Metaheuristic-Driven Feature Selection for IoT Intrusion Detection: A Hierarchical Arithmetic Optimization Approach

Jing GUO*, Dejun ZHU, Qing XU
School of Big Data, Chongqing College of Mobile Communication, Chongqing City, 401520, China
Chongqing Key Laboratory of Public Big Data Security Technology, Chongqing City, 401520, China

***Abstract*—The increasing sophistication of cyberattacks in Internet of Things (IoT) networks requires strong Intrusion Detection Systems (IDS) with optimal feature selection mechanisms. High-dimensional data, computational complexity, and suboptimal detection accuracy hinder conventional IDS mechanisms. To overcome these limitations, in this study, the Hierarchical Self-Adaptive Arithmetic Optimization Algorithm (HSAOA) is introduced as a new metaheuristic method for IDS feature selection. HSAOA combines a stochastic spiral exploration method, an adaptive hierarchical model of leaders and followers, and a differential mutation mechanism to improve exploration-exploitation balance, global search capability, and premature convergence. The NF-ToN-IoT dataset is used to test the model, wherein HSAOA undertakes the feature selection process, and classification accuracy is increased by utilizing Random Forest (RF). The experimental results indicate that the proposed HSAOA is better than other advanced approaches in accuracy, computational efficiency, and convergence speed. These results validate the proposed algorithm as a scalable and effective solution for enhancing cybersecurity in IoT environments by improving IDS performance and reducing feature selection complexity.**

***Keywords*—Intrusion detection; internet of things; feature selection; hierarchical arithmetic optimization; cybersecurity**

I. Introduction

A. Research Problem

The Internet of Things (IoT) is revolutionizing how devices communicate, enabling easy connection of all devices from various industries, from healthcare to smart cities and industrial automation [1]. While the high growth of connected devices opens new avenues in the IoT space, the vast increase in connected devices makes the IoT networks vulnerable to various cyberattacks [2]. These vulnerabilities indicate the need for strong Intrusion Detection Systems (IDSs) to monitor the network's traffic and detect attacks in real-time. Proper IDS are essential to secure IoT environments by detecting and preventing attacks and protecting the integrity and privacy of the sensitive data exchanged over IoT networks [3].

Conventional IDS models face challenges in IoT networks, particularly regarding computational efficiency and feature selection [4]. IoT networks generate abundant data, resulting in high-dimensional data that is challenging to process efficiently through conventional IDS mechanisms [5]. Most IDS mechanisms are marred by high false-positive rates and slow processing times, primarily due to inappropriate feature selection. These schemes either maintain irrelevant features or miss the most critical ones, resulting in low detection accuracy and considerable computational complexity. As such, the need for more adaptive and efficient IDS mechanisms to cope with the distinctive requirements of IoT security is ever-escalating [6].

B. Contribution

This study proposes the Hierarchical Self-Adaptive Arithmetic Optimization Algorithm (HSAOA), an innovative metaheuristic solution for feature selection in IDS. HSAOA incorporates an adaptive hierarchical model, differential mutations, and a stochastic spiral exploration method to reduce computational complexity while enhancing IDS accuracy.

Through the utilization of such novel features, the capability of the proposed method to identify the relevant features is significantly increased, along with a consequent reduction in the dataset's dimensionality. The technique is predicted to be better than conventional methods in accuracy, computation complexity, and performance, making the algorithm an effective instrument for IoT network security. This work primarily contributes to:

- Presenting HSAOA, an adaptive metaheuristic that improves feature selection for intrusion detection on the IoT by combining spiral-guided random walks, hierarchical leader-follower organization, and rank-based differential mutations for enhanced exploration, exploitation, and solution refinement.
- Applying HSAOA on the NF-ToN-IoT dataset and evaluating it with other state-of-the-art methods, demonstrating improvements over detection efficiency, feature salience, and computation efficiency.
- Performing thorough simulations to test the efficiency of HSAOA at providing high intrusion accuracy, fast convergence, and feature dimensionality reduction.

