International Journal of Advances in Applied Sciences (IJAAS)
Vol. 14, No. 2, June 2025, pp. 373~383
ISSN: 2252-8814, DOI: 10.11591/ijaas.v14.i2.pp373-383

Contraction control factor-based gorilla troop optimizer for features in intrusion detection systems

Shalini Sharma[1], Supriya Khaitan[2], Gayatri Hegde[3], Divya Rohatgi[4], Nusrat Parveen Mohammad Rafique[4], Suhas Janardan Lawand[5]

[1]Department of Computer Science, Accenture Solutions, Noida, India
[2]Department of Engineering and Technology, Bharati Vidyapeeth (Deemed to be University), Navi Mumbai, India
[3]Department of Computer Science and Business Systems, Bharati Vidyapeeth (Deemed to be University), Pune, India
[4]Department of Computer Science and Engineering, Bharati Vidyapeeth (Deemed to be University), Pune, India
[5]Department of Computer Science and Engineering, Xavier Institute of Engineering, Mumbai, India

Article Info

Article history:

Received Oct 18, 2024
Revised Feb 5, 2025
Accepted Mar 28, 2025

Keywords:

Contraction control factor
Gorilla troop optimizer
Internet of things
Intrusion detection system
Long short-term memory
Multiple parametric exponential linear unit

ABSTRACT

Internet of things (IoT) has evolved into a large-scale network due to the increasing number of connected devices and massive amount of data they generate. IoT networks produce massive amounts of heterogeneous data from various devices, making it difficult to identify relevant features for intrusion detection. Hence, this research proposes the contraction control factor-based gorilla troop optimizer (CCF-GTO) for feature selection and multiple parametric exponential linear units based long short-term memory (MPELU-LSTM) approach for classification of intrusion detection system (IDS) in IoT. CCF-GTO. It uses adjustable parameters to prioritize relevant information while eliminating unnecessary features, making the model more efficient and resulting in better classification accuracy. The experimental results demonstrate that the MPELU-LSTM approach achieves better accuracy of 99.56% on the UNSW-NB15 dataset as compared to the earlier approaches like convolutional neural network with LSTM (CNN-LSTM) and optimized deep residual convolutional neural networks (DCRNN). These findings suggest that the MPELU-LSTM method significantly enhances the accuracy and robustness of IDS in IoT environments by addressing issues like the identification of relevant features and feature redundancy, contributing to more effective and secure systems. This research has valuable implications for enhancing the security bearing of IoT infrastructure.

Corresponding Author:

Shalini Sharma
Department of Computer Science, Accenture Solutions
DDC4A SeaView Tower Sector 135 Noida, India
Email: shalinisharma1980@gmail.com

1. INTRODUCTION

Internet of things (IoT) integrates various embedded computing devices within the physical environment, enabling seamless connectivity between devices, services, and systems. This seamless connectivity has become an integral part of daily life, supporting a broad range of applications and improving overall efficiency [1], [2]. IoT networks continuously include a greater amount of heterogeneous devices with individual communication protocols and data standards. With the fast growth of the Internet, network intrusion activities happen continuously, and conventional security approaches are typically focused on cloud services [3]. These centralized security measures are complex to meet the fast-developing requirements of network attacks and intrusions. Ensuring that security and privacy remain significant challenges for IoT devices due to their limited computational resources, making conventional security mechanisms inappropriate for these systems [4]. This makes them vulnerable to a broad range of cyberattacks such as denial of service (DoS), distributed DoS (DDoS), and spoofing [5]. To identify these malicious attacks, a network intrusion detection system (NIDS) is important to control the operations of the IoT networks. NIDS has been designed to detect the advanced networks, comprised of IoT networks. Through examining the hints of network intrusions, NIDS has detected the attacks and raised the alarms in real-time [6], [7]. Conventional NIDS approaches have difficulty in managing the complexity and diversity of the IoT network traffic, designing it complex to determine abnormal activities, particularly in IoT devices. Moreover, the conventional approaches suffer from drawbacks such as maximum false alarm rate and minimum detection rate [8]. Various researchers have cooperated on the development of intrusion detection systems (IDS), leveraging the power of artificial intelligence (AI) approaches [9], [10]. Machine learning (ML) is a kind of interdisciplinary cross-functional area that emulates human intelligence. However, the absence of redundant features hampers machine-learning approaches in effectively detecting and addressing novel attacks in the current IoT network [11]–[13].

The NIDS was developed utilizing deep learning (DL) approaches, benchmark datasets are preferred to maximize the detection of intrusions. DL-based NIDS are typically trained by utilizing the recent datasets developed for intrusion detection. The effectiveness of NIDS using DL approaches often improves as the number of features in the dataset enhances. Furthermore, DL methods minimize the size of the feature vector into an ideal number of essential features [14], [15]. IDS have been broadly utilized in various studies because of their complexity in protecting computer networks from cyber threats. This research summarizes DL approaches used in the existing works for the design of IDS. Altunay and Albayrak [16] developed the three various DL approaches convolutional neural network (CNN), long short-term memory (LSTM), and the hybrid method of CNN+LSTM for the IDS in industrial IoT networks. In the pre-processing step, the missing values were solved and the min-max normalization step was performed to enhance the classification performance. The min-max normalization technique had the benefit of accompanying all data connections efficiently. However, the lack of an effective feature selection process to identify key attack features led to poor classification performance. Hnamte *et al*. [17] introduced the two-stage DL approach through the hybrid method of LSTM and autoencoder (AE) for the detection of IDS. The data from the LSTM-AE approach had been filtered with respect to solving the over-fitting and under-fitting problems. The LSTM-AE approach effectively balanced the dimensionality reduction and the feature retention in the highly balanced datasets. Furthermore, the LSTM-AE approach is effective in identifying significant anomalies in network traffic, which can be indicative of future cyber-attacks. However, the selection of important network attack features from the raw data is important to attain better results.

