

INDIA-NET: IOT INTRUSION DETECTION VIA ENHANCED TRANSIENT SEARCH OPTIMIZED ADVANCED DEEP LEARNING TECHNIQUE

K. Anusha^{1,*}, B. Muthu Kumar² and J. Ragaventhiran³

¹Department of School of Computer Science and Engineering at VIT Chennai, Tamilnadu India.

²Department of School of Computing and Information Technology, REVA University, Bengaluru Karnataka 560064 India.

³Department of School of Computing and Information Technology, REVA University Bangalore India.

*Corresponding e-mail: knu.anusha@gmail.com

Abstract The Internet of Things (IoT) has become an increasingly popular study area, with billions of devices deployed globally in recent years. These devices can speak with one another without the need for human involvement because they are all connected to the Internet. But this also gives rise to new security problems that are becoming more and more prevalent and important areas for study. A novel INtrusion Detection in Iot using Advanced deep learning NETwork (INDIA-NET) has been presented in this paper to address these shortcomings. It effectively detects intrusions using this innovative Deep Learning (DL) technique. The dataset's duplicate and redundant data are first eliminated by preprocessing using the Minkowski distance-based closest neighbour algorithm. After preprocessing an enhanced transient search optimization algorithm has been used for feature selection. After selecting the features, the features are sent to the novel convolutional neural network combined Generative Adversarial Network (CNN-GAN) which classifies the output into DOS attack, UR3 attack and normal. The accuracy, precision, recall, and detection rate, among other particular metrics, have all been used to calculate the suggested method's performance. Using the NSL-KDD dataset, the suggested INDIA-NET system has been assessed. Its results show improvements in accuracy, precision, recall, and F1-score, with respective values of 99.02%, 99.38%, 98.29%, and 98.83%.

Keywords – Intrusion detection, Internet of things, Convolutional neural network, Generative Adversarial Network, Enhanced transient search optimization algorithm.

1. INTRODUCTION

The term "Internet of Things" (IoT) refers to the industrial and research trend in the field of information and communications technology (ICT) that has grown to be a part of everyday technological development [1]. The phrase "internet of things" (IoT) describes a new paradigm for communication in which items are equipped with sensors and actuators that allow them to detect their surroundings,

interact with one another, and exchange data over the internet [2].

Currently, there are over 50 billion internet-connected gadgets (IoT), and in the coming years, this number is predicted to increase significantly. Numerous applications can benefit from the massive amount of data generated by these enormous numbers of devices. Food, demotics, smart farming, e-health, agriculture, assisted living, and improved education are just a few examples of the many scenarios where IoT technologies might be used.

IoT devices are prone to several possible security risks since they are linked to the worldwide internet via immature and insecure communication protocols and apps. One instance of this kind of problem happened in October 2016, when DVRs and webcams, two common IoT devices, were taken advantage of and turned into attack tools in order to compromise a major DNS provider (Dyn). A massive Internet outage resulted from this operation, and popular websites like Amazon, Netflix, and Twitter were no longer accessible. For IoT applications, an accurate intrusion detection technique is also necessary.

Numerous intrusion detection systems (IDS) have been put out to shield Internet of Things devices from cybercriminals. However, because the Internet of Things is connected to the world wide web, there is a considerable danger of intrusion by malicious actors who can defeat preventive measures. To overcome these drawbacks, a novel INtrusion Detection in Iot using Advanced deep learning NETwork (INDIA-NET) has been proposed, which uses novel deep learning technique and detects the intrusion efficiently. The major contributions of the proposed INDIA-NET have been given as follows.

- Initially, the data from the dataset are preprocessed using Minkowski distance based nearest neighbor

technique in which the replica and repeated data are detached from the dataset.

- After preprocessing an enhanced transient search optimization algorithm has been used for feature selection.
- After selecting the features, the features are sent to the novel convolutional neural network combined Generative Adversarial Network (CNN-GAN) which the classify the output into DOS attack, UR3 attack and normal.
- The effectiveness of the developed technique has been calculated utilizing specific parameters involving accuracy, precision, recall, and detection rate respectively.