This study addresses the following research questions:

- How can a metaheuristic algorithm be designed to perform adaptive and efficient feature selection for IDS in IoT environments?
- Does the proposed method improve detection accuracy, convergence speed, and computational efficiency

compared to existing metaheuristic-based IDS approaches?

- Can the proposed solution scale effectively and remain robust in detecting a wide variety of attacks across real-world IoT datasets?

II. Related Work

Davahli, et al. [7] introduced a hybrid method, GA–GWO, using a Genetic Algorithm (GA) and Grey Wolf Optimizer (GWO) for IoT network intrusion detection. The approach aims to reduce the dimensionality of a large traffic volume by extracting the most relevant traffic features. The computational efficiency is enhanced using the hybrid method, and intrusions are detected efficiently with high accuracy and low false alarm rates.

Shahapure and Punitha [8] proposed the Water Moth Search Algorithm (WMSA) to train the Deep Recurrent Neural Network (Deep RNN) to identify intrusions in IoT networks. WMSA algorithm combines the Water Wave Optimization (WWO) and the Moth Search Optimization (MSO), which is tailored using a Wrapper approach for the feature selection process.

Gangula and V [9] suggested an Enhanced Flower Pollination Algorithm (EFPA) integrated with an ensemble classification method for IoT network intrusion detection. The algorithm applies EFPA for the best possible feature extraction from UNSW-NB15 and NSL-KDD datasets, followed by classification under Random Forest (RF), Decision Tree (DT), and Support Vector Machine (SVM) classifiers.

Alghanam, et al. [10] suggested a modified form of the Pigeon-Inspired Optimization (PIO) algorithm, augmented with a Local Search Algorithm (LS-PIO), to enhance Network Intrusion Detection Systems (NIDS) in IoT networks. The proposed model utilizes the ensemble learning method and several one-class classifiers to identify intrusions more efficiently.

Saheed, et al. [11] proposed the IoT-Defender model, combining the Modified Genetic Algorithm (MGA) and the Long Short-Term Memory (LSTM) network for IoT network intrusion detection. MGA performs feature selection, optimizes the LSTM parameters, and improves the model's performance in detecting attacks.

Sharma, et al. [12] developed the Multi-Objective Prairie Dog Optimization (MPDA) algorithm to select features in IoT NIDS. The algorithm uses an archive, grid, and non-dominance to enhance diversity and manage several objectives efficiently. On NSL-KDD, CIC-IDS2017, and IoTID20 datasets, the method attained fewer features, better accuracy, and reduced false alarm rates.

Ogunseyi and Thiyagarajan [13] presented a hybrid DL-based IDS for IoT networks, employing statistical methods and a metaheuristic algorithm for feature selection. The model employs LSTM networks and is trained using public datasets such as NF-BoT-IoT-v2 and IoTID20. The model attained excellent detection accuracy of 98.4% and 89.5% on the datasets.

Li, et al. [14] proposed the GA-tuned ensemble of deep transfer learning as a methodology for detecting intrusions in IoT networks. The system leverages pre-trained Convolutional Neural Networks (CNNs) and tuned GA to identify the optimum features. The optimum models are subsequently employed in a soft voting ensemble to improve the robustness of the detector. The framework attained 100% accuracy in 15 attack classes and outperformed other models, most notably in minority attacks like backdoors and ransomware.

Despite the extensive advancements in IDS for IoT networks, existing solutions are often subject to inefficiency in computational performance and feature choice, as highlighted in Table I. Most existing algorithms still depend on manual traffic features, resulting in increased complexity and diminished accuracy. Second, most metaheuristic-based IDS solutions do not effectively handle real-time discovery or handling of scarce resources in IoT settings.

Most solutions are still black-box, non-interpretable, and their practical deployment in cybersecurity is, as such, curtailed. Our research seeks to bridge these limitations by suggesting the HSAOA, aimed at maximizing the optimization of the choice of features, improving computational performance, and increasing the interpretable capability of IDS for IoT networks, considering accuracy and scalability in practical deployments.

