

Network Traffic Attack Detection using Dimensionality Reduction based Dual Strategy-Horned Lizard Optimization algorithm in Internet of Things

Sushma K H^{1*}, Niranjan Mamadapur², Sharath S M³, Shwetha H R⁴, Mohan Naik R⁵

^{1*}Electronics and Communication Engineering, Shri Dharmasthala Manjunathaeshwara Institute Of Technology, Ujire, Dharmasthala Rd, near Siddhavana, Ujire, Karnataka 57424 sushcco47@gmail.com

²Electronics and Communication Engineering, Shri Dharmasthala Manjunathaeshwara Institute Of Technology, Ujire, Dharmasthala Rd, near Siddhavana, Ujire, Karnataka 57424 niranjan@sdmit.in

³Savlanga Rd, J N N College of Engineering, Shivamogga, Navule, Shivamogga, Karnataka 577204, sharathsm@jnnce.ac.in
ORCID: 0009-0005-2374-2845

⁴Savlanga Rd, J N N College of Engineering, Shivamogga, Navule, Shivamogga, Karnataka 577204, shwethahr@jnnce.ac.in
ORCID: <https://orcid.org/0000-0001-6140-666X>

⁵Electronics and Communication Engineering, Shri Dharmasthala Manjunathaeshwara Institute Of Technology, Ujire, Dharmasthala Rd, near Siddhavana, Ujire, Karnataka 57424, mohannaik@sdmit.in ORCID: 0000-0002-7181-703X

Abstract: The network traffic detection for Internet of Things (IoT) includes monitoring and analyzing data traffic patterns to recognize malicious activities like Denial-of-Service (DoS) which helps to ensure the security of IoT networks. However, enhancing volume of network traffic in IoT system makes it difficult to detect cyberattacks accurately because of huge number of features which most are redundant led to inefficiencies in detection performance. To overcome this issue, Dual Strategy- Horned Lizard Optimization Algorithm (DS-HLOA) is proposed for improving attack detection through applying optimization-based feature selection techniques which removes redundant features and enhance accuracy. The strengthened convergence and mutation are the di strategy used to identify the optimal solution and avoid the local optima. The Long Short-Term Memory (LSTM) performed network, detecting traffic attacks through capturing long-term dependencies in sequential data for identifying complex pattern in traffic data. The DS-HLOA achieves accuracy of 98.21%, 96.99% and 99.98 % for BOT-IoT, CICIDS 2018 and 2017 when compared to Rat Swarm Hunter Prey Optimization- Deep Maxout Network (RSHPO-DMN) and hybrid machine learning approach.

Keywords: Deep Maxout Network, Horned Lizard Optimization Algorithm, Internet of Things, Long Short-Term Memory, Network Traffic Detection.

1. Introduction

Intrusion Attack Detection goal of studies in recent years, the attack detection has useful to enhance the security in this technology. In smart networks, the any action planned to steal data or network distribution system servers is known as cyber-attack. In this modern era, the enhancement of technologies like electrical, medical, energy, cyber-attacks are developing and progressing [1]. The enhancement of computer systems and increase the amount of information sent through the internet. Cybersecurity is one of the critical features of future generation computer systems. The enhancement price of system remote availability is the necessary to generate them with a suitable level of information transmission security [2]. The Internet of Things (IoT) networks position individual security challenges because of various range of interconnected devices with different protocols and capabilities of computational [3]. The traditional Intrusion Detection System (IDS), constructed for centralized and standardized network which face limitations in adapting to the dynamic heterogeneous IoT environment. The standardization lacks in IoT measures with resource constraints, hinders the efficiency of traditional ID method [4]. The IoT devices will have contact to the most isolated information and which are predicted to mobile number devices in terms of incidence. Probabilities attack rise



as an output, enhancing the attack surface area. The IoT intrusion detection are formed to protect communications completed by IoT technologies because which need security as a critical auxiliary element [5]. The identification of device in IoT network mostly analyze information of device unseen in gathered network traffic to classify smart device network traffic collection procedure as follows like camera and energy management tools communicate and connect among the gateway to gather the traffic flows gateway. The device information is hidden in the traffic for identification of device [6]. An IDS have the ability of evaluating data packets and produce response rapidly, investigate data in different IoT network layers with various stack protocol and change to different technologies used in IoT environments [7].

