the complexity and interdependencies of the components.

### Importance of Maintainability Metrics

The ease of software maintenance and modification may be measured using maintainability metrics. They provide a numerical evaluation of software systems' maintainability characteristics, helping programmers see future maintenance challenges and make educated choices. However, the complexity of component-based software systems is frequently not well captured by conventional measures of maintainability.

Soft Computing Methodologies:

Fuzzy logic, evolutionary algorithms, and neural networks are just a few examples of the soft computing techniques that provide malleable and adaptable ways to addressing problems. These methods are well-suited for tackling the issues connected with maintainability in component-based software systems because of their ability to deal with imprecise, ambiguous, and incomplete data.

Developing Maintainability MetricModel:

The following procedures may be used to create a maintainability metric model for component-based software using soft computing techniques:

Identify Relevant Metrics:

The first step in improving the maintainability of software systems built from components is to establish a set of metrics that accurately reflect this quality. Complexity, coherence, coupling, documentation, and dependency graphs among other things might be used as such measures.

Fuzzy Logic for Metric Evaluation:

Due to their inherent ambiguity and imprecision, maintainability measurements might benefit fromthe use of fuzzy logic. Each metric's linguistic phrases may be represented by a fuzzy set, and fuzzy rules can be developed to assess the metricsaccording to their linguistic values.

Genetic Algorithms for Optimization:

It is possible to use genetic algorithms to get the optimal value for the maintainability measure. Genetic algorithms may be used to optimise models by repeatedly refining a population of candidate solutions. Models' fitness is measured, and then genetic operators like selection, crossover, and mutation are used to breed better versions of the originals.

Neural Networks for Prediction:

Based on past information, neural networks may foresee problems with maintainability. A neural network may be taught to recognise trends and establish relationships between various metrics andthe incidence of maintenance issues by being trained on a dataset of maintenance records. With this foresight, developers may anticipate problems and prepare accordingly.

## BENEFITS OF THE MAINTAINABILITY METRIC MODEL:

Several advantages are associated with the suggested model of maintainability metrics:

- The model's quantitative evaluation of maintainability helps software teams spend resources wisely for maintenance tasks, leading to better maintenance planning.
- Prospective Maintenance difficulties can be Identified in Advance The model's use of soft computing approaches allows for the early identification of prospective maintenance difficulties, allowing for preventative steps to be performed.
- Developers are better able to make educated judgements about software changes, refactoring, and component selection because to the model's predictive capabilities.
- The model's accuracy and efficiency gradually increase over time thanks to the genetic algorithms used to optimise its evolution.

## II. REVIEW OF LITERATURE

Lo (2019) proposed several types of enhanced traditional software reliability growth models. They not only developed a defect detection mechanism, but also explained the model they used to anticipate the failures. The quality and reliability index is lowered due to the lingering software defects that cause system failure. Consequently, they used a back-propagation neural network to examine the system's previous record of malfunctions.

Sharma et al., (2019) was presented Method for gauging reuse based on artificial neural networks. They looked at four elements that have a significant effect on the reusability of component based systems. To evaluate a product, developers consider readability, transferability, personalization, and interface complexity. The neural network is trained using data sets from the four specified independent variables, and its anticipated reusability is then evaluated. In the simulation of a neural network, the scientists employed a widerange of options for the number of neurons and thenumber of hidden layers.

Sandhu et al. (2018) compared assessment of component reusability using fuzzy, neuro-fuzzy, and fuzzy-GA methods. Independent input criteria like as volume, complexity, coupling, and reuse frequency from the Halstead software science indicator are employed. The suggested model uses the input vectors to assess the degree to which components may be reused. Although the neuro- fuzzy method worked well, the fuzzy-GA method delivered the highest level of accuracy. Based on the validation data, they found that fuzzy-GA, which uses GA's strong decision making but not its learning power or data optimisation, yielded the best outcomes. Based on our review, it seems that fuzzyGA has not been investigated extensively for challenges involving the analysis and prediction of huge datasets, such as software attribute prediction.

Bhattacherjee (2016) presented Five neurons, one hidden layer, seventeen hidden nodes, and a single anticipated output are only some of the architectural features of the ANN method. Software development time was modelled using a variety of programme and staff factors in the provided model. Predicting how long a given project will take todevelop is the primary goal of the validation. Errors were within the acceptable range (25%) in 91% of instances when compared to the training data. Seventy percent of the time, it falls inside the 20% range. They proved that ANN is a good tool for predicting these kinds of software features.

Pomorova and Hovorushchenko (2011) indicated that The way forward in assessing software quality is with the use of neural network approaches. Metrics from the design stage, such as complexity and quality attributes, were analysed and predicted using an artificial neural network. A total of 324 input vectors were utilised for the experiment, and the results showed that, in this particular scenario, the new approach makes no difference in terms of mean square error.

Ardil and Sandhu (2010) used branch count, number of lines, essential complexity, cyclomatic complexity, and design complexity are all inputs. The severity of software maintenance is predicted using a Fuzzy Inference System (FIS) and an Adaptive Neuro Fuzzy Inference System (ANFIS) of the mamdani type. It has been determined that a neuro-fuzzy based implementation is superior than FIS. The authors conducted their experiments using a NASA defect dataset available in the public domain. Simple logistic approaches and trees, they found, outperform other machine learning techniques and can detect faults with an accuracy of 65%.

## III.RESEARCH METHODOLOGY

### 3.1 Research Design

In this paper, we explore two measures for CBSE maintainability that make use of soft computing. Fuzzy logic is used for the first 16 DTCM maintainability metric implementations. The DRCE maintainability model employs Fuzzy logic, neural networks, and the ANFIS (Neuro-Fuzzy) approach to enhance the efficacy of the original model. Values from the model and those from other researchers are compared as well.

### 3.2 DTCM Maintainability Metric Model

This study picks four factors out of a few elements, based on their profound connection with the maintainability metric. The selected factors for the model are Documentation Quality, Testability, Coupling, and Modifiability (DTCM).

Documentations Quality: The documentation enlightens the way to operate the software.

Compared to a poorly documented component, a component with excellent documentation is supposed to be more maintainable.

- Testability: For a product with higher testability, the fault-finding process will remain manageable and thus simple to maintain.
- Coupling: It refers to the extent of association between various components or modules.
- Modifiability: It refers to the simplicity with which improvements can be made and how the Rules based on Fuzzy Crisp Input Fuzzification Module Fuzzy Inference system Defuzzification Crisp Output program responds to these changes, such as changing environments, technical specifications,and criteria.

The inputs are converted into their fuzzy analogues using the Fuzzification procedure. Fuzzification describes the process by which a scalar value is transformed into a fuzzy one. Some examples of fuzzy membership functions included in the MATLAB fuzzy toolbox for fuzzification (TFN) are trapezoidal, gaussian, S function, singleton, and triangular fuzzy number. TFN has been included into the suggested model due to its ease of usage. Maximum, Minimum, and Average membership functions are taken into account in the DTCM model. Fuzzy's 'Maximum' indicates the highest value placed on Document Quality, while 'Minimum'indicates the lowest value. The five TFN membership functions used to rank Documentation Quality are "Worst," "Bad," "Great," "Fair," and "Excellent." Testability Coupling Modifiability Maintainability To Create a Fuzzy Model The outcome of the maintainability maintenance process.

The suggested DTCM model for measuring maintainability is heavily reliant on a Fuzzy Inference System (FIS). FIS makes decisions based on the rules specified in the rule editor. The MATLAB fuzzy toolbox supports both the Mamdani and Sugeno fuzzy inference techniques.

An estimated output is calculated using the averaging technique included in the defuzzification module. It combines several rules into a single one as a final result. The effects of the rules are bound together by aggregation based on the input that satisfies the conditions.

### 3.3 DRCE Maintainability Metric Model

Three soft computing techniques—fuzzy logic, (ANN) Artificial Neural Network, and Artificial Neural Fuzzy Inference System—are used to create this model.

In this part, we'll go through the data that goes into the DRCE fuzzy maintainability metric: Clear Inputs the DRCE model of maintenance metrics looks at several aspects of maintenance.

The following four factors are selected as crisp input to the proposed model.

- Documentations Quality: A component's excellent documentation improves its chances of selection in a component-based framework.
- Reusability: The possibility of a component being selected as a module in many programs is calledits reusability.
- Coupling: The degree of interconnection between software components is referred to as coupling. Maintenance is inversely proportional to coupling.
- Extensibility: An extensible component will accommodate a significant expansion of new functionalities without requiring substantial changes to its core architecture for future development.

The model takes as inputs the membership functions "Maximum," "Average," and "Minimum" (TFN), and as outputs the membership functions "worst," "bad," "good," "fair," and “excellent”.

## IV. RESULTS

### 4.1 DTCM Model Results

For these parameters, the DTCM metric model generates a maintainability score of 0.497. The outcomes for the following input values are shown in Table 1: Document Quality = 0.5, Testability = 0.75, Modifiability = 0.50, and Coupling = 0.70. For these parameters, the DTCM metric model generates a maintainability score of 0.497. Figure 1for an illustration of the rule viewer for the aforementioned data. Figure 2 displays the overall data input maintainability result.

**Table 1: DRCE Maintainability Metric Model Result**

| DTCM Maintainability Metric Model's Inputs | Document Quality | Testability | Coupling | Modifiability |
|---|---|---|---|---|
| | 0.5 | 0.75 | 0.7 | 0.5 |
| **Output** | **Maintainability** | | | |
| | 0.497 | | | |

$$\text{Centre of gravity}=\frac{\int yxdx+\int yxdx}{\int ydx+\int ydx}=.467 \quad \text{.....(1)}$$

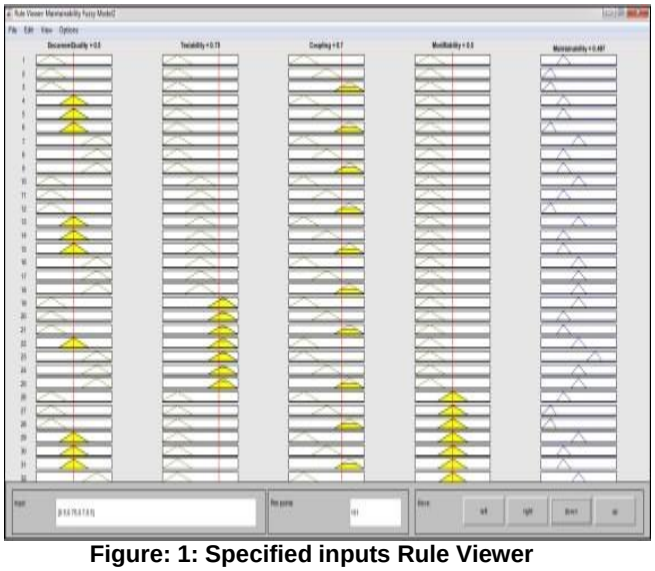


**Figure: 1: Specified inputs Rule Viewer**

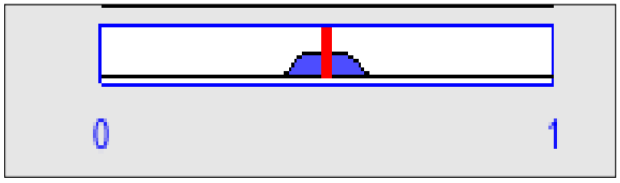



**Figure 2: Result's Fine Point view**

### 4.2 DRCE Implemented Artificial Neural Network(ANN):

The DRCE-ANN maintainability metric model takes its input from the data that was collected and organised for the fuzzy model. Document quality, reusability, coupling, and extensibility are the four inputs to this approach, with maintainability serving as the final output. Figure 3 shows a scatter plot of samples from the training, validation, and test datasets.

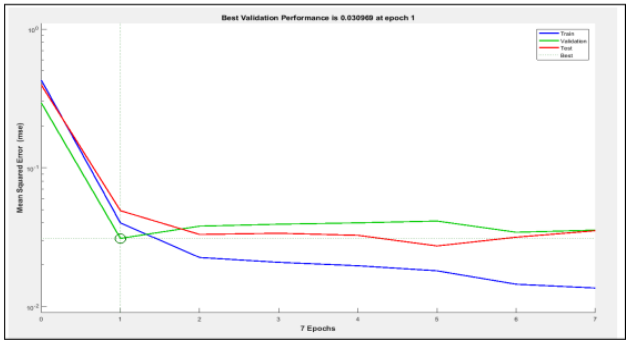
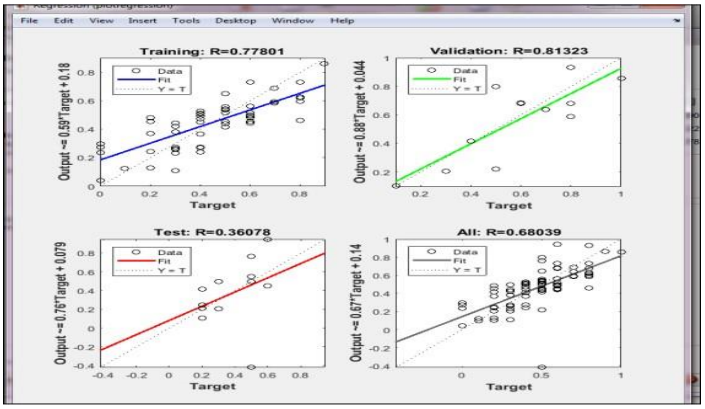



**Figure 3: Data sample's Training, testing and validation**

As can be seen in Figure 4, the DRCE maintainability metric model produced some tangible results. The graphic depicts a variety of different types of graphs, including those used for training, validation, testing, and a hybrid of these three. The data sets have been shown as circles bythe neural network, and in all of these graphs, the well-trained circles are located near the plotted axis. Following the training process, the majority of the circles in the final graph fall quite near to the dotted line. That the DRCE model has been properly trained is indicated. Preliminary results indicate it may be effectively used to evaluate the maintainability of component systems.

**Figure 4: Training, validation and test data sets DRCE model**




Mean absolute error (MAE) is calculated using the formula in equation 2.

$$MAE = 1/n \sum_{i=1}^{n} |xi - x| \quad \text{.....(2)}$$

'xi' =outcome from the DRCE model, 'x' = actual maintainability and 'n' = number of observations

The results of the DRCE artificial neural network metric model for mean absolute error and relative accuracy are shown in Table 2.

$$RE_{accuracy} = \frac{Absoluteerror}{Truevalue} * 100\% \quad \text{.....(3)}$$

**Table 2: DRCE Neural Network Model Validation**

| Model of Validation | MAE | $RE_{accuracy}$ |
|---|---|---|
| Outcome | .028 | .045 |

Since the MAE and RE results are within reasonable bounds, this model might be used to evaluate the maintainability of a component-based system.

### 4.3 ANFIS DRCE Maintainability metric Model:

The ANFIS DRCE metric model's predictions are checked using the Mean Absolute Error (MAE) metric. Relative Error (RE) is also used to verify the accuracy of the result. The formula's effects on the experiments are shown in Table 3.

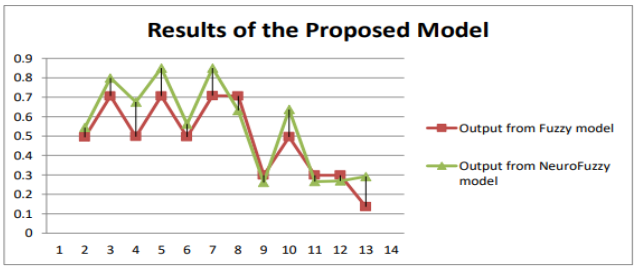
**Table 3: Validation of DRCE Neural Network Model**

| Model of Validation | MAE | $RE_{accuracy}$ |
|---|---|---|
| Outcome | .0029 | .0399 |

The ANFIS model has been shown to accurately forecast the maintainability of a component-based structure, as shown by the supporting evidence.

Figure 5 showing the results of this investigation.

**Figure 5: Comparative analysis of the models**




## V. DISCUSSION

In this investigation, soft computing methods were used to two CBSE maintainability measures. Using fuzzy logic, this research defined the first maintainability measure, the DTCM model. In order to further improve the efficacy of this model, Fuzzy, Neural Network, and ANFIS (Neuro-Fuzzy approach) are combined with an additional measure known as the DRCE maintainability model. A comparison between the model's results and those of other researchers was also given. The DTCM Model and the DRCE Maintainability Model were the subjects of an observational study based on statistical analysis.

A component-based architecture requires upkeep for bug fixes and to keep up with ever-changing environments for software. Quality attribute maintenance is essential in all approaches of software development, including iterative development, agile technology, and Component- based development. In order to automatically anticipate maintainability levels in component- based software, this study presented two maintainability measures.

The following improvements to component-based software creation are made possible by these models' predictions of maintainability:

The paradigm has monetary and time benefits for both the software user and the developer.

The model helps identify which parts of a component-based architecture can be kept up-to-date and which may be thrown out throughout the design process.

The time and energy spent maintaining a gadget may be drastically reduced if its maintainability has been evaluated with the use of a model.

Developers and maintenance engineers will utilise this model to make sure their products are easily repairable.

It will aid the designer in determining whether or not a replacement system is necessary.

Findings from this research recommended using a fuzzy method to make maintenance estimates for component-based software on a scale from excellent to poor. It is also stated that while maintainability is affected by a number of factors, the four chosen inputs have the most effect on it. The results of the model are checked using MATLAB's Fuzzy toolbox, which shows a significant connection between these characteristics and maintainability. Let's pretend we decide to increase the amount of characteristics to get optimal precision. In such instance, the model complexity increases dramatically, necessitating the addition of many more rules to the rule editor. It's possible that we'll have to increase the number of characteristics whose values we extract. When applied to the DTCM, the soft computing method has the potential to increase the model's prowess and depth of understanding. In the future, it may be important to identify the other fundamental properties that affect maintainability. Proper investigation into those aspects will provide the most accurate maintainability evaluation.

## VI. CONCLUSION

Last but not least, the maintainability metric model for component-based software that makes use of soft computing approaches provides a holistic and flexible strategy for coping with the difficulties of maintaining big software systems. The model's quantitative evaluation of maintainability, proactive discovery of maintenance concerns, support for data-driven decision making, and room for continual development are all thanks to the use of fuzzy logic, genetic algorithms, and neural networks. This concept may help make it easier to manage software systems that are built from smaller building blocks, or components. The use of soft computing approaches in maintainability measures will aid in the creation of more stable and long-lasting software systems as the discipline of software engineering develops.