The remaining sections of the paper are organized as follows. Section 2 describes the literature review for intrusion detection frameworks for IOT in detail. Section 3 describes the proposed INDIA-NET for detecting attacks in detail. Section 4 describes the results and discussion and section 5 discusses the conclusion section.

2. LITERATURE REVIEW

In 2023, Sharma, B., et al. [1] introduced a unique anomaly-based IDS solution for IoT networks. The accuracy of the suggested model was 84%. To address the issue of class imbalance in the dataset, Generative Adversarial Networks (GANs) were utilised to produce synthetic data of minority attacks, and with a balanced class dataset, they reached 91% accuracy.

In 2023, Jothi, B. and Pushpalatha, M. [2] proposed a novel intrusion detection system (IDS) that leverages efficient DL frameworks. Driven by the advantages of Long Short-Term Memory (LSTM), whale integrated LSTM (WILS) frameworks have been suggested to project efficient IDS to identify a variety of distinct situations of threats on IoT networks. The accuracy, precision, and recall of the WILS models surpass those of other intelligent models currently in use, demonstrating their suitability for IoT networks.

In 2023, Lazzarini, R., et al [3] introduced the Deep Integrated Stacking for the IoT (DIS-IoT) model, which creates a freestanding ensemble model by integrating 4 dissimilar DL models into a fully linked DL layer. The outcomes were juxtaposed with those of other cutting-edge studies that employed comparable techniques on the identical ToN_IoT dataset and were found in the literature. In binary classification, DIS-IoT performs similarly to other methods, but it surpasses them in multi-class classification.

In 2023, Elnakib, O., et al. [4] suggested an improved anomaly-based Intrusion Detection Deep learning method. A 95% accurate multi-class classification model (EIDM) is able to categorise 15 different traffic behaviours, including 14 different forms of attacks, using data from the CICIDS2017 dataset. Additionally, six kinds of network traffic behaviour are classified using four customized baseline DL frameworks.

In 2023, Thakkar, A. and Lohiya, R., [5] introduced the Bagging classifier that employs a deep neural network

(DNN) as a base estimator as a means of addressing the challenge of class imbalance through the use of ensemble learning. The outcomes of the suggested method are also statistically examined through the use of the Wilcoxon signed-rank test.

In 2023, Bhavsar, M., et al. [6] created an Intrusion Detection System (IDS) that was based on a DL framework known as Pearson-Correlation Coefficient - Convolutional Neural Networks (PCC-CNN). Lastly, they examine and contrast the PCC-CNN model with five conventional PCC-ML models. The suggested DL-based IDS performs better than conventional techniques.

In 2023, Hosseini, S. and Sardo, S.R., [7] presented a hybrid approach that combines deep learning and shallow learning. The suggested technique makes use of a Siamese neural network-based classical to improve data classification after first employing a spider monkey optimisation feature selection framework. The suggested model's accuracy, calculated with a random forest classifier, is 94.69%. Furthermore, the suggested model's minimal computational load makes it a suitable option for Internet of Things devices with limited computing power.

3. INTRUSION DETECTION IN IOT USING ADVANCED DEEP LEARNING NETWORK

In this section a novel INtrusion Detection in Iot using Advanced deep learning NETWORK (INDIA-NET) has been proposed, which uses novel deep learning technique and detects the intrusion efficiently. Initially, the data from the dataset are preprocessed using Minkowski distance based nearest neighbor technique in which the identical and replicated data are detached from the dataset. After preprocessing an enhanced transient search optimization algorithm has been used for feature selection. After selecting the features, the features are sent to the novel convolutional neural network combined Generative Adversarial Network (CNN-GAN) which the classify the output into DOS attack, UR3 attack and normal. The overall architecture of the proposed INDIA-NET has been given in Figure 1.