The IDS are commonly classified into dual classes such as host-based and network-based. While previous monitoring organization operations for malicious sign activity, capturing on internal traffic and behavior of system, after the network monitor traffic and from entire devices on network [8]. The intrusion states to any unexpected activities, that comprise the Confidentiality, Integrity, Availability. The traffic network recognized through packet header fields is difficult for anomaly detection. To extract the relevant features from packets helped to identify abnormal behaviors indicative of unauthorized usage. The vigilant protectors, using a mix of administrative, technological control to improve confidentiality, privacy and security against unauthorized access [9]. IDS retaining anomaly detection are active in recognizing zero-day attacks, addressing through traditional signature-based methods [10]. There are different types of attacks like Distributed Denial of Service (DDos) botnet and phishing because to the enhance in volume and speed of information, the traditional IDS increased computational complexity because of high resources and minimize the capability to detect advanced attack [11]. It operates a supervised technique to detect the prevalent types of network depend on cyber-attacks. The IDS offer three purposes are identifying the types of IoT devices that are connected to the network and profiling their normal behavior [12]. In the second level the monitoring the network for malicious packets during an attack and identifying the type of attack being conducted [13]. Selecting a subset of features that contains related ones decreases false alarms of system. The Deep Learning (DL) learn towards overly smooth functions and more inclined to irrelevant features [14]. While the tree-based model learns irregular patterns and tough to non-informative features. Therefore, development of tree-based IDS models in IoT networks enhance the security and privacy of IoT networks [15]. The main contributions of the research is described in below.

- This paper introduces a novel Dual Strategy-Horned Lizard Optimization Algorithm (DS-HLOA) to enhance attack detection. Through utilizing optimization-based feature selection techniques, DS-HLOA effectively removes redundant features and enhanced accuracy in intrusion detection.
- The strengthened convergence and mutation are the DS which used to identify the optimal solution and avoid the local optimal.
- The Long Short-Term Memory (LSTM) performed in network detecting traffic attacks through capturing sequential data for identifying complex pattern in system.

This research paper is given as follows: section 2 denotes literature survey. Proposed method is discussed in section 3. An experimental results and discussion are showed in section 4. Section 5 represents conclusion.

2.Literature Review

This portion describes the literature on network traffic attack detection in IoT which denotes the advantages and limitations for each literature survey.

Aya G. Ayad *et al.* [16] developed hybrid feature selection technique that integrated with filter and wrapper method. The hybrid feature selection approaches involved four steps such as dataset acquisition, preprocessing, feature selection and classification. In preprocessing stage, the techniques like data cleaning, categorical feature encoding, the balance of the dataset, and feature normalization are utilized. The filter and wrapper approach are combined into a two level of anomaly IDS. In the primary level, the classifiers network which classified the network packets into normal or attack and in the secondary level extra classified the attack to determine the particular category. However, feature selection relies on model utilized in the wrapper phase while the selected model was not well tuned which led to overfitting.

R. Anushiya and V.S. Lavanya [17] implemented Genetic Algorithm and Faster Recurrent Convolutional Neural Network (GA-FR-CNN) help in this framework to categorize intrusions. The Fish Swarm Optimization (FSO) for intruder identification by Assimilated Artificial FSO (AAFSo) for recognizing the recommender system

characteristics. The selected characteristics were considered as input for the FR-CNN used through a GA for the model optimization. The FR-CNN controlled the feature extraction and recurrent layers for sequential pattern learning which enabled the model to detect complex intrusion patterns. Nevertheless, the dataset used for ID was not enough means, the FR-CNN model poorly on unseen attacks which minimized its generalizability.

