

Hierarchical enhanced deep encoder-decoder for intrusion detection and classification in cloud IoT networks

Ramya K M¹, Rajashekhar C. Biradar²

¹Research Scholar, REVA University, BMSCE, Bangalore, India

²Reva University, Bangalore, India

Article Info

Article history:

Received Jul 25, 2024

Revised Apr 18, 2025

Accepted Jul 2, 2025

Keywords:

Cloud computing

HED-EDFD

Intrusion detection

IoT

ABSTRACT

Securing cloud-based internet of things (IoT) networks against intrusions and attacks is a significant challenge due to their complexity, scale, and the diverse nature of connected devices. IoT networks consist of billions of devices, computer servers, data transmission networks, and application computers, all communicating vast amounts of data that must adhere to various protocols. This study introduces a novel approach, termed hierarchical enhanced deep encoder-decoder with adaptive frequency decomposition (HED-EDFD), and is designed to address these challenges within cloud-based IoT environments. The HED-EDFD methodology integrates adaptive frequency decomposition, specifically adaptive frequency decomposition, with a deep encoder-decoder model. This integration allows for the extraction and utilization of frequency domain features from time-sequence IoT data. By decomposing data into multi-resolution wavelet coefficients, the model captures both high-frequency transient changes and low-frequency trends, essential for detecting potential intrusions. The deep encoder-decoder model, enhanced with deep contextual attention mechanisms, processes these features to identify complex patterns indicative of malicious activities. The hierarchical structure of the approach includes a hierarchical wavelet-based attention mechanism, which enhances the accuracy and robustness of feature extraction and classification. To address the issue of imbalanced intrusion data, a cosine-based SoftMax classifier is employed, ensuring effective recognition of minority class samples.

This is an open access article under the [CC BY-SA](#) license.



Corresponding Author:

Ramya K M

Research Scholar, REVA University, BMSCE

Bangalore, India

Email: ramyakm_12@rediff.com

1. INTRODUCTION

The internet of things (IoT) has recently become a major trend due to its extraordinary potential to connect various heterogeneous smart sensors and devices. Currently, IoT devices are applied in diverse domains, including health, smart homes, smart grids, transportation, the environment, infrastructure, and public services. More applications for this technology are being discovered almost daily. However, IoT systems face significant challenges due to limited storage and processing power. They suffer from drawbacks such as security, reliability, integrity, confidentiality, and performance issues. To address these challenges, the integration of IoT with cloud computing, known as the cloud of things (CoT), has emerged as a viable solution. Many researchers have acknowledged that cloud computing helps address the issues associated with IoT by providing reliability, ubiquity, and scalability, along with a high-performance environment for

implementing IoT devices [1], [2]. The cloud environment for IoT networks and platforms offers connectivity among IoT devices and applications, as well as distributed computational resources and storage. Cloud computing is structured into three distinct layers: the system layer, the platform layer, and the application layer. The first two layers primarily address virtual machines (VMs) and operating systems. In contrast, the third layer focuses on applications hosted in the cloud, such as web-based applications. This framework underscores the advantages and extensive use of cloud computing. Despite these benefits, cloud-based IoT networks remain vulnerable to several security threats. As illustrated in Figure 1, common cyber-attacks can be categorized into network-related and other groups, which could seriously harm these networks. Ensuring IoT security, including data security, requires robust methods for detecting and resisting network intrusions and attacks. With the growing prevalence of IoT and smart devices, these systems have become prime targets for hackers. Even simple devices can have numerous exploitable vulnerabilities. While IoT devices offer many advantages and conveniences for accessing the Internet, they also impose higher requirements for network security and data protection [3], [4].

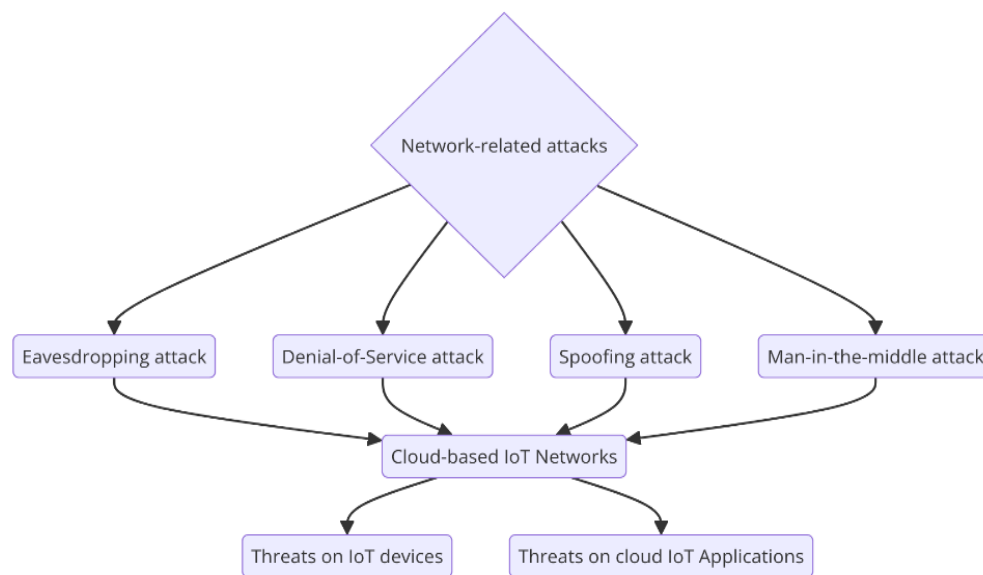


Figure 1. Cloud-based network intrusion in IoT

Figure 1 presents network-related attacks on cloud-based IoT networks, such as eavesdropping, denial-of-service (DoS), spoofing, and man-in-the-middle attacks, which compromise IoT device functionality and cloud applications, leading to data breaches and service disruptions. These attacks target the network layer and are considered a form of intrusion as they involve unauthorized access and interference with network communications. For instance, botnets can control simple terminal devices to form large networks that launch attacks or steal confidential information from the entire IoT network. More alarmingly, hackers can use network intrusions to silently steal or tamper with private information on IoT networks. Consequently, IoT security, particularly data security, has become an urgent research priority. Significant advancements have been made in this field, especially with the application of artificial intelligence and deep learning (DL) technologies in IoT security. Machine learning (ML) and DL techniques are highly effective for processing vast amounts of data, delivering superior computational results. These techniques can either select optimal features from datasets for classification or automatically extract relevant features using DL methods. Within the context of cloud IoT environments, cybersecurity has become a significant concern. This can be addressed by deploying robust intrusion detection systems (IDS) at edge nodes to monitor and secure data traffic across the network. Over the past decades, ML and DL-based network intrusion detection (NID) systems have shown remarkable effectiveness in detecting attacks within IoT networks, making them crucial for enhancing cybersecurity in these environments [5]. Convolutional neural networks (CNN), a DL architecture, have garnered significant attention from researchers due to their exceptional performance in handling image data in areas such as computer vision, image recognition, and segmentation. Consequently, CNNs have been applied in various domains, including networking and medical image processing [6]. Additionally, CNNs have demonstrated efficiency in dealing with numeric tabular datasets used for modeling IDS in cloud IoT environments. Traditional ML algorithms, such as decision tree (DT), random forest (RF),

extra tree (ET), and extreme gradient boosting (XGBoost), have shown strong performance in classifying network traffic that is not represented as image data [7], [8]. Other DL algorithms, like recurrent neural networks (RNN), have been used for time series data analysis with good accuracy. However, RNNs face challenges related to vanishing and exploding gradients, which CNN architectures can overcome. Therefore, CNNs are preferred for image data due to their high performance in image classification, particularly in cloud IoT security scenarios [9], [10].