3.1. Preprocessing

The Minkowski distance is used to calculate the distance between each pair of data, and similarity measures of the dataset's data are used to start this process. Reproduction and redundant records are then eliminated from the dataset and sent into the next stage, where missing values are replaced by computing the nearest neighbour, in order to avoid the classifier from being biased in favour of more common records. Imputation, also known as deficient cost replacement, is a technique for estimating the relevant cost of absent features included in the data. By locating the closest neighbours of the lacking values, the proposed cost of the missing attributes is used to replace the lacking cost in this method of altering lacking records. Suppose there's a dataset with n records and w characteristics per file. We also know that D_1 has a missing attribute. The K closest neighbours of D_1 are then ascertained using the Euclidean distance (E_d) in the following manner:

$$E_d = \sum_{x=1}^n |D_1 - D_x|^2 \quad (1)$$

The K nearest neighbour records for D_1 are found based on the distance value, and D_1 's attribute value is calculated using the attribute values of its k neighbours. The

neighbour records' values are used to calculate the mean value of a, which is then used to replace the missing value. In order to remove all uncertainties from the dataset, all duplicated data is removed in this stage and missing values are restored.

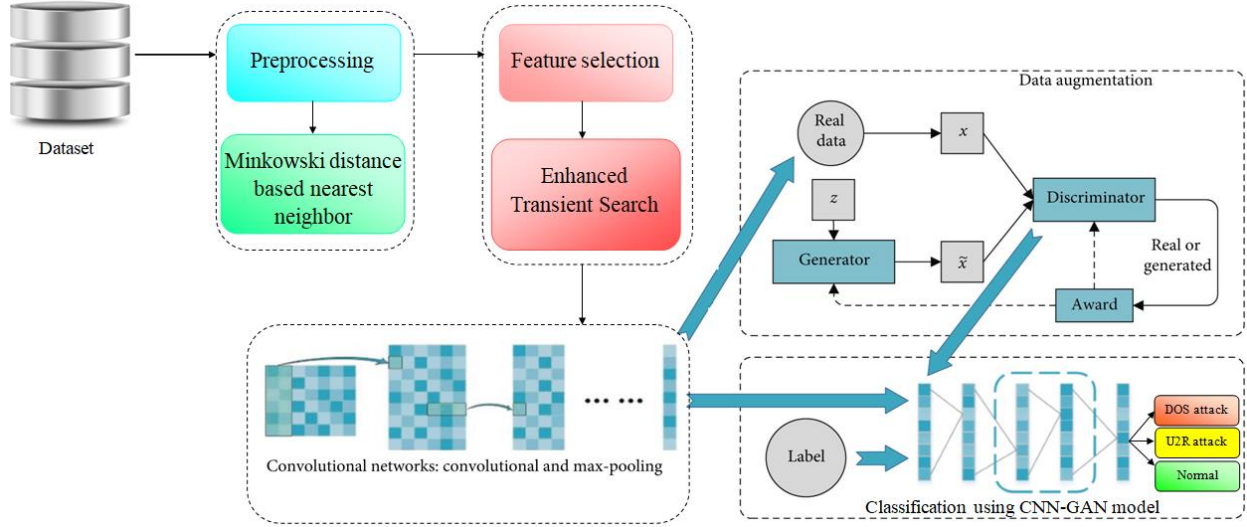


Figure 1. INDIA-NET architecture

3.2. Feature selection phase

To enhance intrusion detection in the IoT, pertinent characteristics are chosen based on their quality at this step of the established FS model, as seen in Figure 1. This is accomplished by utilising a local optimizer technique-based improved version of TSO, which is on the basis of operator of DE. The first step in the established FS approach, called ETSO, is to create the starting population P, which consists of M agents. Next, each agent is transformed into its binary form, and the training set is shrunk by eliminating the characteristics that match zeros inside this binary form. The performance of the chosen feature is then calculated using the KNN classifier's error classification. The agent with the highest fitness value is then assigned. The operators of TSO and DE, together with this best agent, state that the agents in the present population are modernized until they arrive at the best answer.

1) INITIAL POPULATION CONSTRUCTION

The created TSOE begins by separating the dataset into training and testing sets, which are represented by 80% and 20% of the total sample size, respectively. The starting value for a collection of M agents P, that describes the initial population, is then formed using ETSO using Eq. (2).

$$P_x = LL + rand(1, R) \times (UL - LL) \quad (2)$$

where R, the number of retrieved features, denotes each agent's dimension. R values at random from [0, 1] are indicated by the notation rand(1,R). The boundaries of the search domain are denoted by LL and UL.

