scientific reports

Check for updates

OPEN

Advances in IoT networks using privacy-preserving techniques with optimized multi-head self-attention model for intelligent threat detection based on plant rhizome growth optimization

Mimouna Abdullah Alkhonaini[1], Sara Abdelwahab Ghorashi[2], Ghalib H. Alshammri[3], Saied Alshahrani[4], Shouki A. Ebad[5]✉, Sami Saad Albouq[6], Fahad Alzahrani[7] & Menwa Alshammeri[8]

The advances in the Internet of Things (IoT) involve a technology of interconnected devices that interact over the internet, providing convenience and efficiency while also posing significant security risks. Privacy-preserving techniques play a vital role in safeguarding sensitive user data while maintaining system efficiency. The rising tendency of cybersecurity threats and the need to recognize harmful activities in heterogeneous but resource-constrained settings have led to the development of sophisticated intrusion detection systems (IDSs) for quickly identifying intrusion efforts. Conventional IDSs are becoming more inefficient in classifying new attacks (zero-day attacks) whose designs are similar to any threat signatures. To reduce these restrictions, projected IDS depend on deep learning (DL). Due to DL techniques learning from vast amounts of data, they can identify novel, emerging attacks, making them an alternative method to classical cybersecurity. This study proposes an Optimised Multi-Head Self-Attention Model for an Intelligent Intrusion Detection Framework Using Plant Rhizome Growth Optimisation (OMHSA-IDPRGO) method to advance IoT security. The primary focus is on developing an automated cyberattack detection system for an IoT environment by employing advanced techniques. Initially, the mean normalization process is used to measure input data into a structured format. Furthermore, the Crayfish Optimisation Algorithm (COA) is used for optimal feature subset selection, identifying the most relevant features from the dataset. For the cybersecurity detection process, the OMHSA-IDPRGO method uses a hybrid model that encompasses a convolutional neural network and a bidirectional gated recurrent unit with a multi-head self-attention mechanism (CNN-BiGRU-MHSAM) technique. Finally, the hyperparameter selection is performed using the plant rhizome growth optimization (PRGO) approach to enhance classification performance. The experimentation of the OMHSA-IDPRGO model is examined under Edge-IIoT and ToN-IoT datasets. The comparison study of the OMHSA-IDPRGO model showed superior accuracy values of 99.11 and 99.18% compared to existing techniques on the dual datasets.

Keywords Privacy-preserving, Multi-head self attention, Intrusion detection, Cybersecurity, IoT security, Plant rhizome growth optimization, Feature selection

[1]Department of Computer Science, College of Computer and Information Sciences, Prince Sultan University, Riyadh, Saudi Arabia. [2]Department of Computer Sciences, College of Computer and Information Sciences, Princess Nourah Bint Abdulrahman University, P.O. Box 84428, 11671 Riyadh, Saudi Arabia. [3]Department of Computer Science, Community College, King Saud University, 11437 Riyadh, Saudi Arabia. [4]College of Computing and Information Technology, University of Bisha, 61922 Bisha, Saudi Arabia. [5]Center for Scientific Research and Entrepreneurship, Northern Border University, 73213 Arar, Saudi Arabia. [6]Department of Computer and Information Systems, Islamic University of Madinah, 42351 Medina, Saudi Arabia. [7]Department of Information and Computer Science, College of Computing and Mathematics, King Fahad University of Petroleum and Minerals, Dhahran, Saudi Arabia.

[8]Department of Computer Science, College of Computer and Information Sciences, Jouf University, Sakaka, Saudi Arabia. ✉email: shouki.abbad@nbu.edu.sa

Advances in the IoT relate to a system of physical entities, tools, automobiles, and other components equipped with electronic hardware, circuits, sensors, software, and internet connectivity that enable these elements to gather and share information[1]. The IoT permits these items to be monitored and managed remotely through available network setups, generating possibilities for tighter connection of the physical environment with digital platforms, ultimately leading to enhanced performance and precision[2]. The IoT stands as a ground-breaking advancement that signals the future of digital technology and communication, and its progress relies on evolving technical breakthroughs in numerous key areas, including wireless detection systems and nanoscale technologies. The IoT concept offers an infinite amount of availability, accessibility, scalability, integrity, confidentiality, and other benefits related to connected devices[3]. Nevertheless, IoT devices are vulnerable to cyber threats due to a combination of numerous potential vulnerabilities and their relative novelty, resulting in the absence of established security norms and protocols. A broad spectrum of cyber threats is employed against IoT systems, based on the specific component being targeted and the attacker's objectives[4].

However, the deployment of IoT systems has also introduced several significant cybersecurity challenges, where loopholes and unauthorized access to information and critical infrastructure have become major concerns[5]. The IoT represents a global web of smart devices connected to the internet without human intervention, which is beneficial yet prone to cyberattacks, like any conventional system. Studies focus on integrating machine learning (ML)- driven solutions with IoT. An IDS is a reliable method for identifying cyberattacks within any system[6]. Many modern IDS frameworks employ ML models to detect cyber threats in the system. This IDS alerts the network administrator about any doubtful behaviour occurring within the network and therefore serves as an information security mechanism that blocks harmful intrusions[7]. An intrusion occurs when an individual gains unauthorized entry to, or manipulates data assets with malicious intent. A tangible entity aiming to extract information unlawfully, cause harm to others, or perform malicious tasks is referred to as a cybercriminal or an intruder[8]. Figure 1 denotes the general architecture of IDSs in IoT.

Cybersecurity threats have expanded rapidly in several fields, including healthcare, smart homes, agriculture, energy, industrial processes, and automation. Due to its extensive range of services, IoT sensors generate vast volumes of data, which necessitate privacy, security, and authentication[9]. Previously, classical techniques were used to ensure IoT security. The employment of more artificial intelligence (AI) techniques for identifying cybersecurity threats has become increasingly popular over time. Consequently, there is a substantial body of investigation concentrated on IoT cybersecurity[10]. This encompasses AI strategies for safeguarding IoT systems against threats, typically by detecting abnormal activities that could indicate an ongoing attack. As interconnected devices are widely utilized in various sectors, they have created a demand for smarter environments and automated processes. However, this growth is accompanied by increasingly advanced security threats, which may compromise data integrity and user privacy. Efficient threat detection methods that can operate in real-time and on resource-constrained devices are urgently required. Improving security while maintaining system performance is significant to unlocking the full potential of connected technologies. This drives the development of innovative, optimized models that balance accuracy, interpretability, and privacy protection in intrinsic network environments.

This study proposes an Optimised Multi-Head Self-Attention Model for an Intelligent Intrusion Detection Framework Using Plant Rhizome Growth Optimisation (OMHSA-IDPRGO) method to advance IoT security. The primary focus is on developing an automated cyberattack detection system for an IoT environment by employing advanced techniques. Initially, the mean normalization process is used to measure input data into a structured format. Furthermore, the Crayfish Optimisation Algorithm (COA) is used for optimal feature subset selection, identifying the most relevant features from the dataset. For the cybersecurity detection process, the OMHSA-IDPRGO method employs a hybrid model that encompasses a convolutional neural network and a bidirectional gated recurrent unit with a multi-head self-attention mechanism (CNN-BiGRU-MHSAM) technique. Finally, the hyperparameter selection is performed using the plant rhizome growth optimization (PRGO) approach to enhance classification performance. The experimentation of the OMHSA-IDPRGO model is examined under Edge-IIoT and ToN-IoT datasets. The key contribution of the OMHSA-IDPRGO model is listed below.

- The OMHSA-IDPRGO method applies mean normalization during pre-processing to convert raw input data into a uniformly scaled format, thereby improving consistency across features. This enhances training stability and learning efficiency. It also assists in better convergence during model optimization.
- The OMHSA-IDPRGO approach utilizes the COA technique to identify the most relevant features, effectively mitigating dimensionality while preserving crucial data. This selection improves detection accuracy and mitigates computational complexity. As a result, the model becomes more efficient and reliable in processing IoT data.
- The OMHSA-IDPRGO methodology employs a hybrid CNN-BiGRU architecture integrated with the MHSAM model for capturing intrinsic temporal and spatial patterns in IoT network data. This enhances the

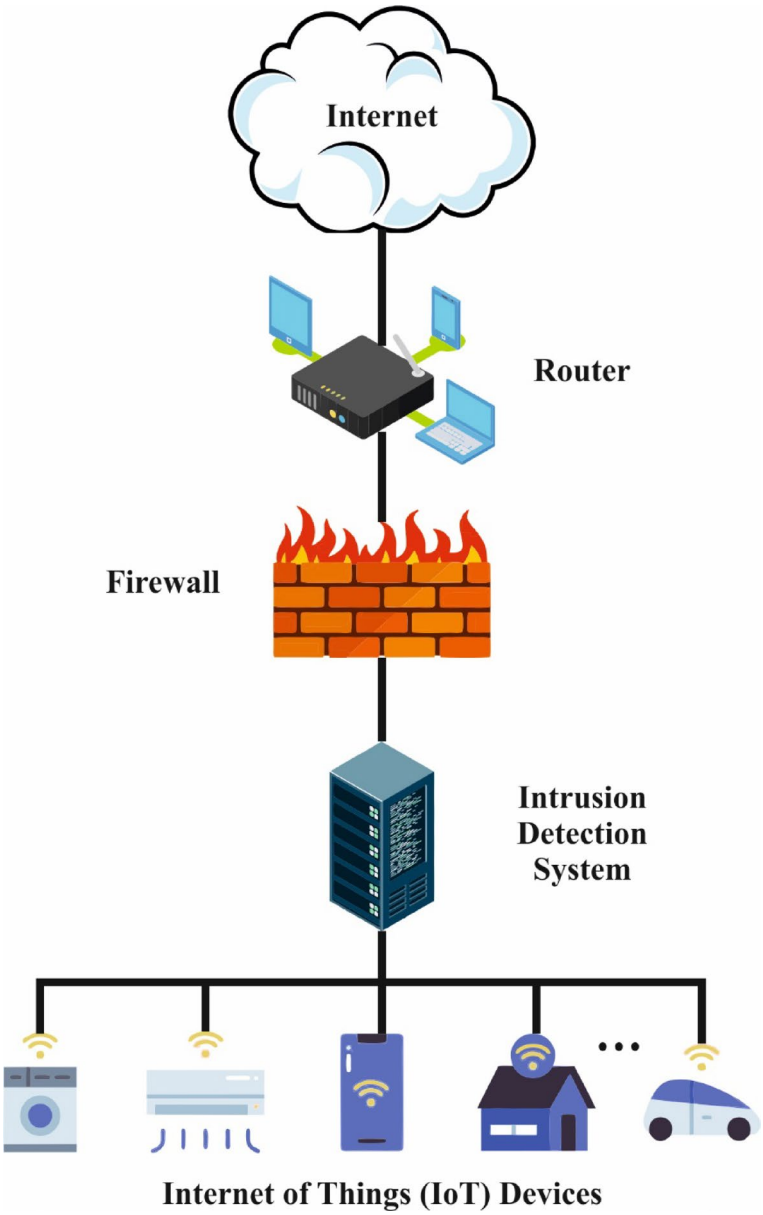

Fig. 1. General structure of intrusion detection in IoT.

model's ability to detect a wide range of cyber threats accurately. It significantly improves intrusion detection performance in dynamic environments.