TABLE I. An Overview of Related Works

Reference	Contribution	Dataset	Weakness
[7]	Hybridization of GA and GWO for dimensionality reduction	AWID	It is limited to wireless intrusion detection and lacks generalization across other types of IoT traffic
[8]	Combination of WMSA and Deep RNN for detecting malicious network activities	Custom	It requires intensive training and may not scale well for large IoT datasets
[9]	EFPA for feature selection and ensemble classification for attack detection	UNSW-NB15 and NSL-KDD	There is a risk of overfitting due to ensemble complexity, and the approach is dependent on large labeled datasets
[10]	LS-PIO for feature selection and ensemble learning	BoT-IoT, UNSW-NB15, and NSL-KDD	It incurs a high computational cost due to the ensemble and may not adapt well to real-time IoT data
[11]	MGA for feature selection and LSTM tuning in IoT-defender framework	BoT-IoT, UNSW-NB15, and N-BaIoT	While focused on resource-constrained devices, it may not perform well in large-scale networks
[12]	MPDA for multi-objective feature selection	NSL-KDD, CIC-IDS2017, and IoTID20	It struggles with real-time detection in dynamic IoT environments
[13]	Deep learning with metaheuristics for feature selection and LSTM-based IDS	NF-BoT-IoT-v2 and IoTID20	It has high complexity and lacks transparency in decision-making
[14]	GA-tuned ensemble of deep transfer learning for NetFlow-based intrusion detection	Custom	It may not generalize well beyond NetFlow datasets and is computationally expensive for real-time use

III. METHODOLOGY

A. Overview of Feature Selection for IDS

Feature selection is required to optimize IDS performance. Feature selection identifies the most informative features from the vast amount of data that IoT networks provide, eliminates redundant or non-informative characteristics, and simplifies computation considerably [15]. Feature selection is essential in conditions of scarcity when the capacity to process large amounts of data efficiently is of the utmost importance. As shown in Fig. 1, the IDS scans network traffic to identify anomalous behavior that may indicate potential security intrusions. The ultimate objective of the feature selection process is to reduce the dimensions so the most relevant features are used to identify the attack while maintaining high classification accuracy.

B. IDS System Architecture and Workflow

The overall architecture of the IDS mechanism, as shown in Fig. 2, involves feature selection followed by machine learning-based classification, including Random Forest. The process uses IoT traffic to detect network attacks, including data from NF-ToN-IoT datasets. The system has two stages: Training and Testing. The initial stage involves using the training set to generate the positions of the features and determine the optimum characteristics by testing them. The optimum classification features are selected to reduce data dimensions. The system identifies traffic as anomalous or normal based on classification features.

IDS identifies anomalous network traffic behavior. As a deviation from normality is detected, the system triggers an alarm so the network can react appropriately. IDS is routinely integrated with firewall subsystems to manage traffic in the network and bolster defenses against attacks. Feature selection is critical in filtering out only the most vital features to be processed, resulting in more accurate attack detection and fewer false alarms.

C. Arithmetic Optimization Algorithm

The Arithmetic Optimization Algorithm (AOA) is a new optimization method based on fundamental arithmetic operators, i.e., addition, subtraction, multiplication, and division [16]. AOA is designed to determine the optimal solution through repeated updates of candidate solution positions based on the operation. The basic phases of AOA, i.e., the initialization, exploration, and exploitation phases, are discussed in this section.

In the initiation process, AOA generates a preliminary population of candidate solutions. The position of each individual in the population is defined in the search space as follows in Eq. (1):

$$X_i^j = UB_j + rand \times (UB_j - LB_j), \quad i = 1, \ldots, N; \quad j = 1, \ldots, D \tag{1}$$

where, D is the spatial dimension of the problem, N is the population size, X_i^j is the position of the j^{th} dimension of the i^{th} individual, UB_j and LB_j are the upper and lower bounds of the j^{th} dimension, and $rand$ is a random value between 0 and 1.

After initialization, AOA proceeds iteratively, where it determines either the exploration stage or the exploitation stage based on the Mathematical Optimization Acceleration (MOA) process, as follows in Eq. (2):

$$MOA = Min + t \times \left(\frac{Max - Min}{T}\right) \tag{2}$$

where, T is the current iteration number, Max and Min are the upper and lower bounds of the objective function, and t represents the current iteration.