## REFERENCES

1. Aggarwal K K, Singh Yogesh, Chandra P., Puri M. (2000), "Measurement of Software Maintainability using a Fuzzy Model", Journal of Computer Sciences, Volume 1 Issue 4, pp. 538.
2. Aggarwal K K, Singh Yogesh, Kaur Arvinder, Malhotra Ruchika (2005), "Software Reuse Metrics for Object-Oriented Systems", Proceedings of the 2005, Third ACIS Int'l Conference on Software Engineering Research, Management and Applications (SERA'05), pp. 11-13.
3. Ali A, Choudhary K, Sharma A (2016) Software quality prediction using random particle swarm optimization (PSO). In: 2016 International conference on circuit, power and computing technologies (ICCPCT). IEEE, pp

1–6
4. Almugrin S, Albattah W, Melton A. Using indirect coupling metrics to predict package maintainability and testability. *J Syst Softw.* 2016;121:298–310.Arar ÖF, Ayan K. A feature dependent Naive Bayes approach and its application to the software defect prediction problem. *Appl Soft Comput.* 2017;59:197–209.
5. Ayon SI (2019) Neural network based software defect prediction using genetic algorithm andparticle swarm optimization. In: 2019 1st international conference on advances in science, engineering and robotics technology (ICASERT).IEEE, pp 1–4
6. Badampudi D, Wohlin C, Petersen K. Software component decision-making: In-house, OSS, COTS or outsourcing—a systematic literature review. *J Syst Softw.* 2016;121:105–124.
7. Bashir K, Li T, Yohannese CW, Mahama Y (2017) Enhancing software defect prediction using supervised-learning based framework. In: 2017 12th international conference on intelligent systems and knowledge engineering (ISKE). IEEE, pp 1–6
8. Bhardwaj Vidhu (2010), "Estimating reusability of software components using fuzzy logic", Thesis, Thapar University, Patiala.
9. Boxall M A S, Araban S (2004), "Interface Metrics for Reusability Analysis of Components". Proceeding of Australian Software Engineering Conference (ASWEC'2004), Melbourne, Australia, pp. 40-46.
10. Cai BL, Zhang RQ, Zhou XB, Zhao LP, Li KQ. Experience availability: tail-latency oriented availability in software-defined cloud computing. *J Comput er Sci Technol.* 2017;32(2):250–257.
11. Chatzipetrou P, Papatheocharous E, Wnuk K, Borg M, Alégroth E, Gorschek T. Component attributes and their importance in decisions and component selection. *Softw Qual J.* 2019
12. Cho, E. S., Kim, M. S., Kim, S. D. (2001), "Component Metrics to Measure Component Quality", 8Asia-Pacific Software Engineering Conference, Macau, pp. 419-426.
13. Gaur Jatain, Gaur Aman, Gaur Deepti (2012), "Estimation of Component Reusability by identifying Quality Attributes of Component: A Fuzzy Approach", Proceedings of the Second International Conference on Computational Science Engineering and Information Technology-CCSEIT 12, pp. 738-742.
14. Gill N S (2003), "Reusability Issues in Component-Based Development". ACM SIGSOFT Software Engineering Notes, Vol. 28, Issue 4, pp. 1-5.
15. Gill N S, Grover P (2004), "Few Important Considerations For Deriving Interface Complexity Metric For Component-Based Systems". ACM SIGSOFT Software Engineering Notes, Vol. 29, No. 2, pp. 1-4.
16. Mili H, Mili F, Mili A (1995), "Reusing Software: Issues and Research Directions". IEEE Transaction on Software Engineering, Vol. 21, Issue 6, pp. 528-561.
17. Poulin J, Caruso J, Hancock D (1993), "The Business Case for Software Reuse". IBM Systems Journal, Vol. 32, Issue 40, pp. 567-594.
18. Roger Jang, Gulley Ned (1995), "Fuzzy Logic Toolbox for MATLAB". User's Guide, the Math Works Inc., USA.
19. Rotaru O P, Dobre M, Petrescu M (2005), "Reusability Metrics for Software components". Proceeding of the 3rd ACS/IEEE International Conference on Computer Systems andApplications (AICCSA-05), Cairo, Egypt, pp. 24-29.
20. Sagar Shrddha, Nerurkar N. W., Sharma Arun (2010), "A Soft Computing Based Approach to Estimate Reusability of Software Components", ACM SIGSOFT Software Engineering Notes, Volume 35 Number 4, pp. 1-4.
21. Sandhu Singh Parvinder, Singh Hardeep (2006), "A Neuro-Fuzzy Based Software Reusability Evaluation System with Optimized Rule Selection". 2nd International Conference on Emerging Technologies, IEEEICET. 2006, Peshawar, Pakistan 13-14 Nov., pp. 664-669.
22. Schach Stephen R, Yang X (1995), "Metrics for Targeting Candidates for Reuse: An Experimental Approach". ACM, SAC, pp. 379-383.
23. Sharma Arun, Grover P. S., Kumar Rajesh (2009), "Reusability Assessment for software component", SIGSOFT Software Engineering Notes, Volume 34 Number 2, pp. 1-6.
24. Sharma Arun, Kumar Rajesh, Grover P S (2009), "Reusability Assessment for Software Components a Neural Network Based Approaches". ACM SIGSOFT Software Engineering Notes, USA, Vol. 34, No. 2, pp. 1-6.
25. Singh Charu, Pratap Amrendra, Singhal Abhishek (2014), "An Estimation of Software Reusability using Fuzzy Logic Technique", International Conference on Signal Propagation and Computer Technology 2014 (ICSPCT 2014) Ajmer, India.
26. Singh Yogesh, Bhatia Pradeep Kumar, Sangwan Omprakash (2009), "ANN Model for Predicting Software Function Point Metric", Published in ACM SIGSOFT, Software Engineering Notes, USA, Vol. 34, No. 1, pp. 1-4.
27. Singh Yogesh, Bhatia Pradeep Kumar, Sangwan Omprakash (2011), "Software Reusability assessment using soft computing techniques", ACM SIGSOFT Software Engineering Notes, Volume 36 Number 1, pp. 1-7.
28. Washizaki H, Y Hirokazu, F Yoshiaki (2003), "A Metrics Suite for Measuring Reusability of Software Components". Proceeding of the 19th International Symposium on Software Metric, pp. 211-223.
29. Zadeh L. A. (2005), "From Computing with Numbers to Computing with Words-From Manipulation of Measurements to Manipulation of Perceptions", International Journal of Applied Mathematics and Computer Science, Volume 1 Issue 4.

# ENHANCED DATA AGGREGATION STRATEGY FOR IOT-BASED WIRELESS SENSOR NETWORK VIA IDEAL CLUSTERING APPROACH

**Dr. Gopinath D**
Assistant Professor, Department of Computer Science,
Kristu Jayanti College (Autonomous), Bengaluru, Karnataka, India.
**Mr. M. Chandru**
Assistant Professor, Department of Computer Science,
VET Institute of Arts and Science (Co-education) college, Tamil Nadu, India.

**Abstract**
**The Internet of Things (IoT), which is built on wireless sensor networks (WSNs), has grown significantly in popularity as a means of providing answers to several real-time applications through communication between sensor nodes and physical world objects. In order to achieve real-time quality of service (QoS), economical operation, and long-term dependability, WSN-based IoT has emerged as a vital technology. The wireless interconnection of numerous items is made possible by IoT-based WSN nodes, which are small, equipped with irreplaceable batteries, and resource-constrained. IoT produces a large volume of data as an outcome of constant sensing and collecting, which may result in significant computing overhead, data redundancy, packet collisions, and high energy usage. Significant research has been done to optimise node energy usage to increase network longevity in order to solve resource constraints. However, these solutions fail to handle data redundancy and have poorer throughput. The majority of present methods concentrate on increasing network lifetime through scheduling methods over duty cycles. The elimination of data redundancy and the generation of energy have both been addressed by cluster-based data aggregation techniques. This paper present an Effective Data Aggregation Strategy (EDAS) for IoT-based WSNs. This scheme takes into account the Enhanced Low Energy Adaptive Clustering Algorithm (E-LEACH) to build the ideal number of Cluster Heads (CH), taking into account the average network energy and node residual energy. Using network coding, which incorporates the special operation and guarantees non-replicated data transmissions, data redundancy is removed.**

**Keywords**: IoT, WSNs, EDAS, E-LEACH, Cluster Heads (CH)

## 1. INTRODUCTION

The Internet of Things (IoT), a novel technology built on wireless sensor networks (WSN), is quickly gaining popularity and is an essential component of time-sensitive applications such as those in the industrial, healthcare, smart city, and automotive sectors [1]. IoT has benefited from cloud computing, which provides flexibility across a range of application domains. IoT-based sensors are tiny, intelligent, battery-powered devices with distributed, self-configured nodes that perform specialised functions in monitoring the physical world without human involvement [2–5]. IoT devices have built-in data analysis capabilities and can connect and share information with actual items in the real world over the internet. Sensors produce enormous amounts of data streams at rapid speeds because they continuously sense the physical environment [6]. To provide consumers with real-time content, the sensed data are analysed and sent between nodes and cloud servers. Nodes need more power, bandwidth, and network resources during data exchange. The task of harvesting network energy is still difficult because several data copies are sent during data connection without redundancy check, increasing energy consumption, computational load, and delay [7–10].

The improvement of node energy consumption and the extension of network lifetime have been addressed by hierarchical cluster-based communication systems [11,12]. Cluster-based communication methods incorporate data aggregation to get rid of redundant copies of data through data redundancy checking and to prevent multiple data transmissions [13].

The network is divided into cluster groups for cluster communication. Each cluster has a cluster head (CH) and cluster members (CM), and the CH is in charge of gathering data from the CM. Data acquired by different cluster heads is computed at the sink node, which is where the cluster head aggregates collected data and forwards to the sink or base station (BS) [14].

The following are the main contributions:

- The EDAS method takes into account the best cluster head formation and rotation by taking into account energy metrics such as node residual energy, average network energy, total network energy, and node beginning energy.
- By using a liner data aggregation technique, which lowers the traffic load and lengthens network lifetime, EDAS removes data redundancy.

## 2. RELATED WORKS

For cluster-based WSN, the author of [15] suggested an energy-efficient data aggregation technique that takes into account mobility agents (MAs), which move close to sensor nodes to collect data and transmit data to sinks. This plan also takes fault tolerance into account by including a secondary cluster head that serves as a backup cluster head in the event of node failure. When moving to various locations, the mobile agent rigorously follows the itinerary plan computed and provided by the sink node. Static itinerary plans are appropriate for monitoring applications, while dynamic itinerary plans are employed for object tracking.

The network lifetime is greatly increased by this technique, however the large computing expense is caused by the itinerary planning. Author [16] suggested the hybrid QoS aware data aggregation (QADA) strategy for IoT, which includes tree-based cluster routing, in order to achieve optimal quality of service (QoS). The QADA process consists of four phases: cluster head selection, cluster creation, tree formation, and data aggregation. By taking into account nodes with high residual energy and distance to sink, sink node chooses cluster head. Author suggested energy balanced high level data aggregation and energy aware compressive sensing aggregation in Ref. [17] to balance the nodes' energy. This approach takes into account mobile sinks when collecting data from various clusters.

The author of Ref. [18] presented an energy-efficient data aggregation approach for cluster-based WSNs, which modified the LEACH clustering stages to choose the cluster head and cluster formation. By delivering control frames to the cluster head, the intra-cluster communication is decreased. The author of Ref. [19] presented a methodical data aggregation strategy for cluster-based WSN. When forming a cluster, active node status is taken into account, and the cluster head is chosen based on the base station's location and the current residual energy level. The author of Ref. [20] suggested a cluster tree-based effective data collection technique for industrial IoT to boost throughput and lengthen network lifetime. Based on the data gathered locally, fuzzy logic is used to choose the cluster head.

## 3. PROPOSED EDAS

### 3.1 Cluster head selection

The cluster head is selected at random in standard LEACH [21] without taking energy factors into account. The threshold value Thresh($n_i$) and random number generator function determine the CH selection. As long as (0, R, 1), the nodes $n_i$ create random number R. If nodes meet the requirement of R Thresh($n_i$).A drawback of typical LEACH is that node with lower energy can also turn into a cluster head, which may perish fast. The suggested EADS technique takes energy parameters into account and chooses the cluster head by computing the threshold value Thresh($n_i$), which is defined as:

$$\text{Thresh(ni)} = \frac{pb_i}{1 - pb_i * \left( r * mod\left(\frac{1}{pb_i}\right)\right)} \; n_i \in v$$

$$\begin{cases} \frac{Pb_i}{1 - Pb_i * (r * mod(1/Pb_i))}, n_i \in V \\ 0, otherwise \end{cases}$$

The energy parameters $Pb_i$ taken into account by EADS are as follows, where $n_i$ is the node and i [1,N]:

$$pb_i = \frac{P * n_i * e_r^i * e_i}{e_i * e_a}$$

P, current residual energy of the node (ei r), starting energy of the node (ei), total network energy (et), and average energy of all nodes (ea) are the variables that determine the probability of choosing the best CH. The suggested threshold value for choosing the cluster head balances the burden on the nodes that distribute energy and

lengthens the lifespan of the network. The proposed threshold value improves the likelihood of that node being chosen as CH if node residual energy exceeds $e_a$. The average energy of all nodes after functioning for r rounds is calculated as follows:

$$a = \frac{e \quad e_{t\left(1-\frac{r}{r_{maximum}}\right)}}{n_i}$$

### 3.2. Cluster formation

After the CH has been chosen, the CHs will broadcast a message inviting non-cluster head nodes to take part in cluster creation. Depending on the received signal strength (RSS) and the distance to the sink node, nodes choose whether or not to take part in cluster formation. In the proposed EDAS, the threshold Thresh(ni) is established in each round after the energy and distance to sink are calculated. Nodes that are closer to CH create random numbers and compare them to Thresh(ni). If Thresh(ni) is less than or equal to Thresh(ni), there is a probability that node will become CH. The calculation of the distance between a node and the sink determines whether the node will participate in cluster formation or communicate directly with the sink to send data.

## 4. PERFORMANCE EVALUATION AND RESULT ANALYSIS

Using the NS2 simulator [22], the proposed EDAS's performance is simulated. Relevant experiments are run to examine the scheme's performance outcomes, and comparisons are made with the LEACH [21] and QADA [16] existing schemes. Analysis is done on performance parameters such packet delivery, energy usage, and the number of nodes that are still active. The following Table 1 contains a list of the simulation parameters.

| Parameters | Value |
|---|---|
| Nodes size | 150 |
| Buffer size | 50 |
| MAC-Protocol | 802.11 collision free |
| Total Simulation Time | 200 s |
| Deployment area | 1000 × 1000 mts |
| Packet Size | 512 bytes |
| Transmission Range | 50 mts |
| Initial Energy | 50 J |
| Sink node | 1 |
| Scheduling | TDMA |

**Table 1 Parameters used in simulation study**

## 5. RESULT ANALYSIS

### Packets received at sink

The performance of the total number of packets delivered at the sink over time. It has been shown that as the time changes, more packets are received at the sink through EADS than through the LEACH and QADA systems that are currently in place. LEACH does not take energy parameters into account, In choosing CH, it makes a random selection of CH based on remaining energy.

Since nodes with lower energy levels can also become CH, the likelihood of a CH becoming faulty is higher due to low energy levels because they do not transmit aggregated data to sink. But, QADA performs better when delivering aggregated data to sink through existing tree-based paths than LEACH in terms of packet delivery at sink node. By figuring out the network delay and delivering packets effectively, EADS outperforms QADA.

### Energy Consumption

The overall energy consumption over time, the suggested EDAS uses less energy and increases network lifetime when compared to LEACH and QADA under the identical simulation conditions. Because QADA takes into account a high residual energy parameter to pick CH and establishes a tree towards the sink through CH, it optimises cluster head selection and increases network lifetime compared to LEACH. In addition to high residual energy, EDAS additionally optimises CH selection compared to LEACH and QADA by taking into account energy metrics including average energy, total network energy, and node starting energy. Through sequence matching of data collected at the sink during data aggregation at CH, EDAS considerably decreases computing functions while reducing redundant data transmissions.

### Number of Alive Nodes

The number of alive nodes over time is depicted. The graph shows that initially LEACH, QADA, and EDAS have more stable networks since there are more of them. Due to the random CH selection, nodes in LEACH begin to die earlier than those in QADA as time goes on. In LEACH, the node with lesser energy can potentially be chosen as the CH, which could cause nodes to die more quickly. Due to the tree-based routing in QADA, which balances CH's energy usage, there are more active nodes than in LEACH. While the sink node uniformly distributes the energy load during cluster formation and CH selection, QADA has a slower rate of dead node decline over time than LEACH.

For reliable data transmission with the fewest hops, EADS bases cluster formation and CH selection on the distance to the sink node. Nodes close to the sink instead of the CH do not take part in cluster formation and speak directly to the sink node. As a result, EDAS has more active nodes in the network than LEACH and QADA do.

**Normalized routing overhead**

The routing overhead of LEACH, QADA, and EDAS is shown over time. It can be seen that when more reliable paths to sink are found, the communication overhead rises over time. Because LEACH conducts more computing functions to determine the shortest path through reliable nodes to construct a reliable path than QADA, it can be seen that the overhead of LEACH increases more quickly than that of QADA. Due to the dependable method that is chosen to sink through the CH tree, the overhead for QADA is lower than for LEACH. The aggregated data is transmitted via a pre-established channel, reducing overhead. By measuring the distance to the sink node and using multihop communication, EDAS determines the best cluster formation and uses the shortest way possible to transmit aggregated data in order to minimise overhead. Due to the fact that CH is familiar with the topology and the distance to the sink, EDAS performs better when computing routes to the sink.

## 6. CONCLUSION

For wireless sensor networks based on the Internet of Things, this research suggests an effective data aggregation strategy (EDAS). To increase network lifetime, EDAS uses improved LEACH cluster-based communication for trustworthy data transmission. When choosing a CH in EDAS, energy characteristics including beginning energy, total network energy, average node energy, and residual energy are taken into account. This helps to determine the best number of clusters. The EDAS system incorporates liner special operation to get rid of duplicated data, which lowers the workload on the network and the amount of computation required. The simulation results show that the proposed EDAS outperforms the current system in terms of packet delivery, energy usage, and overhead. It is compared to the existing E-LEACH, clustering scheme.

## REFERENCES

[1] A. Musaddiq, Y.B. Zikria, O. Hahm, H. Yu, A.K. Bashir, S.W. Kim, A survey on resource management in IoT operating systems, IEEE Access 6 (2018) 8459–8482.

[2] I. UdDin, M. Guizani, S. Hassan, B. Kim, M. Khurram Khan, M. Atiquzzaman, S. H. Ahmed, The internet of things: a review of enabled technologies and future challenges, IEEE Access 7 (2019) 7606–7640.

[3] T. Qiu, N. Chen, K. Li, M. Atiquzzaman, W. Zhao, How can heterogeneous internet of things build our future: a survey, IEEE Commun. Surv. Tutor. 20 (3) (2018) 2011–2027.

[4] Y. Kuo, C. Li, J. Jhang, S. Lin, Design of a wireless sensor network-based IoT platform for wide area and heterogeneous applications, IEEE Sensor. J. 18 (2018) 5187–5197.

[5] M. Lazarescu, Wireless sensor networks for the internet of things: barriers and synergies, in: Components and Services for IoT Platforms, 2017, pp. 155–186.

[6] I. Yaqoob, E. Ahmed, I.A.T. Hashem, A.I.A. Ahmed, A. Gani, M. Imran, M, GuizaniInternet of things architecture: recent advances, taxonomy, requirements, and open challenges, IEEE Wireless Commun. 24 (2017) 10–16.

[7] D. Zhang, G. Li, K. Zheng, X. Ming, Z. Pan, An energy-balanced routing method based on forward-aware factor for wireless sensor networks, IEEE Trans. Ind. Inf. 10 (2014) 766–773.

[8] Y. Zhang, M. Liu, Q. Liu, An energy-balanced clustering protocol based on an improved CFSFDP algorithm wireless sensor networks, Sensors 18 (2018) 881.

[9] M.S. Bahbahani, E. Alsusa, A cooperative clustering protocol with duty cycling for energy harvesting enabled wireless sensor networks, IEEE Trans. Wireless Commun. 17 (2018) 101–111.

[10] Z. Zhao, K. Xu, G. Hui, L. Hu, An energy-efficient clustering routing protocol for wireless sensor networks based on AGNES with balanced energy consumption optimization, Sensors 18 (2018) 3938.

[11] A. Hosen, G. Cho, An energy centric cluster-based routing protocol for wireless sensor networks, Sensors 18 (2018) 1520.

[12] M. Tarhani, Y.S. Kavian, S. Siavoshi, SEECH: scalable energy efficient clustering hierarchy protocol in wireless sensor networks, IEEE Sensor. J. 14 (2014) 3944–3954.

[13] L. Meng, H. Zhang, Y. Zou, Data aggregation transfer protocol based on clustering and data prediction in wireless sensor networks, in: 7th International Conference on Wireless

Communications, networking and Mobile Computing (WiCOM), IEEE, 2011, pp. 1–5.

[14] A. Rajasekaran, V. Nagarajan, Improved cluster head selection for energy efficient data aggregation in sensor networks, Int. J. Appl. Eng. Res. 11 (2) (2016) 1379–1385.

[15] M. El Fissaoui, A. Beni-Hssane, M. Saadi, Energy efficient and fault tolerant distributed algorithm for data aggregation in wireless sensor networks, J Ambient IntellHuman Comput 10 (2019) 569–578, https://doi.org/10.1007/s12652-018-0704-8.

[16] H. Rahman, N. Ahmed, M.I. Hussain, A QoS-aware hybrid data aggregation scheme for Internet of Things, Ann. Telecommun. 73 (2018) 475–486, https://doi.org/ 10.1007/s12243-018-0646-3.

[17] Shima&Avokh PakdamanTirani, Avid, On the performance of sink placement in WSNs considering energy-balanced compressive sensing-based data aggregation, J. Netw. Comput. Appl. 107 (2018), https://doi.org/10.1016/j.jnca.2018.01.012.