The rapid proliferation of IoT devices has led to the emergence of complex, large-scale networks that generate vast amounts of data. These networks, often integrated with cloud computing for enhanced scalability and processing power, are increasingly susceptible to various security threats, including intrusions and attacks. Traditional IDS struggle to keep pace with the dynamic and distributed nature of cloud-based IoT environments, failing to capture the intricate patterns and anomalies present in the data. There is a pressing need for advanced methodologies that can efficiently analyze and detect threats in real time, ensuring the security and integrity of IoT networks. This research aims to address these challenges by developing a robust and scalable intrusion detection framework that leverages the strengths of both adaptive frequency decomposition and DL.

- Hierarchical enhanced deep encoder-decoder with adaptive frequency decomposition (HED-EDFD): we propose a novel HED-EDFD model that integrates adaptive frequency decomposition for feature extraction with enhanced encoder-decoder architecture enhanced by deep contextual attention mechanisms. This hybrid approach enables the effective capture and analysis of both temporal and frequency domain features from IoT data.
- Advanced feature extraction: the use of adaptive frequency decomposition allows for the decomposition of IoT data into multi-resolution wavelet coefficients, capturing both high-frequency transient changes and low-frequency trends. This enhances the ability to detect complex patterns indicative of intrusions.
- Extensive experiments using IoT-23, KDD 99, and TON datasets demonstrate the superior performance of the HED-EDFD model, showing significant improvements in F-score, accuracy, and AUC metrics compared to traditional methods.

2. RELATED WORK

NIDS are designed to monitor large volumes of network traffic and identify malicious activities, making them essential for securing cloud-based IoT environments. Upon detecting abnormal behavior, NIDS sends real-time alerts to administrators to mitigate potential attacks, ensuring the protection of interconnected devices and data in cloud-IoT systems. Tuan *et al.* [11], conducted a comparative study evaluating the performance of various ML methods in classifying Botnet attack traffic. They assessed support vector machine (SVM), multilayer perceptron (MLP), DT, Naive Bayes (NB), and unsupervised methods like K-means clustering on datasets including KDD'99. The study revealed that unsupervised methods achieved the best performance with 98% accuracy. Shao *et al.* [12], an ensemble of hoeffding tree and RF models was created using online learning for both normal and attack traffic. The work in [13] introduced a feature selection technique as a preprocessing step for an ML-based botnet attack detector, ranking features by Pearson correlation coefficients to optimize the detector's performance on the Bot-IoT dataset. Pujar *et al.* [14], developed an attack detection algorithm involving feature extraction from network traffic and ML classifiers such as K-nearest neighbor (KNN), SVM, DT, and MLP, evaluated on a dataset collected in their study. Sreedhara *et al.* [15] focused on an MLP-based Mirai Botnet detector specifically for software defined networks (SDN), feeding five metrics, including communication protocols, into the MLP. The study [16] addressed cloud computing challenges by developing a filter-based ensemble feature selection (FEFS) and deep learning model (DLM) combining RNN with Tasmanian devil optimization (TDO) for intrusion detection, showing improved security based on metrics like F-measure, specificity, sensitivity, and accuracy.

Maheswari *et al.* [17], enhanced IDS performance by integrating teacher learning optimization with deep recurrent neural networks (TL-DRNN) and using modified manta-ray foraging optimization (MMFO) for feature selection, validated with standard datasets to improve false positive and negative rates, accuracy, precision, recall, specificity, and F-measure. Elaziz *et al.* [18] proposed using swarm intelligence algorithms and deep neural networks (DNN) for intrusion detection in IoT-cloud systems, alongside the Capuchin search algorithm (CapSA) for feature selection, demonstrating competitive performance across various datasets. The study [19] introduced the ensemble intrusion detection model for cloud computing using deep learning (EICDL) to tackle issues like privacy, confidentiality, and quantum computing attacks, achieving higher precision, accuracy, and recall compared to other methods. Zhang *et al.* [20] applied DNN for intrusion detection in online music education on public cloud networks, utilizing fuzzy logic-based feature selection, Salp swarm optimization, and the integration of gated recurrent unit (GRU) and CNN, resulting in higher intrusion detection accuracy. Parameswari *et al.* [21] developed the rat swarm hunter prey optimization-deep

maxout network (RSHPO-DMN) model for intrusion detection in computer networks, showing superior performance in accuracy, precision, recall, and F1-score. Joraviya *et al.* [22] explored security in containerized cloud environments using DL methods, specifically CNNs, for anomaly detection in system call sequences and images, enhancing detection accuracy and reducing false positives and negatives. The framework in [23] analyzed and labeled incoming traffic packets using an auto-associative deep random neural network and an online estimate of its statistically measured trustworthiness, enabling IDS to adapt to time-varying network traffic characteristics and eliminating the need for offline data collection. The model in [24] introduced a resilient federated learning architecture for diverse IoT environments and vehicular networks, emphasizing versatility and modularity. Finally, Nakip and Gelenbe [25] developed a novel IDS to address class imbalance in federated learning at both local and global levels, demonstrating improved generalizability in detecting various attacks under both IID and non-IID data settings.

Despite significant advancements in IDS for cloud-based IoT networks, there remains a substantial research gap in effectively capturing and analyzing the complex, high-dimensional data generated by these environments. Traditional methods often fall short in addressing the real-time processing and scalability requirements needed to detect sophisticated and evolving threats. There is a critical need for innovative approaches that leverage advanced signal processing and DL techniques to enhance the accuracy and efficiency of intrusion detection in cloud IoT networks.

3. PROPOSED METHOD

Figure 2 illustrates a hierarchical IoT security system. It shows the flow of data from the application layer through the cloud security layer to the cloud servers. The process begins with flow gathering, followed by detecting symptoms of intrusion. Intrusion detection for the network and modifying core packets are handled next. Managing and monitoring include security management, network infrastructure monitoring, and network function management. This structured approach ensures comprehensive monitoring and security across cloud servers and IoT networks.