Anitha Parabathina *et al.* [18] developed a Rat Swarm Optimization with Improved Self-Attention based Gated Recurrent Unit (RSO-ISAGRU) is proposed for IDS classification. In the preprocessing stage, the hash encoding and min-max normalization was performed and RSO was selected the features through position updating relies on chasing and attack behavior. The ISAGRU was adopt the attacks through self-attention mechanism. The GRU generate a balance among effectiveness, handling long-term dependencies, and focused on suitable data in network traffic. However, RSO depend on the primary feature selection process while the initial features contained noisy or irrelevant data which led to suboptimal intrusion detection performance.

A.Parameswari *et al.* [19] introduced Rat swarm Hunter Prey Optimization Deep Maxout Network (RSHPO-DMN) was constructed to manage different network threats. A Z-score data normalization was applied for data preprocessing through chord distance, data was transformed into usable formats. A transformed data was extracted through CNN while extracted features were converted into vector format for network ID procedure. An Deep Maxout Network (DMN) was utilized to ID in network and RSHPO was performed to enhance the rate of ID in the classifier.

Nasim Soltani *et al.* [20] implemented hybrid ML approach Using K-Nearest Neighbors (KNN) and Random Forest (RF) as supervised classifiers to improve IDS accuracy and reduce the risk of possible attacks. The backward elimination and Linear Discriminant Analysis for feature minimization and lower computational costs. A training stage, when inconsistencies ascended among decision of classifiers. However, KNN was a lazy learning approaches which led to slower prediction time particularly while handle the large-scale network traffic data.

3.Proposed Methodology

The research, DS-HLOA is proposed for IDS classification which classifies network traffic into multiple classes. At the primary stage, the three datasets are utilized for evaluating the proposed DS-HLOA. The hash encoding and min-max normalization are considered in the preprocessing stage, which change categorical features into integer format and normalize features. The HLOA is used for choosing the best features through operating their strengthened convergence and mutation strategy. At last, the LSTM is utilized for classification of network attack network detection. Figure 1 denotes block diagram of proposed methodology.

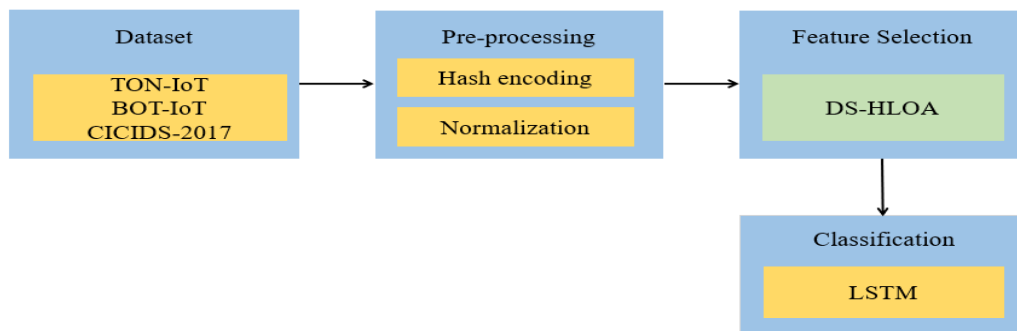


Figure1. Block Diagram of Proposed Methodology

3.1 Datasets Acquisition

The CICIDS 2018 [21] which is generated for estimation of DDOS data through New Brunswick university. The database is considered into different files depend on their dates and unbalanced files are split into balanced for large quality predictions. The database keeps a eight columns are Dst port, protocol, flow duration, label and so on.

The BOT-IoT [22] was recognized through building a realistic environment in the Cyber Range Lab at UNSW Canberra. The BOT-IoT dataset consists of 46 features which split into normal and attacks.

The CICIDS-2017 [23] contains eight distinct files from Canadian Institute of Cybersecurity, about both attack and normal traffic data. It comprises 78 features and 9 classes of BruteForceFTP, BruteForceSSH, DoS, Botnet, Distributed DoS (DDoS), normal and so on.