[18] N.R. Roy, P. Chandra, EEDAC-WSN: energy efficient data aggregation in clustered WSN, in: International Conference on Automation, Computational and Technology Management (ICACTM), 2019, 2019, pp. 586–592, https://doi.org/10.1109/ ICACTM.2019.8776679.

[19] M. Shobana, R. Sabitha, S. Karthik, Cluster-based systematic data aggregation model (CSDAM) for real-time data processing in large-scale WSN, Wireless Pers. Commun. 117 (2021) 2865–2883, https://doi.org/10.1007/s11277-020-07054-2.

[20] Bhanumathi Velusamy Kalaivanan Karunanithy, Cluster-tree based energy efficient data gathering protocol for industrial automation using WSNs and IoT, Journal of Industrial Information Integration 19 (2020), 100156, https://doi.org/10.1016/j.jii.2020.100156. ISSN 2452-414X.

[21] W.R. Heinzelman, A. Chandrakasan, H. Balakrishnan, Energy-efficient communication protocol for wireless microsensor networks, in: Proceedings of the 33rd Annual Hawaii International Conference on System Sciences. Piscataway: IEEE, Jan., 2000, p. 10.

[22] NS, Network Simulator-NS2, 2009. http://www.isi.edu/nsnam/ns.

# BASICS IN MOBILE COMPUTING AND ITS PRINCIPLES & APPLICATIONS: A STUDY

**N.Senthil**
Assistant professor, Department of Computer Science,
Kangeyam Institute of Commerce, Tiruppur, Tamil Nadu, India

**ABSTRACT**

**As people are adapting to the various advancements that mobile computing brings, securing the mobile computing is becoming more and more important. Over the past decades, the technology has reduced the size of the machinery and increased its performance. This development led to the development of a new concept called Mobile Computing. Mobile Computing essentially allows the data to be transmitted through the computers without any connection to any fixed physical parameter. Mobile computing is a collection of IT technologies, products, services, and operational strategies and tactics that allow customers to access computing, data, and related resources and capabilities simultaneously on mobile devices. Mobile computing and sensing combined with cloud computing enable new programs that manipulate big statistics series, saving time and improving productivity. This paper clearly deals with basics in mobile computing and its principles & designs.**

KEYWORDS: Mobile, Computing, Mobile communications, Applications, Mobile security

## 1. INTRODUCTION

Packages of mobile computing have become ubiquitous and ubiquitous in business, consumer, industrial, entertainment and many specialized vertical market sports. While desktop, computer systems offer a higher level of hardware execution, the primary advantage of cellular computing is its convenience, allowing customers to access information and computational resources anywhere, anytime. The beginnings of "mobile computing.

Wireless communication is the foundation of mobile computing. Therefore, it is essential to provide security to this communication or to the data transfer. Since the wireless communication takes place through radio signals instead of wires, the communication channel can be easily intercepted. Therefore, to overcome this problem, it is necessary to adopt some security measures. There are several issues related to security. Some of these issues/problems are: Confidentiality Integrity Availability Legitimacy Accountability In recent years, there has been a major revolution in the area of telecommunication. This revolution has led to the development of many lightweight, portable and compact computing devices such as cell phones, laptops etc.

## 2. OBJECTIVES OF THIS STUDY

- To know about the Mobile computing concepts
- To know about the mobile communications
- To know about the mobile computing applications
- To know about the mobile computing hardware's

## 3. RELATED WORKS

Singh & Gupta (2017). Mobile computing security is becoming increasingly important as people adapt to its various advancements. In later many years innovation has brought about the decrease of the size of apparatus and has expanded execution. The new concept of mobile computing emerged as a result of this development. Versatile Processing essentially permits information move through the PCs, without being associated with any of the decent actual boundaries.

Umai et..al (2015) MCC is an emerging technology. Because of the large number of mobile devices used and the wide range of applications, mobile cloud computing becomes an essential part of mobile devices. It is reliable and portable because data processing and storage

takes place outside the mobile device. In a sense, it helps to save battery and compute power of the mobile device, which is a serious problem in the case of high power mobile device. Mobile cloud computing provides a service where mobile users can access cloud services on their mobile devices and perform calculations. Since mobile Cloud Computing is still in its early stages, it is beneficial to understand the current models and trends.

Vaidehi Tejas Gonjare (2019) The collection of IT tools, services, products, and operational strategies and tactics known as "mobile computing" enables users to access compute, data, and related capabilities and resources simultaneously with cellular. The combination of mobile computing, sensing, and cloud computing enables novel applications that process large data sets, reduce processing time, and boost productivity. This essay discusses trends in technology, mobile hardware, and mobile communication.

Zaslavsky & Tari (1998). Mobile computing has been the subject of intensive research in various fields of computer science & engineering in recent years. Wireless networking, distributed systems, operating systems, distributed databases, software engineering and application development are some of the areas that have been the focus of research in recent years. In this paper, we will provide an overview of what mobile computing is, what it is capable of achieving, what challenges it faces and what opportunities it presents. We will also provide an overview of the current state and ongoing research projects in the field of mobile computing around the world. We will also cover the two Australian workshops on Mobile Computing, Databases and Applications held in 1996 & 1997.

## 4. WHAT IS MOBILE COMPUTING

Mobile Computing is the process of transferring data, voice, and video through a computer or any wireless enabled device without the need to be connected to a physical connection. The main areas of mobile computing are mobile communications, mobile hardware and mobile software.

Mobile Computing is a technology that provides an environment that allows users to transfer data between devices without the need for a physical connection/cable. This means that data can be transmitted wirelessly using wireless devices such as mobile phones, laptops etc. Whenever a device is connected to the network without being connected physically over a link/cable, it can be used to transmit data like messages, voice recordings, videos etc. Mobile Computing technology allows users to access and transfer data from any remote location without being physically present there. With such a large coverage diameter, Mobile Computing is one of the quickest and most reliable sectors of computing technology.

## 5. MOBILE COMMUNICATION

In this case, mobile communication refers to the infrastructure that is put in place to make sure that communication goes smoothly and reliably. This includes things like protocols, services, and bandwidth, as well as portals and data formats. At this stage, the data format is defined so that there is no interference with other systems that provide the same service. Since the media is not guided/unbound, the overlay infrastructure is basically radio wave-oriented, meaning that the signals are transmitted over the air to devices that can receive and send the same kind of signals. Mobile communication is the framework that powers mobile computing technology. It ensures that communication process is consistent and reliable through this framework. The mobile communication framework includes communication devices such as mobile phones, laptops, rules of conduct, etc. These are responsible for providing a smooth communication process.

## 6. MOBILE HARDWARE

A mobile hardware device is a small, portable computing device that is capable of receiving and processing data. Mobile hardware devices such as Smartphone’s, handhelds and wearable fall under the category of mobile hardware. These devices usually have an operating system (OS) built into them and are capable of running application software on them. Mobile hardware devices are equipped with sensors, full duplex data transmission and are capable of operating on wireless networks such as infrared, WiFi and Bluetooth.

## 7. MOBILE SOFTWARE

Mobile Software is a software program that is specially designed to run on mobile hardware. It is typically the operating system of a mobile device. These operating systems offer features such as a touch screen, cellular connection, Bluetooth and Wi-Fi connectivity, GPS mobile navigations, camera (video camera), speech recognition, voice recorders, music players, near field communications and sensors. These sensors and other hardware elements can be accessed through the operating system.

## 8. PRINCIPLES OF MOBILE COMPUTING

### 8.1 PORTABILITY

Mobile computing systems are composed of devices/nodes that are connected within a mobile computing system. Devices/nodes may have a limited number of device capabilities and a limited amount of power supply, but they should have enough processing power and physical transport to work in a mobile environment.

### 8.2 CONNECTIVITY

This defines the Quality of Service (QoS) of the network connectivity. In a mobile computing system, the network availability is expected to be maintained at a high level with the minimal amount of lag/downtime without being affected by the mobility of the connected nodes.

### 8.3 INTERACTIVITY

Nodes in a mobile computing ecosystem are connected to each other to communicate and work together through active data flows.

### 8.4 INDIVIDUALITY

A phone, tablet, or other device that's connected to a cellular network often identifies you as a person; your mobile computing system needs to be able to take advantage of this technology to meet your needs and also get info about each individual node.

## 9. APPLICATIONS OF MOBILE COMPUTING

Mobile computing can be used for a variety of things, like web or internet access, GPS, emergency services, entertainment, and education**.**

## 10. ADVANTAGES OF THIS COMPUTING

Mobile computing offers a range of benefits, including the freedom of the consumer from being tied to a physical location, the increased bandwidth and speed of transmission, the ability to work at the flow into, the potential to save time and increase productivity with a better Return on Investment (ROI), the ability to provide entertainment, news and statistics on the go with the stream of facts, videos and audio, the streamlining of enterprise processes through mobility, the reduction of cumbersome emails, paper processing and delays in communication and transmission, the emergence of new hobby opportunities for IT professionals, and the addition of an added company to the portfolio, all of which will continue to grow in line with the indicative mobile computing trends, as well as enhanced safety.

## 11. DISADVANTAGES OF MOBILE COMPUTING

Battery usage is a problem, there's interference in the shielding, there's not enough bandwidth when sending data, there's connection loss all over the network, there's network stability, there's an interoperability issue, and there's protection issues.

## 12. CONCLUSION

Mobile computing is all about human-computer communication, where a computer is usually carried at the same time as people, which makes it possible to send data, voice, and video. It holds mobile communication, hardware, and software. Nowadays, it's easy to work from anywhere, as long as you have a good connection and security. Plus, with the rise of high speed internet, mobile computing has become even more popular. It's an ever-evolving technology that will soon become a key part of computing, and information and means technology. This article will give you an overview of mobile computing, explain how it's growing, and where it's going in the future, and explain the security issues that come with it.

## REFERENCES

1. Imielinski, T and Korth, H. F. (Eds.). (1996). Mobile computing (Vol. 353). Springer Science & Business Media.
2. Sachin Raj Saxena and Dr. Krishan Pal, (2019). “Mobile Computing: Review”. International Journal of Research Studies in Computer Science and Engineering (IJRSCSE), 6(4), pp.1-6. http://dx.doi.org/10.20431/2349-4859.0604001.
3. Satyanarayanan, M. (2011). Mobile computing: the next decade. ACM SIGMOBILE Mobile Computing and Communications Review, 15(2), 2-10.
4. Singh, M. and Gupta, E. E. (2017). Mobile Computing-A Review. International Journal of Engineering and Management Research (IJEMR), 7(3), 108-110.
5. The data is retrieved from file:///C:/Users/Administrator/Downloads/mobile-computing-leture-Ammar.pdf
6. Umair, S., Muneer, U., Zahoor, M. N. and Malik, A. W. (2015, December). Mobile computing: issues and challenges. In 2015 12th international conference on high-capacity optical networks and enabling/emerging technologies (HONET) (pp. 1-5). IEEE.
7. Vaidehi Tejas Gonjare (2019), Mobile Computing and its Applications. International Research Journal of Engineering and Technology (IRJET), 6(12), 442-444
8. Zaslavsky, A. and Tari, Z. (1998). Mobile computing : overview and current status. Journal of Research and Practice in Information Technology, 30(2), 42–52

# STUDY OF FUTURE 6G AND 5G ACCESS NETWORKS FOR INTERNET OF THINGS APPLICATIONS

**P. Tamilselvan**
Assistant Professor, PG and Research Department of Computer Science, Kaamadhenu Arts and Science College, Sathyamangalam, Tamilnadu, India
**P. Gomathi**
Assistant Professor, PG and Research Department of Computer Science, Kaamadhenu Arts and Science College, Sathyamangalam, Tamilnadu, India
**K. Thulasimani**
Assistant Professor, PG and Research Department of Computer Science, Kaamadhenu Arts and Science College, Sathyamangalam, Tamilnadu, India

**Abstract:**

**The utilisation of modern 5G and upcoming 6G Internet of Things (IoT) protocols/standards, applications, and access networks is thoroughly reviewed in this article. First, to make it easier for readers to comprehend and compare present and future Internet of Things technologies, the majority of IoT protocols/standards and application scenarios are summarised in the form of tables, illustrations, and diagrams. The analysis and discussion of the terrestrial and aerial radio access networks follows. A brief description of the development of 5G terrestrial access networks is provided, along with a quantitative analysis and discussion of its performance constraints. The terrestrial radio access network must contend with substantial route loss brought on by weather conditions such oxygen and water vapour absorption in the atmosphere, rainfall, and cloud/fog attenuation as the operating frequency approaches the sub-millimeter wave range. In order to address the coverage and path loss issues, aerial radio access networks are being developed in anticipation of 6G IoT. The aerial radio access designs and infrastructure are also examined in this assessment. This survey seeks to help readers gain a better understanding of the technical status of 5G IoT and the key performance parameters that must be met in order to transition to 6G IoT in the future.**

Keywords: IoT, or the internet of things Wireless technologies of the fifth (5G) and sixth (6G) generations, LoRaWAN, a long-range wide-area network Millimetre wave (mmWave), sub- millimeter wave, and sub-6 GHz RAN (radio access network)

## 1. INTRODUCTION

The phrase "Internet of Things" started to be used in academic circles and business circles around 2008 and 2009. The "Internet of Things" was first described as [1]: "A global infrastructure for the information society, enabling advanced services by interconnecting (physical and virtual) things based on existing and developing interoperable information and communication technologies." The term is now commonly abbreviated as "IoT."

As illustrated in Fig. 1(a), there are currently eight key sectors where the IoT is being employed, namely smart wearables, smart homes, smart cities, smart agriculture, smart vehicles, health care, industry automation, and smart energy. The development of 5G communication technologies has extended the application of the IoT. Beginning explore on top of the 6G right of entry network have previously in progress, through the intend of in service at frequencies over 100 GHz to 3 THz, on a point in time whilst 5G expertise be human being engaged. For example, a number of research along with commendation pertaining to 6G road and rail network enclose be widely available since 2019, with awake toward a thousand papers being recorded [7]-[33].

Early 6G research will proceed concurrently with the development of 5G, with the first 6G research items anticipated to debut in 2025 and standardisation work beginning in 2030 [29]. In fact, so-called terahertz technology—a type of technology through in service frequencies over 300 to 3 —has only recently appeared and is currently in use. Due en route for the nonionizing sub-millimeter waves' ability to penetrate a wide

range of non-conducting materials, tez knowledge be extensively used during the fields of medicine, with safekeeping, including based isolated sense along with medicinal indicative imaging.

Additionally, appropriate to the millimeter impression brief wave distance end to end, which ranges from 0.1, and their tremendous compassion to tiny ecological change. since of this, millimeter impression be perfect for elevated compassion sense and imaging application, counting biosensors, impressive monitor, workers scanning, cloth width discovery, and nano fabric description plans, as shown.

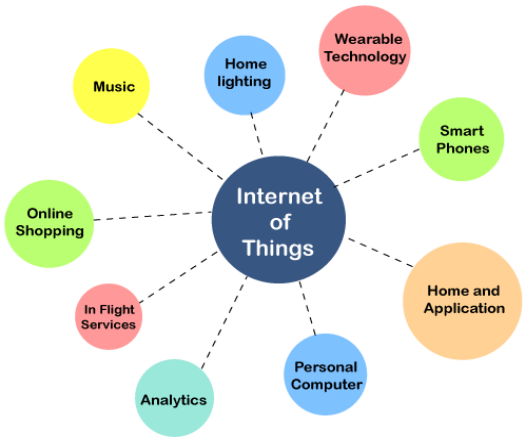



**Fig 1. (a) IoT**

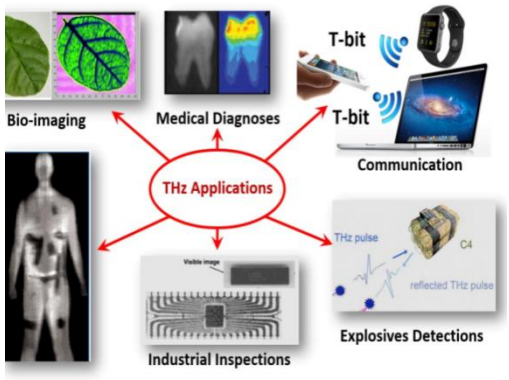



**Fig 2. (b) Terahertz**

The 6G Internet of Things (IoT) will be created at pardon? time these terahertz bottom plans be attached to the outlook internet complex with the aid of AI skin texture. statistics charge of 0.1–1 Tbps, 3–60 bps/Hz spectrum efficiency, 1000 km mobility, 107 devices/km2 of connectivity thickness, and complete automation are all targets of the 6G IoT. As a result, the research for the upcoming 6G IoT could focus on four main technological areas: (1) raising the operating frequencies of the internet access network to the terahertz or sub-millimeter wave band, (2) enhancing multi-antenna technique, (3) implementing artificial intelligence (AI) and machine learning (ML) in the 6 G IoT, and (4) integrate transportation by extra sense capability.

The growth of the sub-millimeter wave monolithic microwave integrated circuits (MMICs) market will have a direct impact on the advancement of 6G IoT research. High sensitivity, high output RF power, high thermal stability, nanoscale integrated circuits, broadband, and low cost are additional requirements for MMICs utilised in 6G IoT in addition to high operating frequencies (> 300 GHz). In actuality, 5G with 6G IoT are very broad and complex knowledge examine areas, including the topic of modulation methods engaged, IoT building, IoT industry model, IoT system sanctuary, and sequence power expenditure difficulty. simply three main IoT sub topics IoT protocol and standards, 5G and 6G IoT application categories, and their corresponding right of entry network are reviewed and examined in this study because they are interconnected.

The third associate theme of 5G and 6G right of entry network, on the other hand, has been thoroughly highlighted and reviewed. For future IoT efficiency and demand, the 6G access networks try to run at frequencies up to 3 THz. However, there are a number of difficulties that traditional terrestrial radio access networks (TRAN) running at frequencies up to terahertz must overcome, including severe route loss problems, energy consumption, and equipment expenses. Currently, a key problem that needs to be fixed is the loss of the network signal up to THz in outdoor applications. Due to the increased route loss of wireless signals at working frequencies > 275 GHz, aerial radio access networks (ARAN) have been suggested and used to replace TRAN.

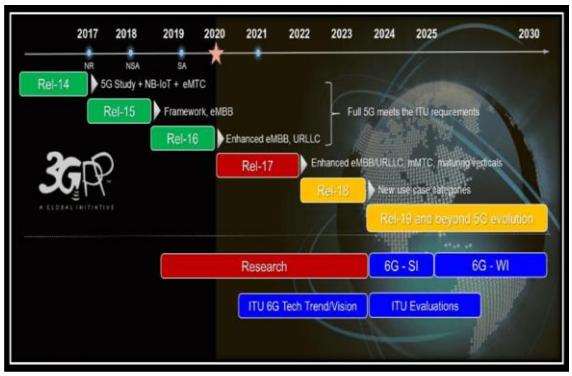
## I. IOT ACCESS STANDARDS AND PROTOCOLS

Wireless technology has allowed the simultaneous usage of several types and applications of IoT as well as human-to-human and device-to-device (D2D) communication in the same setting. This could result in voice quality degradation, data loss, increased interference

and coupling effects between wireless signals, and a decrease in the device's working range and battery life. It is difficult to give each user, device, and application its own frequency spectrum. Because of this, different wireless protocols and standards make use of a range of communication strategies to aid in the peaceful coexistence of communication operations using the same frequency ranges.

As shown in Table 1, there are three categories of coverage distance range applications that can be applied to IoT technology. IoT users can choose a suitable wireless access protocol based on its baud rate, energy consumption, and cost necessary for certain IoT applications in addition to the coverage distance. Currently, Wireless reliability and Long-range wide-area network are the two unlicensed technologies most frequently utilised in the Internet of Things [36]. In fact, Wi-Fi is the undisputed leader in broadband communication, and Table 2 lists its generations and protocols.

**Fig 3. The evolution of 5G**




## II. IOT APPLICATION CATEGORIES FOR 5G AND 6G

The three primary applications for fifth-generation (5G) wireless technology, Enhanced Mobile Broadband (eMBB), Ultra-reliable and Low-latency Communications (uRLLC), and Massive (HD) films, virtual reality (VR), augmented reality (AR), and broadband IoT, in order to meet people's desire for an increasingly digital lifestyle.