2) UPDATING POPULATION

This level begins by using Equation (3) to transform each P_x agent into its Boolean form, AP_x .

$$AP_{xy} = \begin{cases} 1 & \text{if } P_x > 0.5 \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

Eq. (3) states that the large range of functions in the training set is minimized by casting off the ones functions that resemble zeros. The health cost for each agent P_x is then determined using the following equation.

$$f_x = \lambda + \gamma_x + (1 - \lambda) \times \left(\frac{|AP_x|}{R} \right) \quad (4)$$

In Equation (4), γ_x represents the classification error determined by a KNN classifier using the training set. Random weights with $\lambda \in [0,1]$ are employed to maintain equilibrium between the classification error and the ratio of pertinent features $\left(\frac{|AP_x|}{R} \right)$. To ensure clarity, let us assume the following:

$H1_x = [0.0975, 0.2785, 0.5469, 0.9575, 0.9649, 0.1576, 0.9706]$ represents the seven properties (dimension) of the current agent P_x . Once Eq.(4) is used, $AP_x = [0, 0, 1, 1, 1, 0, 1]$. This indicates that the training set is reduced by utilizing the third, fourth, fifth, and seven characteristics, which are selected as relevant features. The evaluation of these selected features is done using Equation (5). Subsequently, the optimal agent P_{best} with the highest fitness value f_{best} is identified. P_{best} is then used to update the current agents with the TSO and DE operators. By enabling these operators to operate in a competitive way, the process is carried out and the variety of the agents is maintained.

3) TERMINAL CONDITIONS

The stopping requirements are examined at this point, and if necessary, action is taken. If this is not the case, the P_{best} is returned and the updating stage is repeated.

3.3. Classification

The convolutional neural network combined generative adversarial networks (CNN-GAN) are utilized to classify the risks into DOS, UR3, and normal. The fs process features are embedded by CNN into fixed-length features,

which are subsequently fed into the backpropagation network (BP network) and generative adversarial network (GAN), respectively. When working with small-scale datasets, the created data not only improves the attack features but also successfully addresses the issues of inadequate samples and single information. Following that, a BP network is trained using a mixture of the produced samples from GAN and the real samples. Because of the modular nature, attack information is shared throughout modules as processed features.

This method allows for the selection of unique and environmentally friendly characteristics for complex information assessment tasks with high dimensionality by utilizing the hierarchical and adaptive properties of CNNs. Furthermore, by using CNNs for characteristic selection, we can automatically extract the most informative functions from the input data without the need for manual characteristic engineering. In order to designate CNN for feature selection, we first instruct the CNN version on how to use the provided dataset. Next, we extract the final convolutional layer from the proficient rendition. Then, since the original version outputs the activations of the closing convolutional layer, we generate a new version that accepts equal inputs. In order to determine the essential functions, we finally compute the mean activation over all test samples. The activations of the convolutional layers represent the learned features in the input data, and features with in height mean activation indicate their significance. The input data's learnt features are represented by the convolutional layer activations, and those with high mean activation suggest their importance. It is important to note that the suggested activations are no longer limited to the final convolutional layer within the CNN network; they can now be generated for any convolutional layer. However, we often focus on the final convolutional layer since it possesses the most informative and discriminative characteristics that aid in identifying IoT attacks.

3.3.1. Data augmentation

The module functions as an adversarial network, wherein the discriminator treats the convolutional attack data as authentic data, and the generator may produce several attack features directly using the similar delivery of processed real-world data. The adversarial network may therefore provide a variety of produced features that comprise varied and instructive assault information assumed the processed attack features as input.

In order to enhance the general framework performance, the data augmentation component has been streamlined. The module's output takes the form of attack characteristics with semantic information instead of phrases or texts. The adversarial network no longer needs to establish a connection with the structure that translates attack traits into words or documents, based on the notions discussed above. In order for the adversarial network to produce attack features that fall under each category, all of the processed features along with their respective categories are fed into it during network training.