3.2 Preprocessing

The three datasets are preprocessed by two techniques like hash encoding and normalization, which translate the categorical features into an integer format and normalize the features. The hash encoding is utilized to convert categorical feature into an integer format which depends on hash function and values of the maps. It efficiently interprets the categorical feature to a high dimensional form and the distance between two vectors of categorical variables is sustained. After the hash encoding, the min-max normalization all features into a scale that ignores certain features have different scale, so normalized into a precise range. An min max normalization is utilized to feature value normalize scale of 0 to 1, which is showed in equation (1).

$$x'_i = \frac{x_i - x_{min}}{x_{max} - x_{min}} \quad (1)$$

Where, x'_i denotes the normalized feature, x_i is an actual feature, x_{max} and x_{min} is the maximum and minimum number of features.

3.3 Feature Selection

The preprocessed features have a noisy attribute and inappropriate features for classification of network traffic attack detection that are need to be removed the redundant features. The optimization algorithm is utilized for feature selection for chose the relevant features from the preprocessed features. Here, the Horned Lizard Optimization Algorithm performed in feature selection stage which is following the six stages which are defined in below.

3.3.1 Initialization

The HLOA starts optimization procedure through a random initialization approach to generate a candidate solution in search space problem by following equation (2).

$$X = LB + rand(UB - LB) \quad (2)$$

Here, LB and UB denotes lower and upper boundaries of variables problem and $rand$ is the uniform random vector.

3.3.2 Crypsis Behavior Strategy

A kingdom of animal grown an ability of remarkable which called as crypsis, includes camouflaging themselves through mimicking elements of their environments like texture and color. Some species level takes it a phase extra through becoming translucent and more obscuring their presence. Inspired through the fascinating phenomenon, below equation (3) operates the color of principles theory to mathematically the crypsis approach performed through HLOA.

$$\vec{X}_i(t+1) = \vec{X}_{best}(t) + \left(\partial - \frac{\partial t}{T}\right) \times \left[c_1 \left(\sin(\vec{X}_{r_1}(t)) - \cos(\vec{X}_{r_2}(t)) \right) - (-1)^\sigma c_2 \left(\cos(\vec{X}_{r_3}(t)) - \sin(\vec{X}_{r_4}(t)) \right) \right] \quad (3)$$

3.3.3 Skin darkening or lightening Strategy

The approach relies thought conservation of energy, related light energy and thermal which shown up the crucial relationship among coloration organism and state of thermal. A darker color enhances the heat absorption through capturing extra light energy and lighter color replicate more solar radiation, minimizing gaining of heat.

3.3.4 Blood Squirting Strategy

HL displays a unique adaption antipredator through pushout blood from their eyes which is a hemoptysis defense mechanism. To expressed equations of projectile motion, comprise blood trajectory into its horizontal (X-axis) and vertical (Y-axis) components.

In horizontal direction, blood shot trails a uniform linear movement, determined through equation (4).

$$\vec{v} = \vec{v}_0 + \int_0^t \vec{g} dt = \vec{v}_0 + \vec{g} t \quad (4)$$

Where, g is the earth gravity ($0.009807 \text{ km} / \text{s}^2$), v_0 is set to 1 second, t shows current iteration.

3.3.5 Move-to-escape strategy

HL existence approach involves a remarkable elusion approach and when susceptible, it quickly and unpredictably changes their direction, confusing predators which is challenge to anticipate its activities. The exploration of local component, showed through $walk\left(\frac{1}{2} - \varepsilon\right) \vec{X}_i(t)$ guides search agent's movement near its present position $\vec{X}_i(t)$. The $\vec{X}_{best}(t)$ presents a global movement aspect, prompting search agents' route towards promising regions within solution space which is shown in equation (5).

$$\vec{X}_i(t+1) = \vec{X}_{best}(t) + walk\left(\frac{1}{2} - \varepsilon\right) \vec{X}_i(t) \quad (5)$$

Where, $\vec{X}_{best}(t)$ - best search agent found.