Machine Type Communications (mMTC), have been established by the International Telecommunication Union (ITU). The eMBB focuses on services with high bandwidth requirements, such as high definition

IoT hardware can now benefit from the latest WiFi protocol generation (Wi-Fi 6), which has improved set presentation, improved outside action, and increased range. LoRaWAN, on the added tender, is utilised for low-power, long choice connection. However, many IoT connectivity technologies—including Extended Coverage GSM IoT (EC-GSM-IoT), Long Term Evolution-Machine Type Communication (LTE-M), Narrowband Internet of Things (NB-IoT), 4G LTE, and 5G New Radio (NR)—continue to rely on cellular networks and licenced spectrum under 3GPP access standards. The major applications of 3GPP standards are long-range high-quality mobile voice and data services, as opposed to non-3GPP standards (LoraWAN and SixFog).

Future 3GPP 6G standardisation is anticipated to begin in 2025 [29]. The short 3GPP standards roadmap for 5G, 6G, and related technologies is depicted in Fig. 2. The International Telecommunication Union (ITU) and International Mobile Communications (IMT-2020) are two other significant international organisations that set 5G technological goals in addition to the 5G specifications established by 3GPP [39]. The foundation of Electrical as well as Electronics Engineers and use millimeter beckon unlicensed band because the 5G technological values, such as IEEE principles, to attain information charge on or after 6.7 Gbps to 20 Gbps, as exposed in table 1.

The uRLLC and mMTC, on the other hand, are suited for vast and critical IoT, respectively. The term "uRLLC" refers to the use of a network for activities like remote management and assisted and automated driving that call for constant and reliable data interchange. As shown in Table 2 [46], the mMTC is used to link to a large number of low-cost, low-power gadgets that are highly scalable and have longer battery lives. Examples of these areas include smart cities and smart agriculture. Figures 4, 5, and 6 depict the applications of the eMBB, uRLLC, and mMTC scenarios, respectively. Table 3 [45] contains a list of the industrial IoT category, applications, technologies, and protocols that are currently available.

As was already indicated, the use of internet-connected gadgets will rise and eventually reach the "Internet of Everything (IoE)" stage. The three 5G scenarios must therefore be enhanced and divided into more sub-scenarios for 6G application scenarios that cover more varied uses. Several 6G application scenarios have been proposed for 2019. Additionally, improved the three 5G tradition state that might be used for 6G applications and gave them the new names uMUB, uHSLLC, and uHDD. While CAeC, EDuRLLC, and COC were identified as 6G usage scenarios by Letaief et al. (2019) [8], they are mentioned in Table 4. Additionally, Table 5 compares the specs, performances, and application requirements for 5G and 6G.

## III.IOT ACCESS NETWORKS IN 5G AND 6G

Although 5G communication can offer a peak data rate, C, of 20 Gbps, the performance of data rate will be significantly decreased (may be decreased by 80% above) (typical 0.1 Gbps of experienced data rate) when the capacity needs to be shared with multiple users and the numerous IoT devices. In light of the ongoing rise in the number of connected IoT devices, it is necessary to take into account the higher intended peak data rate, C, which surpasses the anticipated value of 20 Gbps, and to raise the data rate cap.

According to the Shannon capacity theorem, the channel bandwidth, B, and the received signal-to-noise ratio, SNR have an impact on the peak data rate, C, as follows:

$$C = (n \times BW) \log_2 \left(1 + \frac{S}{N}\right)$$

where n and BW are, respectively, the channel count and bandwidth (in Hertz) of the channel. On the other side, S and N stand for, respectively, the transmit power and noise on the channel (measured in Watts). It is obvious that in order to improve the data rate, C, the channel bandwidth, BW, the number of channels, n, and the power transmission, S, noise on the channel, N, must be minimised. It is necessary to use high operating frequencies up to the mmWave range (> 24 GHz) in order to obtain bandwidth with a BW of more than 800 MHz [40]. A particular percentage of bandwidth offers a larger share of the spectrum as the frequency rises.

### 3.1.TERRESTRIAL RADIO ACCESS NETWORKS (TRAN)

As depicted in Fig. 7, the evolution of the 5G wireless access network moves from the macrocell environment to the small cell coverage region. In actuality, the small cell is a scaled-down variation of the conventional macrocell, which shares all of its traits and properties [48]. The small cell capability, however, is appropriate for 5G deployments, which promise millisecond latency, a million devices per square mile, and extremely high data throughput. The hardware implementation is quicker, simpler, and uses less power (extending battery life) due to the small cell hardware units' reduced complexity.

Due to high loss building materials, such as low emissivity glass, metal, and concrete, indoor wireless access signal performance will deteriorate inside of the building, especially in big buildings with several rooms. Therefore, the distributed antenna system (DAS) is recommended to address the problems with indoor signal degradation by dispersing the available external cell signals over the network of small antennas installed throughout the building to achieve perfect coverage. Additionally, tiny cells and distributed antenna system (DAS) solutions can support a variety of protocols, including 3G/4G cellular and LTE Advanced (LTE-A) systems that use carrier aggregation [48]. Three categories exist.

### 3.2.AERIAL RADIO ACCESS NETWORKS (ARAN)

Non-terrestrial networks, such as unmanned aerial vehicle (UAV) access networks, are currently emerging as an alternative method to address the high loss caused by atmospheric oxygen and water vapour for the transmitted signal of IoT devices up to THz, as shown in Fig. 10. Line-by-line modelling is used to determine attenuation values, which are stated as functions of operating frequency, f, water vapour density,, and ambient temperature, T, pressure, and temperature [55, 56]. In Table 8, the link between the variables (T, P, and ) in the line-by-line model and the sea level height, h, is listed.

Equations suggested by ITU (2017) [57] are used to compute the values of T, P, and. As can be observed from Fig. 10, the attenuation, which is anticipated to decrease due to the oxygen content and water vapour density, will decrease as the height above sea level, h, increases. It should be mentioned that the attenuation, A calculated in Fig. 10 only takes into account the normal atmospheric temperature, T in middle-latitude (22o oN 45o).

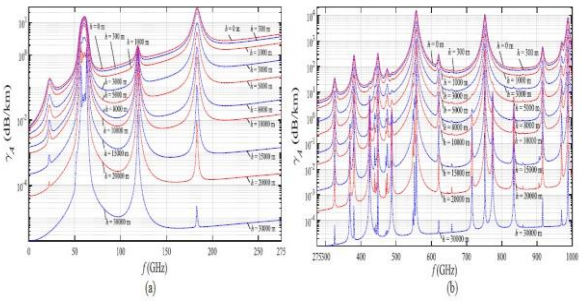


Fig. 10. Specific attenuation, $\gamma_a$ in unit dB/km due to atmospheric gases changes with height above sea level, $h$ from (a) 1 GHz to 275 GHz and (b) 275 GHz to 1 THz in middle-latitude ($22° \leq °N \leq 45°$).

## IV. CONCLUSION

In reality, there are a huge number of wireless standards and protocols that are already in use and that overlap with one another. In the future. As a result, it is challenging to address each technique and standard in a single publication in detail and with clarity Furthermore, it is easier said than done to classify protocols and standards in great element because these guidelines are promulgate with numerous telecommunications committee and international standardization organisations [such as the US Federal transportation charge (FCC), International Telecommunication Union (ITU), European Electronic Communications Committee (ECC), Institute of Electrical and Electronics Engineers (IEEE), and 3rd Generation Partnership Project].

However, in order for readers to compare recent 5G and upcoming 6G IoT, as well as to comprehend their standard protocols, performance, specifications, and applications more easily and quickly, this paper strives to summarise a large number of information and display them in the form of tables and figures. The utilisation scenario for 6G is enlarged from the 5G and classified into more specific applications to infer the future potential, research direction, and key performance indicators (KPIs) for 6G [66].

The rapid advancement of artificial intelligence (AI) and machine learning (ML) will also be indirectly impacted by these new developments in the Internet of Things (IoT), as we are currently and the majority of industry in commission structure be base on effusive usual functional networks [67]–[69]. Large IoT networks also require a lot of electricity to run, therefore recent advances in the study of energy sustainability and energy harvesting have become crucial.

Worldly means of announcement admittance network and aerial radio access networks is the primary emphasis of this article. Future obstacles and difficulties have been explored in terms of attenuation brought on by air oxygen, water vapour, precipitation, clouds, or snow. Future IoT access will be covered by the combo, which will be widely deployed to optimise coverage (>99%).

## REFERENCES

[1] S "6G technologies: key drivers, core requirements, system architectures, and enabling technologies," IEEE Vehicular Technology Magazine, vol. 14, pp. 18–27, 2019.

[2] "The shift to 6G communications: vision and requirements," Human-centric Computing and Information Sciences, vol. 10, pp. 1-27, 2020. M. W. Akhtar, S. A. Hassan, R. Ghaffar, H. J. Jung, S. Garg, and M. S. Hossain.

[3] "Sixth generation (6G) wireless networks: vision, research activities, challenges and potential solutions," Symmetry, vol. 12, pp. 1-27, 2020.

[4] "6G and beyond: the future of wireless communications systems," IEEE Access, vol. 8, 2020, pp. 44983-44998. Ian F. Akyildiz, Ahan Kak, and Shuai Nie

[5] "Survey on aerial radio access networks: Towards a comprehensive 6G access infrastructure," IEEE Communications Surveys & Tutorials, vol. 23, pp. 1193–1225, 2021. N. N. Dao, Q. V. Pham, N. H. Tu, T. T. Thanh, V. N. Q. Bao, D. S. Lakew, and S. Cho.

[6] C. De Alwis et al., "Survey on 6G frontiers: trends, applications, requirements, technologies and future research," IEEE Open Journal of the Communications Society, vol. 2, 2021, pp. 836–886.

[7] "6G internet of things: a comprehensive survey," IEEE Internet of Things Journal, vol. 9, pp. 359-383, 2022. Dinh, C. Nguyen, Ming Ding, Pubudu, N. Pathirana,

Aruna Seneviratne, Jun Li, Dusit Niyato, Octavia Dobre, and H. Vincent Poor.

[8] A. Dáz Zayas, F. J. Rivas Tocado, and P. Rodrguez, "Evolution and testing of NB-IoT solutions," Applied Sciences, vol. 10, no. 7903 (1–17), 2020.

[9] "3GPP 5G new radio system with the ITU IMT-2020 requirements," IEEE Access, vol. 8, pp. 42828-42840, 2020. S. Henry, A. Alsohaily, and E. S. Sousa.

[10] Position on 5G Spectrum from Huawei. Huawei Technologies Co., Ltd., Shenzhen, China, 2017.

[11] "Internet of mobile things: overview of LoRaWAN, DASH7, and NB-IoT in LPWANs standards and supported mobility," IEEE Communications Surveys & Tutorials, vol. 21, no. 5, 2019, pp. 1561–1581. W. Ayoub, A. Samhat, F. Nouvel, M. Mroue, and J-C. Prévotet.

INTERNATIONAL JOURNAL OF INNOVATIVE TECHNOLOGY AND CREATIVE ENGINEERING (ISSN:2045-8711)
VOL.13 NO.08 AUG 2023

# GDSR POLICY LEARNING FOR TARGET OBJECT TRACKING IN WSN

**Dr.P.Suresh Babu**
Associate Professor & Head, Department of Computer Science,
Bharathidasan College of Arts and Science, Erode, Tamilnadu, India
**S.Vasanthkumar**
Research Scholar, Department of Computer Science,
Bharathidasan College of Arts and Science, Erode, Tamilnadu, India

**Abstract -A wireless sensor network (WSN) is a network of small devices, called sensors, used for sensing and monitoring the environmental conditions at various locations through wireless links. Target object tracking is the most important application of WSN where energy conservation plays an important role. Several methods have been developed for target object tracking in WSN with minimal energy consumption. But, the accuracy level was not increased by existing tracking techniques. In order to address these problems, a novel targets object tracking method called Energy Aware Gaussian Distributive Sammon Regularization Policy Learning (GDSRPL) for efficient target object tracking in WSN. The GDSRPL method consists of two major processes namely reference node selection, target detection. First, the energy of the sensor nodes is measured. Then apply Wilcoxon rank-sum test for identifying the reference node based on higher residual energy. When the target node arrived in the network, the reference node transmits the beacon message to all sensor nodes for detecting the target node's location. The sensor node transmits the target information to the base station through the reference nodes. After that, the base station performs target detection by using Akaike Gaussian distributive Generalized Tikhonov Regularization Analysis with the sensed data. In this way, energy-efficient target object tracking is accurately performed in WSN. Experimental evaluation is carried out on factors such as energy consumption, target object tracking accuracy with respect to a different number of sensor nodes. The performance results and discussion reveals that the proposed GDSRPL method efficiently increases the target tracking accuracy and minimizes the energy consumption than the existing techniques.**

***Keywords:*** WSN, target object tracking in WSN, Wilcoxon rank-sum test, Akaike Gaussian distributive Generalized Tikhonov Regularization Analysis, Sammon projective on-policy learning algorithm

## I. INTRODUCTION

Target tracking in WSN is to detect the occurrence of a target and track it continuously. Moving target tracking is a major application of WSNs. Localization and tracking of moving targets mainly depends on energy efficiency. However, moving object tracking consumes much more energy to perform high-frequency sensing and data transmission.

Therefore, a great challenge is energy-efficient target tracking in WSNs. To save energy and extend the lifetime of networks while tracking a moving object effectively, this paper proposes a novel machine learning method that predicts the trajectory of the moving object.

A sequence-to-sequence learning model (Seq2Seq) was developed in [1] to predict the trajectory of a moving object and minimize the computation time. However, the accuracy of target prediction was not improved. A lightweight Auto Regressive Neural Network (ARNN) was developed in [2] for accurate and energy-efficient target tracking. But the time consumption of target tracking was not minimized.

A novel energy-efficient management approach was developed in [3] for target tracking. But it was not efficient to prolong its network lifetime for target tracking. However, it failed to focus on the energy harvest tracking network. Target Detection and Target Tracking (TDTT) model was developed in [5] for continuously tracking the target objects and trajectories. But it failed to track the multiple targets in the sensing region.A partition-based target tracking method was developed in [6] to achieve flexible and energy-efficient tracking. However, the time consumption of energy-efficient tracking was not minimized.

## II.RELATED WORKS

A three-dimensional space target tracking method was developed in [5] for minimizing the error rate of observation. But, the multiple target detection was not performed. A reliable multi-object tracking model was introduced in [1] with the deep learning approach. However, the energy efficiency was not improved in the multi-object target tracking. An enhanced least-square algorithm based on improved Bayesian was developed in [3] for tracking the targets. However, it failed to consider multi-target localization and tracking.

A hybrid filtering algorithm was designed in [8] for multi-target tracking and detection based on WSN. However, the performance of the error rate was not minimized. The multi-target localization and tracking method was introduced in [4] by eliminating abnormal measurements. But the accuracy and stability of target localization and tracking were not improved.

### B. ORGANIZATION OF PAPER

The paper includes four different sections. Section 1, discusses some related energy-efficient object tracking frameworks. In section 2, a brief discussion about the GDSRPL method is presented with a neat diagram. The performance results of this method are discussed insection3.Finally, Section 4 provides the conclusion of the paper.

## II. PROPOSED METHODOLOGY

The main objective of target tracking in WSN is to locate and monitor the movement of the target continuously. The target tracking is carried out by the means of various sensor nodes deployed in the sensing environment. The deployed sensor nodes are typically small in size and are equipped with low-battery power. Therefore energy management is the most significant issue in numerous applications of WSN, especially in target tracking. The higher energy consumption of the dynamic nodes reduces the network lifetime. Therefore, energy-efficient target tracking is a major challenge in WSN to enhance the lifetime of the network. Based on this motivation, a novel machine learning technique called GDSRPL is introduced to track moving targets and saved energy to expand the lifetime of the WSN in the centralized structure.

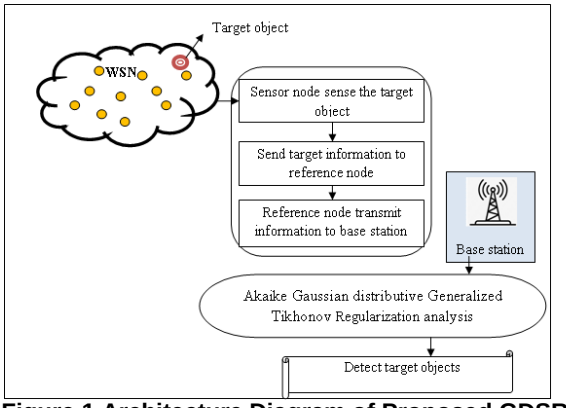



**Figure 1 Architecture Diagram of Proposed GDSRPL Method**

Figure 1 illustrates the basic architecture diagram of the proposed GDSRPL method for target object tracking in WSN. First, the number of sensor nodes $Sn_i \in Sn_1, Sn_2 \ldots Sn_n$ are distributed in a squared network within the transmission range '$T_r$'. In the proposed technique, the energy of all the sensor nodes '$E(SN_i)$' is measured. After that, the Wilcoxon rank-sum statistic analysis is performed to identify the energy-efficient nodes called reference nodes in the network. Wilcoxon rank-sum test is a statistical test used to analyze the relationship between two statistical data (i.e. Energy of the sensor nodes and threshold value). The selected reference node sends the beacon message to the entire sensor node for identifying the target node location in the WSN.

When a target object arrived in a network, the nearby sensor node senses and monitors the target node and transmits the information to the higher energy reference nodes. The reference node receives the sensed information from the sensor node and transmits it to the base station for target tracking. After that, the base station uses the Gaussian distributive

Generalized Tikhonov Regularization analysis with the sensed data to find the target object location with minimum energy consumption. These two processes of the proposed GDSRPL method are briefly explained in the following subsections.

### A. WILCOXONRANK-SUMTEST-BASED REFERENCE NODE SELECTION

The WSN consists of many sensor nodes distributed in the sensing region. However, the entire sensor node does not contribute to the overall process of target tracking due to higher energy consumption. Therefore, the proposed GDSRPL method selects the energy-efficient nodes for target tracking with the help of the Wilcoxon rank-sum test. The Wilcoxon rank-sum test is the statistical process that helps to measure the relationship between the energy of the sensor nodes for selecting the reference node in a distributed network. In the case of WSN, the reference node is one having the maximum energy located in a specified region.

First, the residual energy of the sensor nodes is measured. After that, Wilcox on rank-sum test 'W_t' is applied for measuring the relationship between the residual energy and threshold value. If the correlation returns '1', then the node is selected as the reference node. Otherwise, the energy level of the other sensor node is analyzed. In this way, energy-efficient reference nodes are selected for energy-efficient target tracking in WSN.

### GDGENERALIZED TIKHONOV REGULARIZATION ANALYSIS BASED TARGET DETECTION IN WSN

The second process of the proposed GDSRPL method is to perform the target detection by identifying the position of the target with help of Akaike Gaussian distributive Generalized Tikhonov Regularization Analysis. The Akaike information criterion is applied to identify the probability of target detection by determining the best fit for a given set of data. Then Generalized Tikhonov Regularization Analysis is used to minimize the problem of multicollinearity in linear regression with large numbers of parameters. The multicollinearity refers to a condition in which more than two explanatory variables (i.e. independent variables) are linearly correlated. The independent variables are represented as input data (i.e. target information) that helps to perform target detection.

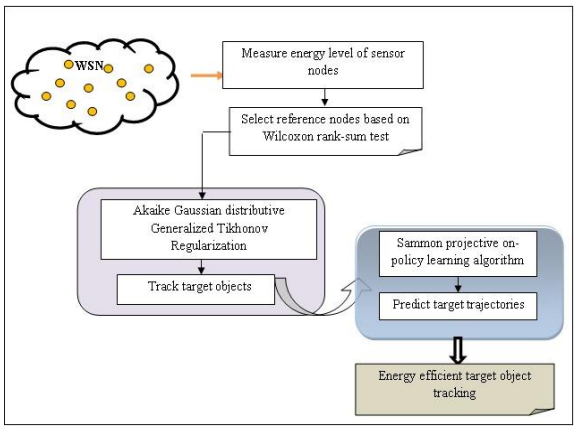



**Figure 2 Akaike Gaussian Distributive Generalized Tikhonov Regularization-based target Tracking**

Figure 2 illustrates the flow process of Akaike Gaussian distributive Generalized Tikhonov Regularization analysis for target tracking in WSN. When the target is entered into the network, the selected reference nodes first transmit the beacon message to the nearest sensor nodes in the network for identifying the target node's location. The nearest sensor node monitors the target movement within the sensing range. The sensing range is defined as the coverage region of a sensor node in meters (m). The sensor node finds the nearest reference node to send the target information using Manhattan distance.