Kumar *et al*. [18] presented the deep residual convolutional neural network (DCRNN) for security enhancement in IDS, which was fine-tuned through an improved gazelle optimization algorithm (IGOA). A novel binary GOA (NBGOA) was used for the feature selection procedure to remove redundant features from data utilized in the hindrance classification procedures. However, NBGOA was complex with feature retention in imbalanced datasets, negatively impacting the IDS accuracy and generalization. Wang *et al*. [19] developed the ResNet, transformer, and bidirectional LSTM (BiLSTM) approach for the IDS, which took out both spatial and temporal features of the network traffic. The spatial feature extraction approach was established through ResNet and the temporal feature extraction approach was established through BiLSTM to extract the features. Eventually, spatiotemporal features were involved to attain the attack detection and classification. However, identifying appropriate cleaning and preprocessing methods for the prevailing network traffic data is crucial for effectively training and testing the classifier on actual network traffic. Halbouni *et al*. [20] introduced the stacked CNN and LSTM approaches based on batch normalization (BN) and dropout layers for the IDS. The CNN could extract the spatial features and LSTM extracted the temporal features to design the hybrid IDS approach. The CNN and LSTM effectively solved the overfitting through the minimization of some trainable parameters and to enhance the generalization. However, the stacked CNN and LSTM approach with BN led to minimizing the interpretability in IDS applications. From this overview, various limitations have been identified: the lack of a feature selection process, challenges with feature retention in imbalanced datasets, and minimized interpretability. However, IoT networks produce massive amounts of heterogeneous data from various devices, making it difficult to identify relevant features for intrusion detection. To overcome this problem, this research proposes the contraction control factor-based gorilla troop optimizer (CCF-GTO) for feature selection and multiple parametric exponential linear unit (MPELU) based LSTM (-) approach for the classification of IDS in IoT. To overcome these problems, this research proposes the CCF-GTO for selecting relevant features, along with the MPELU-LSTM for classifying IDS. MPELU enables LSTM to adaptively learn more effectively over long sequences, which is important for detecting distinctive patterns in time-series data within IDS. The primary highlights of this research are as trails: i) A robust pre-processing step is performed by integrating the data cleaning and solving missing values. Then, the min-max normalization is performed to ensure uniform feature contribution. These techniques enhance the dataset quality, reliability, and classifier effectiveness; ii) The CCF is proposed into the GTO approach for effectively determining the appropriate features while removing redundancies. This optimization approach enhances the classification accuracy; and iii) The MPELU-LSTM approach integrates the sequence learning capabilities of LSTM with the non-linear flexibility of the MPELU activation function for the classification of network attacks. This integration enhances intrusion detection effectiveness for both short and long-term sequences, reaching superior detection rates (DR).

This research is organized as follows: section 2 outlines the proposed methodology. Section 3 presents the feature selection using CCF-GTO. Section 4 demonstrates the results and discussion. Section 5 concludes an overall research.

2. PROPOSED METHODOLOGY

This research proposes the effective CCF-GTO for selecting the relevant features and MPELU-LSTM for the classification of IDS in IoT systems. This research includes four significant phases: data collection, pre-processing, feature selection, and classification. LSTM can recognize patterns and correlations in the time-series data of network traffic or device behavior. Through learning these patterns, the LSTM has distinguished between normal and various types of attacks. Figure 1 determines the working of the proposed method.

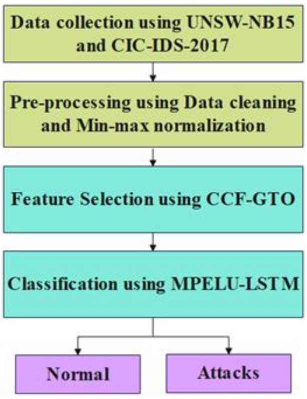
Data collection using UNSW-NB15 and CIC-IDS-2017
Pre-processing using Data cleaning and Min-max normalization
Feature Selection using CCF-GTO
Classification using MPELU-LSTM
Normal
Attacks

Figure 1. Working on the proposed method

2.1. Data collection

The primary focus of this research is data collection, which is considered to estimate the effectiveness of the proposed method. The importance of the proposed method in this research is identified using the two benchmark datasets such as UNSW-NB15 [21] and CIC-IDS-2017 [22]. The detailed description of these datasets is described as follows.

2.1.1. UNSW-NB15

This dataset involves records of benign traffic along with nine types of attacks, including analysis, DoS, backdoor, and fuzzes. It was produced through the Australian Centre for Cyber Security (ACCS) in 2015. The records are acquired from various real-world websites, involving Symantec Corporation, common vulnerabilities and exposures (CVE), and Microsoft Corporation. Table 1 represents the number of samples of the UNSW-NB15 dataset.