- The OMHSA-IDPRGO technique implements the PRGO model to fine-tune hyperparameters, improving learning efficiency and predictive accuracy automatically. This optimization reduces manual tuning efforts and accelerates convergence. Consequently, it enhances overall model performance in complex IoT scenarios.
- The integration of COA-based FS with the CNN-BiGRU-MHSAM model, optimized through PRGO, presents a novel and efficient solution for intelligent threat detection. This incorporation uniquely balances dimensionality reduction, deep temporal-spatial feature extraction, and automated hyperparameter tuning. It significantly improves detection accuracy and computational efficiency in complex IoT environments, setting it apart from existing methods.

Related works

Alzahrani[11] proposed a gorilla troops optimizer and DL-assisted BAD (GTODL-BADC) method. This method utilizes feature selection (FS) in conjunction with fine-tuned DL-aided classification to achieve security within

the IoT landscape. The min–max data normalization technique was employed for data pre-processing, and the GTO framework for FS selects the optimal feature subclasses. Furthermore, the multi-head attention-driven LSTM model is used for BAD. Sekhar et al.[12] provided an effective model for identifying and classifying cyber intrusions. The authors employed a new model with honeybees mating optimization (HBMO) and stochastic gradient boosted distributed decision trees (SGB-DDT) techniques. To advance the recognition precision, an SGD-DDT is a learning approach that is effective and scalable. Alotaibi[13] suggested an innovative structure, which integrates DL, blockchain (BL), and software-defined networking (SDN) technologies for improving IoT cybersecurity. This study aims to identify a potential method for forming a cybersecurity system for IoT-enabled smart settings to safeguard data privacy, identify appropriate threats, and secure financial transactions. The suggested method is an amalgamation of numerous state-of-the-art techniques. The SDN control plane incorporates the SE-driven Bi-LSTM approach for traffic management. In[14], a smart IDS model is presented for identifying cyber threats in the IoAT. The presented model utilizes the downsized kernel partial least squares (DKPLS) and reduced kernel approach for extracting and reducing data features to improve the recognition efficiency. This DRKPLS technique was employed to reduce the dimensions of the kernel matrix produced by the kernel partial least squares (KPLS) model by selecting relevant features. In[15], the Ant Colony-Optimised Artificial Neural-Adaptive Tensorflow (ACO-ANT) approach was recommended for identifying suspicious software. To highlight the importance of tokens in source duplicate data, the noise data underwent processing by weighted attribute and tokenization approaches. DL methods are later deployed for detecting source code duplication. Aljebreen et al.[16] developed a new DDoS attack detection method using the snake optimizer and ensemble learning (DDAD-SOEL) technique in the IoT landscape. The motivation behind this method lies in the effective and automatic detection of DDoS attacks. To fulfil this, the created method leverages the SO procedure for selecting a feature subset.

Kumar et al.[17] combined SDN, digital twin (DT) BC, and DL techniques in the SG network structure. Specifically, a secure communication network was initially established using an authentication model based on BC technology, which mitigates some identified security threats. Then, a different DL design, which contains a softmax classifier, Bi-GRU, and a self-attention mechanism, is proposed to enhance the threat recognition step in SG networks. In[18], an advanced cybersecurity framework assisted in trucking the heuristic process was introduced with three methodologies: improved virtual honeypot (IVHD), IDS, and hidden Markov models (HMM) for segmenting devices into four diverse stages based on their planned task, and observing communication traffic to identify suspicious edge devices. Saheed, Misra, and Chockalingam[19] proposed a model by using an autoencoder with a deep convolutional neural network (DCNN) and long short-term memory (LSTM) for feature reduction and anomaly detection in Industrial Control Systems (ICS), enabling accurate, low-cost, and real-time cyber-attack detection. Thayalan et al.[20] presented a collaborative federated learning framework with edge-cloud architecture using the two-stage attention integrated graph-based multi-source spatio-temporal data fusion (2S-AGMSTDF) network, including Attention-based LSTM, attention-based knowledge graph convolutional network (AKGCN), and graph convolutional network-residual network-based transformer (GCN-ResNet Transformer or GRCMT) method to enhance accurate, scalable, and privacy-preserving predictive security in IoT consumer applications. Saheed, Abdulganiyu, and Ait Tchakoucht[21] presented a framework that integrates a modified genetic algorithm (MGA) model for feature selection with a deep LSTM network, optimized via a genetic algorithm (GA) for hyperparameter tuning, to efficiently detect cyberattacks in IoT networks within an edge computing environment. Kumar, Jolfaei, and Islam[22] proposed a DL-based threat hunting framework (DLTHF) technique by using an LSTM contractive sparse autoencoder (LSTM-CSAE) model for feature extraction and a multi-head self-attention bidirectional recurrent neural network (MhSaBiGRNN) methodology for accurate cyber threat detection in Software Defined-IoT (SD-IoT) networks. Paul et al.[23] presented a model to enhance cybersecurity in deep web environments, utilizing a novel framework that integrates federated learning (FL), graph-based analysis, and a hybrid web crawler with an ontology-based scoring system to detect threats and safeguard sensitive data across cloud, fog, and edge systems.

Kathole et al.[24] developed a secure attack detection framework for Vehicular Ad-Hoc Networks (VANETs) using a modernized random parameter-based green anaconda optimization (MRP-GAO) model for feature selection and an ensemble ML model (EMLM) integrating multi-layer perceptron (MLP), support vector machine (SVM), AdaBoost, and Bayesian network for effective intrusion detection and classification. Amer, Al-Rimy, and El-Sappagh[25] proposed the Modbus-NFA Behaviour Distinguisher (MNBD) model, which applies a non-deterministic finite automaton (NFA) framework to analyze Modbus frame sequences and identify abnormal device behaviour with high accuracy and generalization. Sardar et al.[26] introduced a model by using a graph neural network (GNN) model trained on various datasets and evaluated via NS2 simulations. Kathole et al.[27] presented an ensemble DL model (EDLM) technique that integrates multiple DL models to improve detection accuracy, reduce false alarms, and strengthen network security by averaging prediction scores for robust anomaly detection. Saheed and Chukwuere[28] developed a robust cyber-attack detection system for cyber-physical industrial IoT (CPS-IIoT) that utilizes the Pearson correlation coefficient and agglomerative clustering for privacy preservation, as well as a bidirectional LSTM with scaled dot-product attention (BiLSTM-SDPA) method for accurate threat detection. Kathole et al.[29] developed a secure federated cloud storage system for Internet of Medical Things (IoMT) using a hybrid Mexican axolotl with energy valley optimizer (HMO-EVO)-based attribute-based encryption (ABE) for secure data encryption and multi-scale bi-long short-term

memory and gated recurrent unit (MBiLSTM-GRU) technique with FL for accurate disease prediction. Saheed, Omole, and Sabit[30] developed a model using a genetic algorithm with an attention mechanism and a modified adaptive moment estimation optimized LSTM (GA-mADAM-IIoT) methodology, incorporating explainability via Shapley Additive Explanations (SHAP). Saheed and Chukwuere[31] developed an explainable AI (XAI) ensemble transfer learning (TL) model using SHAP and a hybrid bidirectional long short-term memory with autoencoders (BiLAE) technique for zero-day botnet attack detection, optimized by barnacle mating optimizer (BMO). Saheed and Misra[32] presented an explainable and privacy-preserving deep neural network (DNN) framework with SHAP for accurate and interpretable anomaly detection in Cyber-Physical Systems enabled IoT (CPS-IoT) networks. Ullah et al.[33] developed SecNet-FLIDS, a Blockchain-based FL model with a TOP-K Node selection scheme and context-aware transformer networks, incorporated with synthetic minority over-sampling

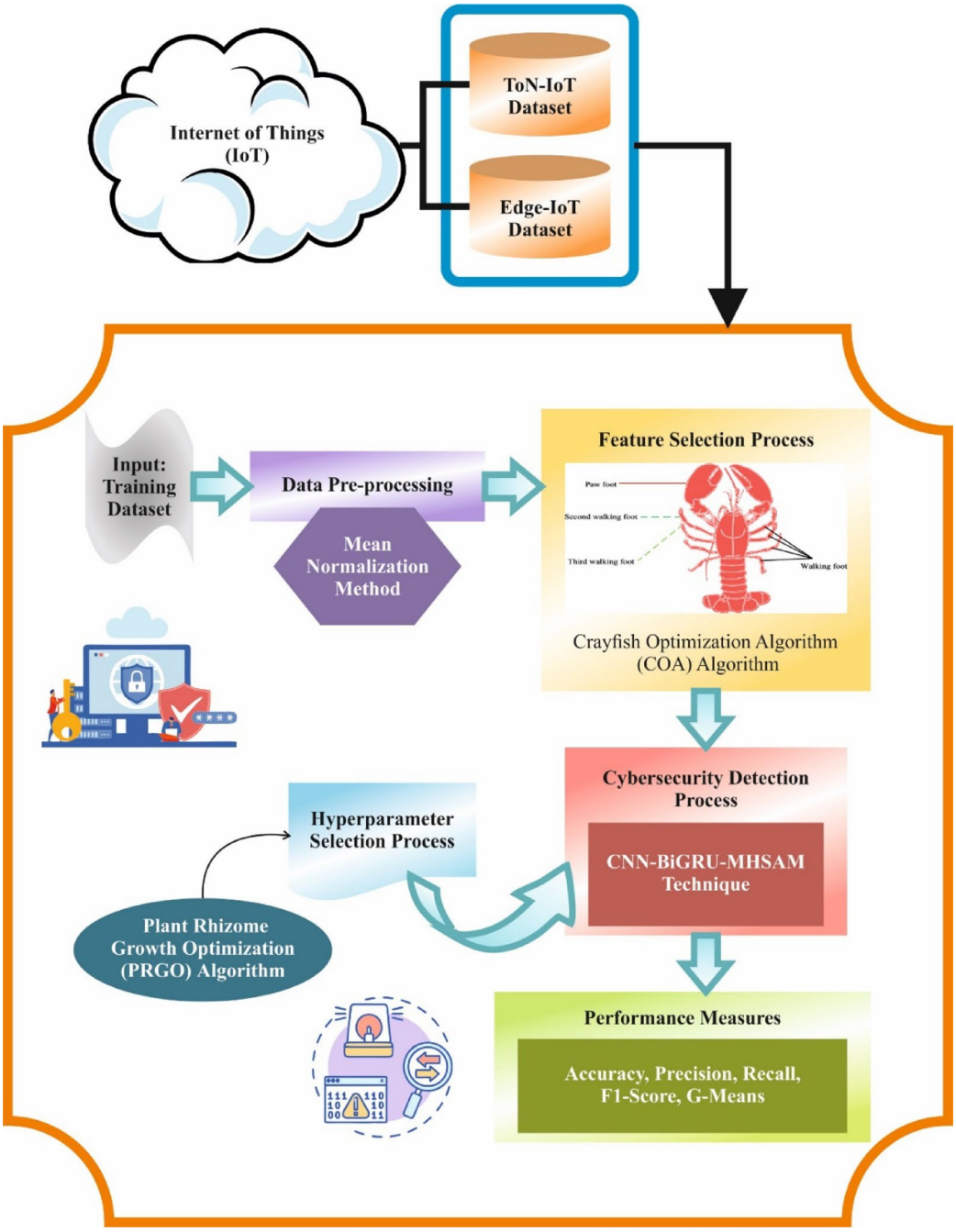

Fig. 2. Working procedure of the OMHSA-IDPRGO model.

technique (SMOTE) and edited nearest neighbours (ENN) for imbalanced data handling, to enable accurate, privacy-preserving, and scalable cyberattack detection in the Internet of Vehicles (IoV).