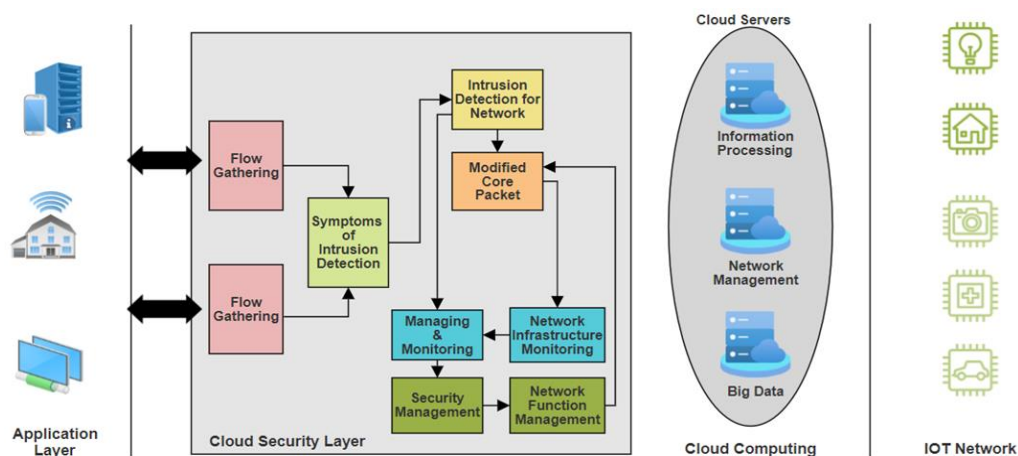


Figure 2. Proposed cloud IoT framework for intrusion detection

As shown in Figure 2, the IoT is a complex distributed hierarchical system comprising billions of devices, computer servers, data transmission networks, and application computers. The data transmitted within IoT networks are complex due to various transmission protocols, making intrusion and attack prediction challenging. The primary challenge in this study is recognizing and extracting useful features from this transmitted data. Given that IoT data are time sequences, extracting frequency domain features is crucial for effective classification, warranting in-depth exploration. The proposed model HED-EDFD is designed as a hierarchical cloud IoT security model considering the three hierarchical i.e., function-based, infrastructure-based, and management-based; the management module ensures the data security along with data transmission. The functional security module ensures the intrusion data packets and the fault detection module ensures network intrusion.

3.1. Hierarchical enhanced deep encoder-decoder with adaptive frequency decomposition

This section of the paper introduces the detection of intrusion for the transmitted information in the IoT models. The proposed approach includes the distinct wavelet enhanced encoder-decoder model that is used in the detection of IoT models. The framework of the proposed system is specified in Figure 3.

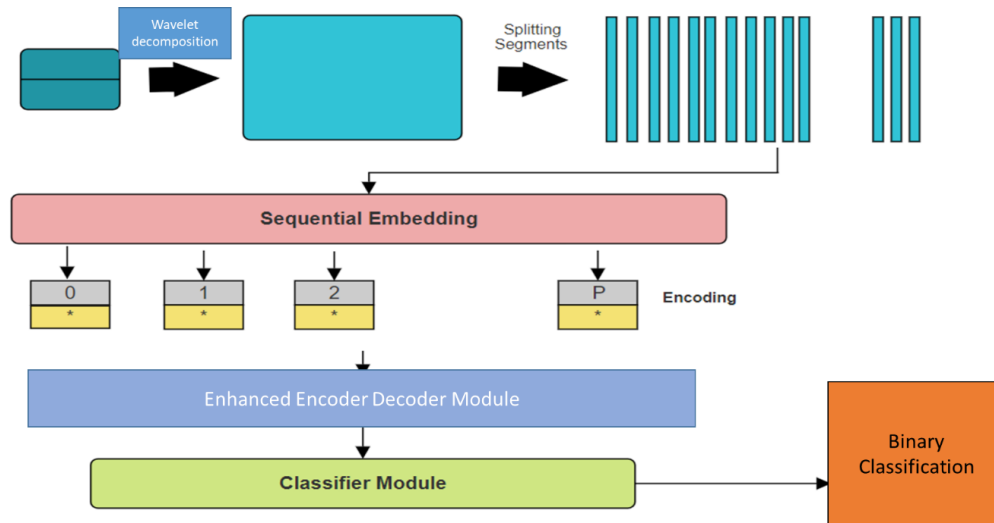


Figure 3. Proposed intrusion detection model

The wavelet transform mechanism is embedded into the enhanced encoder decoder system for retrieval of data through the frequency domain which is essential for the detection of intrusions using actions. Time signals have various applications considering real-time. The data that is used from frequency-time spectrograms are proved to be more helpful in comparison to the original signals of vibration. The short-time Fourier transform is generally used for image processing tasks as well as time sequential processing, however, this technique does not possess flexibility while considering time sequential processing, as the solution is constant at all frequencies as well as periods. Therefore, an analytical transform mechanism called adaptive frequency decomposition that has increased flexibility is utilized. This mechanism grasps the localization idea from the short-time Fourier transform that resolves the challenge of the window dimension not changing with the frequency. The major characteristics are that it can highlight the features of some parts of the problem using transformation as well as frequency localization can be analyzed. Gradually the signal is refined using the operations of translation as well as expansion. In conclusion, the sub-division of time can be attained at increased frequencies as well as the sub-division of frequency can be attained at lower frequencies. The method automatically adapts to the needs of frequency time signal which increases the focus on the signal details.

3.1.1. Optimized data embedding and processing

Assume a sequential signal denoted as $z(v)$, we consider the wavelet to be a series of convolution computations between the sequential signal $z(v)$ and the main wavelet function that is denoted as $\omega(v)$. This results in a series of coefficients for the wavelets that is given as (1). From the (1), $\omega^*(\cdot)$ is the conjugate of $\omega(\cdot)$. The main base function that is termed as $\omega(\cdot)$ has a form as given in (2).

$$Y_z(\mathfrak{Z}, u; \omega) = \int_{-\infty}^{\infty} z(v) \omega_{u,\mathfrak{Z}}^*(v) dv \quad (1)$$

$$\omega_{u,\mathfrak{Z}}(v) = \left(u^{\frac{-1}{2}}\right) \omega((v - \mathfrak{Z})(u)^{-1}) \quad (2)$$

Here, the scale factor is given u and the translating factor is denoted as \mathfrak{Z} . Once the pre-processing of information is completed, the data dimension is given as $T^{J \times Y}$, the length of direction for frequency and time is denoted using J and Y respectively. A one-dimensional sequence is utilized for embedding. The sample information is firstly divided into patches or segments, where every spectrogram of Z belongs to $T^{J \times Y}$ is

split for only the direction of time. This results in time sequences that have length denoted as p which is expressed as $z^r = [z_1, z_2, \dots, z_p]$, where z_k belongs to $T^{J \times r}$, every segment has a width of r and $p = \frac{Y}{r}$.