Table 1. Number of samples of the UNSW-NB15 dataset

Attacks	Number of samples
Benign	56,000
Generic	40,000
Worms	130
DoS	12,264
Analysis	2,000
Exploits	33,393
Backdoor	1,746
Reconnaissance	10,491
Fuzzers	18,184
Shellcode	1,133

2.1.2. CIC-IDS-2017

The CIC-IDS-2017 dataset is envisioned to solve the lack of a real-time network traffic dataset for the estimation of IDS. This dataset comprises of most present and appropriate data for testing the security systems. This dataset involves 2,830,000 samples, of which, 19.70% are attacks and the remaining 80.30% are benign. There are 14 various assault types and 1 normal class. The 84 features are recovered from a developed network traffic and the final column of the dataset involves the multiclass label. Table 2 represents the number of samples of the CIC-IDS-2017 dataset.

Table 2. Number of samples of CIC-IDS-2017

Attacks	Number of samples
Benign	2,260,360
Bot	1,943
DDoS	127,082
DoS GoldenEye	10,289
DoS Hulk	229,198
FTP-Parator	7,894
DoS slow loris	5,771
Dos Slowhttptest	5,485
Infiltration	34
Heartbleed	11
PortScan	157,703
SSH-Parator	5,861
Web Attack-XSS	648
Web Attack-Brute Force	1,497
Web Attack-SQL Injection	21

2.2. Pre-processing

The input from the UNSW-NB15 and CIC-IDS-2017 datasets is provided for the pre-processing step. Here, data cleaning and min-max normalization are performed to enhance the classification performance. The detailed information of these techniques is described in the following.

2.2.1. Data cleaning

Data cleaning is the procedure of determining or eliminating errors, irregularities, and discrepancies in data before it is utilized for modeling. It is an important step in data preprocessing, particularly for DL applications. The collected datasets contain the number of missing values in a few feature columns. In this phase, all unfilled cells in a feature column are occupied with “0”. Every categorical value is demonstrated as a particular numerical value and an alteration procedure is employed. In this phase, the data that contains missing values are removed, and then, the min-max normalization is performed.

2.2.2. Min-max normalization

The min-max normalization technique is performed to support the development of neural networks most dependably. This approach has the benefit of performing all data connections effectively. However, the feature values are provided in a range between 0 and 1 individually [23]. This approach is expressed in (1).

$$X_{new} = \frac{X-X_{min}}{X_{max}-X_{min}} \tag{1}$$

Where X_{new} demonstrates the normalized data; X illustrates the actual value of the feature; and X_{max} and $X_{min} illustrate$ the feature’s maximum and maximum values. This approach supports making sure that all features perform uniformly to the learning process of the model. Then, normalized data are provided for the further process.

3. FEATURE SELECTION USING CCF-GTO

After data pre-processing, the appropriate features are selected by utilizing the meta-heuristic optimization algorithm. This process supports a classifier to enhance classification performance. GTO algorithm is proposed for the feature selection process. The CCF-GTO method enhances the conventional GTO by integrating an adaptive mechanism that fine-tunes the balance between exploration and exploitation, resulting in improved performance in identifying optimal or near-optimal solutions. GTO is a nature-encouraged approach that pretends a social behavior of the gorillas. This algorithm is inspired by the natural intelligence of the gorillas. GTO comprises two significant phases such as exploration and exploitation.

Various operatives compete with optimization operations for the gorilla's behavior in this approach. During the exploration phase, three operations are taken out such as shifting to an unexplored position, shifting toward other gorillas, and moving to a familiar location. Following a silverback as well as striving with adult females are assumed to enhance search effectiveness [24], [25].

Communication among the silverbacks and other gorillas is a significant part of decision-making. Hence, to enhance the exploration capability of GTO, this research proposes the CCF strategy to simulate this association. To solve the local optimum problem, this research moves a single solution to a position of another optimal solution, hence that the local space is effectively explored. The CCF improves the exploration and exploitation capabilities of the GTO, allowing it to effectively select the most relevant features for IDS, reducing redundancy. Through selecting only the most informative features, CCF-GTO minimizes the dimensionality of the dataset, leading to faster training and inference processes. Furthermore, randomness is utilized to move the solution to an exploration area that is non-obtained through an approach. This strategy approach contains the capability to take away of local optimal solution, hence it obtains the most realistic and significant solution. Hence, simulates an arbitrary movement procedure of a gorilla to enhance a solution's quality. U pretends the degree of expertise controlled through the gorillas in this phase. There are binary types of exploration of unidentified areas. If $|CCF| \geq 0.5$, then the behavior of the gorillas to explore unidentified areas is simulated effectively based on their perception. In the initialization stage, the GTO approach randomly generates the population X_i from the normalized features and the position of the silverback gorilla.

Another form of exploration involves navigating unfamiliar positions based on the gorillas' conversational experiences with one another, aiming to reduce the blind nature of exploration. The parameter CCF maintains how the gorillas select among these strategies, significantly extending the gorilla's exploration of unfamiliar areas and enhancing the search space of the algorithm. The gorillas still need to connect to ensure consistent experiences and minimize the limitations of this research. The specific update is expressed in (2).

$$U = \frac{F_i - Silverback\_Score}{Mean - Silverback\ Score} \tag{2}$$

Where F_i demonstrates a fitness (accuracy) value of ith gorillas; $Silverback\_Score$ demonstrates the fitness value of silverback; and $Mean$ illustrates average fitness value of whole gorillas. When $U > 1$, the update (3) and (4). Where $rand$ demonstrates the random number among 0 and 1; $rand(1, dim)$ illustrates the random vector by the dimension ranging between 0 and 1 through unchanging distribution; dim illustrates the dimensionality issue; X_{rl} depicts the arbitrary gorilla individual; and D illustrates the arbitrary vector with the problem dimension produced in an interval $[1 - |CCF|, |CCF|]$ with the static distribution. An estimation of D is formulated in (5).