$\vec{X}_i(t+1)$ - new position of search agent within solution space for generation $t+1$.

$walk$ - random variable produced among -1 and 1.

3.3.6 σ - Melanophore Stimulating Hormone Rate Strategy

A HL shows an adaption of thermoregulatory and adjust the skin coloration, darkening or else whitening handle its solar thermal energy absorption dynamically. The dynamic color changing procedure is mainly obsessed through temperature influence on σ - Melanophore Stimulating Hormone concentrations.

3.3.7 Improved Horned Lizard Optimization Algorithm

The difficulties contain vulnerability to receiving surrounded in suboptimal solutions and ineffective search space exploration which led ignoring possibly optimal regions. While overcome these, the proposed method integrates two key enhancements.

3.3.8 Strengthened Convergence Strategy

In HLOA, population is showed through individual optimal and because of lack prior knowledge which is difficult define if individual optimal is optimal position of global. While, local optimal position, population will coverage to local led to suboptimal solutions. To overcome these limitations, strengthened convergence strategy is proposed. This strategy permits individuals to increasingly move towards from global optimal solution through learned from fitness witness in population which is denoted in equation (6).

$$\{\vec{X}_i(t+1) = \frac{\vec{X}_{r_1}(t) + \vec{X}_{r_2}(t) + \vec{X}_{r_3}(t)}{3}, \frac{t}{T} < 0.4 \vec{X}_i(t+1) = \frac{\vec{X}_{r_1}(t) + \vec{X}_{r_2}(t)}{2} - Distance(i), \frac{t}{T} \geq 0.4 \quad (6)$$

3.3.9 Mutation Strategy

The HLOA is identified its strong abilities and easiness but which is difficult from minimized population diversity and premature convergence to local optima in advanced iterations. A proposed mutation strategy reports problem through integrating conditions from multiple individuals to produce new solutions, so conserving diversity and enabling to exploring a broader solution space. While combined with further algorithms, this technique leverages their global search strengthens mitigating optimal solution risk and improving algorithms performance in solving complex optimization problems which is described in equation (7).

$$\vec{X}_i(t+1) = \vec{X}_i(t) + \beta \times rand \times \left(\vec{X}_{r_1}(t) - \vec{X}_i(t) \right) \quad (7)$$

3.4 Classification using LSTM

The LSTM is a method of kind of Recurrent Neural Network (RNN) which is generally utilized in DL. The LSTM module contains of cell, output, input and forgot gates. A cell identifies a value completed casual time intermissions and lasting gates accomplish data flow and from the cell.

In forgot gate, a memory is managed through an individual RNN and utilizes a sigmoid activation to calculate output depend on input. Equation (8) to (10) is described the function. The output gate produces the output which is denoted in equation (11) and (12):

$$fn_j = h(w[x_j, s_{j-1}, m_{j-1}] + b_{fn}) \quad (8)$$

$$ip_j = h(w[x_j, s_{j-1}, m_{j-1}] + b_{ip}) \quad (9)$$

$$C_j = fn_j \times C_{j-1} + ip_j \times \tanh([x_j, s_{j-1}, m_{j-1}] + b_c) \quad (10)$$

$$Op_j = h(w[x_j, s_{j-1}, m_{j-1}] + b_{op}) \quad (11)$$

$$g_j = Op_j \times \tanh(C_j) \quad (12)$$

Where, h is sigmoid function, s is previous unit output and m is gate memory, w denotes weight associated with it and b is the bias.

4. Experimental Results

A proposed Di-HLOA approaches is evaluated on python 3.8 and system specifications windows10, 128-bit OS, Intel Core i7. The proposed method is estimated through various evaluation metrics such as F1-score, accuracy, recall and precision. An equation of each metric are followed in below equations (13) to (16).

$$Accurcay = \frac{TP+TN}{TP+TN+FP+FN} \quad (13)$$

$$Recall = \frac{TP}{TP+FN} \quad (14)$$

$$Precision = \frac{TP}{TP+FP} \quad (15)$$

$$F1\ score = 2 * \frac{precision*recall}{precision+recall} \quad (16)$$

Where, TP - True Positive, TN - True Negative, FP - False Positive, FN - False Negative.