The Generator

The vanishing and expanding gradient problem in backpropagation are the reason for the widespread usage of recurrent neural networks (RNNs) in natural language processing (NLP) due to the discontinuous distribution of textual data.

$$u_t = \sigma(Z_{u_t} \cdot [s_{t-1}, a_t]) \quad (5)$$

where the vector concatenation is $[s, a]$. The amount of data that is written from the previous state to the present candidate set \tilde{s}_t is managed by the reset gate e_t . Less information is written from the prior state the smaller the e_t . The reset gate serves the following purposes:

$$e_t = \sigma(Z_{e_t} \cdot [s_{t-1}, a_t]) \quad (6)$$

The GRU determines how the generative network is set up. The update gate u_t controls how much data is sent from the previous state into the current state. The higher the value of u_t , the more state information is brought at the previous time. The modified gate carries out the subsequent tasks:

$$\tilde{s}_t = \tanh(Z_{\tilde{s}_t} \cdot [e_t \odot s_{t-1}, a_t]) \quad (7)$$

$$s_t = (1 - u_t) \odot s_{t-1} + u_t \odot \tilde{s}_t \quad (8)$$

The elementwise product is represented by \odot . To create the mapping from noise space to text semantic space, we feed random noise into the generator.

The Discriminator

Convolutional networks are built to distinguish between the source of the text and the generated text since both may be quantified as features of a given length. The following is how the textual feature $\{a_1, a_2, \dots, a_m\}$ is processed:

$$k_x = f(a_x \otimes z + i) \quad (9)$$

where $z \in \mathbb{R}^n$ is a 1D kernel to create a new feature map, i is a bias term, a_x is the 1-dim textual features, and \otimes is the convolution process. Different characteristics are extracted using different numbers of kernels with varied window widths. In particular, the textual feature that was extracted using window size n by kernel z is expressed as follows:

$$k_x = [k_{x,1}, k_{x,2}, \dots, k_{x,l-n+1}] \quad (10)$$

The max-pooling procedure is used to shift all pooling features from different kernels to a fully linked softmax layer, which is then applied to the feature map $\tilde{k} = \max(s)$ to estimate the probability that a given feature is real.

Classifier Module

The multilayer-feedforward network created by dense networks is a classifier that uses the error backward propagation technique. The idea is to iteratively compute the variance between the actual and predicted outputs, and then the network modifies the weights based on that difference.

The network receives $Z_{l:g}$ and produced features $Z^*_{l:g}$ as input for training. ReLU serves as the activation in this case. A dense network with three to five completely linked layers is built in order to investigate the best possible network structure. In addition, the dropout probability is 0.5 and the dropout modules are put between the fully linked layers to prevent overfitting. The network is trained using the produced data $Z^*_{l:g}$ and the real data $Z_{l:g}$. The following are the activation function and its derivative:

$$f(a) = \max(0, a) = \begin{cases} 0, & a < 0 \\ 1, & a \geq 0 \end{cases} \quad (11)$$

4. RESULTS AND DISCUSSION

The experimental design and content, including datasets, baselines, parameter setup, and analysis of the experimental results, are covered in detail in this part. We evaluate the suggested INDIA-NET system's performance using important performance indicators including F-score, accuracy, precision, and recall.

4.1. Dataset Description

We utilised the NSL-KDD benchmark dataset, the most recent iteration of the KDD'99 dataset 27, to assess INDIA-NET's performance. This dataset, which has 22,554 records for testing and 125,973 records for training, includes DoS and U2R cyberattacks. There are 41 features in all, which include source bytes, destination bytes, flag, protocol, service, and duration.

4.2. Performance comparison

The efficacy of the suggested method has been assessed utilizing specific parameters involving accuracy, precision, recall, f1 score, and detection rate which are represented as follows. Performance comparison shown in Figure 2.