$$GX_i = \frac{[(UB - LB) \times (|CCF| - rand) \times rand(1, dim)]}{2} + LB, |CCF| \geq 0.5 \tag{3}$$

$$GX_i = (X_i - X_{rl}) \times D, |CCF| < 0.5 \tag{4}$$

$$D = unifrnd([|CCF|, |CCF|, 1, dim]) \tag{5}$$

Similarly, to enhance the performance of exploration when $U \leq 1$, the current gorilla individuals are fused with randomly selected gorilla individuals. This process increases experimental parameters while integrating the influence of the current gorilla individuals. A position update is formulated in (6).

$$GX_i = X_{r2} + (X_i - X_{r2}) \times U + [X_i / (X_i \times rand(1, dim))] \times (1 - U) \tag{6}$$

This phase significantly moved a solution of the present individual to an arbitrary individual solution. A parameter U defines a small range of movement, allowing for significant exploration of the local space between two solutions to identify the best solution. This improves the capability to explore significantly while broadly eliminating blind searches. Then, the selected features are fed into the further process. The parameters of the proposed CCF-GTO approach include a population size ranging from 50 to 100, a number of iterations between 100 and 500, a CCF between 0.1 and 1.0, and an exploration parameter that is adjusted according to the CCF value.

3.1. Classification

The selected features from the input data are provided as input to the classification process to classify the data into two categories normal and attack. The MPELU-LSTM activation function introduces multiple parameters which allows the LSTM network to capture complex non-linear relationships within the data more effectively. This capability is important for detecting intricate patterns in network traffic which represents possible intrusions. A detailed explanation of the MPELU-LSTM is provided as follows.

The primary purpose of LSTM is to solve long-term dependency problems. In the conventional LSTM approach, there are 4 layers: 2-input, 1 forgot, and 1 output gate. The input gates work combined to choose an input that is extended to their state. According to the present cell state, a forget gate identifies which past cell states should be discarded. Then, an output gate decides which data will be transmitted by these gates. A memory cell unit is designed with input, output, and forgot gates utilized to effectively evoke and forget input data. Once an input x_t is sent through a memory cell unit, data is significantly forgotten as well as stored. Once an input is expressed as $x = (x_1, x_2, \dots, x_{t-1}, x_t)$ and output is $(S_1, S_2, \dots, S_{t-1}, S_t)$, the gates are expressed in (7).

$$g(x) = MPELU(W_x + b) \tag{7}$$

The initial phase in the LSTM procedure is to send by forgotten gate f_t. Data in a memory unit of an earlier cell is identified when sent by this gate if is sent to further progression or rejected. A forgot gate f_t is formulated in (8).

$$f_t = MPELU\left(W_f.[S_{t-1}, x_t] + b_f\right) \tag{8}$$

Where W_f denotes the weight matrix of forgot gate and b_f demonstrates a bias. Another phase is to update the data by a constituent input gate i_t extended to a memory unit. In this procedure, a value to be updated is identified through a sigmoid function. Moreover, a probable regeneration vector cell state c_t is produced in $tanh$ layer. The input i_t and cell state c_t is formulated in (9) and (10).

$$i_t = MPELU(W_i.[S_{t-1}, x_t] + b_i) \tag{9}$$

$$c_t = f_t.c_{t-1} + i_t.tanh(W_c.[S_{t-1}, x_t] + b_c) \tag{10}$$

Where the value of i_t as the vector is acquired from $[0,1]$; W_i, $[S_{t-1}, x_t]$ and b_i as learned parameters are utilized in input gates; W_c demonstrates the cell state matrix weight; and b_c demonstrates bias. In this procedure, the update of the cell state c_t is restructured after determining which portions of the data are retained and which are discarded. An output gate in this procedure will identify recent update data in a cell, then it will be handled as LSTM output. An output gate o_t is estimated in the final phase of the LSTM procedure through a sigmoid function by remote weight matrix output gate represented through W_o and $[S_{t-1}, x_t]$ and b_o denoting a bias. In this procedure, an output S_t is acquired from multiplied o_t and $tanh$ output c_t is consequential with output S_t is formulated in (11) and (12).

$$o_t = MPELU(W_0.[S_{t-1}, x_t] + b_o) \tag{11}$$

$$S_t = o_t.tanh(c_t) \tag{12}$$

MPELU is an activation function that proposes to simplify and unify a rectified linear unit (ReLU) and ELU. The significant aim is to better classification effectiveness. MPELU is capable of flexibly changing among the ReLU and ELU, making α hyperparameter learnable to further enhance its realistic capability and tune the function shape. MPELU is formulated in (13).

$$MPELU(x) = \lambda \begin{cases} x\, x > 0 \\ \alpha\left(e^{\beta x} - 1\right) x \leq 0 \end{cases} \tag{13}$$

Where $\beta > 0$. MPELU allows the LSTM to adapt to various types of data, enabling it to capture complex, non-linear patterns more effectively. MPELU integrates the benefits of parametric activation functions with the advantages of exponential linear units, providing a flexible and effective activation function. This flexibility improves the LSTM capability to learn intricate patterns in the data, leading to the most efficient classification.