4.1 Performance Analysis

The performance estimation of DS-HLOA is evaluated through three datasets and IDS malicious attack identification is challenging because of huge number of features which most are redundant led to inefficiencies in detection performance. Tables 1 denotes the proposed DS-HLOA performance for BOT-IoT, CICIDS 2018 and CICIDS 2017 attains improves outcomes in IDS.

Table1. Proposed DS-HLOA performance for BOT-IoT, CICIDS 2018 and CICIDS 2017 in feature selection

Datasets	Methods	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
BOT-IoT	POA	91.35	92.18	93.59	92.87
	FOA	93.49	93.46	95.37	94.40
	BWKA	95.17	94.27	96.72	95.47
	Proposed DS-HLOA	98.21	95.37	98.18	96.75
CICIDS 2018	POA	92.58	92.89	93.17	93.02
	FOA	93.52	93.64	96.28	94.94
	BWKA	95.27	95.89	97.65	96.76
	Proposed DS-HLOA	99.99	99.99	99.99	99.99

CICIDS2017	POA	93.17	93.54	95.17	94.34
	FOA	94.25	94.78	95.28	95.02
	BWKA	96.37	96.51	96.19	96.34
	Proposed DS-HLOA	99.98	99.91	99.91	99.91

Table 1 shows the proposed DS-HLOA performance for the BOT-IoT, CICIDS 2018 and CICIDS 2017 in feature selection process. The different feature selection techniques like Parrot Optimizer Algorithm (POA), Firefly Optimization Algorithm (FOA) and Binary Killer Whale Algorithm (BKWA). The better result of DS-HLOA attains accuracy of 98.21 and 99.98 in BOT-IoT and CICIDS 2017. Table 2 shows analysis of classification with actual features.

Table2. Performance Analysis of classification with actual features

Datasets	Methods	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
BOT-IoT	RNN	87.25	87.34	87.46	87.39
	DNN	88.46	91.28	91.28	91.28
	ANN	89.34	93.94	95.37	94.64
	Proposed DS-HLOA	92.18	95.37	98.18	96.75
CICIDS 2018	RNN	88.28	89.25	89.96	89.60
	DNN	91.58	91.48	91.27	91.37
	ANN	95.63	96.58	95.49	96.03
	Proposed DS-HLOA	99.99	99.99	99.99	99.99
CICIDS2017	RNN	89.63	89.97	91.28	90.62
	DNN	92.56	93.48	93.65	93.56
	ANN	93.17	97.21	96.17	96.68
	Proposed DS-HLOA	99.98	99.91	99.91	99.91

Table 2, depicts the analysis of classification with actual features and different various classification methods like RNN, Deep Neural Network (DNN) and Artificial Neural Network (ANN). An DS-HLOA achieves 95.37 % of precision, 99.91% of recall in BOT-IoT and CICIDS 2017 respectively. Table3 depicts the Performance Analysis of classification with optimized features.

Table3. Performance Analysis of classification with optimized features

Datasets	Methods	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
BOT-IoT	RNN	95.16	89.24	91.87	90.53
	DNN	96.37	91.38	94.56	92.94
	ANN	97.18	93.67	96.77	95.19
	Proposed DS-HLOA	98.21	95.37	98.18	96.75
CICIDS 2018	RNN	96.31	96.29	97.24	96.76
	DNN	97.85	97.23	97.91	97.56
	ANN	98.02	98.11	98.26	98.18
	Proposed DS-HLOA	99.99	99.99	99.99	99.99
CICIDS2017	RNN	97.28	96.98	97.21	97.09
	DNN	98.01	97.19	97.35	97.26
	ANN	99.62	98.03	98.56	98.29
	Proposed DS-HLOA	99.98	99.91	99.91	99.91

Table 3 describes the analysis of classification with optimized features and different classification approaches are RNN, DNN, ANN utilized. The CICIDS 2017 dataset accomplished recall of 99.91% for DS-HLOA. Table 4 denotes the K-fold validation of proposed DS-HLOA.