Despite various advanced methods, such as GTO, HBMO, Bi-LSTM, and FL, being applied across IoT, CPS-IIoT, IoV, and SDN environments, several limitations still exist. Various models rely on large, labelled datasets, which are often scarce or imbalanced, thereby affecting detection accuracy and generalizability. Trust among users and security experts is reduced as few models suffer from interpretability issues. Furthermore, various solutions lack scalability or are computationally intensive, which restricts their real-time deployment on resource-constrained edge devices. The integration of multiple technologies, such as BC, DL, and optimization algorithms, remains complex, resulting in increased system overhead. The research gap in addressing these challenges lies in developing lightweight, explainable, and privacy-preserving IDS that efficiently handle imbalanced data, optimize feature selection, and adapt to heterogeneous IoT and CPS environments with minimal human intervention.

Methodology

This study designs and develops an OMHSA-IDPRGO model to advance IoT security. The primary focus of this novel is to enhance automatic cyberattack detection in IoT environments by employing advanced techniques. To achieve this, the OMHSA-IDPRGO technique contains numerous stages, namely data pre-processing, FS, classification model, and parameter tuning process. Figure 2 signifies the overall working flow process of the OMHSA-IDPRGO technique.

Mean normalization

To achieve this, the proposed method initially executes a mean normalization method to transform the input data into a structured format[34]. This normalization is chosen due to its effective capability in scaling features to have a mean of zero, which assists in stabilizing and speeding up the training process of ML models. This method mitigates the impact of outliers by centring the data around zero without strictly bounding it within a fixed range. This enhances gradient descent convergence, especially in DL methods, by preventing features with large magnitudes from dominating the learning process. Additionally, mean normalization maintains the relative distribution of data, which is significant when dealing with complex IoT datasets. Overall, it presents a balanced approach that enhances model stability and accuracy compared to simpler scaling techniques.

This method signifies a frequently implemented data normalization model in data analysis and ML domains. It comprises scaling the data so that the mean (average) of the feature becomes 0. This procedure focuses on the data around the mean, effectually extracting some bias in the value of features. A feature x' is specified:

$$x' = \frac{x - mean\,(x)}{\max\,(x) - \min\,(x)} \tag{1}$$

Now, $\min\,(x)$ and $\max\,(x)$ signify the minimal and maximal values of x, and $mean(x)$ depicts the mean of feature x.

FS using COA model

Next, the COA approach is employed for the optimal selection of feature subsets[35]. This model is chosen for its robust global search capability and effective balancing. This technique also avoids local optima, ensuring the selection of the most relevant and informative features. Its bio-inspired nature allows it to adapt dynamically to complex, high-dimensional IoT datasets, enhancing feature relevance and mitigating redundancy. Furthermore, COA requires relatively low computational resources, making it appropriate for real-time applications. Overall, COA improves detection accuracy and model efficiency better than many conventional FS models. The mathematical model for the optimization issue is developed by considering the natural behaviours of crayfish at diverse phases. COA is a metaheuristic swarm intelligence model. The significant stages of COA are given.

Initialize population

During this initial stage of the model, the crayfish population is described as a matrix X, where X_i ($X_i = \{X_{i,1}, X_{i,2}, \ldots, X_{i,dim}\}$, with dim depicting the dimension) indicates the location of $thel^{th}$ crayfish. Every X_i is constrained within a predefined boundary and represented as a $1 \times dim$ vector, equivalent to a possible solution. The function $f\,(\cdot)$ is presented to assess X_i, offering its fitness value. According to the size of population N and dimension dim, the COA initializes the procedure by arbitrarily generating a set of candidate solutions X. The location of individual i in dimension $X_{\mathrm{i,j}}\cdot$ is specified:

$$X_{i,j} = lb_j + (ub_j - lb_j) \times rand \tag{2}$$

Here, ub_j and lb_j refer to the upper and lower bounds, respectively, while $rand$ indicates an arbitrary value.

Define temperature

Temperature variations influence the behaviour of crayfish, prompting them to enter various phases. The temperature is specified:

$$temp = rand \times 15 + 20 \tag{3}$$

While the ambient temperature ranges from 15 to 30 °C, particularly at 25 °C, crayfish display searching behaviour. The feeding amount is assessed by employing the expression specified in Eq. (4).

$$p = C_1 \times \left[\frac{1}{\sqrt{2\pi} \times \sigma} \times \exp\left(-\frac{(temp - \mu)^2}{2\sigma^2} \right) \right] \tag{4}$$

Here, μ depicts the optimum feeding temperature. The crayfish's food intake is influenced by the parameters C_1 and σ, and set to 3 and 0.2, respectively.

Summer resort and competition stage

The temperature $temp$ is greater than 30 during these dual stages, and the cave X_{shade} is deliberated as:

$$X_{shade} = (X_G + X_L)/2 \tag{5}$$

Now, X_L and X_G refer to the optimal locations attained through iterative calculations for individuals and populations.

While $rand < 0.5$, no other individuals participate in the competition, and the crayfish immediately enters the cave, upgrading its location.

$$X_{i,j}^{t+1} = X_{i,j}^{t} + C_2 \times rand \times \left(X_{shade} - X_{i,j}^{t} \right) \tag{6}$$

Now t and $t+1$ indicate the next and existing iteration stages. C_2 is specified:

$$C_2 = 2 - \left(\frac{t}{T} \right) \tag{7}$$

Here, T specifies the upper limit of iterations. While rand ≥ 0.5, other individuals are also interested in the cave, inducing competition, and their locations are upgraded accordingly.

$$X_{i,j}^{t+1} = X_{i,j}^{t} - X_{z,j}^{t} + X_{shade} \tag{8}$$

Now, z refers to an arbitrarily selected individual from the population.

$$z = round\left[rand \times (N-1) \right] + 1 \tag{9}$$

Foraging stage

During this phase, the temperature does not exceed 30 °C, and the crayfish will evaluate the dimensions of the food after positioning it and select diverse feeding approaches. The position of food X_{food} is specified:

$$X_{food} = X_G \tag{10}$$

The size of food Q is given:

$$Q = C_3 \times rand \times (fitness_i / fitness_{food}) \tag{11}$$

Here, $fitness_i$ and $fitness_{food}$ represent the fitness values of the crayfish and the food, respectively. C_3 represents the food factor, set to a value of 3, correspondingly. The crayfish will assess the size of the food based on its type. While $Q > (C_3 + 1)/2$, it specifies that the food is huge, inducing the crayfish to tear it apart, utilizing its primary claw foot.

$$X_{food} = \exp\left(-\frac{1}{Q} \right) \times X_{food} \tag{12}$$

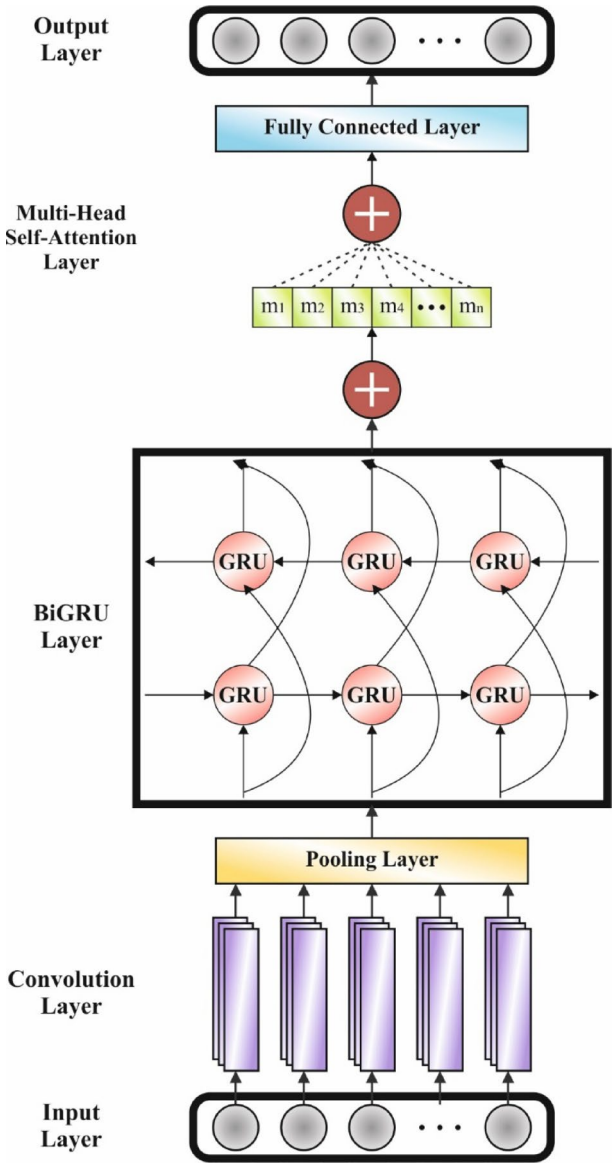

Fig. 3. Structure of the CNN-BiGRU-MHSAM technique.

Once the food is torn apart and turns small, the preceding dual paws will alternatively pinch the food to consume it. This alternative approach is designed by integrating $sine$ and $cosine$ functions. Additionally, food diversity is achieved by simulating the feeding behaviour of the animals.

$$X_{i,j}^{t+1} = X_{i,j}^{t} + Xfood \times p \times [\cos(2 \times \pi \times rand) - \sin(2 \times e \times rand)] \tag{13}$$

While $Q \leq (C_3 + 1)/2$, the crayfish might instantly move towards the nutrition and upgrade its location:

$$X_{i,j}^{t+1} = \left(X_{i,j}^{t} - X_{food}\right) \times p + p \times rand \times X_{i,j}^{t} \tag{14}$$

Due to its simplicity, fast convergence speed, and higher computational efficiency, the COA is used to address the SRB optimizer concern.

The fitness function (FF) determines the classifier's precision and the various features chosen. It increases the classifier's precision and decreases the size of the desired feature set. Thus, the subsequent FF is used to evaluate different solutions presented in Eq. (15).

$$Fitness = \alpha * ErrorRate + (1 - \alpha) * \frac{\#SF}{\#All\_F} \tag{15}$$

Here, $ErrorRate$ signifies the rate of classification errors using the chosen features. $ErrorRate$ is computed as the percentage of improperly classified instances to the total number of classifications made, expressed as a value between 0 and 1. $\#SF$ is the number of chosen aspects, and $\#All\_F$ is the total number of attributes in the original dataset. α is utilized to control the significance of subset length and classifier quality.

Hybrid classification model

For the cybersecurity detection process, the hybrid model, named CNN-BiGRU-MHSAM, is used[36]. This technique is chosen for its ability to capture both spatial and temporal patterns in complex IoT network data. CNN layers outperform in extracting local features, while BiGRU handles long-term dependencies in sequential data. The addition of the MHSAM improves the capability of the technique in concentrating on the most relevant parts of the input, enhancing detection accuracy. Compared to standalone models, this hybrid approach mitigates overfitting and increases robustness against diverse cyber threats. It also presents better interpretability and scalability, making it superior to conventional ML and simpler DL methods. Figure 3 represents the structure of the CNN-BiGRU-MHSAM technique.

CNN reduces the parameter counts and sizes of the data by removing spatial features from the input data through methods such as pooling, weight sharing, and local connections. These processes improve the analytical capabilities and computational efficiency of the CNN. The network includes pooling, input, output, convolutional, and fully connected (FC) layers. As a fundamental module of CNNs, the convolutional process focuses on capturing spatial relationships and various patterns. This framework combines dual convolutional components—the first layer uses 16 3x3 kernel filters, whereas the subsequent layer uses 32 3x5. dimensional matrices to analyze the combined input resources. The pooling layer decreases the input data dimensionality, improving CNN calculation speed and alleviating overfitting. Normal pooling techniques include stochastic pooling, max pooling, and mean pooling. The FC layer connects every neuron to each neuron in the previous layer.