4. RESULTS AND DISCUSSION

The proposed method is implemented on Python 3.10.12 software tools and system specification with Windows 10 64-bit OS, Intel core i5 processor, and 16 GB RAM. The proposed IDS classification approach utilizes various performance metrics to validate the system's effectiveness. Performance metrics like accuracy, precision, recall, F1-score, and detection rate are utilized for estimating the proposed method. The mathematical expression for every metric is described as the following (14) to (17). Where TP denotes the true positive, TN signifies the true negative, FP is the false positive, FN refers to the false negative.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{14}$$

$$Recall = \frac{TP}{TP+FN} \tag{15}$$

$$Precision = \frac{TP}{TP+FP} \tag{16}$$

$$F1-score = \frac{2TP}{2TP+FP+FN} \tag{17}$$

4.1. Performance analysis

The importance of the proposed method is estimated using various performance metrics based on UNSW-NB15 and CIC-IDS-2017 datasets. Table 3 demonstrates the analysis of the feature selection approaches. The existing optimization-based feature selection approaches like Golden Jackal optimization (GJO), Honey Badger optimization (HBO), whale optimization algorithm (WOA), and GTO are compared and estimated with the CCF-GTO approach. CCF-GTO employs randomness to move solutions into unexplored regions, helping to prevent premature convergence and ensuring a more thorough exploration of the search space. This flexibility enables the CCF-GTO to adapt to various optimization problems, making it appropriate for complex and dynamic feature selection scenarios. In the UNSW-NB15 dataset, the proposed CCF-GTO approach attains a better accuracy of 99.56% and DR of 99.45%. In the CIC-IDS-2017 dataset, the proposed CCF-GTO approach reaches a better accuracy of 99.94% and DR of 99.80% respectively.

Table 3. Analysis of feature selection approaches

Dataset	Method	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)	DR (%)
UNSW-NB15	GJO	91.20	85.34	85.09	86.92	87.29
	HBO	93.12	87.11	87.62	88.29	91.29
	WOA	95.48	89.29	90.87	90.82	93.02
	GTO	97.39	91.70	93.12	92.12	95.63
	CCF-GTO	99.56	93.29	95.20	94.25	99.45
CIC-IDS-2017	GJO	92.45	87.65	86.78	89.30	91.32
	HBO	94.21	89.98	88.38	91.33	94.21
	WOA	96.29	90.12	92.09	93.42	95.64
	GTO	98.33	92.38	94.36	95.67	97.87
	CCF-GTO	99.94	99.69	99.71	99.70	99.80

Table 4 demonstrates the analysis of the different classifiers. The existing classifiers like CNN, gated recurrent unit (GRU), recurrent neural network (RNN), and LSTM are compared and estimated with the MPELU-LSTM approach. MPELU integrates the advantages of ReLU and ELU functions, providing a more flexible activation function. This flexibility enhances the LSTM's capability to model intricate dependencies, which is crucial for detecting patterns in IDS. In the UNSW-NB15 dataset, the proposed MPELU-LSTM approach attains a better accuracy of 99.56%, precision of 93.29%, recall of 95.20%, F1-score of 94.25%, and DR of 99.45%. In the CIC-IDS-2017 dataset, the proposed MPELU-LSTM approach attains a better accuracy of 99.94%, precision of 99.69%, recall of 99.71%, F1-score of 99.70%, and DR of 99.80% respectively.

Table 4. Analysis of different classifiers

Dataset	Method	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)	DR (%)
UNSW-NB15	CNN	92.44	88.01	88.23	87.45	92.34
	GRU	95.63	90.22	90.76	90.39	94.56
	RNN	96.49	91.32	92.18	92.45	96.32
	LSTM	98.45	92.23	94.22	93.56	97.54
	MPELU-LSTM	99.56	93.29	95.20	94.25	99.45
CIC-IDS-2017	CNN	93.56	91.32	91.30	91.29	94.29
	GRU	95.35	93.12	92.32	93.24	95.26
	RNN	97.34	95.38	94.20	94.32	96.22
	LSTM	98.45	97.39	95.33	96.47	98.76
	MPELU-LSTM	99.94	99.69	99.71	99.70	99.80

Table 5 demonstrates the performance analysis of LSTM with different activation functions. The LSTM is compared and estimated with the different activation functions like ELU, ReLU, Parametric ReLU (PReLU), Leaky ReLU (LReLU), and MPELU. Through adjusting these parameters, MPELU captures a broad range of non-linear relationships in the data, allowing it to model the subtle patterns that other activation functions often miss. This adaptability enables the MPELU with LSTM to attain better results in determining the underlying patterns in the data.

Table 5. Analysis of LSTM with different activation functions

Dataset	Method	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)	DR (%)
UNSW-NB15	ELU-LSTM	93.12	89.90	90.12	90.32	91.47
	ReLU-LSTM	95.43	90.32	91.78	91.82	94.23
	PReLU-LSTM	97.39	91.31	93.21	92.16	95.23
	LReLU-LSTM	96.19	92.91	94.23	93.23	97.65
	MPELU-LSTM	99.56	93.29	95.20	94.25	99.45
CIC-IDS-2017	ELU-LSTM	93.58	93.76	94.37	92.12	94.32
	ReLU-LSTM	95.29	94.21	95.29	94.22	95.43
	PReLU-LSTM	96.42	97.53	97.46	95.29	96.54
	LReLU-LSTM	97.12	98.42	98.32	97.67	98.75
	MPELU-LSTM	99.94	99.69	99.71	99.70	99.80

4.2. Comparative analysis

In this section, the effectiveness of the proposed MPELU-LSTM approach is compared with the existing methods using UNSW-NB15 and CIC-IDS2017 datasets. Table 6 demonstrates the comparative analysis of the proposed method. The existing methods such as CNN+LSTM [16], Optimized DRCNN [18], Res-TranIDS [19], and CNN-LSTM [20] are compared and estimated with the proposed MPELU-LSTM method in terms of various performance metrics. The learnable parameters of MPELU provide flexibility, improving the LSTM's capability to identify non-linear relationships in IDS data for enhanced classification accuracy.