Table 4. K fold validation of proposed DS-HLOA

Datasets	k-fold	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
BOT-IoT	k=3	92.35	89.12	91.38	90.23
	k=5	98.21	95.37	98.18	96.75
	k=7	95.61	90.25	94.19	92.17
	k=9	96.78	92.78	95.67	94.20
CICIDS 2018	k=3	92.58	93.28	96.12	94.67
	k=5	99.99	99.99	99.99	99.99
	k=7	94.29	94.25	97.58	95.88
	k=9	95.44	96.37	98.02	97.18
CICIDS2017	k=3	95.18	97.19	99.32	98.24
	k=5	99.98	99.91	99.91	99.91
	k=7	96.36	98.36	98.47	98.41
	k=9	97.89	99.17	97.67	98.41

4.2 Comparative Analysis

The proposed DS-HLOA model is evaluated with existing method like AAFSO with GA-FR-CNN [17], RSO-ISAGRU [18], RSHPO-DMN [19] and Hybrid Machine Learning Approach [20]. Table 5 represents the comparative analysis of BOT-IoT, CICIDS 2018 and CICIDS2017 datasets with existing approaches.

Table5. Comparative Analysis

Dataset	Methods	Accuracy (%)	Precision (%)	Recall (%)	F1-Score
BOT-IoT	AAFSO with GA-FR-CNN [17]	93.77	86.60	95.87	93.38
	Proposed DS-HLOA	98.21	95.37	98.18	95.27
	RSHPO-DMN [19]	86.77	93.90	95.79	
	Proposed DS-HLOA	98.21	95.37	98.18	95.27
CICIDS 2018	RSHPO-DMN [19]	83.57	89.87	92.57	91.20
	Proposed DS-HLOA	96.99	95.99	97.99	
CICIDS2017	RSO-ISAGRU [18]	99.74	99.71	99.71	99.71
	Hybrid Machine Learning Approach [20]	99.96	99.86	99.85	99.85
	Proposed DS-HLOA	99.98	99.91	99.91	99.96

5.Conclusion

In this research, the proposed DS-HLOA utilized for classifying IDS and removes redundant features from the dataset. The DS-HLOA selects the best relevant features through strengthened convergence and mitigation strategy to avoid local optimal position. The LSTM performed a detecting traffic attacks through capturing long-term dependencies in sequential data for identifying complex in the features in the network traffic. If attack traffic is depending on majority class, lesser attention is provided, and if it is the minority class, more attention is provided. The dataset is preprocessed through hash encoding and min-max normalization which change the categorical feature into an integer format and normalizes features. A strengthened convergence and mitigation strategy allow to adoption of attack patterns and improves classification accuracy. A proposed DS-HLOA achieves accuracy of 98.21%, 96.99%

and 99.98 % for BOT-IoT, CICIDS 2018 and 2017 when compared to RSHPO-DMN and hybrid machine learning approach. In future, improved or hybrid optimization algorithm can be employed for IDS to enhance classification performance.