The GRU method effectively transfers its capability to process time-series data, successfully addressing problems such as exploding and vanishing gradients. In comparison with LSTM, the GRU requires fewer parameters, provides faster training, and reduces overfitting, making it particularly effective for time-series prediction. This method is primarily composed of two segments: the gate of reset r_t and the gate of update z_t. The data processing process executed by the GRU model is as demonstrated:

(i) The update gate z_t controls the extent to which the data from the preceding instant is preserved at present moments. The greater the update gate's output value z_t, the additional data from the prior instant is preserved. The evaluation equation of the update gate z_t. is as stated in Eq. (16).

$$z_t = \sigma\left(w_z \cdot [h_{t-1'} x_t] + b_z\right) \tag{16}$$

Here, $\sigma(\cdot)$ refers to the sigmoid function, w_z denotes the update gate weight, h_{t-1} signifies the hidden layer (HL) of the preceding moment, x_t stands for the present moment input, and b_z. Means the bias value.

(ii) The reset gat r_t. Controls the extent to which the data from the preceding moment is maintained on the present moment candidate HL $\tilde{h}_t$. The greater the reset gate's output value, s r_t. The more additional data from the prior instant is kept in the candidate HL, $\tilde{h}_t$. The prediction formula for the reset gate, r_t, is represented in Eq. (17).

$$r_t = \sigma\left(w_r \cdot [h_{t-1'} x_t] + b_r\right) \tag{17}$$

N, w_r. denot reset gate weight, and b_r signifies bias value.

(iii) The upgrade of the ndidate HL information $\tilde{h}_t$ is as stated in Eq. (18).

$$\tilde{h}_t = tanh\left(w_h \cdot [r_t \odot h_{t-1}, x_t] + b_h\right) \tag{18}$$

Whereas, $tanh(\cdot)$ symbolizes activation function, w_h characteristics ndidate HL weight, and b_h denotes bias value of the candidate HL.

Component	Hyperparameters	Values
CNN	Filters, kernel size, activation	64, 3 × 3, ReLU
BiGRU	Units, layers, dropout	128, 2, 0.3
MHSAM	Attention heads, attention dimension	8, 64
Training settings	Optimizer, learning rate, batch size, epochs	Adam, 0.001, 64, 50

Table 1. Key hyperparameters used in the proposed attention-based DL model for intelligent threat detection in IoT networks.

(iv) Compute the present moment HL h_t according to $\tilde{h}_t$ and h_{t-1}:

$$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \tilde{h}_t \tag{19}$$

This model successfully takes the forward characteristics. Nevertheless, time series data are also affected by previous data, not only upcoming values. To address either forward or backwards dependencies, the BiGRU model is presented. This method combines either forward or backwards GRU elements to well-examine the relationships in time-series data. This model comprises a forward GRU, an input, an output and a reverse GRU. The forward and reverse GRUs extract data from the relevant reverse and forward sequences of the input data.

$$\vec{h} = f\left(x_{t'} \vec{h}_{t-1}\right) \tag{20}$$

$$\underset{t}{\overleftarrow{h}} = f\left(x_{t'} \underset{t-1}{\overleftarrow{h}}\right) \tag{21}$$

$$Q_t = \vec{w}\vec{h} + \overleftarrow{w} \underset{t}{\overleftarrow{h}} + b_t \tag{22}$$

Here, $\vec{w}$ and $\overleftarrow{w}$. represent output weights of the forward and backwards GRU HLs, individually, and b_t denotes the bias value of the Bi-GRU method.

The MHSA model creates numerous data sequences by using dissimilar linear projections. Every prediction, or "head", captures different data features. These sequences are then concatenated and passed through additional linear projections to form the final sequence. For the input sequence of data $X = [x_1, x_2, x_3, \cdots, x_m]$, multiply the put sequence of data by the matrices of linear transformation W_j^Q, W_j^K, and W_j^V, correspondingly, to get the vectors Q_j, K_j, and V_j, $j \in (1, 2, 3, \cdots, N)$, and N characterizes head counts. The particular computation procedure is presented in Eqs. (23–25).

$$Q_j = W_j^Q \cdot X \tag{23}$$

$$K_j = W_j^K \cdot X \tag{24}$$

$$V_j = W_j^V \cdot X \tag{25}$$

Formerly, the attention computation equation of the jth head is presented in Eq. (26).

$$head_i = Attention\left(Q_i, K_i, V_i\right) = softmax\left(\frac{Q_i \cdot K_i^T}{\sqrt{N}}\right) \cdot V_i \tag{26}$$

The output is measured by Eq. (27).

$$Y = Concat\left(head_1, head_2, head_3, \cdots, head_N\right) \cdot W^o \tag{27}$$

Now, W^o refers to the output weighted matrix. Table 1 describes the key parameters utilized in the CNN-BiGRU-MHSAM model.

PRGO-based parameter tuning model

Finally, the hyperparameter range is implemented by the PRGO approach[37]. This model is chosen for its effective global search capability, motivated by the natural rhizome growth patterns. This model helps avoid local minima and find optimal hyperparameter values, thereby efficiently balancing the exploration and exploitation phases. Compared to conventional tuning methods, such as grid or random search, PRGO requires fewer iterations and computational resources while achieving better convergence. Its adaptability to complex, high-dimensional search spaces make it ideal for optimizing DL techniques in IoT environments. Overall, PRGO enhances model accuracy and training speed, outperforming many conventional optimization techniques.

In this paper, a method is required to model plant rhizome growth and explain the fundamental meaning of five terms: soil space, adventitious, fibrous, lateral, and primary roots, in the context of root system development. Indeterminate and fibre roots serve as searching agents within the solution area and represent primary elements to discover the solution area. Lateral roots stimulate development by combining food absorbed by earlier fibre roots with nutrients taken up by the fibre roots. The primary root stimulates development by combining nutrition formerly absorbed by lateral and fibre roots with food intake by now evolving fibre roots.

Numerical growth method for taproot plants

Assume that taproot plants arbitrarily have N fibrous roots, signified by $(\{X_1, X_2 \ldots, X_N\})$. Fibrous roots additionally absorb nutrients from the soil, apart from exploring the nutrition-rich parts of the soil individually; therefore, all fibrous root is considered as an experimental solution to the optimization issue, and the consistent soil region is observed as the solution area of the targeted problem.

1. The numerical representation of fibrous root growth.

$$Seed_1 = X_i^t + \alpha_1 \left(X_{best} - X_{r1}^t\right) \tag{28}$$

whereas $Seed_1$ denote ith fiber root at the tth growth time, X_{best} signifies the optimal fiber root , r_1 refers to randomly generated numbers amongst $(1, N)$, X_{r1}^t signifies the fiber root ars, $\alpha_1 \in 2 \times rand\,(0, 1) - 0.5$ denote randomly generated number applied to produce $(-0.5, 1.5)$ intervals, if $\alpha_1 \in (0, 1.5)$, it means that the arbitrarily chosen fiber root X_{r1}^t advance to the historic finest fiber root X_{best}, and if $\alpha_1 \in (-0.5, 0)$, then X_{best} develops to X_{r1}^t. The model of this searching mechanism, in addition, guarantees that the historic finest fibre root X_{best} plays an integral part in directing each fibre root to mature in the best way. It is additionally guaranteed that a few arbitrary fibre roots X_{r1}^t direct the remaining fibre roots to carry out arbitrary searching that may prevent the search from dropping into a local best to the particular area.

2. The numerical modelling of lateral root growth in the soil area.

$$Seed_2 = X_{mean}^t + \alpha_2 \left(X_{mean}^t - X_{r2}^t\right) \tag{29}$$

whereas $Seed_2$ means the nutrition capture by lateral roots at the $t + lst$ development, X_{mean}^t represents average value, r_2 refers to a randomly generated number among $(1, N)$, X_{r2}^t signifies the fiberous root arbitrarily chosen from the N fibre roots inside the tth development, $\alpha_2 \in 2 \times rand\,(0, 1) - 1$ signifies randomly generated numbers applied to create the interval $(-1, 1)$. If $\alpha_2 \in (0, 1)$, it means that the root absorbs nutrients, and if $\alpha_2 \in (-1, 0)$, it indicates that the root permits organic mixtures to the nearby atmosphere.

3. The numerical modelling of primary root growth in the soil area.

$$Seed_3 = Seed_2 + \alpha_3 \left(Seed_1 - R_l^u\right) \tag{30}$$

Here $Seed_3$ embodies the food absorbs by the primary root at the $t + 1st$ growing, R_l^u refers to complete soil area (for example: the targeted problem search area), u and characterize lower and upper limits of the problem area, $\alpha_3 \in \varepsilon \times rand\,(0, 1) + \tilde{\varepsilon}, \varepsilon$ denote randomly generated integer of 0 or 1 produced by $randi\,(0, 1)$ function, $\tilde{\varepsilon}$ stands for reverse of ε. If ε is selected as 1, $\alpha_3 \in (0, 1)$, which specifies that fibre roots do not get sufficient food intake, and if ε is set to 0, $\alpha_3 = 1$, which stipulates that the fibrous roots can get the nutrition.

Edge-IIoT dataset	
Type of event	**Data record**
"Normal"	2500
"DDoS-UDP"	2500
"DDoS-ICMP"	2500
"SQL injection"	2500
"DDoS-TCP"	2500
"Password"	2500
"DDoS-HTTP"	2500
"Uploading"	2500
"Backdoor"	2500
"XSS"	2500
"Ransomware"	2500
"Fingerprinting"	2500
Total records	**30,000**

Table 2. Details of edge-IIoT dataset.