Table 6. Comparative analysis of the proposed method (NA=not applicable)

Dataset	Method	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)	DR (%)
UNSW-NB15	CNN+LSTM [16]	NA	92.91	93.10	93.00	NA
	Optimized DRCNN [18]	99.06	NA	NA	NA	98.99
	CNN-LSTM [20]	93.78	NA	NA	NA	94.53
	Proposed MPELU-LSTM	99.56	93.29	95.20	94.25	99.45
CIC-IDS-2017	Res-TranIDS [19]	99.15	NA	NA	NA	NA
	CNN-LSTM [20]	99.64	NA	NA	NA	99.70
	Proposed MPELU-LSTM	99.94	99.69	99.71	99.70	99.80

4.3. Discussion

This section discusses the limitations of existing works and explains how the proposed MPELU-LSTM approach addresses these limitations, along with its advantages. The limitations of the existing works such as lack of performing the feature selection process, challenging with feature retention in imbalanced datasets, and minimized interpretability. Hence, this research aims to propose the CCF-GTO for selecting the informative features and MPELU-LSTM for the classification of IDS for solving the above-mentioned limitations from the literature survey. CCF-GTO maintains a better balance between exploration and exploitation. This balance is important in feature selection processes to solve local optima problems and make sure a comprehensive search for the most relevant features. The adaptive mechanism introduced by the CCF improves the convergence rate of the optimization process, avoiding local optima and achieving global optimization for feature selection. The MPELU activation function produces supplementary parameters learned during the training process, enabling more flexibility in the activation behavior. These parameters allow the activation function to adapt its shape based on data distribution, designing it well appropriate for capturing complex, non-linear relationships in IDS data. The utilization of the MPELU activation function improves the non-linear modeling capability of LSTM, enabling it to better capture patterns in IDS data. This flexibility enhances the LSTM’s capability to model complex relationships in IDS data, resulting in better classification performance. The proposed MPELU-LSTM approach attains a better DR of 99.45% and 99.80% on UNSW-NB15 and CIC-IDS-2017 datasets. However, the existing method of optimized DRCNN [18] and CNN-LSTM [20] attained the less DR of 98.99% and 94.53% in the UNSW-NB15 dataset, whereas, in the CIC-IDS-2017 dataset, CNN-LSTM [20] attained the less DR of 99.70% respectively. These results demonstrate that the proposed CCF-GTO approach attains better results as compared to the existing methods by selecting the most relevant features. The CCF-GTO approach ensures that the IDS model is not overwhelmed by irrelevant data, leading to enhanced detection rates, accuracy, and overall model performance.

5. CONCLUSION

IDS is crucial in the realm of data protection for IoT, playing a vital role in securing user data and protecting intellectual devices. Nevertheless, traditional IDS based on statistics and expert systems are complex to meet security requirements for rapid network development and the continuous growth of large data. Hence, this research proposes the CCF-GTO for selecting the informative features and MPELU-LSTM for the classification of IDS. The additional parameters of MPELU provide significant flexibility in activation functions, enabling the LSTM to adapt better to various data patterns in IDS, and leading to better classification performance. The experimental results illustrate that the proposed MPELU-LSTM method attains better accuracies of 99.56% and 99.94% on UNSW-NB15 and CIC-IDS-2017 datasets as compared to the existing methods such as optimized DRCNN and CNN-LSTM. The future work involves a hybrid DL approach to enhance the overall classification performance for IDS in IoT.

ACKNOWLEDGEMENT

The authors would like to thank our supervisor, Prof. Syed Zeeshan Hussain, for his continuous motivation and support throughout this research.

FUNDING INFORMATION

Authors state no funding involved.

AUTHOR CONTRIBUTIONS STATEMENT

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
Shalini Sharma	✓	✓	✓	✓	✓	✓		✓	✓	✓			✓	
Supriya Khaitan		✓				✓		✓	✓	✓	✓	✓		
Gayatri Hegde	✓		✓	✓			✓			✓	✓		✓	✓
Divya Rohatgi		✓				✓		✓	✓	✓	✓	✓		
Nusrat Parveen Mohammad Rafique					✓		✓			✓		✓		✓
Suhas Janardan Lawand	✓		✓	✓			✓			✓	✓		✓	✓

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis

I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting

Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

DATA AVAILABILITY

The data that support the findings of this study are openly available in [UNSW-NB15 and CIC-IDS2027] at https://research.unsw.edu.au/projects/unsw-nb15-dataset and https://www.kaggle.com/datasets/chethuhn/network-intrusion-dataset.