## III. PERFORMANCE RESULTS AND DISCUSSION

The simulation results of the GDSRPL technique and Seq2Seq model [1] and ARNN [2] are discussed in this section with different performance metrics such as energy consumption, target tracking accuracy, with respect to a number of sensor nodes. The performance of proposed and existing methods is discussed with help of a table and graph.

### A. IMPACT OF ENERGY CONSUMPTION

Energy is the most significant metric to enhance the lifetime of the network. The minimum energy consumption of the sensor node enhances the network lifetime and performs accurate target tracking. Energy consumption is referred to as the amount of energy consumed by the sensor nodes to perform the target tracking in the network.The formula for calculating the energy consumption is given below

$$Conp_E = n * E_c\ (SSn)$$

Where, 〖Conp〗_E indicates energy consumption, n denotes the number of sensor nodes, E_c indicates energy consumed by the single sensor nodes (SSn). The performance of energy consumption is measured in terms of a joule (J).

**Table 1 Comparison of Energy Consumption**

| Number of Sensor Nodes | Energy Consumption (%) | | |
|---|---|---|---|
| | GDSRPL | Seq2Seq model | ARNN |
| 50 | **21.5** | **24** | **27.5** |
| 100 | **24** | **27** | **30** |
| 150 | **28.5** | **34.5** | **37.5** |
| 200 | **32** | **36** | **40** |
| 250 | **37.5** | **40** | **43.75** |

Table 1report the performance assessment results of energy consumption with a number of sensor nodes using three different methods GDSRPL method and Seq2Seq model [1] and ARNN [2]. For the simulation purposes, the number of sensor nodes is taken in the ranges from 50 to 250. For each method, five different results are observed with respect to the number of sensor nodes. The overall comparison results confirm that the performance of energy consumption using the GDSRPL method is comparatively reduced by 10% when compared to [1] and 17% when compared to [2] respectively. The simulation results are plotted in the graph.

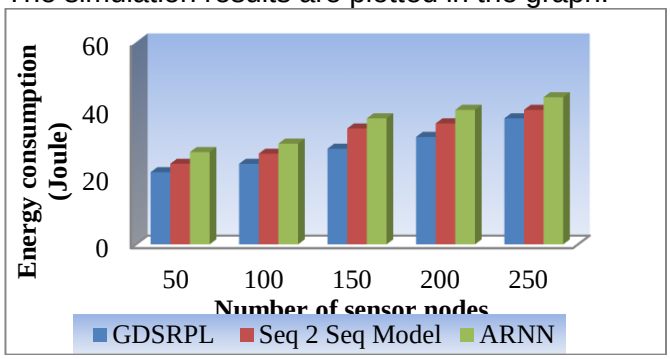



**Figure 3 Graphical Representation of Energy Consumption**

Figure 3 reveals the performance analysis of energy consumption against a number of sensor nodes using three different methods namely the GDSRPL method and existing [1] and [2].

As shown in figure 5, the energy consumption of all three methods gets increased while increasing the number of sensor nodes. But comparatively, the GDSRPL method decreases the performance of energy consumption when compared to the two existing methods.

## B. IMPACT OF TARGET TRACKING ACCURACY

The accuracy of target tracking is calculated based on the number of (No. of) sensor nodes that correctly transmit the target object information to the base station. Therefore, the accuracy is mathematically calculated as given below,

$$Tar_{Trac}Tar\_Trac_{Acc} = \left(\frac{No.ofSns\ correctlytransmitsthe\ information}{n}\right) * 100$$

Where, 〖Tar_Trac〗_Acc denotes target tracking accuracy, n denotes the number of sensor nodes. It is measured in terms of percentage (%).

**Table 2 Comparison of Target Tracking Accuracy**

| Number of sensor nodes | Target Tracking Accuracy (%) | | |
|---|---|---|---|
| | GDSRPL | Seq2Seq model | ARNN |
| 50 | 94 | 90 | 88 |
| 100 | **93** | **89** | **87** |
| 150 | 92.66 | 88.66 | 86.66 |
| 200 | **92.5** | **88.5** | **86.5** |
| 250 | 92.4 | 88 | 86 |

Table 2 reports the simulation results of the target tracking accuracy versus the number of sensor nodes. The number of sensor nodes considered for simulation varied from 50,100 …250. The table value indicates the performance results of target tracking accuracy using three methods GDSRPL method and Seq2Seq model [1] and ARNN [2]. Among the three methods, the GDSRPL outperforms well in terms of achieving higher accuracy. With the consideration of 50 sensor nodes in the first iteration, the observed target tracking accuracy is94%. Followed by, nine different results are obtained for each method.

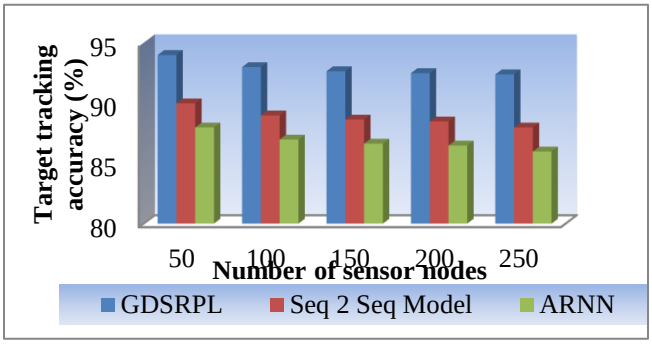



**Figure 4 Graphical representation of Target tracking accuracy**

The above figure 4 illustrates the performance of target tracking accuracy with respect to a number of sensor nodes. As shown in the graph, the performance of the tracking accuracy is observed at the y-axis and the number of sensor nodes on the x-axis. The graphical results indicate that the performance of the GDSRPL provides improved performance than the other two existing methods [1] [2]

## 4. CONCLUSION

In this paper, a novel GDSRPL method is developed for accurately tracking a target with minimum time. First, the energy-efficient sensor nodes are identified for increasing the network lifetime. After selecting the energy-efficient nodes, the target object detection process is carried out at the base station with help of Akaike Gaussian distributive Generalized Tikhonov Regularization Analysis. Finally, the base station predicts the target trajectory by means of the Sammon projective on-policy learning algorithm based on state transition property. In this way, energy-efficient target tracking is performed in WSN. Experimental evaluation is carried out on factors such as energy consumption, target object tracking accuracy and target object tracking time, and prediction Error with respect to a different number of sensor nodes. The simulation is conducted on various performance metrics such as energy consumption, target tracking accuracy, target tracking time, and prediction error. The discussed results have revealed that the GDSRPL technique has considerably improved the target tracking accuracy with minimum time as well as error rate than the existing methods.

## REFERENCES

[1]. Xiaofei Cao and Sanjay Madria, “An efficient moving object tracking framework for WSNs using sequence-to-sequence learning model”, Peer-to-Peer Networking and Applications, Springer, Volume 14, 2021, Pages 3915–3928. https://doi.org/10.1007/s12083-021-01157-8

[2]. Jayesh Munjani and Maulin Joshi, “A non-conventional lightweight Auto Regressive Neural Network for accurate and energy efficient target tracking in Wireless Sensor Network”, ISA Transactions, Elsevier, Volume 115, September 2021, Pages 12-31. https://doi.org/10.1016/j.isatra.2021.01.021

[3]. Ce Pang, Gong-Guo Xu, Gan-lin Shan, Yun-Pu Zhang, “A new energy-efficient management approach for wireless sensor networks in target tracking”, Defence Technology, Elsevier, Volume 17, Issue 3, 2021, Pages 932-947. https://doi.org/10.1016/j.dt.2020.05.022

[4]. Zhiyi Qu, Huihui Xu, Xue Zhao, Hongying Tang, Jiang Wang, Baoqing Li, “A fault-tolerant sensor scheduling approach for target tracking in wireless sensor networks”, Alexandria Engineering Journal, Elsevier, Volume 61, 2022, Pages 13001–13010. https://doi.org/10.1016/j.aej.2022.07.008

[5]. P. Leela Rani and G. A. Sathish Kumar, “Detecting Anonymous Target and Predicting Target Trajectories in Wireless Sensor Networks”, Symmetry, Volume 19, 2021, Pages 1-22. https://doi.org/10.3390/sym13040719

[6]. Tingyang Wei, Jinghui Zhong, Jun Zhang, “An Energy-Efficient Partition-Based Framework With Continuous Ant Colony Optimization for Target Tracking in Mobile Sensor Networks” IEEE Transactions on Emerging Topics in Computational Intelligence, Volume 5, Issue 4, 2021, Pages 700 – 713. DOI: 10.1109/TETCI.2019.2940978

[7]. Seyoung Kang, Taehyun Kim, and Wonzoo Chung, “Multi-Target Localization Based on Unidentified Multiple RSS/AOA Measurements in Wireless Sensor Networks”, Sensors, Volume 21, Issue 13, 2021, Pages 1-15.https://doi.org/10.3390/s21134455

[8]. Mohammad Anvaripour; Mehrdad Saif; Majid Ahmadi, “A Novel Approach to Reliable Sensor Selection and Target Tracking in Sensor Networks”, IEEE Transactions on Industrial Informatics, Volume 16, Issue 1, 2020, Pages 171 – 182. DOI: 10.1109/TII.2019.2916091

[9]. Linghua Zhao and Zhihua Huang, “A Moving Object Detection Method Using Deep Learning-Based Wireless Sensor Networks”, Complexity, Hindawi, Volume 2021, April 2021 Pages 1-12. https://doi.org/10.1155/2021/5518196

[10]. Thulio G. S. Amorim, Leonardo A. Souto, Tiago P. Do Nascimento, Martin Saska, “Multi-Robot Sensor Fusion Target Tracking With Observation Constraints”, IEEE Access, Volume 9, 2021, Pages 52557 – 52568. DOI: 10.1109/ACCESS.2021.3070180

# ANALYTICAL METHOD TO PREDICT BREAST CANCER USING MACHINE-LEARNING ALGORITHMS

**Dr.Krithika.D.R.**
Assistant Professor,
Vels Institute of Science Technology and Advanced Studies,Chennai,Tamilnadu,India

**Abstract: Every day each 13 minutes one women dies cause of breast cancer. Breast cancer is one type of cancer, which is growing cells in our body unconditionally. Mostly Women only affects. Four stages of disease, the chances of survival of stage 3 & 4 is very less but the stage 1 & 2 survival rate is high and the disease will be curable. Breast cancer results of damage of DNA and genetic Mutations. In UK and USA breast cancer is common disease. Aware of breast cancer is celebrated all over the world every month of October. It causes lack of exercise, obesity, weight gain, hormonal changes, and Family history. In this paper using Machine learning Algorithm like decision tree, KNN, Logistic and random Forest to predicted breast cancer.**

Keywords: Breast cancer, Obesity, family history, hormonal changes, machine learning, Decision Tree, Random Forest

## 1.INTRODUCTION

Breast cancer is a second leading cause of deaths. Breast cancer is the result of mutations in the genes of the breast cells. It causes them to divide and multiply uncontrollably. It can occur in women and rarely in men. Symptoms of breast cancer include a lump in the breast, bloody discharge from the nipple and changes in the shape or texture of the nipple or breast. Its treatment depends on the stage of cancer. It may consist of chemotherapy, radiation, hormone therapy and surgery. Its difficult to choose which therapy is suitable for affected person.

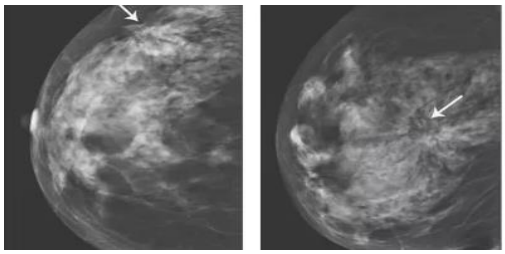


**Fig 1 Breast Cancer image**

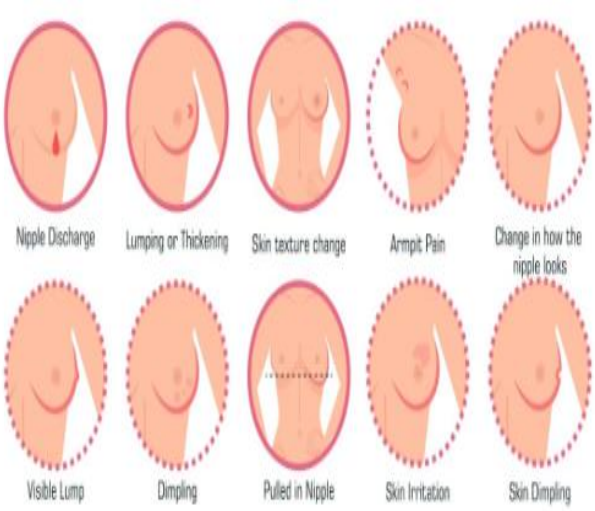



**Fig 1a. Breast cancer changes**

In this paper we discussed machine learning algorithms like decision tree and random forest. Random forest is a bagging method and bootstrap aggregation is done in this algorithm. Random forest splits the data set in small multiple sets of data and gets results from each data set, then it will be aggregated majority of voting. So bootstrap aggregation is done in this algorithm.

## 2.LITERATURE REVIEW

In this paper, they applied “SVM and ANN techniques for prediction of the classification of breast cancer to find which machine learning methods performance is better [1]. This work is the proposed an ensemble machine learning method for diagnosis breast cancer, in which we can see in the table and graph that proposed method is showing with the 98.50% accuracy. In this paper we used only 16 features for diagnosis of cancer [2]. This paper discussed about comparative study of different machine learning algorithms, for the detection of breast cancer and considered Accuracy, AUC and performance parameters for comparison of these ML algorithms, carried out with the WDBC data set. Five classifiers are observed and presented to all WDBC subsets resulting in qualified models.

It has been observed that each of the algorithms had an accuracy of more than 90% and AUC is more than 95 apart from kNN. Supervised machine learning algorithms are suitable for early detection of breast cancer [3]. The paper concludes that SVM and Random Forest give the maximum accuracy with an accuracy of 96.5%. they try to strengthen our work in future by handling a comparatively large dataset and incorporating some more functions such as breast cancer phase detection and so on [4]. The paper used to develop treatments for epilepsy and other neurological disorders. As such, neuro Bi is a valuable tool for translating clinical research into therapeutic devices[5]. In this survey they have briefly described the ways for obtaining energy-efficiency in cooperative sensing. The relevant methods can be applied in local (single node) spectrum sensing procedures, in selection of cooperating, sensing, reporting and relaying nodes, in the application of the appropriate fusion rule, and finally in the proper network organisation "[6].

## 3. DATA AND IMPLEMENTATION

Fig 1.1 show the data set of breast cancer sample values. Fig 2 explores the correlation matrix. Fig3 explores the mean medium, Fig 4 explores the comparison of algorithms for which gives better accuracy. comparison made following machine learning algorithm like logistic regression, KNN, Decision tree and Random Forest. Random forest is showing the better accuracy.

| | mean_radius | mean_texture | mean_perimeter | mean_area | mean_smoothness | diagnosis |
|---|---|---|---|---|---|---|
| 0 | 17.99 | 10.38 | 122.80 | 1001.0 | 0.11840 | 0 |
| 1 | 20.57 | 17.77 | 132.90 | 1326.0 | 0.08474 | 0 |
| 2 | 19.69 | 21.25 | 130.00 | 1203.0 | 0.10960 | 0 |
| 3 | 11.42 | 20.38 | 77.58 | 386.1 | 0.14250 | 0 |
| 4 | 20.29 | 14.34 | 135.10 | 1297.0 | 0.10030 | 0 |

**Fig 1.1 Dataset**

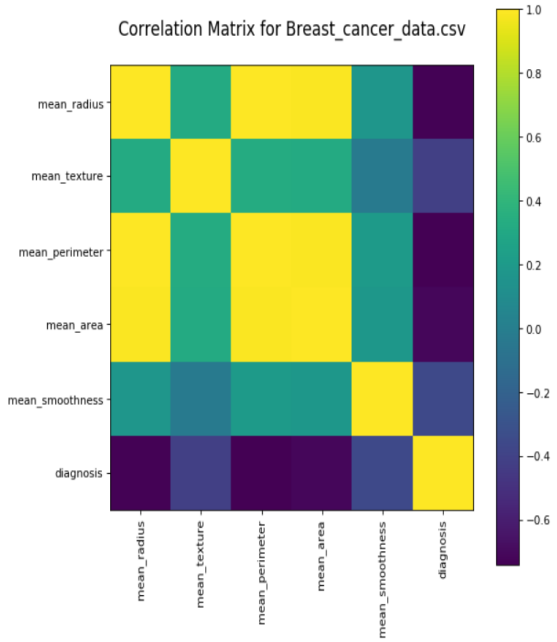



**Fig 2. Correlation Matrix**

| | mean_radius | mean_texture | mean_perimeter | mean_area | mean_smoothness | diagnosis |
|---|---|---|---|---|---|---|
| mean_radius | 1.000000 | 0.323782 | 0.997855 | 0.987357 | 0.170581 | -0.730029 |
| mean_texture | 0.323782 | 1.000000 | 0.329533 | 0.321086 | -0.023389 | -0.415185 |
| mean_perimeter | 0.997855 | 0.329533 | 1.000000 | 0.986507 | 0.207278 | -0.742636 |
| mean_area | 0.987357 | 0.321086 | 0.986507 | 1.000000 | 0.177028 | -0.708984 |
| mean_smoothness | 0.170581 | -0.023389 | 0.207278 | 0.177028 | 1.000000 | -0.358560 |
| diagnosis | -0.730029 | -0.415185 | -0.742636 | -0.708984 | -0.358560 | 1.000000 |

**Fig 3 mean, median code generates a correlation matrix using the Pearson correlation coefficient from a pandas Data Frame.**

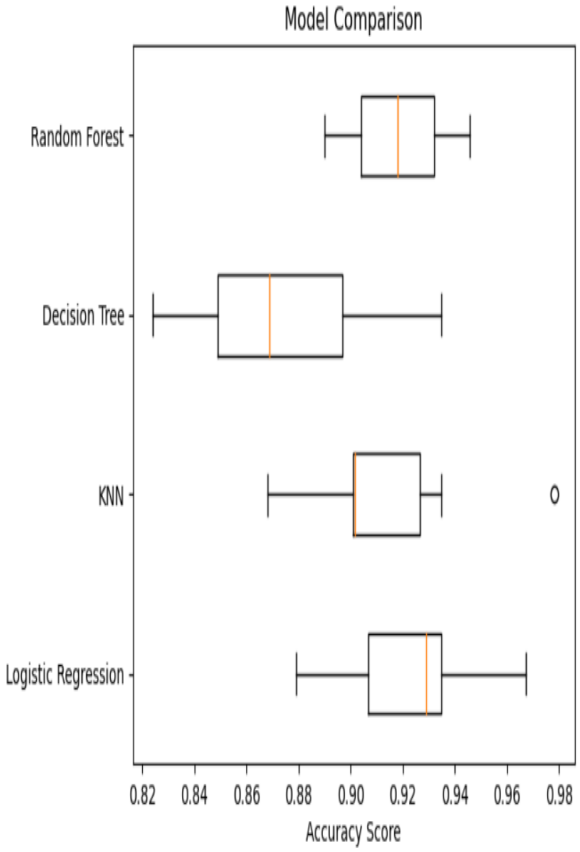



**Fig4. Model Comparison**

### 3.1. MODEL RESULTS

```
Logistic Regression train Set Accuracy: 0.9289617486338798
Logistic Regression Test Set Accuracy: 1.0
KNN train Set Accuracy: 0.936247723132969
KNN Test Set Accuracy: 1.0
Decision Tree train Set Accuracy: 1.0
Decision Tree Test Set Accuracy: 0.9
Random Forest train Set Accuracy: 1.0
Random Forest Test Set Accuracy: 1.0
```

**Fig. 5. Logistic Random Forest Results**

```
Cross-validation scores: [0.94545455 0.93636364 0.91818182 0.94545455 0.88073394]
Average cross-validation score: 0.925
Evaluation results:
--------------------
Confusion matrix:
[[ 9  0]
 [ 0 11]]
Precision: 1.000
Recall: 1.000
F1 score: 1.000
Accuracy: 1.000
```

**Fig 6. Evaluation results**

## 4. CONCLUSION

In this paper breast cancer and breast cancer data set is implemented to machine learning algorithms like decision tree, random forest. Random forest shows the better accuracy. In future plan to predict more about cancer disease.