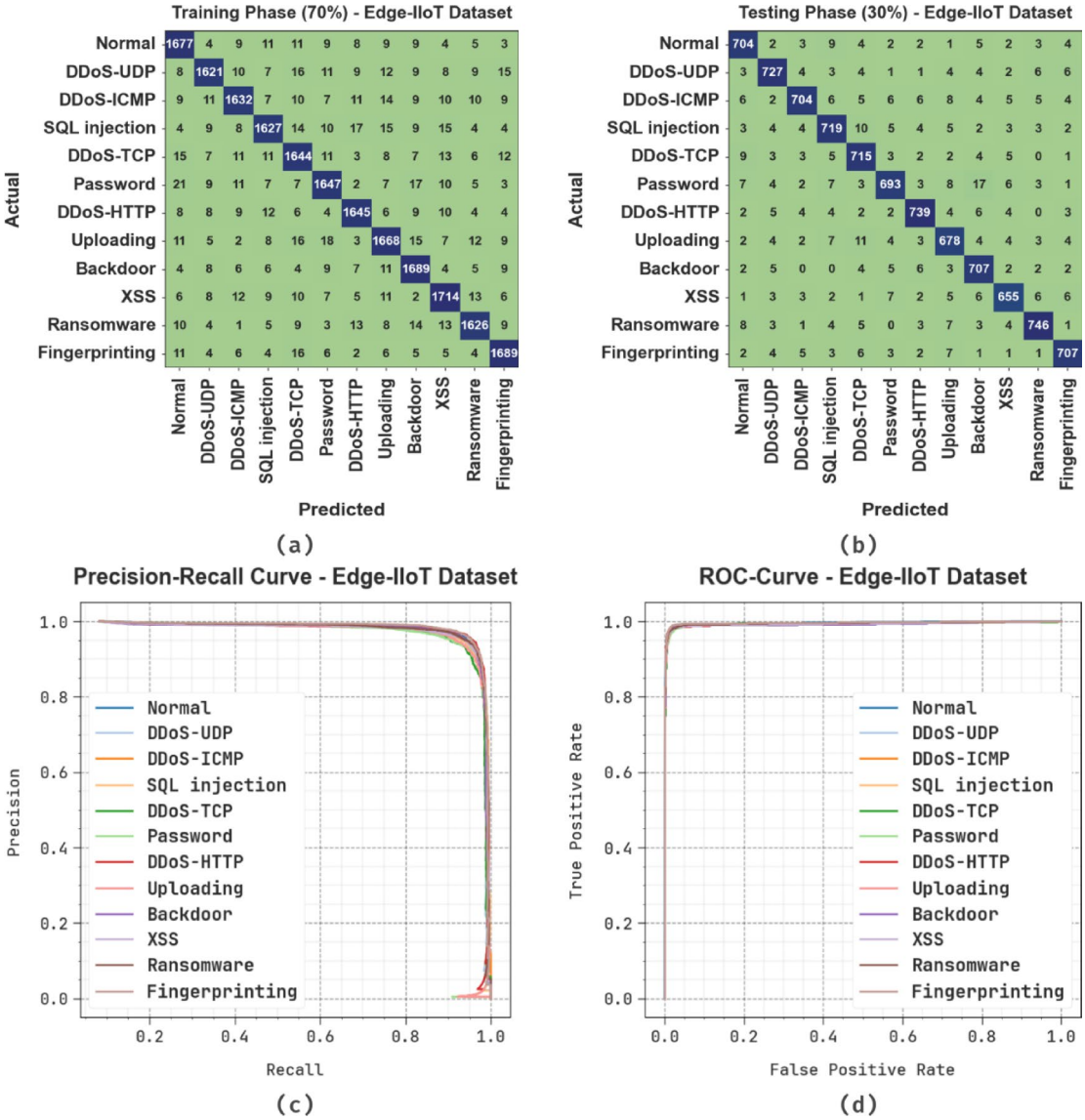

Fig. 4. Edge-IIoT dataset (**a**–**b**) confusion matrices, (**c**) curve of PR, and (**d**) curve of ROC.

This randomly generated number rule ensures that α_3 is an arbitrary number within the range of (0, 1), which accurately reflects the actual food intake by the plant root system. $Seed_1 - R_l^u$ characterizes the food intake by the fibrous roots.

Numerical model of plant growth in the fibre root system

The primary root of the fibrous root system deteriorates during seed germination, and simultaneously, the adventitious or fibrous root matures at the bottom of the stem. The adventitious root's growing point X_i^{t+1} at the following instant is jointly defined by the present X_i^t, the thick (best) X_{lbest}, X_{r3} and X_{r4} inside the soil area, and this searching method guarantees that X_i^t emerges from the local optimal with a higher possibility. Besides, while the growth of adventitious roots within the fibre root system is highly arbitrary, each of the adventitious roots maintains a particular number of aggregations, which allows the adventitious roots to discover additional nutrition-rich regions. According to the preliminary study, the root growth of fibrous-root plants is demonstrated as shown.

$$X_c = \frac{\alpha_4 \left(X_{r3} + X_{r4} + X_i^t\right)}{3} \tag{31}$$

$$Seed_4 = X_c + \left(X_{lbest} - X_i^t + X_{r3} + X_{r4}\right) \tag{32}$$

Class labels	$Accu_y$	$\Pr ec_n$	$Reca_l$	$F1_{Score}$	$G_{Measure}$
TRPHE (70%)					
Normal	99.10	94.00	95.34	94.67	94.67
DDoS-UDP	99.09	95.47	93.43	94.44	94.44
DDoS-ICMP	99.09	95.05	93.85	94.44	94.45
SQL injection	99.07	94.92	93.72	94.32	94.32
DDoS-TCP	98.94	93.25	94.05	93.65	93.65
Password	99.08	94.55	94.33	94.44	94.44
DDoS-HTTP	99.24	95.36	95.36	95.36	95.36
Uploading	98.99	93.97	94.02	94.00	94.00
Backdoor	99.15	94.15	95.86	94.99	95.00
XSS	99.10	94.54	95.06	94.80	94.80
Ransomware	99.21	95.48	94.81	95.14	95.14
Fingerprinting	99.28	95.32	96.08	95.69	95.69
Average	**99.11**	**94.67**	**94.66**	**94.66**	**94.66**
TSPHE (30%)					
Normal	99.09	93.99	95.01	94.50	94.50
DDoS-UDP	99.14	94.91	95.03	94.97	94.97
DDoS-ICMP	99.02	95.78	92.51	94.12	94.13
SQL injection	98.94	93.50	94.11	93.80	93.80
DDoS-TCP	98.98	92.86	95.08	93.96	93.96
Password	98.90	94.80	91.91	93.33	93.34
DDoS-HTTP	99.22	95.60	95.35	95.48	95.48
Uploading	98.87	92.62	93.39	93.00	93.00
Backdoor	99.03	92.66	95.80	94.20	94.22
XSS	99.11	94.52	93.97	94.24	94.24
Ransomware	99.21	95.89	95.03	95.46	95.46
Fingerprinting	99.23	95.41	95.28	95.35	95.35
Average	**99.06**	**94.38**	**94.37**	**94.37**	**94.37**

Table 3. Intrusion detection of OMHSA-IDPRGO model on edge-IIoT dataset.

whereas X_c signifies the food intake by the present adventitious root X_i^t, directing the remaining portion of the rhizomes, apart from the thick adventitious root, to mature extensively within the soil, X_{lbest} refers to the global best for the present fiber root. $\alpha_4 \in \sigma \times rand\,(0,1)$, σ signifies randomly generated integer of 0 or 1 made by utilizing the $rand\,i\,(0,1)$ function, if $\sigma = 1$, $\alpha_4 \in (0,1)$ specifies that X_c has initiated the area comprising the nutrition by developing extendedly within the soil, if $\sigma = 0$, $\alpha_3 = 0$ designates that X_c failed to detect the region consisting of the nutrition in Eq. (5) defines the stochastic searching method, with r3 and r4 denote randomly chosen integers.

The PRGO model creates an FF to achieve enhanced classification performance. It calculates a positive integer to depict the enhanced performance of the candidate solution. In this manuscript, the minimization of the classification error rate is deliberated as the FF, as shown in Eq. (33).

$$\begin{aligned} fitness\,(x_i) =& Classifier\ Error\ Rate\,(x_i) \\ =& \frac{no.\ of\ misclassified\ samples}{Total\ no.\ of\ samples} \times 100 \end{aligned} \tag{33}$$

Experimental outcome

The experimental study of the OMHSA-IDPRGO technique is examined under the Edge-IIoT dataset[38]. The dataset comprises 30,000 total records with 12 types of events, as summarised in Table 2. The complete no. of features is 63, but only 35 were chosen.

Figure 4 represents the classifier results of the OMHSA-IDPRGO technique on the Edge-IIoT dataset. Figure 4a, b portray the confusion matrices, which precisely classify all classes under a 70:30 ratio. Figure 4c exhibits the PR investigation, demonstrating maximum performance for each class. Finally, Fig. 4d exemplifies the ROC examination, signifying capable outcomes with greater values of ROC.

Table 3 and Fig. 5 denote the intrusion detection result of the OMHSA-IDPRGO model on the Edge-IIoT dataset. With 70%TRPHE, the OMHSA-IDPRGO model provides average $accu_y$, $prec_n$, $reca_l$, $F1_{Score}$, and $G_{Measure}$ of 99.11%, 94.67%, 94.66%, 94.66%, and 94.66%, respectively. Moreover, under 30%TSPHE, the

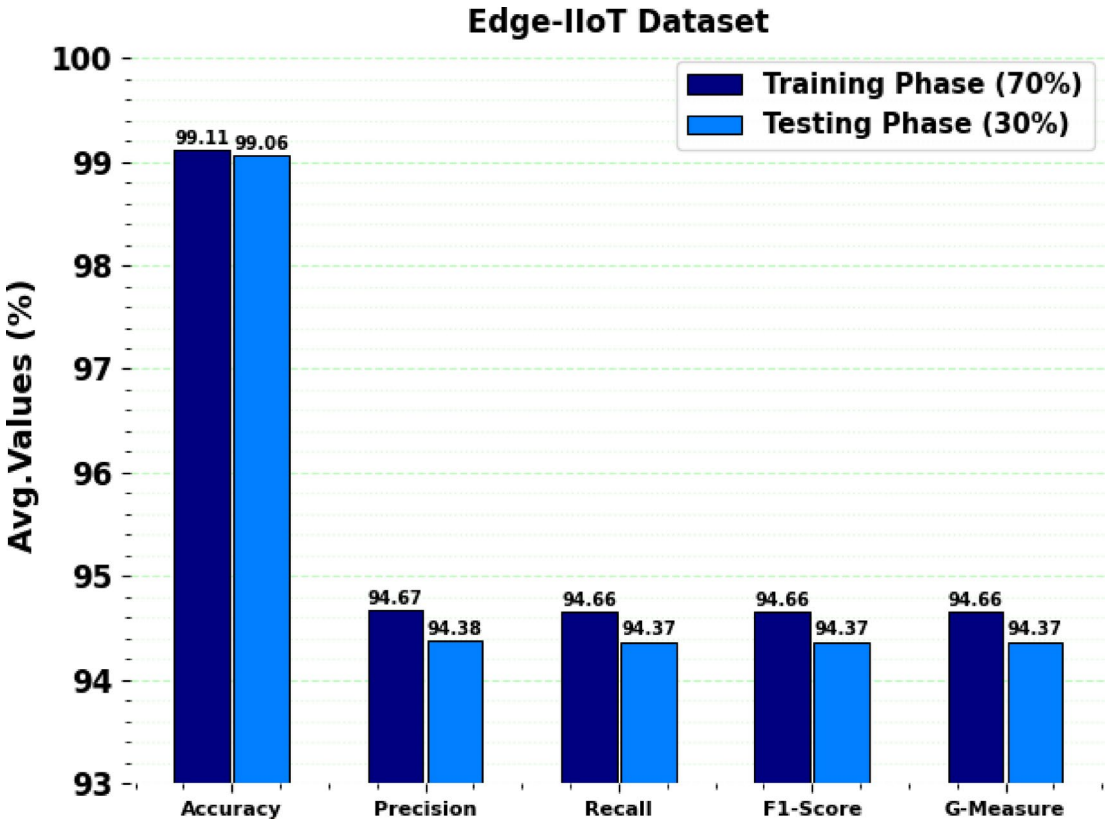

Fig. 5. Average values of the OMHSA-IDPRGO model on the Edge-IIoT dataset.

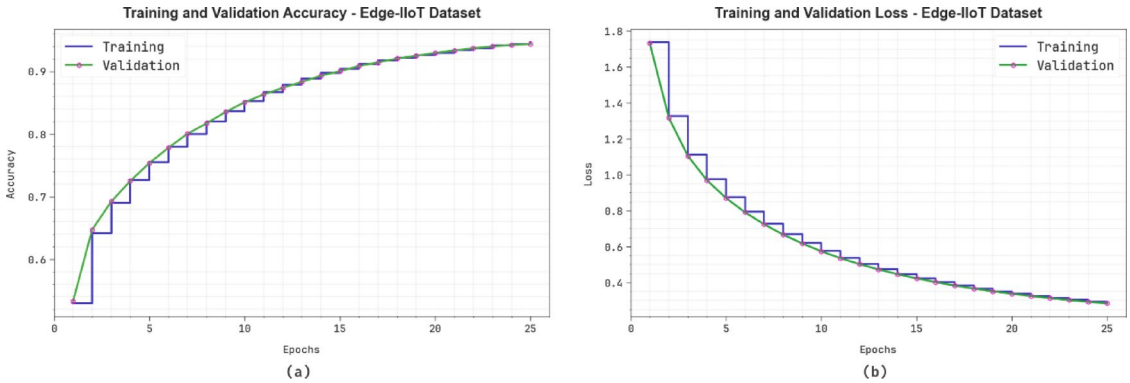

Fig. 6. (**a**) Accuracy curve and (**b**) loss curve on Edge-IIoT dataset.