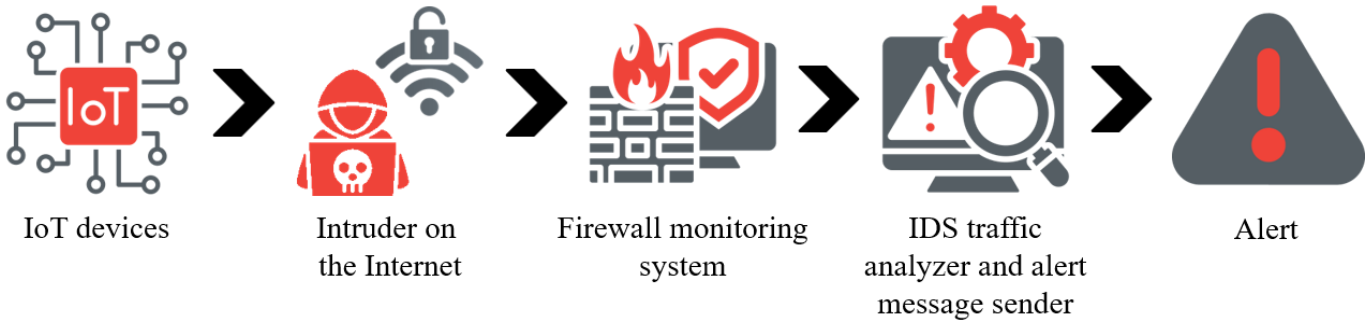

Fig. 1. Feature selection process in IDS.

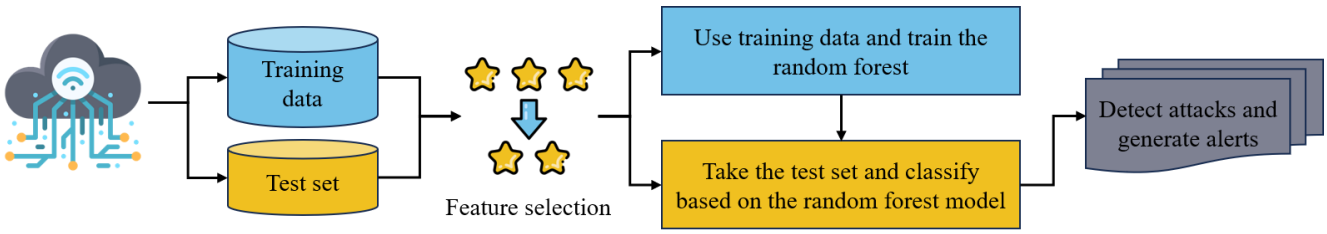

Fig. 2. Overview of the proposed IDS framework.

In the exploration period, AOA modifies individual positions through the application of multiplication or division operations. The update rule is as follows in Eq. (3):

$$X_i^j(t+1) = \begin{cases} X_i^j(t) \div (MOP + \epsilon) \times \left(UB_j - LB_j\right) + LB_j, & r_2 > 0.5 \\ X_i^j(t) \times \left(MOP \times \left(UB_j - LB_j\right) + LB_j\right), & r_2 \leq 0.5 \end{cases} \quad (3)$$

MOP is the modified optimization parameter, defined as in Eq. (4):

$$MOP = 1 - \left(\frac{t}{T}\right)^{\frac{1}{\alpha}} \quad (4)$$

r_2 is a stochastic value uniformly distributed between (0, 1), and ϵ is a small value to prevent division by zero.

In the exploitation stage, AOA updates the positions of individuals by applying addition or subtraction operators. The update rule for the exploitation phase is given by Eq. (5):

$$X_i^j(t+1) = \begin{cases} X_{best}^j - MOP \times \left(UB_j - LB_j\right) + LB_j, & r_3 > 0.5 \\ X_{best}^j + MOP \times \left(UB_j - LB_j\right) + LB_j, & r_3 \leq 0.5 \end{cases} \quad (5)$$

where, X_{best}^j is the position of the best solution found so far and r_3 is a random value uniformly distributed over the interval (0, 1).