During the embedding process for the segments, the mechanism used focuses on a decreased dimension constant vector representation. Firstly, the information input is stored in the linear projection that is given as Y belongs to $T^{r \times f}$, the size of the input information is denoted as f . Once the sequence of the projected segment is obtained, a class token that is expressed as a_k is added. The spectrogram representation is obtained at this stage. Once the segment embeddings is completed, it is essential to perform the process of position embedding. After the position as well as segment embedding is completed, the information embedding form is given as (3). Here, a_k belongs to T^f represents the class token while $G_{position}$ belongs to $T^{(p+1)f}$ that represents the position embedding.

$$z_{gk} = [a_k; z_1 Y, z_2 Y, \dots, z_p Y] + G_{position} \quad (3)$$

3.1.2. Hierarchical enhanced encoder decoder

The information of the IoT systems is taken into account with the time sequence. The information has to be retrieved from the sequence in applications of natural language processing as well as data of lesser frequencies from the time sequence while they are being processed. Therefore, the adaptive frequency decomposition model is embedded as the enhanced encoder decoder system for the retrieval of data having a lesser frequency. The enhanced encoder-decoder model is modified in the proposed architecture as compared to the traditional model. For the decoder phase, attention learning is utilized in the adaptive frequency decomposition model. Here, the model consists of two stages, namely the decomposition and the reconstruction. These two stages are used together as well as can be used separately. The model majorly involves three phases, decomposition, attention, and the reconstruction phase. The modified adaptive frequency decomposition model phases include the reconstruction as well as the decomposition phase. The attention phase is a part of cross-attention.

The adaptive frequency decomposition model as well as the transform for multiple wavelets is essential and used widely for digital applications for signal processing mechanisms. They are applied for the retrieval of data from the frequency domain for digital signals that are essential to what the present applications of DL require. Prior studies have proposed to use of the learning algorithm for hierarchical wavelets. However, only one signal can undergo the processing process using this algorithm. Therefore, in this study, the attention-learning algorithm needs to grasp the process of learning various input signals to attain attention-learning.

The adaptive frequency decomposition model transforms a signal of one dimension to a two-dimensional time-scale expression that is generally applied to uninterrupted signal processes. The form that is transformed is expressed using (2). Considering the adaptive frequency decomposition model, the base wavelet for the adaptive frequency decomposition model is as expressed in (4).

$$\omega_{l,m}(v) = \left((u_0^l)^{-\frac{1}{2}} \right) \omega \left((v - m \beth_0 u_0^l) (u_0^l)^{-1} \right) \quad (4)$$

In the (4), l and m are integers, u_0 is greater than 1 and \beth_0 is a constant parameter. For this experiment, the values of u_0 is set to 2 and \beth_0 is set to 1. While the adaptive frequency decomposition model is applied, the base of the wavelet function has to be chosen. These functions should possess some characteristics that include regularity constraints, admissibility constraints as well as vanishing situations. Here, the base function for wavelets is taken as orthogonal polynomials, specifically Legendre polynomials for this study. The Legendre polynomials are denoted as $R_k(z)$ having weight function as $y_N(z)$ is equal to 1 where -1 is less than equal to z is less than equal to 1 which satisfies as given in (5) and (6).

$$\int_{-1}^1 R_k(z) R_l(z) dz = (2(2k+1)^{-1}) \delta_{kl} \quad (5)$$

$$\delta_{kl} = \{1, k \text{ is equal to } l, 0, k \text{ is not equal to } l\} \quad (6)$$

The importance of the adaptive frequency decomposition model is shown by its ability to decompose signals at various scales as well as choosing various scales based on various targets. The multiple wavelet transform combines the benefits of both the adaptive frequency decomposition model as well as the orthogonal polynomials. Consider a function $h(z)$, the coefficients used for multiple wavelets for scale p for measure ρ are given as (7).

$$\begin{aligned} u_n^p &= [\langle h, \sigma_{kn}^p \rangle_{\vartheta}]_{k=0}^{m_f-1} \\ f_n^p &= [\langle h, \omega_{kn}^p \rangle_{\vartheta}]_{k=0}^{m_f-1} \end{aligned} \quad (7)$$

For the (7), u_n^p, f_n^p belongs to $T^{m \times 2^p}$, the wavelet orthogonal basis of polynomials is represented as σ_{kn}^p and ω_{kn}^p . The count of dimensions that are transformed or the count of orthogonal basis is represented as m_f . The scales for reconstruction or decomposition are given as (8).

$$\begin{aligned} u_n^p &= J^{(0)} u_{2n}^{p+1} + J^{(1)} u_{2n+1}^{p+1} \\ u_{2n}^{p+1} &= \Sigma^{(0)} (J^{(0)V} u_n^p + I^{(0)V} f_n^p) \\ u_{2n+1}^{p+1} &= \Sigma^{(0)} (J^{(0)V} u_n^p + I^{(0)V} f_n^p) \\ f_n^p &= I^{(0)} u_{2n}^{p+1} + J^{(1)} u_{2n+1}^{p+1} \\ u_{2n+1}^{p+1} &= \Sigma^{(1)} (J^{(0)V} u_n^p + I^{(1)V} f_n^p) \end{aligned} \quad (8)$$

$$(J^{(0)} J^{(1)} I^{(0)} I^{(1)}) (J^{(0)} J^{(1)} I^{(0)} I^{(1)})^V = (\Sigma^{(0)-1} \ 0 \ 0 \ \Sigma^{(1)-1}) \quad (9)$$

Here, the linear coefficients are expressed as $(J^{(0)}, J^{(1)}, I^{(0)}, I^{(1)})$ that are constant matrices for the decomposition of wavelets and have to satisfy the (9). The inverse matrices of $\Sigma^{(0)}$ and $\Sigma^{(1)}$ are given as $\Sigma^{(0)-1}$ and $\Sigma^{(1)-1}$ respectively. On computation of the multiple wavelet tensor product basis as well as the multiple scale, we obtain the multiple wavelet expression. The multiple wavelet application possesses the benefits of both the orthogonal polynomials as well as the wavelets. The multiple wavelets portray the function on a subspace of polynomials by utilizing the Legendre polynomials given in (5).

To introduce the attention model based on the multiple wavelets, the framework of the multiple wavelet system is applied. The proposed system utilizes attention learning elements as units A, B, and C rather than using neural networks in general. The system includes reconstruction as well as decomposition. The decomposition behaves as a RNN. Where, for every iteration, u^{p+1} denotes the input including l, s , and x as sub-scores. Every sub-score is computed separately about the mechanism specified further. On using (8), the multi-level and multiple wavelet coefficients at a superficial stage are denoted as u_p and f_p is computed respectively. Furthermore, after storing the values in the four attention networks, namely, C, D, E , and V the resulting W 's multi-level and multiple wavelet coefficients are obtained. An operator $Vc = w$ has to be mapped, the mapping based on the multiple wavelet form is given as (10). From the (10), one separate layer of perceptrons is denoted as \underline{V} that is used to process the signal remaining post-decomposition stages denoted as N .

$$\begin{aligned} W_{f_n}^p &= C_p f_n^p + D_p u_n^p \\ W_{u_n}^p &= E_p f_n^p \\ W_{u_n}^N &= \underline{V}_n^N \end{aligned} \quad (10)$$

This mechanism involves top-down as well as bottom-up methods. For the top-down approach, for every sequential input having dimension O , the signal is retrieved by the decomposition that has a fixed sequence by computing at least N iterations. In the next phase, the reconstruction model utilizes a bottom-up operation for the output that results in the decomposition step for the computation of coefficients at multi-level at a more refined stage. Finally, the constituent parameters are gathered at the reconstruction stage as given in (10). The process of training is executed until the required output is attained. The means of this implementation is the transformation of the output to a hierarchical wavelet space for every iteration. During execution, the networks of attention learning that are C, D , and E possess the same attributes in architecture. The parameters that are used in the reconstruction model as well as the decomposition model are constant matrices J and I , which implies that the models do not require prior training.