## REFERENCES

1. E. A. Bayrak, P. Kırcı and T. Ensari, "Comparison of Machine Learning Methods for Breast Cancer Diagnosis," *2019 Scientific Meeting on Electrical-Electronics & Biomedical Engineering and Computer Science (EBBT)*, Istanbul, Turkey, 2019, pp. 1-3, doi: 10.1109/EBBT.2019.8741990.
2. N. Khuriwal and N. Mishra, "Breast cancer diagnosis using adaptive voting ensemble machine learning algorithm," *2018 IEEMA Engineer Infinite Conference (eTechNxT)*, New Delhi, India, 2018, pp. 1-5, doi: 10.1109/ETECHNXT.2018.8385355.
3. V. A. Telsang and K. Hegde, "Breast Cancer Prediction Analysis using Machine Learning Algorithms," *2020 International Conference on Communication, Computing and Industry 4.0 (C2I4)*, Bangalore, India, 2020, pp. 1-5, doi: 10.1109/C2I451079.2020.9368911.
4. S. Ara, A. Das and A. Dey, "Malignant and Benign Breast Cancer Classification using Machine Learning Algorithms," *2021 International Conference on Artificial Intelligence (ICAI)*, Islamabad, Pakistan, 2021, pp. 97-101, doi: 10.1109/ICAI52203.2021.9445249.
5. P. Tumuluru, C. P. Lakshmi, T. Sahaja and R. Prazna, "A Review of Machine Learning Techniques for Breast Cancer Diagnosis in Medical Applications," *2019 Third International conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)*, Palladam, India, 2019, pp. 618-623, doi: 10.1109/I-SMAC47947.2019.9032427.
6. Naveen, R. K. Sharma and A. Ramachandran Nair, "Efficient Breast Cancer Prediction Using Ensemble Machine Learning Models," *2019 4th International Conference on Recent Trends on Electronics, Information, Communication & Technology (RTEICT)*, Bangalore, India, 2019, pp. 100-104, doi: 10.1109/RTEICT46194.2019.9016968.

INTERNATIONAL JOURNAL OF INNOVATIVE TECHNOLOGY AND CREATIVE ENGINEERING (ISSN:2045-8711)
VOL.13 NO.08 AUG 2023

# EDGE COMPUTING - BEYOND THE CLOUD

**Tashita Maria**
Student, M.Sc (Information Technology),
Hindusthan College of Arts and Science, Coimbatore, Tamil Nadu, India
**Tabitha Maria**
Student, M.Sc (Information Technology),
Hindusthan College of Arts and Science, Coimbatore, Tamil Nadu, India
**Ragavi. A**
Student, M.Sc (Information Technology),
Hindusthan College of Arts and Science, Coimbatore, Tamil Nadu, India

**Abstract- This article provides a guide to Edge Computing. Recently, industry investment and research interest in Edge Computing has grown dramatically, where the computing moves storage and computing to the internet from the cloud, where the internet has a closer proximity to the devices and sensors. Hence, Edge Computing makes it very much easier and possible to provide a highly responsive delivery of services for mobile computing, privacy-policy enforcement, and scalability of big data-based AIprocessing.**

Keywords: Edge, Cloud, range

## I. INTRODUCTION

Edge computing is an emerging computing standard that refers to a group or a range of networks and devices thatare with or near the user. Edge, in simpler terms, is that, is used to process data at a close range to where it is being generated. This enables the processing of the data to be very efficient, as it is done at greater speeds and volumes, which leads to better results, that is action - ledin real-time.

**Real-time:** If numerous edge devices are included, the terminal data given is equal to the number of devices transmitted for processing to the cloud. the volume of the intermediate data transmission will be greatly increased and the performance of the data transmission will be reduced, which results in network transmission bandwidth generating in large loads, resulting in the delay in data transmission. In some application scenarios that require real-time feedback, such as traffic, monitoring, etc., cloud computing will not be able to meet business real-time requirements.

**Security and privacy:** Every application which is available or used on a smartphone, requires user data, which includes private information as well. The risk of privacy leakage or attack on this data when uploaded to the cloud center, is considerably very high.

**Energy consumption:** The amount of power consumption of data centers of the world has increased significantly as the number of smart devices continue to increase. Improving the use efficiency of cloud computing energy consumption [3] cannot meet the increasing demand for data energy consumption. The rapidly developing intelligent society will have higher requirements for the energy consumption of cloud computing.

## II. EDGE AND CLOUD COMPUTING

This section presents the fundamental concepts of Edge computing and Cloud computing and its main principals and differences.

### EDGE COMPUTING

Edge computing is a networking concept or method that focuses on bringing computing as close as possible to the source of data in order to reduce latency and use of the bandwidth. In simpler terms, edge computing means limiting the running of processes and functions in the cloud server, by moving those processes to local places, such as on a user's computer, an IoT device, or an edge server. Bringing computation to the network's edge has a main goal, that is to reduce the amount of long- distance communication that has to happen between a client and server.

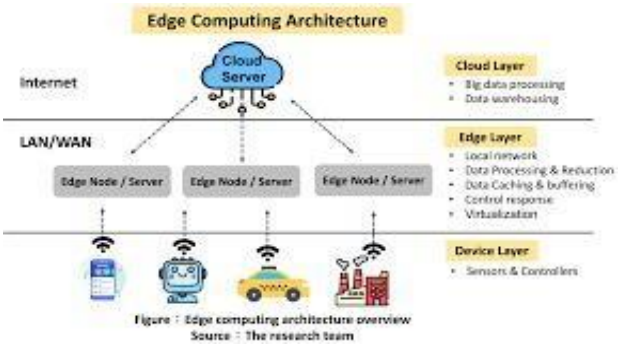



Figure : Edge computing architecture overview
Source : The research team

Edge computing is a distributed computing paradigm that brings computation and data storage closer to the sources of data. This is expected to improve response times and save bandwidth. Edge computing can be termed as an architecture rather than a specific technology, and a topology- and location-sensitive form of distributed computing.

The original creation and implementation of Edge computing was based on content - distributed networks. These were created in the late 1990s and it's main purpose was to web and video content to the users with the edge servers being placed close to the users. In the early 2000's the previously known networks developed in to host applications and application devices on edge servers, resulting in the first commercial edge computing services that hosted processes and applications such as dealer locators, shopping carts, real- time data aggregators, and ad insertion engines. The Internet of Things (IoT) is an example of edge computing. A common misconception is that edge and IoT are synonymous.

## CLOUD COMPUTING

Cloud computing is a the general term that refers to services that deliver hosted services over the internet. These services are divided into three main categories or types of cloud computing:

1. Infrastructure as a service (IaaS)
2. Platform as a service (PaaS)
3. Software as a service (SaaS).


A cloud can be private or public. A public cloud provides services which is opened to the public on the internet. A private cloud is a service that is opened only to a limited number of people, and these individuals can access the content and the various services it provides. The cloud service includes applications, software, storage, tools, etc, with certain access and permissions settings. Private or public, the goal of cloud computing is to provide easy, scalable access to computing resources and IT services.

For the implementation of Cloud computing model, certain hardware and software components are required for a proper implementation. This structure is known as cloud infrastructure. Cloud computing can also be thought of as utility computing or on-demand computing. The origin of the name cloud computing was the cloud symbol. It served as an inspiration as it is often used to represent the internet in flowcharts and diagrams.

## IV.EDGE COMPUTING METHODOLOGYEDGE COMPUTING AND AI

Artificial Intelligence and cloud based solution have always crated various application, implementations and have worked hand in hand in creating a number of problem solving outcome. The cloud provides resource for training complex IoT systems on the datasets. And when combined with the concepts of Artificial Intelligence it brings the concepts to life. It perfects the design and working, updates the algorithm which streamlines the structure of the outcome.

A combination of edge computing and AI technologies, is also known as Edge AI. This concept by using AI algorithms and methodologies allows developers to process and work on the data that is generated by the on location devices or on the server near them. As a result, devices make decisions in a matter of milliseconds subjecting to real time functions and processes. The main advantage is that there is no need to connect to the cloud, or transfer data to other network edges.

There is large demand and utilization of edge devices— sensors, mobile phones, edge nodes. This results in a shift from the network core,. The storage space is shifted from the. The cloud storage space is not sufficient to hold an process the huge volumes of data. Hence the space is shifted to network edge to process the huge volumes of data in proximity.

In an edge computing architecture, data processing is accomplished on devices (end nodes) and at the gateways (edge nodes).

Edge AI is initially deployed in hybrid mode. Then it combines with the decentralized network edge. Followed by the training mode where the system or devices is subjected to training with the AI algorithms. And then finally the revised mode, where the algorithms are revisited and updated. The nodes at the upper level collect the data that is being generated by sensors and aggregators. Later this data is used for training the AI systems or devices.

Machine Learning algorithms are run locally on end nodes and embedded systems For example, the wake words or phrases are simple Machine Learning models, like “Alexa” for the AI device or “OK Google” for the Google search engine that are stored locally on the device speaker. Whenever the smart speaker hears the wake word/phrase, it will begin “listening”, i.e., streaming audio data or sending the audio data in real time to a remote server where it can process the full request.

Model updates are sent back to the cloud, which processes, and enhances them. And finally sends the perfected model back to the end devices.

## USE CASE OF EDGE COMPUTING AND AI ALGORITHM

Edge AI employed in video cameras may be used to track customer footfall and analyze buying patterns. For example, NVIDIA’s EGX AI platform built for physical stores and warehouses empowers video analytics applications, which digitize shopping locations, analyze anonymous shopper behavior inside stores, and optimize store layouts.

In another use case, video analytics can help identify pickpockets and track criminals using facial recognition software. With pattern and facial recognition done at the edge, not in a centralized cloud, retailers get immediate alarm notifications and reduced latency and costs.

## CHARACTERISTICS OF EDGE COMPUTING

Edge computing possesses several characteristics similar to cloud computing but extends the cloud by its specific architecture:

**Geographical Distribution** - IoT applications based on sensor networks highly benefit from processing data locally through edge computing platforms. Big data analytics can be performed rapidly with better accuracy. Edge systems enable real-time analytics and AI processing on a large scale. Examples include sensor networks to monitor the environment, for example, pipeline monitoring or collision avoidance systems.

**Mobility Support -** As the number of mobile devices is rapidly growing, Edge computing also supports mobility to communicate directly with mobile devices.

**Location Awareness** allows the use of technologies such as GPS to find the location of devices. Hence, location awareness can be used by Edge computing applications such as Fog-based vehicular safety applications and edge-based disaster management.

**Proximity -** The availability of the computational resources and services in the local vicinity allows the users to leverage the network context information for making offloading decisions and service usage decisions.

**Low Latency** - The low latency of Edge computing enables the users to execute their resource-intensive and delay-sensitive applications on the edge device. Such applications include connected vehicles, remote health monitoring, warehouse logistics, and industrial control systems.

**Heterogeneity** refers to the existence of different platforms, architectures, computing, and communication technologies used by the Edge computing elements (end devices, edge servers, and networks).

**Bandwidth-intensive use-cases -** An increasing amount of data generated by IoT deployments today is bandwidth-intensive, especially the data from video from surveillance cameras (video analytics). Placing computational resources as close as possible to high-bandwidth data sources implies that much less data need to be sent to distant cloud data centers. For instance, videos and sensor data from hazardous locations can be processed locally to provide real-time information to responders in public safety applications.

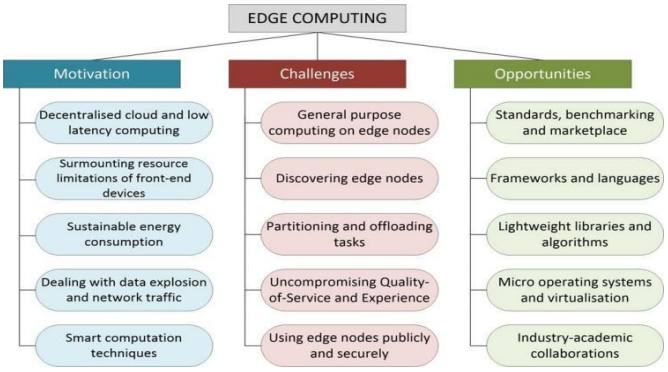
## IMPORTANCE OF EDGE COMPUTING

In the present day, most of the computing processes already take place at the edge or base in places like hospitals, factories, and retail locations, processing and working with the most sensitive data, and powering critical systems that must function on the important concepts of reliably and safely. These places require solutions that do not need a network connection. What makes Edge so instigating is the potential it has for transforming business across every industry and function, from client engagement and marketing to production and back-office operations. In all cases, edge helps make business functions visionary and adaptive—often in real-time—leading to new, optimized experiences for people.

Edge allows businesses to bring the digital realm into the physical world. Bringing online data and algorithms into physical structure stores to improve retail experiences. Creating systems that workers can train and situations where workers can acquire knowledge and understanding from machines. With this in mind, people have created smart environments that have a main focus on the safety and comfort for people. All these examples have one thing in common and that is edge computing, which enables companies to run applications with the most critical reliability, safety,

real-time process, and functions and data requirements directly available on- site.

Ultimately, this allows companies to innovate at a greater speed, create new products, and provide services more quickly, and opens up possibilities for the creation of new streams of revenue.




## ADVANTAGES OF EDGE COMPUTING

Edge computing systems store all the collected data on the edge devices which is present locally near the system without uploading it to a cloud computing platform. This method provides various important advantages to lower the pressure of network bandwidth and other disadvantages of cloud computing.

**Performance -** Edge computing technology provides quick and fast data processing and analysis. It has a noticeably fast response speed which supports real-time services and functions. It provides users with a variety of fast response services, especially in the section of automatic driving, intelligent manufacturing, video monitoring, and other location awareness. Most importantly it provides a fast feedback response. For example, Edge computing enables applications in real- time computer vision.

**Privacy and security** - All the data has to be uploaded onto the cloud for a unified processing, during the traditional cloud implementation. This method unfortunately resulted in exposure to significant risks of data loss and data leakage. Edge computing there is no need to upload data to the cloud as it works on local data storage. This avoids the risks brought by network transmission. As a result, the failures and attacks only affect local data, not the entire data that has been stored. Edge computing and AI inference make it possible to ensure data security and safety.

**Efficiency -** When computing on-site devices, it canresult in a reduced amount of data that is to be transmitted on the network, which in turn reduces the transmission cost and pressure on the network bandwidth. It reduces the energy consumption of local equipment or devices and improves computing efficiency. Edge computing supports the development of highly efficient AI devices such as deep learning accelerators (e.g., Vision Processing Units).

**Reliability** - Edge computing technology provides methods to make services more reliable, efficient, and easily accessible. There is a high dependence on connected on-device systems is an especially important project - critical applications that can be subjected to failures due to network disconnections which may lead to consequences that cannot be altered.

## OPPORTUNITIES IN EDGE COMPUTING

Cloud research has been rapidly progressing which in turn has paved the way for many competitors in this crowded field/ marketplace. Many cloud applications are user-driven, which has resulted in opportunities for large-scale data analytics. However, using the cloud as the main server simply increases the flow of communication between user devices, such as smartphones, tablets, wearables and gadgets, and other devices that we refer to as edge devices, and geographically distant cloud data centers.

The main outcome is to explore the various methods of performing computations on nodes through which network traffic is directed, such as routers, switches, and base stations. These components or devices are referred to as edge nodes. The motivation, challenges, and opportunities in edge computing are summarized.

## USE CASES AND EXAMPLES IN EDGE COMPUTING

Edge computing techniques are used to collect, segregate, process, and analyze data that is near the network edge. It's a powerful means of using data that can't be first moved to a centralized location because of the great volume of the data, which adversely does not support functions that are expensive, processes that are practically challenging or might otherwise go against certain basic mandatory requirements such as data sovereignty. This can be further explained by the variousexamples listed below:

**1. Manufacturing** - An industrial manufacturing company utilized edge computing to access and acquire real-time analysis, monitor manufacturing processes, and machine learning at the local sectors to locate any sort of production or manufacturing errors and improve the quality of manufacturing. Edge computing supported the application of environmental sensors throughout the production plant, which provides an insight on

how each product component is assembled, packed, and stored -- and how long each of the components remains in stock. The manufacturer can now make quicker and more accurate business decisions regarding the factory facility and manufacturing methods and applications.

**2. Farming** - Consider a business that grows crops indoors without sunlight, soil, or pesticides. The growth of the crop is reduced to about more than 60%. The business owners can track the amount of water used, and the nutrient density and in turn determine the required harvest methods, by using sensors. Data is collected and analyzed to find the effects of environmental factors, which is then used to improve the crop growing methodology and ensure that crops are harvested in there blooming condition.

**3. Network optimization** - Edge computing can help optimize network performance by measuring performance for users across the internet and then employing analytics to determine the most reliable, low- latency network path for each user's traffic. In effect, edge computing is used to navigate" traffic through the network for best traffic function, which is time cautious Workplace safety. Edge computing can combine and analyze data from on-site cameras, safety devices of the workers, and various other sensors to help business owners to oversee workplace activities, working conditions or ensure that employees follow provided or illustrated safety protocols -- especially when the workplace is remote or unexpectedly dangerous, such as construction sites, mines or oil rigs.

**4. Improved Healthcare** - The healthcare industry has dramatically expanded the amount of patient data collected from devices, sensors, and other medical equipment. That enormous volume of data and information requires edge computing to utilize automation and apply machine learning to access the data, ignore the correct or normal data and identify problematic data so that clinicians can take immediate action to find a solution to help the patients avoid healthcare accidents in real-time.

**5. Transportation** - Autonomous vehicles gather information about location, speed, vehicle conditions, road conditions, traffic conditions, and other vehicles. These vehicles require and produce anywhere from 5 TB to 20 TB per day. And the data must be segregated and analyzed in real time, while the vehicle is moving. This requires significant onboard edge computing, that is each autonomous vehicle becomes an "edge." In addition, the data can help authorities and business owners to manage vehicle fleets based on present real- time conditions on the ground.

**6. Retail** - Retail businesses can also produce enormous data volumes from surveillance, stock tracking, sales data, and other real-time business details. Edge computing can help understand and process this data. And in turn helps to identify business opportunities, such as an effective campaign, predict sales and optimize market placing orders, and so on. Since retail businesses can vary drastically in local environments, edge computing can be an effective solution for local market functioning at each store.

## CHALLENGES IN EDGE COMPUTING

Although edge computing has the ability and resources to provide quite a number of benefits across multiple use cases, the technology has not reached in peak in perfection. Beyond the common problems of network limitations faced by the industry, there are several key points that can affect the utilization of edge computing:

**Limited capability** - Part of the allure that cloud computing brings to edge, computing is the variety and scale of the resources and services. Bringing into application the infrastructure at the edge can be quite effective, but only when the scope and purpose of the edge deployment are clearly defined -- even with edge computing deployment is expected to perform certain functions even if it performs the basic functions using the limited resources and services available

**Connectivity** - Edge computing can be able to tackle most of the common network limitations and restrictions, but even the simplest edge deployment requires some basic level of connectivity. It's highly necessary to design an edge deployment that can work with poor or unstable connectivity and if we look into what happens at the edge when connectivity is cut off. Failure planning is crucial in connectivity problems to ensure a successful edge computing process.

**Security -** IoT devices are notoriously insecure, so it's vital to design an edge computing deployment that will emphasize proper device management, such as policy- driven configuration enforcement, as well as security in the computing and storage resources -- including factors such as software patching and updates -- with special attention to encryption in the data at rest and in flight. IoT is used in major cloud providers which involves secure communications, but this function cannot be programmed to perform automatically when building an edge site from the base.