OMHSA-IDPRGO method provides average $accu_y$, $prec_n$, $reca_l$, $F1_{Score}$, and $G_{Measure}$ of 99.06%, 94.38%, 94.37%, 94.37%, and 94.37%, correspondingly.

Figure 6 depicts the classifier outcomes of the OMHSA-IDPRGO model on the Edge-IIoT dataset. Figure 6a presents the accuracy study of the OMHSA-IDPRGO model. The figure indicates that the OMHSA-IDPRGO model yields increasing values across successive epochs. Furthermore, the rising validation over training demonstrates that the OMHSA-IDPRGO method proficiently learns from the test dataset. Finally, Fig. 6b illustrates the loss analysis. The findings indicate that the OMHSA-IDPRGO method achieves similar validation and training loss values.

Also, the OMHSA-IDPRGO technique is examined under the ToN-IoT dataset[39]. The dataset contains a total of 119,957 samples across nine classes. The complete details are presented in Table 4 below. The no. of attributes present in this dataset is 42, but only 31 were selected.

ToN-IoT dataset	
Class	**No. of instances**
"Normal"	78,369
"MiTM"	336
"DoS"	5440
"DDoS"	5987
"Password"	6016
"Injection"	5867
"XSS"	5951
"Ransomware"	5976
"Backdoor"	6015
Total instances	**119,957**

Table 4. Details of ToN-IoT dataset.

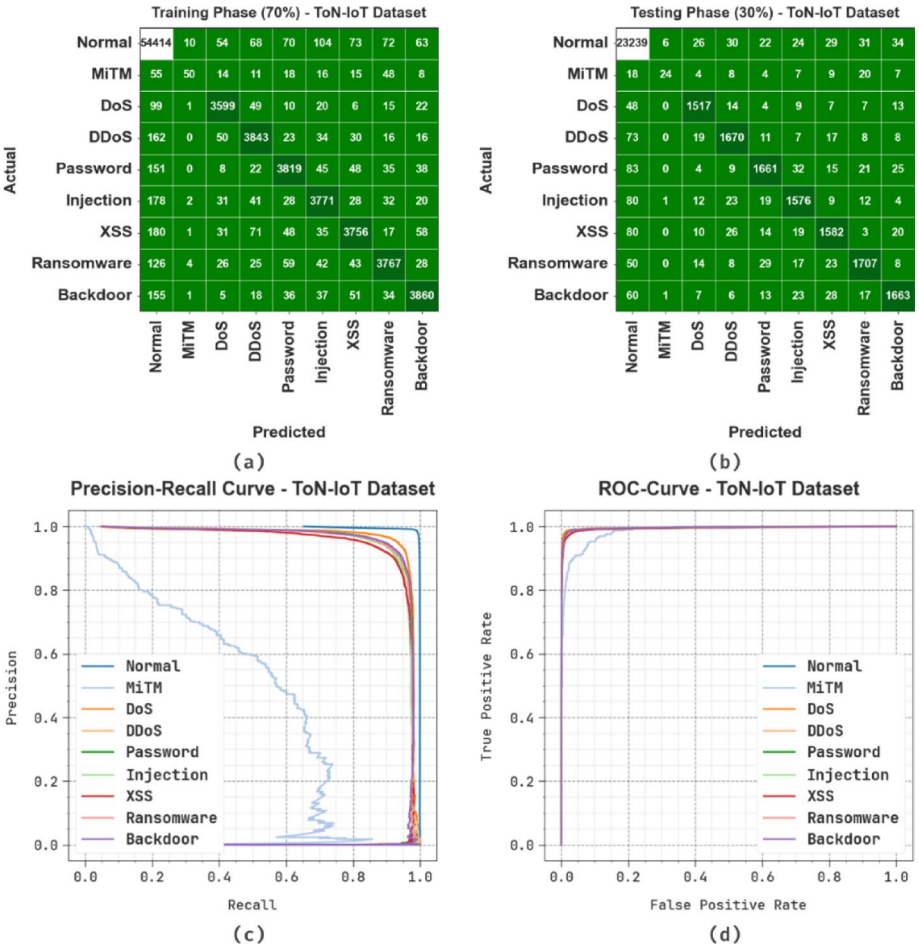

Fig. 7. ToN-IoT dataset (**a**–**b**) confusion matrices, (**c**) curve of PR, and (**d**) curve of ROC.

Figure 7 displays the classifier results of the OMHSA-IDPRGO model on the ToN-IoT dataset. Figure 7a, b illustrate the confusion matrices with precise classification across all classes under a 70:30 ratio. Figure 7c illustrates the PR investigation, which specifies the maximum performance for all class labels. Lastly, Fig. 7d exemplifies the ROC study, representing skilful outcomes with increased ROC values.

Table 5 and Fig. 8 signify the intrusion detection result of the OMHSA-IDPRGO technique on the ToN-IoT dataset. Under 70%TRPHE, the OMHSA-IDPRGO model gives average $accu_y$, $prec_n$, $reca_l$, $F1_{Score}$, and $G_{Measure}$ of 99.18%, 91.34%, 84.72%, 86.47%, and 87.18%, respectively. Furthermore, depending on 30%TSPHE,

Class labels	$Accu_y$	$\Pr ec_n$	$Reca_l$	$F1_{Score}$	$G_{Measure}$
TRPHE (70%)					
Normal	98.07	98.01	99.06	98.53	98.53
MiTM	99.76	72.46	21.28	32.89	39.27
DoS	99.47	94.26	94.19	94.23	94.23
DDoS	99.24	92.65	92.07	92.36	92.36
Password	99.24	92.90	91.67	92.28	92.28
Injection	99.17	91.89	91.29	91.58	91.59
XSS	99.12	92.74	89.49	91.09	91.10
Ransomware	99.26	93.33	91.43	92.37	92.38
Backdoor	99.30	93.85	91.97	92.90	92.90
Average	**99.18**	**91.34**	**84.72**	**86.47**	**87.18**
TSPHE (30%)					
Normal	98.07	97.93	99.14	98.53	98.53
MiTM	99.76	75.00	23.76	36.09	42.22
DoS	99.45	94.05	93.70	93.87	93.87
DDoS	99.26	93.09	92.11	92.60	92.60
Password	99.15	93.47	89.78	91.59	91.61
Injection	99.17	91.95	90.78	91.36	91.36
XSS	99.14	92.03	90.19	91.10	91.11
Ransomware	99.26	93.48	91.97	92.72	92.72
Backdoor	99.24	93.32	91.47	92.39	92.39
Average	**99.17**	**91.59**	**84.77**	**86.70**	**87.38**

Table 5. Intrusion detection outcome of the OMHSA-IDPRGO model on the ToN-IoT dataset.

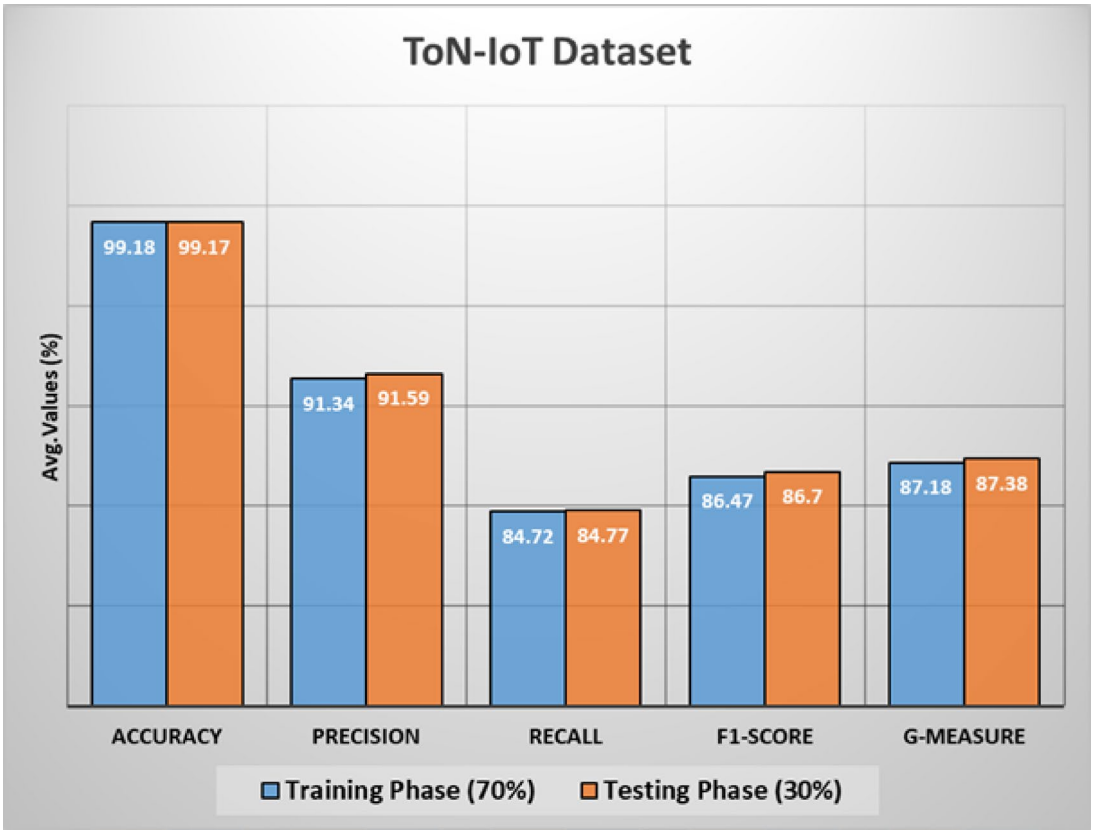

Fig. 8. Average values of the OMHSA-IDPRGO model on the ToN-IoT dataset.

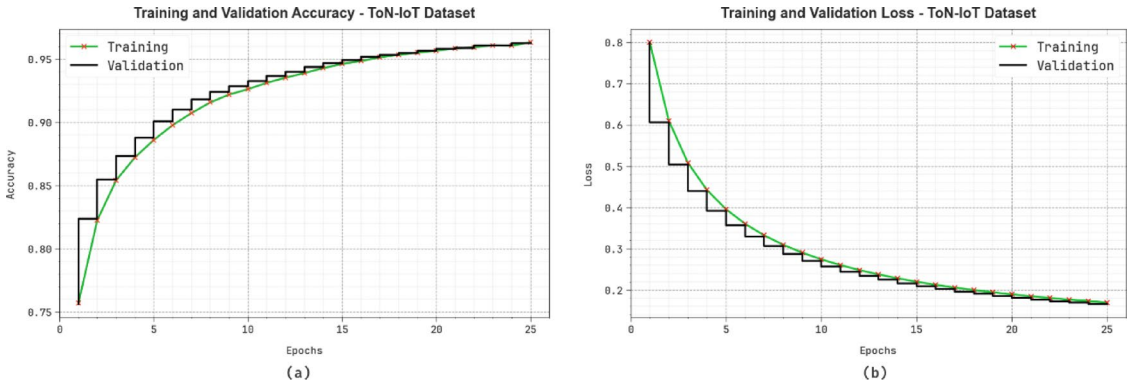

Fig. 9. (**a**) Accuracy curve and (**b**) loss curve on ToN-IoT dataset.