As shown in Fig. 3, the AOA process is divided into three phases: initialization, exploration, and exploitation. The algorithm iterates through these phases until a stopping criterion (the maximum number of iterations) is attained, ultimately finding the optimum solution to the problem.

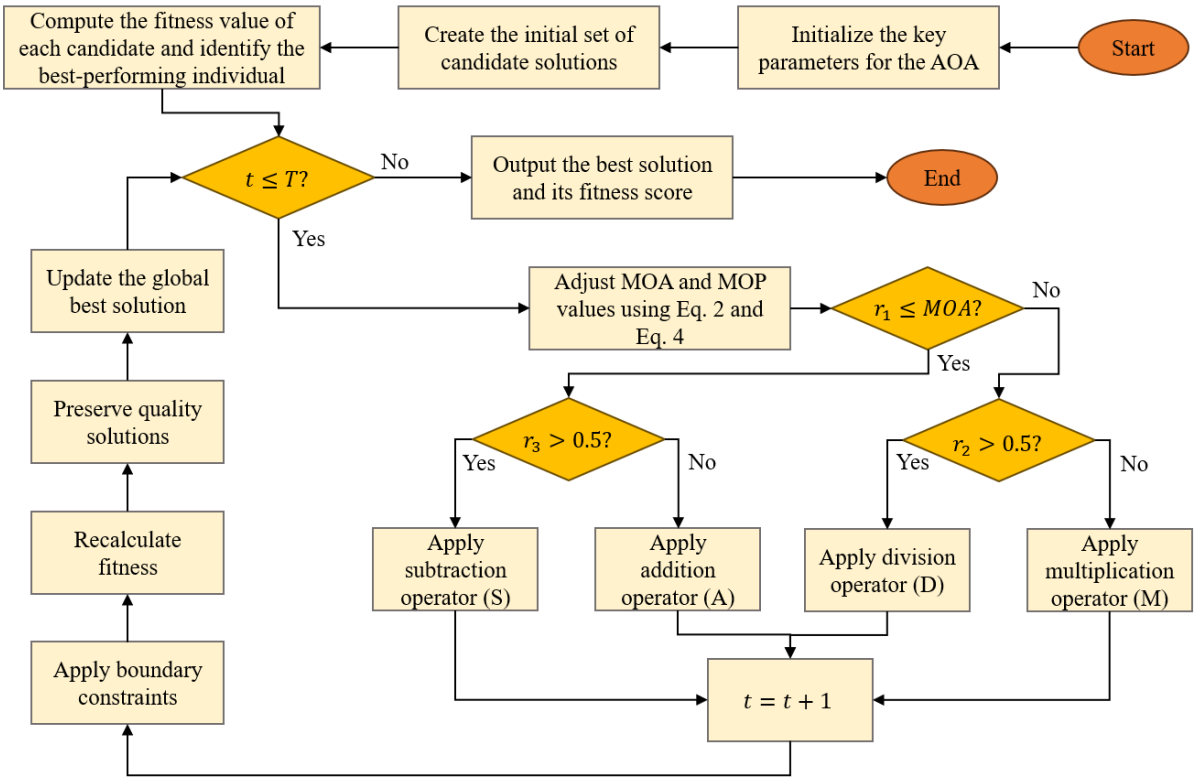

Fig. 3. Flowchart of AOA.

D. Enhancements in HSAOA

HSMAOA is an extension of the basic AOA able to overcome premature convergence and local optimum trapping, common in basic AOA. HSMAOA adds new mechanisms, such as a stochastic spiral exploration method, an adaptive hierarchical model of leaders and followers, and a differential mutation mechanism to enhance exploration and exploitation phases.