REFERENCES

[1] S. Roy, J. Li, B. J. Choi, and Y. Bai, "A lightweight supervised intrusion detection mechanism for IoT networks," *Future Generation Computer Systems*, vol. 127, pp. 276–285, 2022, doi: 10.1016/j.future.2021.09.027.

[2] A. Momand, S. U. Jan, and N. Ramzan, "ABCNN-IDS: attention-based convolutional neural network for intrusion detection in IoT networks," *Wireless Personal Communications*, vol. 136, no. 4, pp. 1981–2003, 2024, doi: 10.1007/s11277-024-11260-7.

[3] Y. Yang, J. Cheng, Z. Liu, H. Li, and G. Xu, "A multi-classification detection model for imbalanced data in NIDS based on reconstruction and feature matching," *Journal of Cloud Computing*, vol. 13, no. 1, 2024, doi: 10.1186/s13677-023-00584-7.

[4] Y. Kayode Saheed, A. Idris Abiodun, S. Misra, M. Kristiansen Holone, and R. Colomo-Palacios, "A machine learning-based intrusion detection for detecting internet of things network attacks," *Alexandria Engineering Journal*, vol. 61, no. 12, pp. 9395–9409, 2022, doi: 10.1016/j.aej.2022.02.063.

[5] M. A. Talukder *et al.*, "Machine learning-based network intrusion detection for big and imbalanced data using oversampling, stacking feature embedding and feature extraction," *Journal of Big Data*, vol. 11, no. 1, 2024, doi: 10.1186/s40537-024-00886-w.

[6] M. Sarhan, S. Layeghy, N. Moustafa, M. Gallagher, and M. Portmann, "Feature extraction for machine learning-based intrusion detection in IoT networks," *Digital Communications and Networks*, vol. 10, no. 1, pp. 205–216, 2024, doi: 10.1016/j.dcan.2022.08.012.

[7] S. Layeghy, M. Baktashmotlagh, and M. Portmann, "DI-NIDS: domain invariant network intrusion detection system," *Knowledge-Based Systems*, vol. 273, 2023, doi: 10.1016/j.knosys.2023.110626.

[8] N. Dat-Thinh, H. Xuan-Ninh, and L. Kim-Hung, "MidSiot: a multistage intrusion detection system for internet of things," *Wireless Communications and Mobile Computing*, vol. 2022, 2022, doi: 10.1155/2022/9173291.

[9] J. Li, H. Zhang, Y. Liu, and Z. Liu, "Semi-supervised machine learning framework for network intrusion detection," *Journal of Supercomputing*, vol. 78, no. 11, pp. 13122–13144, 2022, doi: 10.1007/s11227-022-04390-x.

[10] M. Vishwakarma and N. Kesswani, "A new two-phase intrusion detection system with naïve Bayes machine learning for data classification and elliptic envelop method for anomaly detection," *Decision Analytics Journal*, vol. 7, 2023, doi: 10.1016/j.dajour.2023.100233.

[11] M. Sajid *et al.*, "Enhancing intrusion detection: a hybrid machine and deep learning approach," *Journal of Cloud Computing*, vol. 13, no. 1, 2024, doi: 10.1186/s13677-024-00685-x.

[12] V. Gowdhaman and R. Dhanapal, "An intrusion detection system for wireless sensor networks using deep neural network," *Soft Computing*, vol. 26, no. 23, pp. 13059–13067, 2022, doi: 10.1007/s00500-021-06473-y.

[13] Y. Wang, G. Sun, X. Cao, and J. Yang, "An intrusion detection system for the internet of things based on the ensemble of unsupervised techniques," *Wireless Communications and Mobile Computing*, vol. 2022, 2022, doi: 10.1155/2022/8614903.

[14] V. Ravi, R. Chaganti, and M. Alazab, "Deep learning feature fusion approach for an intrusion detection system in SDN-based IoT networks," *IEEE Internet of Things Magazine*, vol. 5, no. 2, pp. 24–29, 2022, doi: 10.1109/IOTM.003.2200001.

[15] S. A. Bakhsh, M. A. Khan, F. Ahmed, M. S. Alshehri, H. Ali, and J. Ahmad, "Enhancing IoT network security through deep learning-powered intrusion detection system," *Internet of Things (Netherlands)*, vol. 24, 2023, doi: 10.1016/j.iot.2023.100936.

[16] H. C. Altunay and Z. Albayrak, "A hybrid CNN+LSTM based intrusion detection system for industrial IoT networks," *Engineering Science and Technology, an International Journal*, vol. 38, 2023, doi: 10.1016/j.jestch.2022.101322.

[17] V. Hnamte, H. Nhung-Nguyen, J. Hussain, and Y. Hwa-Kim, "A novel two-stage deep learning model for network intrusion detection: LSTM-AE," *IEEE Access*, vol. 11, pp. 37131–37148, 2023, doi: 10.1109/ACCESS.2023.3266979.

[18] G. S. C. Kumar, R. K. Kumar, K. P. V. Kumar, N. R. Sai, and M. Brahmaiah, "Deep residual convolutional neural network: an efficient technique for intrusion detection system," *Expert Systems with Applications*, vol. 238, 2024, doi: 10.1016/j.eswa.2023.121912.

[19] S. Wang, W. Xu, and Y. Liu, "Res-TranBiLSTM: an intelligent approach for intrusion detection in the internet of things," *Computer Networks*, vol. 235, 2023, doi: 10.1016/j.comnet.2023.109982.