**Data life cycles - The most common and noteworthy issue is that there is** so much of that data is unnecessary. Consider a medical monitoring device, it is necessary that only a certain amount of

the data is to be stored for a certain period of time and there is also a certain time limit to save the normal patient data. Most of the data involved in real-time analytics is considered short-term data that isn't kept for a longer period of time. A business must decide which data to keep and what to discard once analysis processes are performed. All the data that is retained, must be saved under secure conditions and be protected in accordance with business and regulatory protocols and policies

## EDGE COMPUTING MAINTENANCE

Edge computing is quite a simple concept that can seem easy to understand and put on paper, but developing a professional workflow layout and implementing efficient deployment at the edge can be a challenging endeavor. No edge implementation would be complete without looking into the concept of edge maintenance:

1. **Security** - Physical and logical security precautions are vital and should involve tools that emphasize vulnerability management and intrusion detection and prevention. Security must also cover the physical components, that is the sensor and IoT devices, as every device is a network that can be accessed or hacked -- which can result in a number of vulnerabilities and threats that can result in violent attacks.
2. **Connectivity** - Connectivity is another issue, and provisions must be made for access to control and reporting even when connectivity for the actual data is unavailable. Some edge deployments use alternate
resource backup connectivity and control. They are known as secondary connections.
3. **Management** - When working on deploying the Edgecomputing in remote locations which are often inhospitable, remote provisioning and management are essential concepts that have to be implemented. IT managers must be able to monitor and know what's happening at the edge remote locations and be able to control the deployment functions and processes when necessary.
4. **Physical maintenance** - Physical maintenance requirements can't be overlooked. The lifespan of the devices depends upon their battery and device components and hence it is observed that the IoT devices have a limited life span. When any component fails, they will eventually require maintenance and replacement.

## V.CONCLUSION

Edge computing is an upcoming computing methodology that works on the concept of close range, on-site computing where a range of computer networks and devices closer to the user. Edge computing is all about processing data that has been generated by a user by having the data stored closer to the user's location. This ensures quicker decision making, real-time process and avoids any delays caused due to data transfer.

## VI.REFERENCES

1. Cao, K., Liu, Y., Meng, G., & Sun, Q. (2020). An . Overview on Edge Computing Research. IEEE Access, 8, 85714–85728. doi:10.1109/access.2020.2991734.
2. Varghese, B., Wang, N., Barbhuiya, S., Kilpatrick, P.,& Nikolopoulos, D. S. (2016). Challenges and Opportunities in Edge Computing. 2016 IEEE International Conference on Smart Cloud (SmartCloud). doi:10.1109/smartcloud.2016.18
3. Dolui, K., & Datta, S. K. (2017). Comparison of edge computing implementations: Fog computing, cloudlet and mobile edge computing. 2017 Global Internet of Things Summit (GIoTS). doi:10.1109/giots.2017.8016213
4. Khan, W. Z., Ahmed, E., Hakak, S., Yaqoob, I., & Ahmed, A. (2019). Edge computing: A survey. Future Generation Computer Systems. doi:10.1016/j.future.2019.02.050

# A PERSPECTIVE STUDY ON TWEET SENTIMENT ANALYSIS USING DATA MINING, MACHINE LEARNING AND DEEP LEARNING PARADIGMS

**G.Priyadarshini**
Assistant Professor, Department in Computer Applications
KG College Of Arts and Science, Coimbatore, Tamilnadu, India
**Dr.D.Karthika**
Associate Professor, Department in Computer Science
VET Institute of Arts and Science College, Erode, Tamilnadu, India

**Abstract- In recent years there has been an increase in interest in collecting and studying text from social networks, review websites, blogs, forums and other forms of user-generated information. The text offers a vast array of ideas from people of diverse profiles, including education, age and their perspectives, region of residence, on how they see goods and services, policy opinions, etc. The analysis of judgments, responses, and emotions drawn from texts is known as sentiment analysis. The sentiment categorization procedure establishes whether a text is subjective or objective, or whether it provokes both positive and negative responses. The most popular method of classification is based on polarity or orientation for accomplishing tweet sentiment analysis. In this paper, a detailed survey on various algorithms used for performing opinion mining, sentiment analysis, tweet sentiment analysis is discussed in detail. The study shows that text preprocessing, data mining, machine learning algorithm and deep learning paradigms plays a vital role in categorization of people's feeling on a specific topic or a product. In this study, the existing challenges in optimizing the process of tweet sentiment analysis is also discussed and the suggestions for improving is also discussed.**

Keywords: opinion mining, tweets, sentiment analysis, machine learning, deep learning, polarization

## 1. INTRODUCTION

Although data mining mostly focuses on finding patterns in numerical data, language can also be used to convey information. Contrary to mathematical data, text is usually unorganized and challenging to manage. A subset of text mining is referred known as data mining [1]. The goal of text mining, a young field, is to extract informational value from written or unstructured data. As a result, the term "text mining" describes the act of looking through textual material to discover pertinent patterns. Since text databases are unstructured, using them might be difficult. The practice of extracting keywords, concepts, and other data from various text documents is known as text mining [2]. A common method in opinion mining for identifying sentiments, subjectivities, and sensitive states in online writings is sentiment analysis. The process of product evaluation was finished by organizing the product attributes. Currently, sentiment polarity assessment is employed in a variety of areas like banking, politics, sports, education, entertainment, etc.

This concentrates on the review of content with a direction, such as opinions or views. A text's subjective or objective nature or whether it stimulates both positive and negative feelings are determined by the sentiment categorization process [3]. There are several significant elements of this classification methodology, such as different procedures, vocations, methodologies, qualities, and application domains. There are many occupations in the classification of sentiment polarity. The three main components of categorization are a class, a level, and an assumption regarding the sentiment sources and objectives.

The distinguishing two-class problem entails categorizing emotions as positive or negative. Sentiment analysis focuses on defining a user's point of view in relation to a given area. Assessment, perception, and even emotional stages are all part of the point of view. The categorization of the polarity of specific text at the levels of characteristics, documents, phrases, and so on is the most important task in sentiment analysis. Applying the classification of polarities, emotional stages such as happy, angry and sad are determined [4].

This classification also identifies the benefits and drawbacks of statements in online reviews, assisting in the more accurate evaluation of products. Agreement detection is another type of binary emotion categorization.

## II. RELATED WORK

More people are using the internet and social media to communicate their views and opinions. As a result, the number of user-generated sentences including sentiment data increased. It's unavoidable to experiment with new ways in order to obtain a better understanding of how people feel and react in various scenarios.

Abd et al [5] in their work evaluated the performance of various machine learning and deep learning methods, as well as providing a new hybrid system for sentiment categorization that combines text mining and neural networks. The dataset used in this study contains over one million tweets from five different domains. 75 percent of the dataset was used to train the system, while the remaining 25% was used to test it.

Bing et al. [6] devised a two-step automated method for sorting tweets. To make the task of creating classifiers easier, they used a raucous tutoring set. They classified tweets into subjective and objective categories first and foremost. Subjective tweets are now referred to as "immense" and "negative" tweets.

Zulfadzli et al [7] done a detailed review about sentiment analysis in social media that looked into the methodology, platforms used, and applications. Users have submitted a vast amount of raw material to social media in the form of text, videos, photographs, and audio. The following trusted and credible databases were used to conduct a systematic review of papers published between 2014 and 2019. The publications were evaluated in light of the study's objectives. The findings suggest that the majority of publications used the opinion-lexicon method to analyse text sentiment in social media, extracting data from microblogging sites, mostly Twitter, and using sentiment analysis such as healthcare, business, events and politics.

Dorababu et al [8] contributed sentiment analysis based on the assumption assessment for customers assessment class, which is used to evaluate data in the form of a collection of tweets, where investigations are extremely unstructured and are either high fine or dreadful, or somewhere in the between. For this, initially the dataset is preprocessed dataset, then extracted the adjective from the dataset that has a couple of significance, referred to as the capacity vector, selected the component vector posting, and then performed device examining using Naive bayes, maximum entropy, and SVM laterally the edge of semantic overview based on word net, which extracts synonyms and similarity for the content.

Luciano et al [9] offered a method for automatically detecting feelings in Twitter messages (tweets) that considers specific features of how tweets are written as well as meta-information about the words that make up these messages. In addition, they use sources of noisy labels as training data. A few sentiment detection websites provided these noisy labels based on twitter data.

Hagen et al [10] categorize the features represented in a tweet as either positive, negative or neutral, we repeat three classification algorithms using different feature sets. The replicated techniques are also merged in an ensemble, with individual classifiers' confidence scores for the three classes averaged and sentiment polarity determined based on these averages.

Deep learning and neural networks have been increasingly important in sentiment analysis in recent years, and they are now widely regarded as the most advanced way for analyzing a variety of languages. Tamil is one of the Indian languages where a state-of-the-art sentiment analysis model is still required. Tamil language presents greater obstacles due to its unique features, grammar structure, and agglutinative nature.

To analyse Tamil tweets, Anbukkarasi et al [11] developed a combination of character-based Deep Bidirectional long short-term memory neural networks.

Gangula et al [12] designed a corpora for telugu text and assigning polarities to them is described in this study. Following the establishment of corpora, they trained classifiers to produce accurate classification results. A sentiment classifier is usually trained on data from the same domain that it will be evaluated on. However,there may not be enough data in the same area, and combining data from several sources and domains may aid in the development of a more universal sentiment classifier that can be used across multiple domains. Sentiment data is used from the above corpus from several domains to develop this generalist classifier. For both in-domain and cross-domain categorization, initially examined sentiment analysis models developed

with a single data source. Then, using data samples from several areas to construct a sentiment model and validated their performance based on its accuracy of classification.

Hughes et al [13] developed a system for automatically classifying clinical literature at the sentence level to handle complicated text features, a Deep Convolutional Neural Network is used. They used extensive classification of health data to train the model.

Kudo [14] improved the neural network performance by addressing the issue of segmentation vagueness. The regularization method is used to train sub words with a simple regularization. The sampling process is aided by multiple sub word segmentation using the unigram language technique.

Devlin et al [15] anticipated Bidirectional Encoder Representation model, which is a new language representation paradigm. To comprehend the pattern of unlabeled text, the method pretrains bidirectional representation in all layers, both left and right.

Sultana et al. [16] developed a deep learning model which performs sentiment analysis of education data. They used Multilayer perceptron and Support vector machine for training education dataset to predict the sentiment analysis.

Amir Hamzah et al. [17] on their wok used Hidden Markov Model along with POS TAG to determine it is a positive or negative opinion. Automatic detects the orientation of the opinion with its target label.

Wook et al. [18] performed lexicon-based opinion mining which explores the assessment of teaching results. In this work sentiment tendency is analyzed over the student's feedback which is used for extracting intensifier words. Teaching assessment is described in terms of positive, neutral or negative opinions.

Kumar Ravi et al. [19] stated in their survey about the various natural language processing techniques, machine learning and sub task performance to achieve the sentiment analysis.

Leary et al. [20] reported in their work about blog mining, which comprised of mining and blog searching. This work analyzed the blog types, its unit and type of opinions to be extracted from blogs.

Montoyo et al. [21] in their work discussed about the problems involved in sentiment and subjectivity analysis and how the machine learning approaches are used to solve the problems in a better manner.

Liu et al. [22] designed a various task conceivable and works based on opinion mining and sentiment analysis. The activities discussed in this work are sentiment classification and subjectivity, sentiment analysis based on aspect, lexicon generation, comparative opinions, summarization of opinion and spam detection opinion and reviews quality are analyzed.

Tsytsarau and Palpanas [23] performed a survey on opinion mining for spam detection, contradiction analysis and aggregation of opinion. Several opining mining approaches are compared with few of the general dataset.

Ali Hasan et al. [24] in their paper designed a hybrid approach which comprised of sentiment analyzing along with machine learning models. This work also performed political view-based sentiment analysis by comparing the support vector machine and naïve bayes.

Kanavos et al. [25] explore an algorithm which handles the emotions from tweets by handling huge volume of dataset to perform sentiment analysis. The author in another work [26] determines the social communities with significant by conveying metric value for users' emotional posts which is collected from twitter profiles.

Alcober et al [27] concentrated on creating a cutting-edge community detection model which will take use of user emotion. The user's tweets are analyzed using Ekman emotional scale, which uses three variants measures and deployed community modularity detection technique.

After more than a year of adjusting to distance learning techniques that are now thought of as the new norm, Mohana et al [28] constructed an opinion mining on the education level of Filipinos. They employed three distinct classification methods to assess opinion mining's accuracy. Final results reveal that deep learning algorithms are most suitable for opinion mining.

Vohra et al [29] designed a finely tuned convolutional network with annotated set of tweets using valence aware dictionary. They adopted fast text embedding model is adopted to training the text corpus dataset. The authors stated that effective extraction of useful text is very challenging to produced best sentiment analysis. work about sentiment analysis on twitter dataset using deep learning and machine learning algorithms. The voting-based ensemble technique is used for binary classification of polarity of the tweet.

Habib et al [30] devised a hybrid deep learning algorithm is used to discover tweet as

negative or positive sentence. The long short-term memory improves the classification of either multi-class or binary class categorization by using different word embedding strategies. The work expresses the attitudes, emotions from the opinion of publics.

**Limitation in existing algorithms for sentiment analysis and opinion mining**

- Class imbalance among the training dataset and testing dataset of tweet and sentiment analysis.
- Overfitting problem when the volume of dataset is huge in opinion mining
- Recognition of accurate user comments is very difficult in case of polarity detection in opinion or sentiment analysis.
- While using conventional Deep learning algorithms the weights are assigned in trial and error basis, then the prediction model produces less accuracy with higher error rate. So, assignment of weight values is very important during classification or prediction process
- The unknown pattern of tweets cannot be categorized by the classification model accurately.

## 3. CONCLUSION

In this paper, the detailed investigation of the existing algorithms used in tweet sentiment or polarity analysis in various domains, such as product, education, books, etc., The most of the classification algorithms focused on dataset with labels. They strongly depend on the training phase of the dataset in sentiment analysis. The text preprocessing, extraction of features and utilizing the significant part of the text are the major challenges to improve the quality of the sentiment analysis. The objective of this survey is to explore the major challenges in text classification based on polarity or opinion mining. The exiting problems can be overcome by focusing on

Developing classification model with the ability of handling class imbalance

Construction deep learning models with optimized hyperparameter control

Designing unsupervised learning model to attain global optimization by fusing optimization algorithms.

## REFERENCES

1. M. H. Abd El-Jawad, R. Hodhod, Y. M. K. Omar, "Sentiment Analysis of Social Media Networks using Machine Learning," 2018 14th International Computer Engineering Conference (ICENCO), 2018, pp. 174-176.
2. Yi, Shanshan, Xiaofang Liu, Machine learning based customer sentiment analysis for recommending shoppers, shops based on customers review, Complex & Intelligent Systems 6, no. 3 (2020): 621-634.
3. Asare, A. O., Yap, R., Truong, N. & Sarpong, E. O. The pandemic semesters: Examining public opinion regarding online learning amidst COVID-19. J. Comput. Assist. Learn. 37, 1591–1605, 2021.
4. Kumar Ravi, A survey on opinion mining and sentiment analysis: Tasks, approaches and applications. Knowledge-Based Systems. 89. 14-46, 2015.
5. M. H. Abd El-Jawad, R. Hodhod, Y. M. K. Omar, "Sentiment Analysis of Social Media Networks using Machine Learning," 2018 14th International Computer Engineering Conference (ICENCO), 2018, pp. 174-176.
6. L. Bing, K. C. Chan, C. Ou, Public sentiment analysis in twitter data for prediction of a company's stock price movements, in EBusiness Engineering (ICEBE), 2014 IEEE 11th International Conference on, 2014, pp. 232-239.
7. Zulfadzli Drus, Haliyana Khalid, Sentiment Analysis in Social Media and Its Application: Systematic Literature Review, The Fifth Information Systems International Conference 2019, Procedia Computer Science 161 (2019) 707–714.
8. Dorababu Sudarsa, Siva kumar.P, L. Jagajeevan Rao, Sentiment Analysis for Social Networks Using Machine Learning Techniques, International Journal of Engineering & Technology, 7 (2.32) (2018) 473-476.
9. Luciano Barbosa, Junlan Feng, Robust sentiment detection on Twitter from biased and noisy data, COLING '10: Proceedings of the 23rd International Conference on Computational Linguistics, Pages 36–44, 2010
10. Hagen M., Potthast M., Büchner M., Stein B. (2015) Twitter Sentiment Detection via Ensemble Classification Using Averaged Confidence Scores. In: Hanbury A., Kazai G., Rauber A., Fuhr N. (eds) Advances in Information Retrieval. ECIR 2015. Lecture Notes in Computer Science, vol 9022.
11. S. Anbukkarasi, S. Varadhaganapathy, Analyzing Sentiment in Tamil Tweets using Deep Neural Network, 2020 Fourth International Conference on Computing Methodologies and Communication (ICCMC), 2020, pp. 449-453, ICCMC-00084.
12. Gangula Rama Rohit Reddy, Radhika Mamidi, Resource Creation Towards Automated Sentiment Analysis in Telugu (a low resource language) and Integrating Multiple Domain Sources to Enhance Sentiment Prediction, Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018), 2018, pp 624-634.

13. M. Hughes, I. Li, S. Kotoulas, and T. Suzumura, “Medical Text Classification Using Convolutional Neural Networks,” Studies in Health Technology and Informatics, 2017.
14. T. Kudo, “Subword regularization: Improving neural network translation models with multiple subword candidates,” in ACL 2018 - 56th Annual Meeting of the Association for Computational Linguistics, Proceedings of the Conference (Long Papers), 2018.
15. J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, “Bert: Pre-training of deep bidirectional transformers for language understanding,” arXiv preprint arXiv:1810.04805, 2018.
16. J. Sultana, N. Sultana, K. Yadav and F. AlFayez, "Prediction of Sentiment Analysis on Educational Data based on Deep Learning Approach," 2018 21st Saudi Computer Society National Computer Conference (NCC), Riyadh, 2018, pp. 1-5, doi: 10.1109/NCG.2018.8593108.
17. Amir Hamzah, Opinion Mining and Sentiment Analysis Application for Opinion Classification from Education Questionnaire, International Journal of Engineering Research & Technology, Volume 06, Issue 11, 2017
18. Wook, M., Razali, N.A.M., Ramli, S. et al. Opinion mining technique for developing student feedback analysis system using lexicon-based approach (OMFeedback). Educ Inf Technol 25, 2549–2560 (2020). https://doi.org/10.1007/s10639-019-10073-7
19. Kumar Ravi,Vadlamani Ravi, A survey on opinion mining and sentiment analysis: tasks, approaches and applications, Knowledge-Based Systems 89 (2015) 14–46
20. D.E. O'Leary, Blog mining-review and extensions: “From each according to his opinion”, Decision Support Systems 51 (2011) 821–830.
21. Montoyo, P. Martínez-Barco, A. Balahur, Subjectivity and sentiment analysis: An overview of the current state of the area and envisaged developments, Decision Support Systems 53 (2012) 675–679
22. Liu, Sentiment analysis: A multi-faceted problem, IEEE Intelligent Systems 25, no. 3 (2010): 76-80.
23. M. Tsytsarau, T. Palpanas, Survey on mining subjective data on the web, Data Min Knowl Disc (2012) 24:478–514,
24. Ali Hasan, Sana Moin , Ahmad Karim, Shahaboddin Shamshirband, Machine Learning-Based Sentiment Analysis for Twitter Accounts Math and computational applications, 2018, 23, 11; pp 1-15
25. Kanavos, Andreas Nodarakis, Nikolaos Sioutas, Spyros Tsakalidis, Athanasios Tsolis, Dimitrios Tzimas, Giannis, Large Scale Implementations for Twitter Sentiment Classificatio, Algorithms . Vol 10 pg 33,2017.
26. Kanavos A, Perikos, I, Hatzilygeroudis, I, Tsakalidis A, Emotional community detection in social networks, Computers & Electrical Engineering Volume 65, Pages 449-460, 2017.
27. G. M. I. Alcober and T. F. Revano, "Twitter Sentiment Analysis towards Online Learning during COVID-19 in the Philippines," 2021 IEEE 13th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management (HNICEM), 2021, pp. 1-6.
28. R. S. Mohana, S. Kalaiselvi, K. Kousalya, Twitter based sentiment analysis to predict public emotions using machine learning algorithms, Third International 1. Conference on Inventive Research in Computing Applications (ICIRCA), 2021, pp. 1759-176.
29. Vohra A., Garg R, Deep learning based sentiment analysis of public perception of working from home through tweets. J Intell Inf Syst (2022).
30. Habib Md. Ahsan. (2021). Tweet Sentiment Analysis using Deep Learning Technique. IJITEE (International Journal of Information Technology and Electrical Engineering). 10. 27-36.