Technique	$Accu_y$	Prec_n	Reca_l	$F1_{Score}$
Edge-IIoT dataset				
LSTM-CSAE	96.35	89.58	93.79	93.55
MhSaBiGRNN	98.19	91.63	93.21	89.67
FL	95.32	94.00	93.62	89.92
EECA-LSTM	91.35	90.47	92.98	93.90
LSTM-KPCA	95.81	89.02	93.23	93.04
ML-PCC and IF	98.60	90.95	92.53	89.09
Shallow ANN	94.76	93.81	92.97	89.34
Baseline DNN	98.65	93.75	92.18	89.49
DAE-LSTM	91.56	93.87	90.05	91.58
XCT-DF	89.04	90.59	92.02	93.61
OMHSA-IDPRGO	99.11	94.67	94.66	94.66
ToN-IoT dataset				
MNBD	90.56	90.76	83.71	82.34
NFA	98.26	89.83	82.88	80.69
GNN	90.46	91.36	84.31	82.84
CNN method	90.05	90.24	83.20	81.60
DNN algorithm	97.59	89.13	82.36	80.08
LSTM	89.76	90.59	83.64	82.19
Decision tree	98.15	89.03	80.75	80.29
kNN algorithm	97.18	90.03	80.76	84.44
PCA model	89.23	90.01	80.14	80.54
Naïve Bayes	96.29	89.92	82.53	85.11
OMHSA-IDPRGO	99.18	91.34	84.72	86.47

Table 6. Comparative analysis of OMHSA-IDPRGO model on Edge-IIoT and ToN-IoT datasets[22,23,25,26,40–43].

the OMHSA-IDPRGO model provides an average $accu_y$, $prec_n$, $reca_l$, $F1_{Score}$, and $G_{Measure}$ of 99.17%, 91.59%, 84.77%, 86.70%, and 87.38%, respectively.

Figure 9 establishes the classifier outcomes of the OMHSA-IDPRGO method on the ToN-IoT dataset. Figure 9a determines the accuracy of the OMHSA-IDPRGO method. The figure indicates that the OMHSA-IDPRGO method yields increasing values over successive epochs. Additionally, the growing validation over training exhibits that the proposed approach learns efficiently on the testing dataset. Lastly, Fig. 9b illustrates the loss examination of the OMHSA-IDPRGO model. The outcomes denote that the OMHSA-IDPRGO model accomplishes close values of training and validation loss.

Table 6 presents a comparative study of the OMHSA-IDPRGO model on Edge-IIoT and ToN-IoT datasets, evaluating various measures against existing techniques[22,23,25,26,40–43].

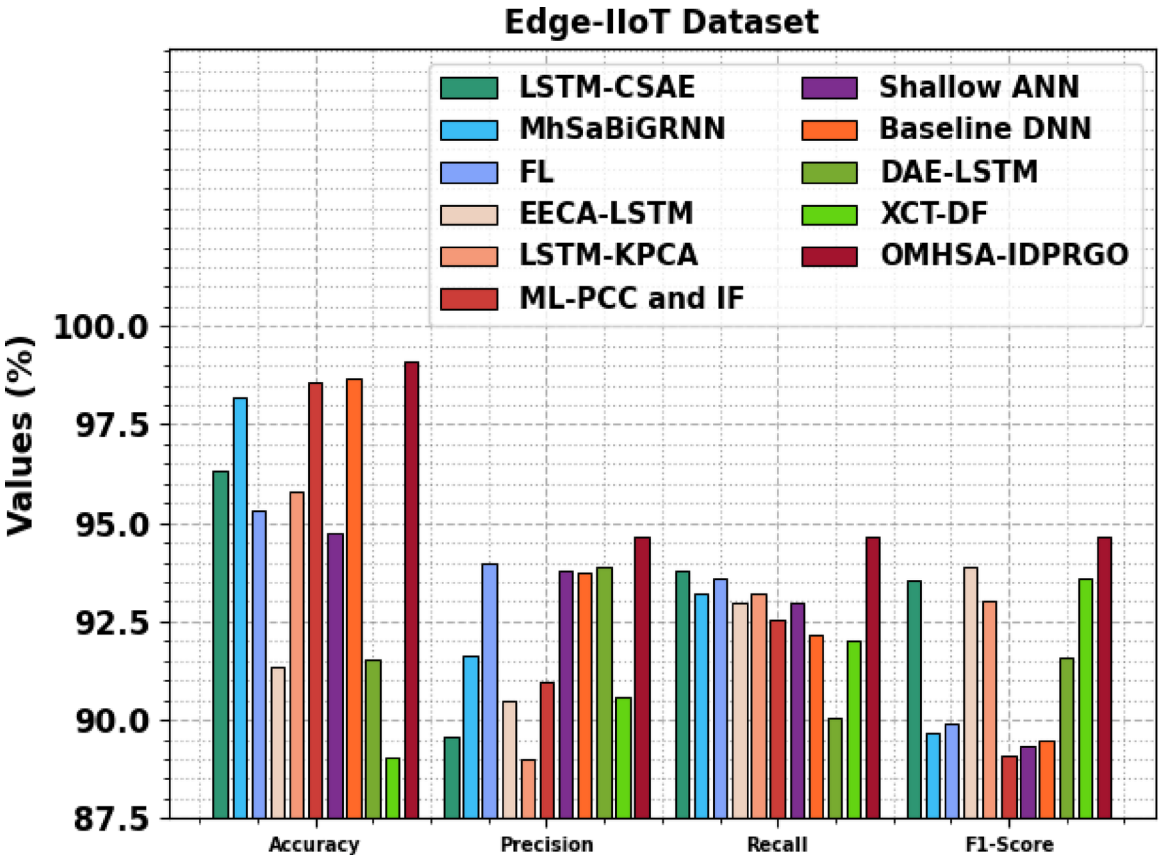

Fig. 10. Comparative analysis of the OMHSA-IDPRGO model on the edge-IIoT dataset.

Figure 10 shows the comparison findings of the OMHSA-IDPRGO model on the Edge-IIoT dataset. The results highlighted that the LSTM-CSAE, MhSaBiGRNN, FL, EECA-LSTM, LSTM-KPCA, ML-PCC, IF, Shallow ANN, Baseline DNN, DAE-LSTM, and XCT-DF methods performed the worst. Additionally, the OMHSA-IDPRGO model demonstrated enriched performance with maximum $accu_y, prec_n$, $reca_l$, and $F1_{score}$ of 99.11%, 94.67%, 94.66%, and 94.66%, respectively.

The comparative exploration of the OMHSA-IDPRGO approach on the ToN-IoT dataset with current approaches is displayed in Fig. 11. The OMHSA-IDPRGO approach achieves the highest performance, with $accu_y, prec_n$, $reca_l$, and $F1_{score}$ of 99.18%, 91.34%, 84.72%, and 86.47%, correspondingly. Whereas the present models, namely MNBD, NFA, GNN, CNN, DNN, LSTM, DT, kNN, PCA, and NB, obtain lesser values across diverse metrics.

Conclusion

This study designs and develops an OMHSA-IDPRGO model to advance IoT security. The primary focus of this novel is improving the automatic cyberattack detection in an IoT environment by employing advanced techniques. To achieve this, the OMHSA-IDPRGO method executes a mean normalization method to transform the raw data into a structured format. Following this, the COA approach is employed to select the optimal feature subset, thereby identifying the most relevant features from a dataset. For the cybersecurity detection process, the OMHSA-IDPRGO method uses a hybrid model named CNN-BiGRU-MHSAM. Finally, the hyperparameter selection is implemented by the PRGO approach to enhance the classification performance. The experimentation of the OMHSA-IDPRGO model is examined under Edge-IIoT and ToN-IoT datasets. The comparison study of the OMHSA-IDPRGO model yielded superior accuracy values of 99.11 and 99.18% compared to existing techniques on the dual datasets. The limitations of the OMHSA-IDPRGO model comprise the use of static datasets, which may not fully capture the real-time dynamics of evolving network environments. The model also faces scalability threats when deployed in large-scale or heterogeneous networks. Additionally, the absence of real-world deployment restricts practical validation. Environmental factors, such as varying mobility patterns or interference, were not extensively considered. The framework may require frequent updates to keep pace with newly emerging threats. Future work may explore real-time adaptive systems with continuous learning capabilities. Expanding the model to handle multi-domain and cross-platform data sources will also improve its robustness and applicability.

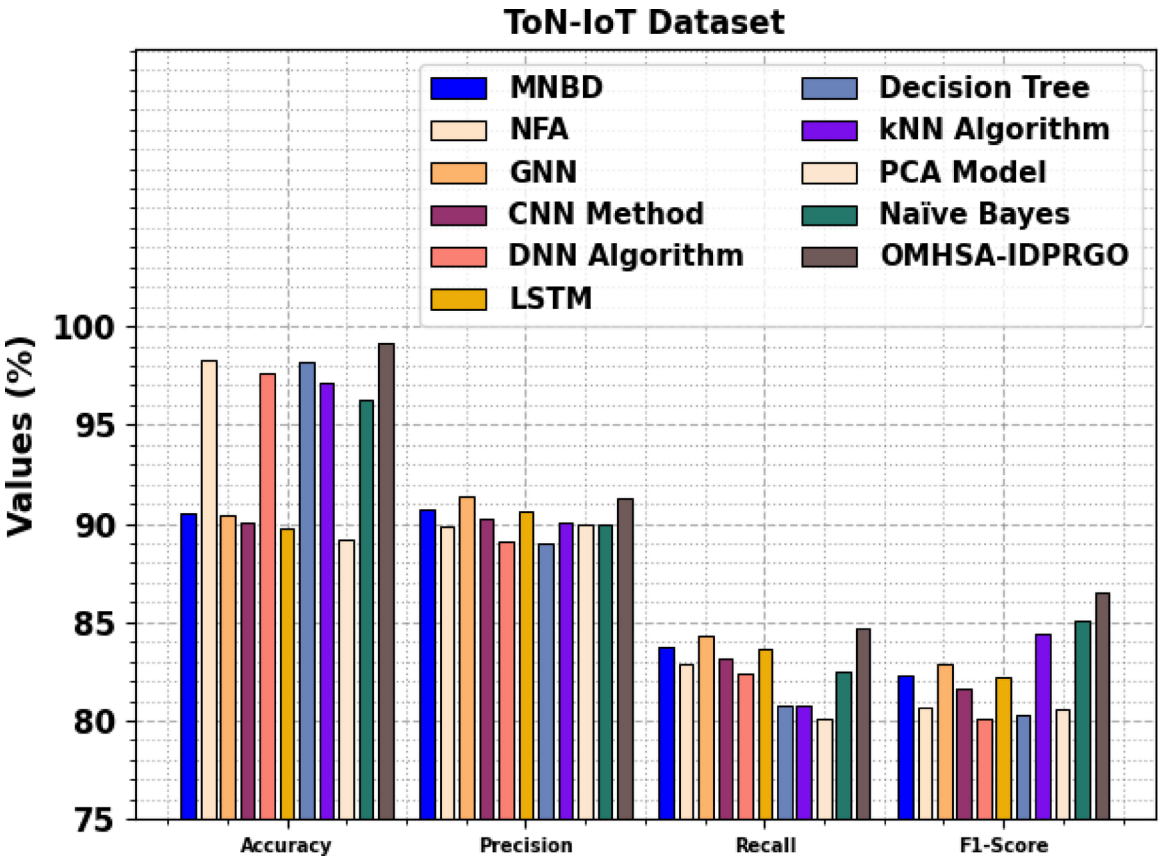

Fig. 11. Comparative analysis of the OMHSA-IDPRGO model on the ToN-IoT dataset.