[20] A. Halbouni, T. S. Gunawan, M. H. Habaebi, M. Halbouni, M. Kartiwi, and R. Ahmad, "CNN-LSTM: hybrid deep neural network for network intrusion detection system," *IEEE Access*, vol. 10, pp. 99837–99849, 2022, doi: 10.1109/ACCESS.2022.3206425.

[21] N. Moustafa, "UNSW-NB15 dataset," UNSW Sydney, 2021. (accessed Aug. 20, 2024) [Online]. Available: https://research.unsw.edu.au/projects/unsw-nb15-dataset.

[22] H. N. Chethan, "CIC-IDS2027 dataset," Kaggle, 2017. (accessed Aug. 20, 2024) [Online]. Available: https://www.kaggle.com/datasets/chethuhn/network-intrusion-dataset.

[23] S. Songma, T. Sathuphan, and T. Pamutha, "Optimizing intrusion detection systems in three phases on the CSE-CIC-IDS-2018 dataset," *Computers*, vol. 12, no. 12, 2023, doi: 10.3390/computers12120245.

[24] T. Wu *et al.*, "A modified gorilla troops optimizer for global optimization problem," *Applied Sciences*, vol. 12, no. 19, 2022, doi: 10.3390/app121910144.

[25] E. S. Ghith and F. A. A. Tolba, "Tuning PID controllers based on hybrid arithmetic optimization algorithm and artificial gorilla troop optimization for micro-robotics systems," *IEEE Access*, vol. 11, pp. 27138–27154, 2023, doi: 10.1109/ACCESS.2023.3258187.

BIOGRAPHIES OF AUTHORS

Shalini Sharma completed Ph.D. in Computer Science from Jamia Millia Islamia University, New Delhi, India. She acquired a Master of Technology in Information Technology from Guru Gobind Singh University, Delhi in 2011. She worked as Assistant Professor at Sharda University during 2011-2015. She has served IT organizations info gain India, IRIS Software and Accenture from 2015 onwards. She can be contacted at email: shalinisharma1980@gmail.com.

Dr. Supriya Khaitan Ph.D. in Computer Science and Engineering is presently working as an Assistant Professor in the Department of Engineering and Technology Navi Mumbai. She received her Master of Technology degree from Guru Gobind Singh Indraprastha University, Delhi. Dr. Supriya Khaitan has worked as an Assistant Professor and visiting faculty in the Computer Science and Engineering Departments of different Engineering Colleges for more than 20 years. She had guided many M.Tech, MCA, BCA, and B.Tech Students in their Dissertation and final Project work. Her areas of interest are artificial intelligence, machine learning, deep learning, and network security. She is associated with a few reputed SCI/Scopus journals as a reviewer. She has authored 7 book chapters; she has published many SCI/Scopus-indexed papers in high-impact peer-reviewed international journals. She can be contacted at email: Supriyakhaitan21@gmail.com.

Dr. Gayatri Hegde currently serves as the Head of the Department of Computer Science and Business Systems at Bharati Vidyapeeth Deemed University, DET. With 24 years of teaching experience, she has over 30 publications in prestigious journals and conferences. Dr. Hegde has also reviewed and chaired numerous research sessions for reputed academic forums. Her contributions to intellectual property include 5 patents and 8 copyrights. Dr. Hegde's expertise spans various domains, including blockchain technology, big data analysis, artificial intelligence and machine learning, internet of things, and cloud computing. Additionally, she is a certified Innovation Ambassador under the Institution's Innovation Council (IIC). She can be contacted at email: gayatri.hegde@bvucoep.edu.in.

Dr. Divya Rohatgi has a total of 18 years of teaching experience in various reputed Engineering Colleges of Mumbai University, and Amity University for B.Tech, M.Tech, and Ph.D programs. She is Gate Qualified, Microsoft Technology Associate (MTA) certified, and All India Topper of NPTEL certification for Software Testing by IIT Kharagpur. She has published research papers in the field of AI-ML and Software Testing in SCI/Scopus-indexed high-impact peer-reviewed international journals and conferences. She has also authored a book Cyber Law and Cyber Crimes published by Whytes and Co. New Delhi. She has organized various national and international conferences and was session chair in BDA-2021 by IIIT Allahabad and ICRACT-2022 Amity University. Apart from this she has developed MOOC and served as a reviewer in various national and international conferences and journals. She has also published patents in the field of AI-ML and completed technical consultancy projects. Her area of interest includes software engineering, AI-ML, and deep learning. She can be contacted at email: divya.rohatgi@bvucoep.edu.in.

Nusrat Parveen Mohammad Rafique is completed her Ph.D. in Computer Technology (Machine Learning). She has 24 years of teaching experience. She is good in various subjects such as machine learning, web applications, and databases. Nusrat’s research is mainly focused on medical diagnosis using machine learning. She has published 24 papers in international conferences, international journals, and national conferences and one chapter has been published in a book under Tailor & Francis (CRC-press). She won five times cash prize in Indo-Korean festive competitions for outstanding innovators. She can be contacted at email: nusrat.parveen@bvucoep.edu.in.

Suhas Janardan Lawand currently working as an Assistant Professor in the Department of Computer Science and Engineering at Xavier Institute of Engineering, Mahim Mumbai. He is pursuing his Ph. D in Information Technology from the University of Mumbai. He has 16 years of teaching experience. His research interests include supply chain management, artificial intelligence, blockchain, and cyber security. His research papers are published in various reputed journals. He can be contacted at email: suhas.l@xavier.ac.in.