# EFFICIENT CLASSIFICATION TECHNIQUE FOR MANGO LEAF DISEASE DETECTION

**B. Sowmya**[1]
Assistant Professor, Department of Computer Science,
K.S.Rangasamy College of Arts and Science (Autonomous)
Tiruchengode, Tamilnadu, India.

**Dr. P. R. Tamilselvi**[2]
Assistant Professor, Department of Computer Science,
Government Arts and Science College,
Komarapalayam, Tamilnadu, India.

**Abstract— Agriculture is the basic foundation for Indian economy. In India most of the people depend agriculture to run their life. Based on nature life cycle the productivity of plants differs and it also affected by various environmental condition. Most of the time diseased and affected areas have been easily detected by normal vision. This method of inspection prolong the period to identification of the disease and most of the time fruit flowers fall down before it goes to the stage of fruit. Immediate identification may save the production of plants and it leads to more yield. Still now many farmers have no clear knowledge about Computer Aided Analysis (CAD). Usability of Machine learning and Deep Learning spread their branch all over the field of agriculture. This paper experiment the mango tree disease particularly Anthracnose disease detection and classification with the help of Machine learning and Deep Learning algorithms such as Support Vector Machine (SVM), Convolution Neural Network (CNN). The accuracy of CNN classifier is higher when compared to SVM classifier. The result comparison shows that CNN provides better accuracy than SVM classification. Early and timely analysis may help the farmers to protect the mango fruit decay by this disease and support to improve the production of mango.**

Keywords: Machine Learning, Deep Learning, Anthracnose, SVM, CNN

## 1. INTRODUCTION

In agriculture fruit cultivation is main source of income for farmers. Today's environment condition shows that humans need to consume more fruits, vegetables and minerals to protect their healthy life. Mango is one of the fruit which is the king of all other fruit and it is our national fruit. Plant disease is a dangerous threat to agriculture. Recent year's mango yield greatly affected by many diseases due to weather changes and environment changes. Anthracnose, Bacterial canker, Cutting Weevil, Die Back, Gall Midge, Sooty Mould, Powdery Mildew were diseases which is possibly occur in mango tree. Anthracnose is a mango leaf disease which affects the mango fruit yield and which creates economic loss for farmers. It destroys the young shoots, flowers and fruits. Leaf spot, twig blight, blossom blight, wither tip and fruit rot were symptoms produced by anthracnose. Early detection of diseased leafs can protect the mango fruit blossoms and healthy mango fruit. Modern technologies provide greater hope for agriculture to detect and identify the diseased portion.

Machine learning algorithms consider more features to detect the disease while Deep learning algorithms not need external feature extraction and it works more intelligent manner than machine learning algorithms. This research utilize both algorithms to detect and classify anthrocnose disease in mango leaf and recommend which algorithms helps to detect and classify easily. For that Support Vector Machine (SVM) and Convolution Neural Network (CNN) has been experimented. The accuracy of both methods shows that CNN produce high accuracy in disease detection, As a result the farmers can easily identified the anthrocnose disease and protect the falling of mango flower blossoms from anthrocnose disease and their by increase the productivity and economic background of farmers

.

## II.LITERATURE REVIEW

Debasish et al., proposed a model to develop the automatic detection algorithm for leaf diseases. Support Vector Machine (SVM), Random Forest, Random Forest were experimented to classify the leaf diseases. SVM attained 87.6% of accuracy, Random Forest attained accuracy of 70.05% and Random Forest attained accuracy of 67.03% and recommend SVM produced better result than other algorithms.

Kumar et al, experimented a Alternaria Alternata, Anthracnose, Bacterial Blight and Cercospora Leaf Spot diseases in plant. Image enhancement, segmentation, feature extraction and classification by SVM were carried out to detect the diseased portion and achieved the accuracy of 98.38% and also found affected area 15.01%.

Kaur et al, mentioned various diseases present in wheat crop and also presents the application of SVM classifier for disease identification and prediction. This paper presents the rate of affected area and found accuracy of 80.02%.

Mohammed et al, proposed knowledge based data for SVM classification. For disease detection, the crops Wheat, Tomato, Cucumber were used and nearly 12 diseases were analysed and suggest by increasing of training samples improves the accuracy. This experiment found 90.61% of disease case, 77.4% of healthy case and accuracy of total detection was 88.1% at testing.

Arshiya et al., performed segmentation, classification for Grape leaf disease and executed image enhancement by CLAHE, segmentation by fuzzy c-means algorithm, features selected by PCA and classification done by PSO SVM, BPNN, Random Forest Algorithm. This experiment record PSO SVM achieves higher accuracy of 100% while compared to other algorithms.

Sunil et al., experimented machine learning algorithms for Tomato leaves and kmeans clustering used to partitioning the voronoi cells. Feature extraction was performed with Discrete Wavelet Transform, Principal Component Analysis and Grey Level Cooccurrence Matrix and classification done by Support Vector Machine (SVM), Convolution Neural Network (CNN), K-Nearest Neighbor (K-NN). SVM achieved the accuracy of 88%, K-NN achieved the accuracy of 97% and CNN achieved the accuracy of 99.6% and recommend CNN better for Tomato leaf disease detection.

Sanath et al., experimented Grapes and Mango leaf diseases by Convolution Neural Network (CNN). This paper analyse presence and absence of diseases with the accuracy of 99% and 89% respectively and also developed Android Smartphone app for the same analysis in the name of JIT CROPFIX.

Vijai et al., presented an segmentation algorithm and automatic detection of classification algorithm for plant leaf diseases. Banana, Beans, Lemon, Rose leafs were experimented. This paper also presents various related survey about plant diseases and the classification experimented with Minimum Distance Criterion with K-Means Clustering and attained accuracy of 86.54%, SVM classification achieved the accuracy of 95.71%.

Jennifer et al., Introduced the Arabica coffee leaf dataset named JMuBEN and JMuBEN2 and various diseases of coffee leaf was analysed and data augmentation was carried out to increase the dataset.

Suna et al., proposed a CNN with FL-Efficient Net ( Focal loss Efficient Net) and features extracted with moving flip bottleneck convolution and attention mechanism, CrossEntropy loss function was replaced by Focal loss function. ResNet-50, DenseNet-169 and Efficient Net were experimented for new plant disease dataset (NPDD). Efficient Net attained accuracy of 99.72% while compared others and it takes short training time.

## III. METHODOLOGY

The following figure1.1 shows the proposed method used for anthrocnose disease detection and classification.

A. Dataset

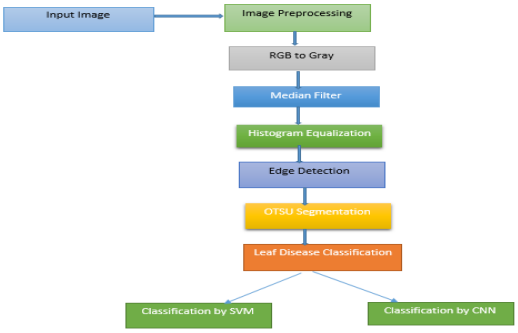
The dataset was taken from Plant Village dataset. It contains many mango leaf diseases, and this research only analysed the Mango anthrocnose leaf images. It contains nearly 1000 images. Healthy and anthrocnose images were analysed for execution.

B. Preprocessing

Preprocessing is the preliminary step to remove noises, bad illumination from images and improve the quality of image which helps for better visualization. In this research, preprocessing of anthracnose involves following steps.

1. The given image was converted to Rgb2gray.
2. Noise was removed by using Median Filter.
3. Histogram equalization executed to improve the contrast of the image.

**Figure 1. Overall Methodology of Anthracnose Leaf Disease Detection**



C. Edge Detection and OTSU Segmentation:

Edge detection is the basic segmentation technique for image processing. Normally it transform the given images into grey based images and it detects the boundaries, background in the image. It identify the edges based on brightness changes and discontinuities in the image. Sobel edge detection technique used to identify finite edges in the anthrocnose. leaf.

Segmentation used to detect the diseased portion of the image. OTSU multilevel thresholding method processed to get diseased leaf portion. It is one of the prominent method of segmentation. Following figure. 2 shows the diseased anthrocnose leaf detection process.

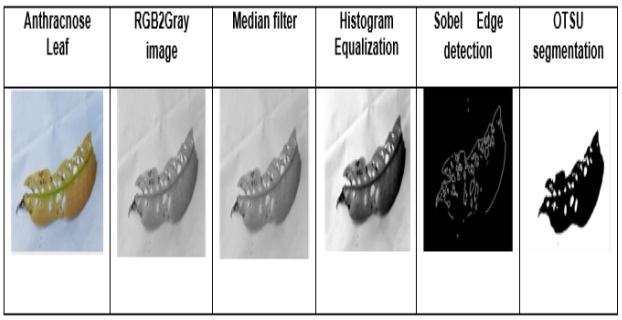



*Figure 2. Edge Detection and OTSU Segmentation*

D. Feature Extraction and Classification:

Features are important for classification and Grey-level Co-occurence Matrix (GLCM) is used to get features from the leaf image. The dimension and pattern of image has been easily identified by these methods. In this experiment contrast, correlation, Energy and Homogeneity features were extracted and used for SVM and CNN classification.

Classification is used to group the similar nature of pixels. The dataset has been trained and tested to accomplish classification. In this research SVM, CNN have been trained and tested for anthrocnose leafs.

E. Support Vector Machine

It is one of the supervised learning technique and it can be used for classification and pattern recognition. It is one of the non-linear algorithm and it is not depend on the dimensionality of the data. It produces better result than query based searching algorithms. It also reduces the information which not needed and it predict the results more accurately. The boundary created by SVM is called maximum margin classifier and it has various kernel functions to process the data which are Linear, Non-linear, Polynomial, Radial Basis Function (RBF), and Sigmoid. Kernels are used to find hyper plane, decision boundry and solve non-linear problems. Among these kernels Linear works well and it produce better accuracy for anthracnose classification.

i.RBF Kernel

RBF kernel is one of the common kernel function used in SVM and it can be applied where no previous knowledge about data. It also analyse relation between two points. RBF kernel calculated by

$$K(X_1, X_2) = exp\left(\frac{||X_1 - X_2||^2}{2\sigma^2}\right)$$

F. Convolution Neural Network

CNN is the primary working method of deep learning. The features has been extracted with the help of convolution operation and thereby produce the feature map. These deep learning algorithms not need any external feature extraction methods while compared to machine learning models. In CNN binary classification with max pooling, sigmoid function, ReLu activation function and with Binary cross entropy loss function, RMSProp optimizer with learning rate of 0.001. Model loss and accuracy is also predicted

Where $\sigma$ is hyperparameter and it is variance.

$||X_1 - X_2||$ - Euclidean distance between two points $X_1$ $and$ $X_2$.

If point a are similar and no distance between them means they are similar points.

If points are distance and kernel value close to 0 (less than 1) means they are dissimilar points

i. ***Linear kernel***

It is one of the prominent kernel function and it will perform well with large number of features. It is faster than other kernels and it works based on regularization parameter.

$$\mathbf{F(x, xj) = sum( x.xj)}$$

Where x,xj represents the classified data.

ii. **Polinomial kernel**

It is generalized indication of linear kernel.

$$k(\mathbf{x_i}, \mathbf{x_j}) = (\mathbf{x_i} \cdot \mathbf{x_j} + 1)^d$$

Where d represents the degree F(x,xj) decision boundry

## III. RESULT AND DISCUSSION

Following figure 3 shows the few images of anthracnose disease detection.

The following figure 4 shows Training accuracy and loss and Validation accuracy, loss for CNN

**Figure. 3 Anthracnose disease detection**

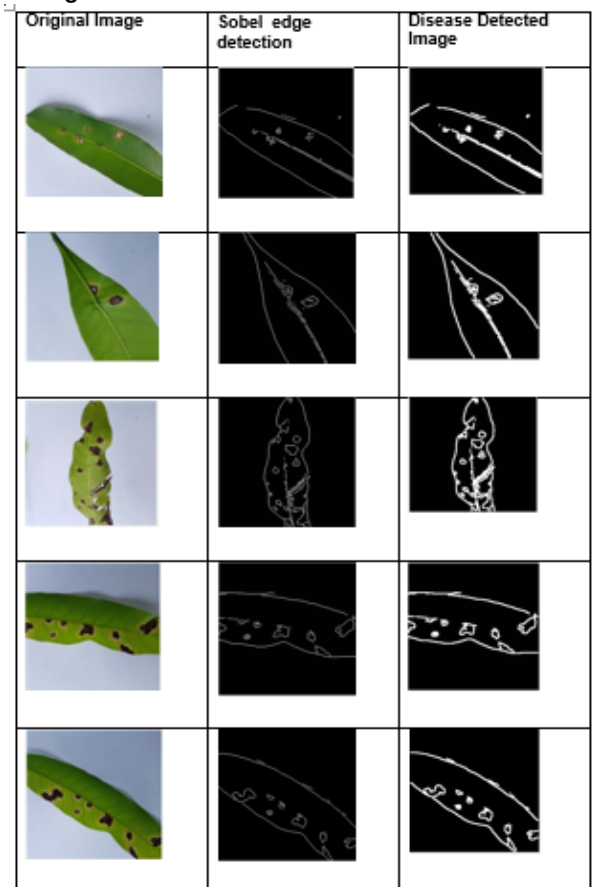



**Figure 4 . Training accuracy and loss and Validation accuracy, loss for CNN**

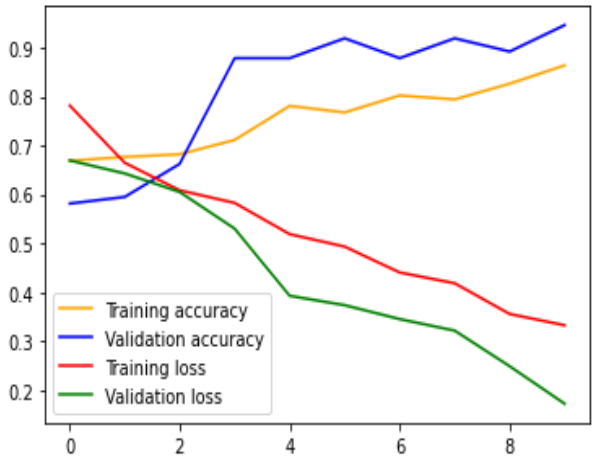



**Table 1. values of Training accuracy and loss and Validation accuracy, loss .**

| Training accuracy (%) | Training loss (%) | Validation accuracy (%) | Validation loss (%) |
|---|---|---|---|
| 0.8625 | 0.3446 | 0.9519 | 0.2515 |

The following table 2 shows the accuracy of Linear Kernel, RBF Kernel, Polynomial Kernels.

**Table 2. Accuracy of Linear Kernel, RBF Kernel, Polynomial Kernels**

| Linear Kernel (%) | RBF Kernel (%) | Polynomial Kernels (%) |
|---|---|---|
| 0.84 | 0.80 | 0.71 |

The following figure 5 shows the accuracy performance comparison of SVM kernels.

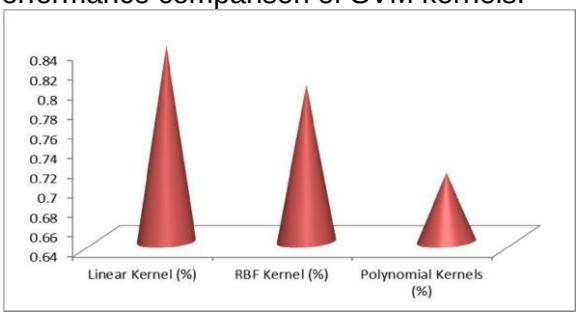



**Figure 5. Accuracy performance Comparison of SVM kernels**

The following table3 and figure 6 shows the accuracy performance comparison of SVM and CNN.

**Table 3. Accuracy performance Comparison of SVM and CNN.**

| Linear Kernel (%) | CNN accuracy (%) |
|---|---|
| 0.84 | 0.95 |

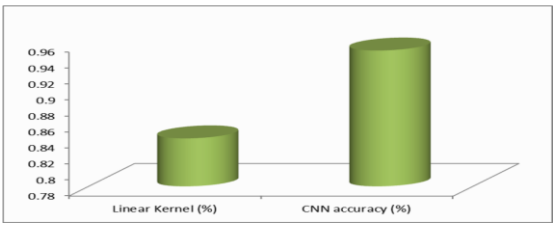



**Figure 5. Accuracy performance Comparison of SVM and CNN**

## IV.CONCLUSION

Fruit yield one of the main sources of income for farmers which need correct mango production in time. Even a single mango can fulfil the 40% of daily fibre needs and mangoes available at 1000 varieties in India. Diseases present in mango plant can reduce the mango yield. Now a day's lot of Computer Aided Techniques helps agriculture to identify timely manner and thereby help to control many diseases. This paper analyses the detection and classification of Anthracnose disease of Mango plant with the help of OTSU segmentation and SVM, CNN algorithms. In SVM linear kernel produces higher accuracy than other kernels and while comparing both classification algorithms, CNN achieves accuracy of 98% than SVM. This research recommend CNN for Anthracnose disease detection and the farmers will benefited by early and timely detection.

## REFERENCES

1. Debasish Das, Mahinderpal Singh, Sarthak Swaroop Mohanty and S. Chakravarty, "Leaf Disease Detection using Support Vector Machine", International Conference on Communication and Signal Processing, July 28 - 30, 2020.
2. R.Bharath Kumar, K. Balakrishna, M.Shreyas, S.Sonu, H.Anirudh, and B.Abhishek, "SVM Based Plant Diseases Detection using Image Processing", Vol-7, Issue- 5, May 2019.
3. Er.Varinderjit Kaur , Dr.Ashish Oberoi, "Wheat Disease Detection Using SVM Classifier" Journal of Emerging Technologies and Innovative Research (JETIR). Vol-5, Issue- 8, August 2018.
4. Mohammed A. Hussein, Amel H. Abbas, "Plant Leaf Disease Detection Using Support Vector Machine", Al-Mustansiriyah Journal of Science, Volume 30, Issue 1, 2019.
5. Arshiya S. Ansari , Malik Jawarneh , Mahyudin Ritonga ,Pragti Jamwal , Mohammad Sajid Mohammadi , Ravi Kishore Veluri , Virendra Kumar , and Mohd Asif Shah, "Improved Support Vector Machine and Image Processing Enabled Methodology for Detection and Classification of Grape Leaf Disease", Hindawi, Journal of Food Quality, Vol- 2022, PP: 1-6,
6. Sunil S. Harakannanavar , Jayashri M. Rudagi , Veena I Puranikmathb, Ayesha Siddiquaa, R Pramodhini, "Plant leaf disease detection using computer vision and machine learning algorithms", Global transform proceedings, 3(2022), PP: 305-310.
7. Sanath Rao, R.Swathi , V.Sanjana, L.Arpitha, K.Chandrasekhar, Chinmayi , Pramod Kumar Naik, "Deep Learning Precision Farming: Grapes and Mango Leaf Disease Detection by Transfer Learning", Global transform proceedings, 2(2021), PP:535-544.
8. Vijai Singh , A.K. Misra, "Detection of plant leaf diseases using image segmentation and soft computing techniques", Information Processing in Agriculture, 4 (2017), PP: 41–49.
9. M. Alessandrini , R. Calero Fuentes Rivera, L. Falaschetti , D. Pau, V. Tomaselli , C. Turchetti, "A grapevine leaves dataset for early detection and classification of esca disease in vineyards through machine learning", Elsevier Data in Brief, 35(2021), PP: 19.
10. Jennifer Jepkoech, David Muchangi Mugo, Benson K. Kenduiywo , Edna Chebet Too, "Arabica coffee leaf images dataset for coffee leaf disease detection and classification", Elsevier Data in Brief, 36(2021), PP: 1-8.
11. Xuewei Suna, Guohou Li , Peixin Qua, Xiwang Xie , Xipeng Pan, Weidong Zhang, "Research on plant disease identification based on CNN", Cognitive Robotics 2(2022), PP:155-163


August 2023
Vol - 13 No - 8
@IJITCE Publication