The attention phase is the same as the enhanced encoder decoder module. The input of the attention phase includes keys denoted as m , queries that are given as s , and values expressed as w . The queries are taken from the decoder and are attained as $s = z_{encoder} \cdot y_s$. The encoder gives the values as well as the keys and it can be received as $x = z_{decoder} \cdot y_x$ and $z_{decoder} \cdot y_m$. The expression of attention can be formulated as given (11).

$$Attention(s, m, x) = softmax \left((sm^V) (f_s)^{-\frac{1}{2}} \right) x \quad (11)$$

The last stage of the proposed study for intrusion detection includes a classifier that aids in decision-making for general information as well as intrusive information. The model proposed includes completely linked layers along with a SoftMax classifier that is implemented. Normally, the samples are discriminated for information classification using Euclidean distance. However, there are constraints while considering certain applications that could include environments having high-dimension. We look towards modifying the performance of the model by using a classifier based on cosine. The loss function of this classifier is as in (12).

$$N = -\log \left(\left(e^{u \cdot \cos(\phi_{k,l+1\{l=a_k\},o})} \right) \left(\sum_{l=1}^p e^{u \cdot \cos(\phi_{k,l+1\{l=a_k\},o})} \right)^{-1} \right) \quad (12)$$

For the (12), the angle between the l – th weights is denoted as $\phi_{k,l}$ for the final completely linked layer Y_l and the output is expressed as h_k . The hyperparameter scale is given as u and the angular marginal penalty is expressed as o between z_k and Y_l that is utilized simultaneously for improvement of the discrepancy as well as the compactness inside the classes. For the proposed study, the values of u is fixed to 32 and o is set to 0.5. Considering real-time scenarios, the count of normal data is generally more in comparison to that of data having fault or intrusion. This could result in an imbalance in classification that leads to difficulties while considering intrusion detection. To resolve this issue, we include a higher number of intrusive samples compared to normal samples while considering the training process. We use inverse proportion for particular classes concerning the training data as the weight. The lesser the count of data samples in the training dataset for a particular class, the higher the weight for that class while training, and the higher the count of data samples in the training dataset for a particular class, the lower the weight for that class while training. Simultaneously, the normal values of the weights are between 0 and 1.

4. PERFORMANCE EVALUATION

This section evaluated the HED-EDFD model considering the different datasets, further model is compared with the existing model based on meta-learning [26], this employs a lightweight meta-learning ensemble approach, combining weak learners like RF and MLP to enhance detection accuracy. It is optimized to run efficiently on resource-limited devices, achieving high accuracy and low false positive rates. The model uses a stacked ensemble with a meta-estimator for robust performance across various datasets.

4.1. Dataset details

The IoT-23 [27] dataset, sponsored by Avast and captured by Czech Technical University's Stratosphere Lab, spans from 2018 to 2019. This includes 20 IoT malware traffic scenarios with 5931K flows, alongside three benign traffic scenarios with 2645K flows, featuring real and labeled IoT malware infections and benign traffic. The other two datasets are KDD99 [28] and TON [29].

4.2. Results

In Table 1, the HED-EDFD model achieved an accuracy of 0.998 and a false positive rate (FPR) of 0.030, significantly outperforming traditional models such as super-learner, subsemble-learner, and sequential-learner, which had accuracies of 0.994, 0.993, and 0.993, respectively. The HED-EDFD model also maintained high recall and precision rates at 0.995 each, resulting in a robust F1-score of 0.995. Furthermore, the HED-EDFD model demonstrated efficient processing with a time of 2.5 seconds, making it highly suitable for real-time applications. These results underscore the HED-EDFD model's exceptional capability in providing accurate and efficient intrusion detection compared to conventional ensemble methods.

Table 1. Binary classification on IoT-23 dataset

Model type	Acc	FPR	Recall	Pre	F1-s	P-v	Time(s)
Super-learner	0.994	0.038	0.994	0.994	0.994	0.0099	3.95
Subsemble-learner	0.993	0.045	0.993	0.993	0.993	0.0099	7.62
Sequential-learner	0.993	0.04	0.993	0.993	0.993	0.0099	19.9
Bagging	0.993	0.038	0.993	0.993	0.993	0.0099	264
Boosting	0.988	0.099	0.989	0.989	0.988	0.0099	0.15
Stacking	0.892	0.923	0.892	0.877	0.851	0.0099	0.38
HED-EDFD (proposed model)	0.998	0.030	0.995	0.995	0.995	0.0098	2.5

In Table 2, the HED-EDFD model achieved perfect scores in accuracy (1.000), FPR (0.001), recall (1.000), precision (1.000), and F1-score (1.000), outperforming other models like super-learner, subsemble-

learner, sequential-learner, and stacking. Notably, while the bagging model also achieved perfect scores, its processing time was significantly longer (999.0 seconds) compared to the HED-EDFD model's efficient 1.5 seconds. These results highlight the HED-EDFD model's exceptional accuracy and efficiency, making it an optimal solution for real-time intrusion detection applications.

Table 2. Binary classification comparison on KDD99

Model type	Acc	FPR	Recall	Pre	F1-s	P-v	Time(s)
Super-learner	0.998	0.002	0.998	0.998	0.998	0.0099	2.69
Subsemble-learner	0.998	0.002	0.998	0.998	0.998	0.0099	7.29
Sequential-learner	0.998	0.001	0.998	0.998	0.998	0.0099	17.3
Bagging	1	0.001	1	1	1	0.0099	999.0*
Boosting	0.973	0.119	0.997	0.973	0.972	0.0099	0.13
Stacking	0.999	0.001	0.999	0.999	0.999	0.0099	0.23
HED-EDFD (proposed model)	1	0.001	1	1	1	0.0098	1.5

In Table 3 the HED-EDFD model achieved the highest accuracy at 0.990 and the lowest FPR at 0.002. Additionally, the HED-EDFD model maintained high recall and precision rates at 0.987 each, resulting in a robust F1-score of 0.987. The processing time of the HED-EDFD model was also efficient at 0.5 seconds, making it suitable for real-time applications. In comparison, other models like super-learner, subsemble-learner, and sequential-learner exhibited lower accuracy and higher FPR, while the Bagging model, despite its high accuracy, had a significantly longer processing time.