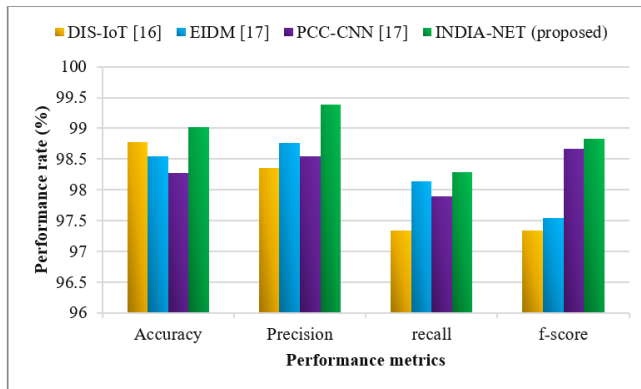


Figure 2. Performance comparison

An essential metric for assessing classifiers' efficacy in intrusion detection is accuracy. It is evident that the INDIA-NET strategy outperforms earlier methods like DIS-IoT [16], EIDM [17], and PCC-CNN [19]. Our suggested work's accuracy is compared to earlier research, which further supports the idea that the INDIA-NET framework performs better. This is because earlier research, which used unclear datasets and non-optimal feature processing, was unable to forecast the class of data. By overcoming these problems, we were able to get a high level of accuracy. The

other performance metric we studied was the capacity to recall previously published and suggested works. It was found that the INDIA-NET framework has strong performance in recall metrics as well, a crucial aspect of intrusion detection. Therefore, all anomalies can be detected by the suggested INDIA-NET framework with little errors.

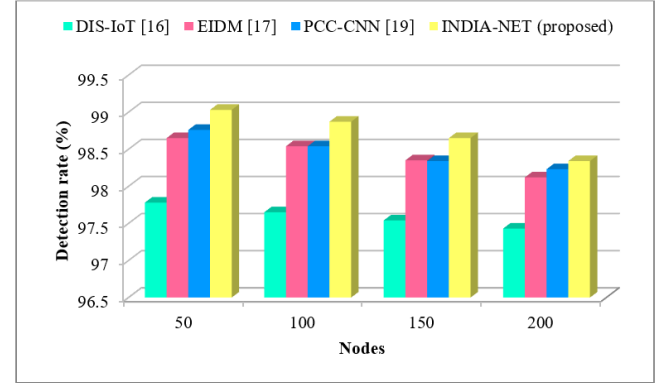


Figure 3. Detection rate comparison

Figure 3 shows the detection rate comparison of suggested INDIA-NET with the existing techniques such as DIS-IoT [16], EIDM [17], and PCC-CNN [19]. The comparison has been made with varying number of nodes ranging from 50 to 200. As the number of nodes increases, the detection rate decreases. However, the proposed INDIA-NET achieves higher detection rate than existing techniques.

5. CONCLUSION

In this paper, a novel INtrusion Detection in Iot using Advanced deep learning NETWORK (INDIA-NET) has been proposed, which uses novel deep learning technique and detects the intrusion efficiently. The proposed method uses a novel convolutional neural network combined Generative Adversarial Network (CNN-GAN) which classifies the output into DOS attack, UR3 attack and normal. The efficacy of the suggested technique has been calculated employing specific parameters involving accuracy, precision, recall, and detection rate respectively. Experimental results shows that the developed method achieves higher detection rate than other existing techniques such as DIS-IoT, EIDM, and PCC-CNN respectively. The proposed INDIA-NET system has been evaluated utilizing the NSL-KDD dataset, and the suggested work attains better performance in accuracy, precision, recall and F1-score of 99.02%, 99.38% , 98.29% and 98.83% repectively. In the future, we want to assess INDIA-NET using a variety of classifiers, like random forests, decision trees, and Naive Bayes, on datasets like KDD-99 and UNSW-NB 15.

CONFLICTS OF INTEREST

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

FUNDING STATEMENT

Not applicable.

ACKNOWLEDGEMENTS

The author would like to express his heartfelt gratitude to the supervisor for his guidance and unwavering support during this research for his guidance and support.