References

1. Maazalahi, M. and Hosseini, S., 2025. Machine learning and metaheuristic optimization algorithms for feature selection and botnet attack detection. *Knowledge and Information Systems*, pp.1-49.
2. Olszewski, D., Iwanowski, M. and Graniszewski, W., 2024. Dimensionality reduction for detection of anomalies in the IoT traffic data. *Future Generation Computer Systems*, 151, pp.137-151.
3. Gaber, T., El-Ghamry, A. and Hassanien, A.E., 2022. Injection attack detection using machine learning for smart IoT applications. *Physical Communication*, 52, p.101685.
4. Dai, Q., Zhang, B. and Dong, S., 2022. Eclipse attack detection for blockchain network layer based on deep feature extraction. *Wireless Communications and Mobile Computing*, 2022(1), p.1451813.
5. Walling, S. and Lodh, S., 2024. Network intrusion detection system for IoT security using machine learning and statistical based hybrid feature selection. *Security and Privacy*, 7(6), p.e429.
6. Zheng, H., Yin, H. and Li, H., 2024. FE-DIoT: IoT Device Classification Through Dynamic Feature Selection and Adaptive Cross-Network Model. *IEEE Access*.
7. Alrayes, F.S., Zakariah, M., Amin, S.U., Khan, Z.I. and Helal, M., 2024. Intrusion detection in IoT systems using denoising autoencoder. *IEEE Access*.
8. Patel, N.D., Mehtre, B.M. and Wankar, R., 2024. A computationally efficient dimensionality reduction and attack classification approach for network intrusion detection. *International Journal of Information Security*, pp.1-31.
9. Wang, W., 2024. Abnormal Traffic Detection for Internet of Things Based on An Improved Residual Network. *Physical Communication*, p.102406.
10. Abbasi, F., Naderan, M. and Alavi, S.E., 2025. Dimensionality reduction with deep learning classification for botnet detection in the Internet of Things. *Expert Systems with Applications*, 267, p.126149.
11. Houkan, A., Sahoo, A.K., Gochhayat, S.P., Sahoo, P.K., Liu, H., Khalid, S.G. and Jain, P., 2024. Enhancing Security in Industrial IoT Networks: Machine Learning Solutions for Feature Selection and Reduction. *IEEE Access*.
12. Saif, S., Yasmin, N. and Biswas, S., 2023. Feature engineering-based performance analysis of ML and DL algorithms for Botnet attack detection in IoMT. *International Journal of System Assurance Engineering and Management*, 14(Suppl 1), pp.512-522.
13. Alzubi, O.A., Alzubi, J.A., Qiqieh, I. and Al-Zoubi, A.M., 2025. An IoT Intrusion Detection Approach Based on Salp Swarm and Artificial Neural Network. *International Journal of Network Management*, 35(1), p.e2296.
14. Nguyen, D.T. and Le, K.H., 2023. The robust scheme for intrusion detection system in internet of things. *Internet of Things*, 24, p.100999.
15. Choudhary, V., Tanwar, S. and Choudhury, T., 2024. Evaluation of contemporary intrusion detection systems for internet of things environment. *Multimedia Tools and Applications*, 83(3), pp.7541-7581.
16. Ayad, A.G., Sakr, N.A. and Hikal, N.A., 2024. A hybrid approach for efficient feature selection in anomaly intrusion detection for IoT networks. *The Journal of Supercomputing*, 80(19), pp.26942-26984.
17. Anushiya, R. and Lavanya, V.S., 2023. A new deep-learning with swarm-based feature selection for intelligent intrusion detection for the Internet of things. *Measurement: Sensors*, 26, p.100700.
18. Parabathina, A., Dabburu, M. and Kasukurthi, V.R., 2024. Rat Swarm Optimization with Improved Gated Recurrent Unit for Intrusion Detection System. *International Journal of Intelligent Engineering & Systems*, 17(5).
19. Parameswari, A., Ganeshan, R., Ragavi, V. and Shereesha, M., 2024. Hybrid rat swarm hunter prey optimization trained deep learning for network intrusion detection using CNN features. *Computers & Security*, 139, p.103656.
20. Soltani, N., Rahmani, A.M., Bohloul, M. and Hosseinzadeh, M., 2024. Robust intrusion detection for network communication on the Internet of Things: a hybrid machine learning approach. *Cluster Computing*, pp.1-17.
21. CICIDS 2018 dataset Link: <https://www.kaggle.com/datasets/dhoogla/csecicids2018> (Accessed on 31 January 2025)
22. BOT-IoT dataset Link: <https://www.kaggle.com/datasets/vigneshvenkateswaran/bot-iot> (Accessed on 31 January 2025)
23. CICIDS 2017 dataset link: <https://www.kaggle.com/datasets/chethuhn/network-intrusion-dataset> (Accessed on 31 January 2025).