Table 3. Binary classification comparison on TON dataset

Model type	Acc	FPR	Recall	Pre	F1-s	P-v	Time(s)
Super-learner	0.979	0.019	0.979	0.994	0.994	0.0099	0.57
Subsemble-learner	0.981	0.016	0.981	0.993	0.993	0.0099	0.89
Sequential-learner	0.981	0.011	0.983	0.993	0.993	0.0099	3.34
Bagging	0.986	0.008	0.986	0.986	0.986	0.0099	80
Boosting	0.963	0.004	0.963	0.965	0.963	0.0099	0.19
Stacking	0.97	0.005	0.97	0.971	0.97	0.0099	0.4
HED-EDFD (proposed model)	0.99	0.002	0.987	0.987	0.987	0.0098	0.5

The evaluation of models on the IoT-23 dataset shows that the HED-EDFD (proposed model) outperforms others with an accuracy of 99.80% and an FPR of 3.00%. Super-learner, subsemble learner, sequential learner, and bagging ensemble also demonstrated high accuracies (99.30%-99.40%) but varied in FPRs, with bagging ensemble achieving the lowest at 3.84%. In contrast, models like stacking ensemble and NB had high FPRs (92.34% and 92.30%), despite moderate accuracies, making them less reliable. The MLP model achieved perfect accuracy (100%) but an impractically high FPR (100%). Overall, the HED-EDFD model's superior accuracy and low FPR make it the most effective choice for real-time intrusion detection in IoT environments as shown in Figure 4.

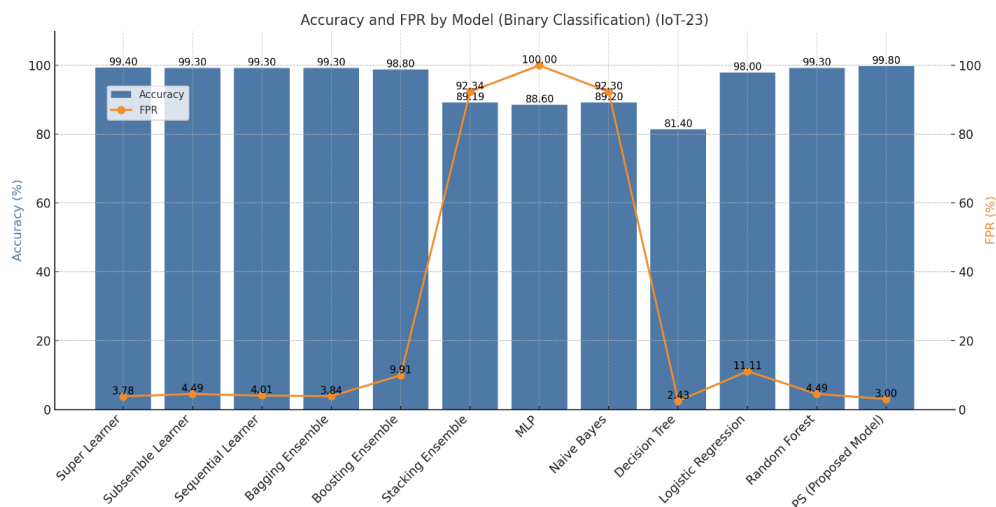


Figure 4. Accuracy and FPR comparison on IoT-23 dataset

The evaluation of models on the KDD99 dataset highlights the HED-EDFD (proposed model) as the top performer, with perfect accuracy of 100% and the lowest FPR of 0.001%. While the bagging ensemble also achieved 100% accuracy, its FPR was higher at 0.030%. Super-learner, subensemble learner, sequential learner, and stacking ensemble all maintained high accuracies around 99.80%-99.90% but had slightly higher FPRs ranging from 0.001% to 0.160%. Boosting ensemble, despite achieving a lower accuracy of 97.26%, showed an FPR of 11.940%, indicating a higher rate of false positives. Notably, NB had a poor performance with an FPR of 37.700% despite an accuracy of 91.70%. MLP, DT, logistic regression, and RF also performed well in terms of accuracy (99.70%-100%) and had low FPRs (0.020%-0.040%). Overall, the HED-EDFD model's perfect accuracy and minimal FPR make it the most effective model for intrusion detection in the KDD99 dataset. Figure 5 shows the accuracy and FPR comparison on various models.

The evaluation of models on the TON dataset demonstrates the HED-EDFD (proposed model) as the superior performer with an accuracy of 99.00% and the lowest FPR of 0.002%. Among other models, subensemble learner, sequential learner, and bagging ensemble also showed high accuracies ranging from 98.10% to 98.60% but had higher FPRs between 0.780% and 1.590%. The boosting ensemble model, despite its lower accuracy of 96.30%, managed an FPR of 0.370%. Notably, NB performed poorly with an FPR of 49.840% and an accuracy of 65.10%. MLP, DT, logistic regression, and RF displayed high accuracies from 90.90% to 97.90%, but varied significantly in their FPRs, with DT achieving the highest FPR of 8.570%. Overall, the HED-EDFD model's combination of high accuracy and exceptionally low FPR positions it as the most effective solution for real-time intrusion detection on the TON dataset. Figure 6 shows the accuracy and FPR comparison with various models.