REFERENCES

- [1] B. Sharma, L. Sharma, C. Lal, and S. Roy, "Anomaly based network intrusion detection for IoT attacks using deep learning technique", *Computers and Electrical Engineering*, vol. 107, pp.108626, 2023. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [2] B. Jothi, and M. Pushpalatha, "WILS-TRS—A novel optimized deep learning-based intrusion detection framework for IoT networks", *Personal and Ubiquitous Computing*, vol. 27, no. 3, pp.1285-1301, 2023. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [3] R. Lazzarini, H. Tianfield, and V. Charissis, "A stacking ensemble of deep learning models for IoT intrusion detection", *Knowledge-Based Systems*, vol. 279, pp.110941, 2023. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [4] O. Elnakib, E. Shaaban, M. Mahmoud, and K. Emara, "EIDM: deep learning model for IoT intrusion detection systems", *The Journal of Supercomputing*, pp.1-21, 2023. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [5] A. Thakkar, and R. Lohiya, "Attack classification of imbalanced intrusion data for IoT network using ensemble learning-based deep neural network", *IEEE Internet of Things Journal*, 2023. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [6] M. Bhavsar, K. Roy, J. Kelly, and O. Olusola, "Anomaly-based intrusion detection system for IoT application", *Discover Internet of Things*, vol. 3, no. 1, pp.5, 2023. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [7] S. Hosseini, and S.R. Sardo, "Network intrusion detection based on deep learning method in internet of thing", *Journal of Reliable Intelligent Environments*, vol. 9, no. 2, pp.147-159, 2023. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]

AUTHORS



K. Anusha is an Associate Professor at the School of Computer Science and Engineering at VIT Chennai, Tamilnadu India. She has received her Doctor of Philosophy from VIT University, Vellore, Tamil Nadu, India and has completed her MTech (Software Engineering) from Bharathidasan University Tamilnadu, India. Her research interests include Network Security, Wireless Networks, Information security and Mobile Adhoc networks.

She has number of research Publications in National, International Journals Conferences. She is also Editorial Board member for various Journals. She has chaired many international conferences' and delivered invited, technical lectures along with keynote addresses.



B. Muthu Kumar, Professor, School of Computing and Information Technology, REVA University, Bengaluru received his BE (CSE) degree from Anna University, Chennai in the year 2005, MTech (CSE) (Gold Medalist) received from Dr. MGR. University, Chennai in the year 2007 and Doctoral degree from St. Peter's University, Chennai in the year 2013. He is having more than 16 years of teaching experience in reputed

engineering colleges. He has published more than 40 peer reviewed International Journals, 50 International/National Conference and attended more than 150 Workshops/FDPs/Seminars etc., He organized many events like Conference/FDPs/Workshops/Seminars/Guest Lecture. He has

published more than 10 patents in various fields like Wireless Sensor Networking, Image Processing, Optimization Techniques and IoT. He received nearly 5.67 Lakhs funding from various agencies like AICTE, ATAL and IEL. He has written 2 books from reputed publishers. He received Best Researcher Award in the year 2021 and Innovative Research and Dedicated Professor Award in Computer Science and Engineering in the year 2018. He has professional membership on ISTE, CSI, IEL, IACSIT, IAENG, CSTA, and SIAM. He has invited as Guest Lecture / Chairperson / Examiner / Reviewer/Editorial Board Member in various Institutions/Journals/Conferences. He is a recognized supervisor in Anna University, Chennai and currently guiding 4 research scholars. His areas of interest are Image Processing, Wireless Networks, IOT and Computing Techniques.



J. Ragaventhiran received his PhD from Anna University, Chennai, India, M.E degree in Computer Science and Engineering from Anna University, Chennai from Madurai Kamaraj University and Anna University, Tamil Nadu, India in 2002 and 2008 respectively. He is currently pursuing PhD in the Department of Information and Communication Engineering in Anna University, Chennai, Tamil Nadu, India and B.E Computer Science and Engineering from Madurai Kamaraj University, Tamil Nadu, India. He is currently working as a Professor in School of Computing and Information Technology, REVA University, Bengaluru, India. His research interest includes Data Mining, Big Data, Machine Learning. He is a life member in Computer Society of India(CSI) and Institution of Engineers(IEI) India. He has reviewed and chaired various national and international conferences including IEEE conferences.

Arrived: 27.11.2023

Accepted: 25.01.2024