Data availability

The data that support the findings of this study are openly available in the Kaggle repository at https://www.kaggle.com/datasets/mohamedamineferrag/edgeiiotset-cyber-security-dataset-of-iot-iiot, https://www.kaggle.com/datasets/dhoogla/cictoniot reference numbers[38,39].

Received: 13 May 2025; Accepted: 12 August 2025
Published online: 01 October 2025

References

1. Lee, I. Internet of things (IoT) cybersecurity: Literature review and IoT cyber risk management. *Future Internet* **12**(9), 157 (2020).
2. Kuzlu, M., Fair, C. & Guler, O. Role of artificial intelligence in the internet of things (IoT) cybersecurity. *Discov. Internet Things* **1**(1), 7 (2021).
3. Andrade, R. O., Yoo, S. G., Tello-Oquendo, L. & Ortiz-Garcés, I. A comprehensive study of the IoT cybersecurity in smart cities. *IEEE Access* **8**, 228922–228941 (2020).
4. Raimundo, R. J. & Rosário, A. T. Cybersecurity in the internet of things in industrial management. *Appl. Sci.* **12**(3), 1598 (2022).
5. AlSalem, T. S., Almaiah, M. A. & Lutfi, A. Cybersecurity risk analysis in the IoT: A systematic review. *Electronics* **12**(18), 3958 (2023).
6. Altulaihan, E., Almaiah, M. A. & Aljughaiman, A. Cybersecurity threats, countermeasures and mitigation techniques on the IoT: Future research directions. *Electronics* **11**(20), 3330 (2022).
7. Jun, Y., Craig, A., Shafik, W. & Sharif, L. Artificial intelligence application in cybersecurity and cyberdefense. *Wirel. Commun. Mob. Comput.* **2021**(1), 3329581 (2021).
8. Abdullahi, M. et al. Detecting cybersecurity attacks in internet of things using artificial intelligence methods: A systematic literature review. *Electronics* **11**(2), 198 (2022).
9. Zeadally, S., Adi, E., Baig, Z. & Khan, I. A. Harnessing artificial intelligence capabilities to improve cybersecurity. *IEEE Access* **8**, 23817–23837 (2020).
10. Abdalla Musa, A.I. & Al-Hagery, M.A., Integrating machine learning with two-person intuitionistic neutrosophic soft games for cyberthreat detection in blockchain environment. *Int. J. Neutros. Sci. (IJNS)*, **25**(2) (2025).
11. Alzahrani, A., An optimized approach to deep learning for botnet detection and classification for cybersecurity in internet of things environment. *Comput. Mater. Continua* **80**(2) (2024).
12. Sekhar, J.C., Priyanka, R., Nanda, A.K., Josephson, P.J., Ebinezer, M.J.D. & Devi, T.K., Stochastic gradient boosted distributed decision trees security approach for detecting cyber anomalies and classifying multiclass cyber-attacks. *Comput. Secur.*, p. 104320. (2025).
13. Alotaibi, J. A hybrid software-defined networking approach for enhancing IoT cybersecurity with deep learning and blockchain in smart cities. *Peer-to-Peer Netw. Appl.* **18**(3), 123 (2025).
14. Zidi, K., Abdellafou, K. B., Aljuhani, A., Taouali, O. & Harkat, M. F. Novel intrusion detection system based on a downsized kernel method for cybersecurity in smart agriculture. *Eng. Appl. Artif. Intell.* **133**, 108579 (2024).

15. Sadu, V.B., Abhishek, K., Al-Omari, O.M., Nallola, S.R., Sharma, R.K. & Khan, M.S., Enhancement of cyber security in IoT based on ant colony optimized artificial neural adaptive tensor flow. *Netw. Comput. Neural Syst.*, pp.1–17. (2024).
16. Aljebreen, M. et al. Enhancing DDoS attack detection using snake optimizer with ensemble learning on internet of things environment. *IEEE Access* **11**, 104745–104753 (2023).
17. Kumar, P. et al. Digital twin-driven SDN for smart grid: A deep learning integrated blockchain for cybersecurity. *Sol. Energy* **263**, 111921 (2023).
18. Lamir, I.M., Gital, A.Y.U., Ibrahim, K.M., Lawal, M.A., Danlami, M. and Yakubu, I.Z., Improved cybersecurity framework based on truckingc heuristics algorithm for detecting malicious devices in fog computing and internet of things (IoT) environments. In: 2023 *IEEE Fifth International Conference on Advances in Electronics, Computers and Communications* (ICAECC) pp. 1–9. IEEE. (2023).
19. Saheed, Y.K., Misra, S. and Chockalingam, S., Autoencoder via DCNN and LSTM models for intrusion detection in industrial control systems of critical infrastructures. In: *2023 IEEE/ACM 4th International Workshop on Engineering and Cybersecurity of Critical Systems (EnCyCriS)* pp. 9–16. IEEE. (2023).
20. Thayalan, S., Radhakrishnan, N., Ramana, T.V., Devarajan, G.G., Karuppiah, M. and Al Dabel, M.M., Real-time threat detection and AI-driven predictive security for consumer applications. *IEEE Trans. Consumer Electron*. (2025).
21. Saheed, Y. K., Abdulganiyu, O. H. & Ait Tchakoucht, T. Modified genetic algorithm and fine-tuned long short-term memory network for intrusion detection in the internet of things networks with edge capabilities. *Appl. Soft Comput.* **155**, 111434 (2024).
22. Kumar, P., Jolfaei, A. & Islam, A. N. An enhanced deep-learning empowered threat-hunting framework for software-defined internet of things. *Comput. Secur.* **148**, 104109 (2025).
23. Paul, S., Mitra, A., Shambhavi, S., Ghosh, S. & Bandyopadhyay, A. Holistic cyber security framework for deep web using federated learning in healthcare and distributed computing. *Proc. Comput. Sci.* **258**, 2405–2414 (2025).
24. Kathole, A. B. et al. Enhanced security mechanism in vehicular networks using ensemble machine learning to detect malicious activity in VANETs. *J. Discret. Math. Sci. Cryptogr.* **27**(7), 2005–2014 (2024).
25. Amer, E., Al-rimy, B. A. S. & El-Sappagh, S. Strengthening ICS defense: Modbus-NFA behavior model for enhanced anomaly detection. *J. Inf. Secur.* **89**, 103990 (2025).
26. Sardar, T. H. et al. Enhancing security in MANETs with deep learning-based intrusion detection. *Proc. Comput. Sci.* **259**, 120–129 (2025).
27. Kathole, A. B. et al. A novel approach to IoT security for intrusion detection system using ensemble network and heuristic-assisted feature fusion. *J. Discret. Math. Sci. Cryptogr.* **27**(7), 2207–2217 (2024).
28. Saheed, Y.K. and Chukwuere, J.E., 2025. *CPS-IIoT-P2 attention: Explainable privacy-preserving with scaled dot-product attention in cyber physical system-industrial IoT network. IEEE Access*.
29. Kathole, A.B., Vhatkar, K.N., Goyal, A., Kaushik, S., Mirge, A.S., Jain, P., Soliman, M.S. and Islam, M.T., 2024. *Secure federated cloud storage protection strategy using hybrid heuristic attribute-based encryption with permissioned blockchain. IEEE Access*.
30. Saheed, Y. K., Omole, A. I. & Sabit, M. O. GA-mADAM-IIoT: A new lightweight threats detection in the industrial IoT via genetic algorithm with attention mechanism and LSTM on multivariate time series sensor data. *Sens. Int.* **6**, 100297 (2025).
31. Saheed, Y. K. & Chukwuere, J. E. Xaiensembletl-iov: A new explainable artificial intelligence ensemble transfer learning for zero-day botnet attack detection in the internet of vehicles. *Results Eng.* **24**, 103171 (2024).
32. Saheed, Y. K. & Misra, S. CPS-IoT-PPDNN: A new explainable privacy preserving DNN for resilient anomaly detection in cyber-physical systems-enabled IoT networks. *Chaos Solitons Fractals* **191**, 115939 (2025).
33. Ullah, I., Deng, X., Pei, X., Mushtaq, H. & Khan, Z. Securing internet of vehicles: A blockchain-based federated learning approach for enhanced intrusion detection. *Clust. Comput.* **28**(4), 256 (2025).
34. Kim, Y. S. et al. Investigating the impact of data normalization methods on predicting electricity consumption in a building using different artificial neural network models. *Sustain. Cities Soc.* **118**, 105570 (2025).
35. Chen, S., Huang, X., Jiang, Z. and Ma, M., *Reliability-Based Design Optimization of Spherical Roller Bearing for Wind Turbines. Available at SSRN 5222291*.
36. Mao, J., Zhao, J., Zhang, H. & Gu, B. A novel hybrid deep learning model for day-ahead wind power interval forecasting. *Sustainability* **17**(7), 3239 (2025).
37. Zhang, J., Yan, F. & Yang, J. Binary plant rhizome growth-based optimization algorithm: An efficient high-dimensional feature selection approach. *J. Big Data* **12**(1), 13 (2025).
38. https://www.kaggle.com/datasets/mohamedamineferrag/edgeiiotset-cyber-security-dataset-of-iot-iiot
39. https://www.kaggle.com/datasets/dhoogla/cictoniot
40. Bukhari, S. M. S. et al. Enhancing cybersecurity in Edge IIoT networks: An asynchronous federated learning approach with a deep hybrid detection model. *Internet Things* **27**, 101252 (2024).
41. Izuazu, U. U., Nwakanma, C. I., Kim, D. S. & Lee, J. M. Explainable and perturbation-resilient model for cyber-threat detection in industrial control systems networks. *Discov. Internet Things* **5**(1), 9 (2025).
42. Kalaria, R., Kayes, A. S. M., Rahayu, W., Pardede, E. & Salehi, A. IoTPredictor: A security framework for predicting IoT device behaviours and detecting malicious devices against cyber attacks. *Comput. Secur.* **146**, 104037 (2024).
43. Li, J., Chen, H., Shahizan, M. O. & Yusuf, L. M. Enhancing IoT security: A comparative study of feature reduction techniques for intrusion detection system. *Intell. Syst. Appl.* **23**, 200407 (2024).

Acknowledgments

Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2025R755), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia. Ongoing Research Funding program, (ORF-2025-537), King Saud University, Riyadh, Saudi Arabia. The authors extend their appreciation to Northern Border University, Saudi Arabia, for supporting this work through project number "NBU-CRP-2025-1564. The authors are thankful to the Deanship of Graduate Studies and Scientific Research at University of Bisha for supporting this work through the Fast-Track Research Support Program. The author would like to acknowledge the support of Prince Sultan University for paying the Article Processing Charges (APC) of this publication and for their support.

Author contributions

Mimouna Abdullah Alkhonaini: Conceptualization, methodology, validation, investigation, writing—original draft preparation, funding Sara Abdelwahab Ghorashi: Conceptualization, methodology, writing—original draft preparation, writing—review and editing Ghalib H. Alshammri: methodology, validation, writing—original draft preparation Saied Alshahrani: software, visualization, validation, data curation, writing—review and editing Shouki A. Ebad: Project administration, validation, original draft preparation, writing—review and editing Sami Saad Albouq: methodology, validation, Conceptualization, writing—review and editing Fahad Alzahrani: methodology, validation, original draft preparation. Menwa Alshammeri: validation, original draft preparation,

writing—review and editing.

Declarations

Competing interests
The authors declare no competing interests.

Additional information
Correspondence and requests for materials should be addressed to S.A.E.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.