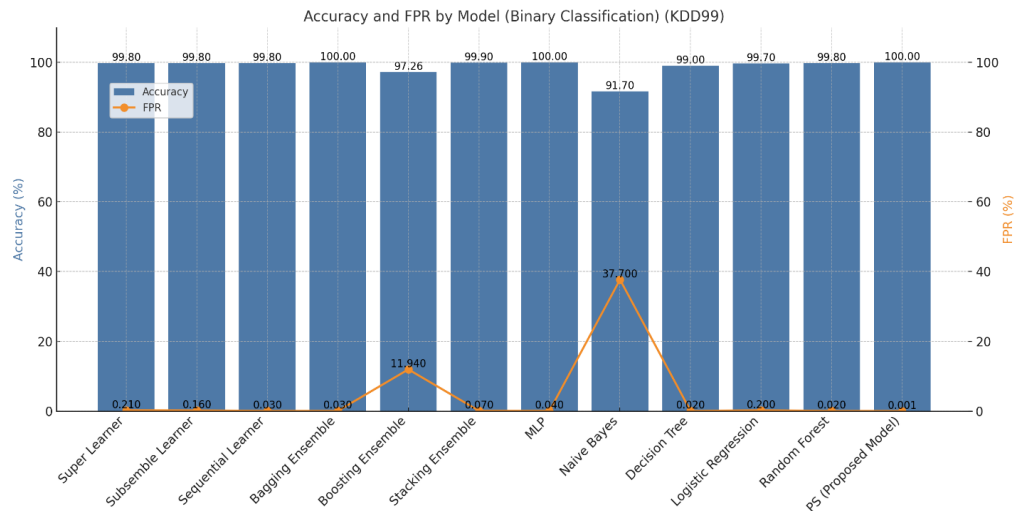


Figure 5. Accuracy and FPR comparison on various model

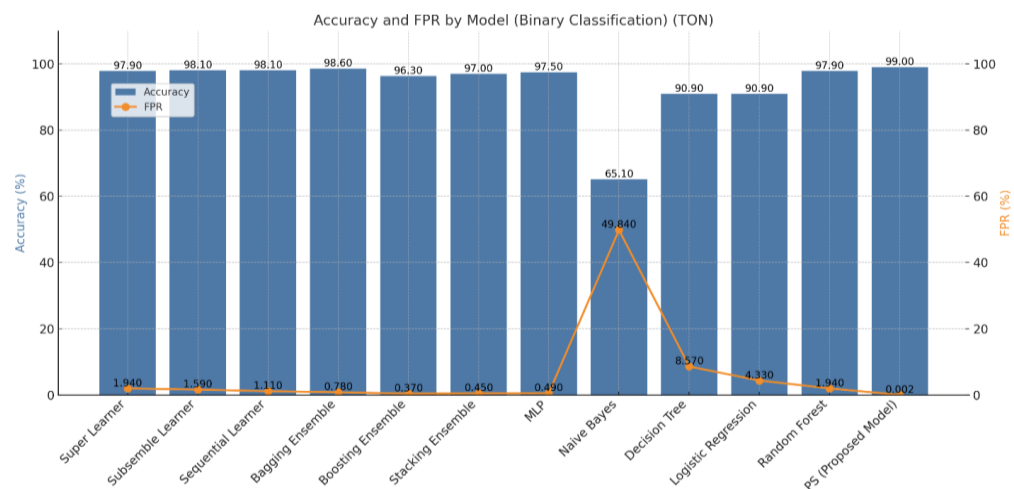


Figure 6. Accuracy and FPR comparison with various model

4.3. Comparative analysis

The comparative analysis of ensemble and DL models across the IoT-23, KDD99, and TON datasets reveals the HED-EDFD (proposed model), a DL-based model, as the most effective for binary classification in IDS. On the IoT-23 dataset, the HED-EDFD model achieved an accuracy of 99.80% with an FPR of 3.00%, outperforming ensemble models like super-learner and bagging ensemble, which had higher FPRs despite high accuracy. For the KDD99 dataset, the HED-EDFD model attained perfect accuracy (100%) and the lowest FPR (0.001%), significantly surpassing ensemble models such as super-learner and sequential-learner, and even matching bagging ensemble’s accuracy but much lower processing time. On the TON dataset, the HED-EDFD model again demonstrated superior performance with an accuracy of 99.00% and an exceptionally low FPR of 0.002%, outperforming ensemble models like subensemble learner, bagging ensemble, and boosting ensemble. This consistent high performance across all datasets underscores the HED-EDFD model’s robustness, accuracy, and efficiency, making it the optimal choice for real-time intrusion detection in diverse environments.

5. CONCLUSION

The HED-EDFD model introduced in this study offers a robust and innovative solution for intrusion detection in cloud-based IoT environments by integrating adaptive frequency decomposition with an enhanced deep encoder-decoder architecture and deep contextual attention mechanisms, the HED-EDFD model effectively captures both high-frequency transient changes and low-frequency trends in IoT data. This dual-domain feature extraction capability is crucial for identifying complex patterns indicative of malicious activities, which traditional IDS often fail to recognize. Performance evaluations across IoT-23, KDD99, and TON datasets highlight the model’s superiority. The HED-EDFD achieved 99.80% accuracy and a 3.00% false positive rate on the IoT-23 dataset, and perfect scores on the KDD99 dataset (100% accuracy, 0.001% FPR). Its efficient processing time further supports real-time application suitability.

ACKNOWLEDGEMENT

We would like to express our sincere gratitude to all those who have supported and contributed to this research project. Primarily, we extend our heartfelt thanks to our guide for his unwavering guidance, invaluable insights, and encouragement throughout the research process.

FUNDING INFORMATION

No funding is raised for this research.

AUTHOR CONTRIBUTION

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Ramya K. M.	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓			✓	
Rajashekhar C. Biradar								✓	✓	✓	✓	✓		

C : Conceptualization	I : Investigation	Vi : Visualization
M : Methodology	R : Resources	Su : Supervision
So : Software	D : Data Curation	P : Project administration
Va : Validation	O : Writing - Original Draft	Fu : Funding acquisition
Fo : Formal analysis	E : Writing - Review & Editing	

CONFLICT OF INTEREST

Author declares no conflict of interest.




DATA AVAILABILITY

Dataset is utilized in this research mentioned in reference [27]-[29].




REFERENCES

- [1] M. Adam, M. Hammoudeh, R. Alrawashdeh, and B. Alsulaimy, "A survey on security, privacy, trust, and architectural challenges in IoT systems," *IEEE Access*, vol. 12, pp. 57128–57149, 2024, doi: 10.1109/ACCESS.2024.3382709.
- [2] A. Dunmore, J. Jang-Jaccard, F. Sabrina, and J. Kwak, "A comprehensive survey of generative adversarial networks (GANs) in Cybersecurity Intrusion Detection," *IEEE Access*, vol. 11, pp. 76071–76094, 2023, doi: 10.1109/ACCESS.2023.3296707.
- [3] H. Jmila, G. Blanc, M. R. Shahid, and M. Lazrag, "A survey of smart home IoT device classification using machine learning-based network traffic analysis," *IEEE Access*, vol. 10, pp. 97117–97141, 2022, doi: 10.1109/ACCESS.2022.3205023.
- [4] H. Kim *et al.*, "Panop: mimicry-resistant ANN-based distributed NIDS for IoT networks," *IEEE Access*, vol. 9, pp. 111853–111864, 2021, doi: 10.1109/ACCESS.2021.3103015.
- [5] X. Zhou, W. Liang, W. Li, K. Yan, S. Shimizu, and K. I. K. Wang, "Hierarchical adversarial attacks against graph-neural-network-based IoT network intrusion detection system," *IEEE Internet of Things Journal*, vol. 9, no. 12, pp. 9310–9319, Jun. 2022, doi: 10.1109/JIOT.2021.3130434.
- [6] A. A. M. Teodoro *et al.*, "An analysis of image features extracted by CNNs to design classification models for COVID-19 and Non-COVID-19," *Journal of Signal Processing Systems*, vol. 95, no. 2–3, pp. 101–113, Mar. 2023, doi: 10.1007/s11265-021-01714-7.
- [7] P. Lin, K. Ye, and C. Z. Xu, "Dynamic network anomaly detection system by using deep learning techniques," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 11513 LNCS, 2019, pp. 161–176.
- [8] O. D. Okey *et al.*, "BoostedEnML: efficient technique for detecting cyberattacks in IoT systems using boosted ensemble machine learning," *Sensors*, vol. 22, no. 19, p. 7409, Sep. 2022, doi: 10.3390/s22197409.
- [9] Y. Fu, F. Lou, F. Meng, Z. Tian, H. Zhang, and F. Jiang, "An intelligent network attack detection method based on RNN," in *Proceedings - 2018 IEEE 3rd International Conference on Data Science in Cyberspace, DSC 2018*, Jun. 2018, pp. 483–489, doi: 10.1109/DSC.2018.00078.
- [10] M. Masum and H. Shahriar, "TL-NID: deep neural network with transfer learning for network intrusion detection," in *2020 15th International Conference for Internet Technology and Secured Transactions, ICITST 2020*, Dec. 2020, pp. 1–7, doi: 10.23919/ICITST51030.2020.9351317.
- [11] T. A. Tuan, H. V. Long, L. H. Son, R. Kumar, I. Priyadarshini, and N. T. K. Son, "Performance evaluation of Botnet DDoS attack detection using machine learning," *Evolutionary Intelligence*, vol. 13, no. 2, pp. 283–294, Jun. 2020, doi: 10.1007/s12065-019-00310-w.
- [12] Z. Shao, S. Yuan, and Y. Wang, "Adaptive online learning for IoT botnet detection," *Information Sciences*, vol. 574, pp. 84–95, Oct. 2021, doi: 10.1016/j.ins.2021.05.076.
- [13] M. Shafiq, Z. Tian, A. K. Bashir, X. Du, and M. Guizani, "CorraUC: a malicious Bot-IoT traffic detection method in IoT network using machine-learning techniques," *IEEE Internet of Things Journal*, vol. 8, no. 5, pp. 3242–3254, Mar. 2021, doi: 10.1109/JIOT.2020.3002255.
- [14] P. Pujar, A. Kumar, and V. Kumar, "Plant leaf detection through machine learning based image classification approach," *IAES International Journal of Artificial Intelligence*, vol. 13, no. 1, pp. 1139–1148, Mar. 2024, doi: 10.11591/ijai.v13.i1.pp1139-1148.
- [15] S. H. Sreedhara, V. Kumar, and S. Salma, "Efficient big data clustering using adhoc fuzzy c means and auto-encoder CNN," in *Lecture Notes in Networks and Systems*, vol. 563, 2023, pp. 353–368.
- [16] C. Kavitha, M. Saravanan, T. R. Gadekallu, K. Nimala, B. P. Kavin, and W. C. Lai, "Filter-based ensemble feature selection and deep learning model for intrusion detection in cloud computing," *Electronics (Switzerland)*, vol. 12, no. 3, p. 556, Jan. 2023, doi: 10.3390/electronics12030556.
- [17] K. G. Maheswari, C. Siva, and G. Nalinipriya, "Optimal cluster based feature selection for intrusion detection system in web and cloud computing environment using hybrid teacher learning optimization enables deep recurrent neural network," *Computer Communications*, vol. 202, pp. 145–153, Mar. 2023, doi: 10.1016/j.comcom.2023.02.003.
- [18] M. A. Elaziz, M. A. A. Al-qaness, A. Dahou, R. A. Ibrahim, and A. A. A. El-Latif, "Intrusion detection approach for cloud and IoT environments using deep learning and capuchin search algorithm," *Advances in Engineering Software*, vol. 176, p. 103402, Feb. 2023, doi: 10.1016/j.advengsoft.2022.103402.
- [19] D. B. Salvakkam, V. Saravanan, P. K. Jain, and R. Pamula, "Enhanced quantum-secure ensemble intrusion detection techniques for cloud based on deep learning," *Cognitive Computation*, vol. 15, no. 5, pp. 1593–1612, Sep. 2023, doi: 10.1007/s12559-023-10139-2.
- [20] J. Zhang, J. D. Peter, A. Shankar, and W. Viriyasitavat, "Public cloud networks oriented deep neural networks for effective intrusion detection in online music education," *Computers and Electrical Engineering*, vol. 115, p. 109095, Apr. 2024, doi: 10.1016/j.compeleceng.2024.109095.
- [21] A. Parameswari, R. Ganeshan, V. Ragavi, and M. Shereesha, "Hybrid rat swarm hunter prey optimization trained deep learning for network intrusion detection using CNN features," *Computers and Security*, vol. 139, p. 103656, Apr. 2024, doi: 10.1016/j.cose.2023.103656.
- [22] N. Joraviya, B. N. Gohil, and U. P. Rao, "DL-HIDS: deep learning-based host intrusion detection system using system calls-to-image for containerized cloud environment," *Journal of Supercomputing*, vol. 80, no. 9, pp. 12218–12246, Jun. 2024, doi: 10.1007/s11227-024-05895-3.
- [23] M. Nakip and E. Gelenbe, "Online self-supervised deep learning for intrusion detection systems," *IEEE Transactions on Information Forensics and Security*, vol. 19, pp. 5668–5683, 2024, doi: 10.1109/TIFS.2024.3402148.
- [24] M. H. Bhavsar, Y. B. Bekele, K. Roy, J. C. Kelly, and D. Limbrick, "FL-IDS: federated learning-based intrusion detection system using edge devices for transportation IoT," *IEEE Access*, vol. 12, pp. 52215–52226, 2024, doi: 10.1109/ACCESS.2024.3386631.
- [25] G. Singh, K. Sood, P. Rajalakshmi, D. D. N. Nguyen, and Y. Xiang, "Evaluating federated learning-based intrusion detection scheme for next generation networks," *IEEE Transactions on Network and Service Management*, vol. 21, no. 4, pp. 4816–4829, Aug. 2024, doi: 10.1109/TNSM.2024.3385385.
- [26] C. A. Fadhilla, M. D. Alfikri, and R. Kaliski, "Lightweight meta-learning BotNet attack detection," *IEEE Internet of Things Journal*, vol. 10, no. 10, pp. 8455–8466, May 2023, doi: 10.1109/JIOT.2022.3229463.
- [27] S. Garcia *et al.*, "A labeled dataset with malicious and benign IoT network traffic (version 1.0.0)." 2020. [Online]. Available: <https://www.stratosphereips.org/>
- [28] D. Dua and C. Graff, "UCI machine learning repository." 2017. [Online]. Available: <http://archive.ics.uci.edu/ml>
- [29] T. M. Booi, I. Chiscop, E. Meeuwissen, N. Moustafa, and F. T. H. D. Hartog, "ToN_IoT: the role of heterogeneity and the need for standardization of features and attack types in IoT network intrusion data sets," *IEEE Internet of Things Journal*, vol. 9, no. 1, pp. 485–496, Jan. 2022, doi: 10.1109/JIOT.2021.3085194.

BIOGRAPHIES OF AUTHOR

Ramya K M    earned her Bachelors of Engineering B.E. degree in CSE from VTU, Belagavi in 2013. She has obtained her master's degree in M.Tech. (CSE) from Dayananda Sagar College of Engineering, Bangalore in 2015. And currently she is a research scholar at REVA University doing her Ph.D. in computer science and engineering and also working as an Assistant Professor in BMS College of Engineering. She has attended many workshops and FDPs conducted by industry and other academic institutions. She also carries industry experience working at Cognizant Technology Solutions as a Programmer Analyst/ Adobe Experience Manager Developer. Her areas of interest are cyber security, IoT, and cloud computing. She can be contacted at email: ramyakm_12@rediff.com.



Dr. Rajashekhar C. Biradar    is currently working as Pro Vice Chancellor at Reva University, Bangalore. He has completed his Ph.D. in 2012 from VTU, Belgaum. He has 32 years of teaching experience. 132 research publications are under his name. His areas of interest are wireless sensor network, network security and communication system. He can be contacted at this email: provc@reva.edu.in.