1) Stochastic spiral exploration method. During the exploration phase, AOA relies on multiplication and division to reposition individuals, potentially causing a loss of diversity following the initial population generation. To mitigate this issue, HSMAOA introduces a spiral-based random walk function as a replacement for conventional arithmetic operators, aiming to enhance global search capability. This mechanism leverages the Archimedean spiral to influence the step size of the random walk, taking advantage of its isometric growth properties, which have proven effective in diverse optimization applications. The Cartesian coordinates of the spiral are defined as Eq. (6):

$$x = cos(\theta) \times (r_0 + r_1 \times \theta) \quad (6)$$

$$y = sin(\theta) \times (r_0 + r_1 \times \theta)$$

$$\theta = \frac{1}{e} \times \left(\frac{n}{N}\right)$$

where, r_0 is the initial radius of the spiral (distance from the center), r_1 specifies the incremental growth of the radius per angular rotation, θ is the angle, n is the current iteration, and N is the maximum number of iterations.

In HSMAOA, the position update for an individual in the exploration phase is determined by Eq. (7):

$$X_i^d(n+1) = \begin{cases} X_{best}^d + sign(rand() - 0.5) \times cos(2\pi\theta) \times \\ \quad \left(\omega \times \left(X_a^d - X_\beta^d\right) + b \times 2\pi\theta\right), r_2 > 0.5 \\ X_{best}^d + sign(rand() - 0.5) \times sin(2\pi\theta) \times \\ \left(\omega \times \left(X_a^d - X_\beta^d\right) + b \times 2\pi\theta\right), \ otherwise \end{cases} \quad (7)$$

where, X_a^d and X_β^d represent the best and worst solutions for the d^{th} dimension, and ω is the step factor that is updated according to a Gaussian distribution with decreasing variance as in Eq. (8):

$$\omega = Gaussian\left(0, cos\left(\frac{\pi}{2} \times \frac{n}{N}\right)\right) \quad (8)$$

This spiral-guided mechanism enhances randomness in the search process and retains diversity after initialization, thereby improving the algorithm's ability to explore previously unvisited regions of the search space.

2) Adaptive hierarchical leader-follower structure. In HSMAOA, the adaptive hierarchical leader and follower scheme is employed to facilitate optimization processes by promoting cooperation and the exchange of information among individuals in the population. As shown in Fig. 4, the hierarchy is split into several levels. The individuals in the higher levels are assigned as leaders, and those in the lower levels are followers. The algorithm adapts the hierarchy through branching degree adjustment as follows in Eq. (9):

$$m = floor\left(\frac{1}{5} \times P - \left(\frac{1}{5} \times P - 2\right) \times \frac{n}{N}\right) \quad (9)$$

where, N, n, and P correspond to the maximum iteration limit, current iteration, and population size, respectively.

The hierarchical structure enables better interaction among individuals by maximizing the exchange of information between the leaders and followers. The leaders search the solution space independently using addition and subtraction to update positions, whereas the followers collaborate with the leaders and their parent individuals to improve solutions. This process helps overcome local optima and ensures better convergence in the exploration phase. Hierarchical adaptation and information exchange improve the algorithm's global search ability while maintaining local optimization.

3) Rank-Based differential mutation strategy. To enhance diversity and information exchange, HSMAOA incorporates a ranked selection-based differential mutation approach. This method prioritizes better individuals, allowing them to influence the next generation more significantly. The ranking of individuals is based on their fitness values, and the rank r_i for each individual X_i is used to calculate the mutation probability as in Eq. (10):

$$p_i = \frac{r_i}{P}, \ i = 1,2,\ldots,P \quad (10)$$

This ranked selection method prioritizes individuals with superior fitness, increasing their likelihood of influencing the mutation process. The mutation for the selected individuals is calculated as follows in Eq. (11):

$$V_i = X_i + \lambda \times (X_{r1} - X_{r2}) + (1 - \lambda) \times (X_{best} - X_{r3}) \quad (11)$$

where, X_{best} is the best individual, X_{r1}, X_{r2}, and X_{r3} are individuals selected based on their rank in the population, and λ is a control parameter set to 0.5 after extensive testing.

Finally, the position update for individual X_i is given by Eq. (12):

$$X_i = \begin{cases} V_i, & if\ fun(V_i) > fun(X_i) \\ X_i, & otherwise \end{cases} \quad (12)$$

where, $fun(V_i)$ and $fun(X_i)$ are the fitness values of the mutated individual and the current individual.

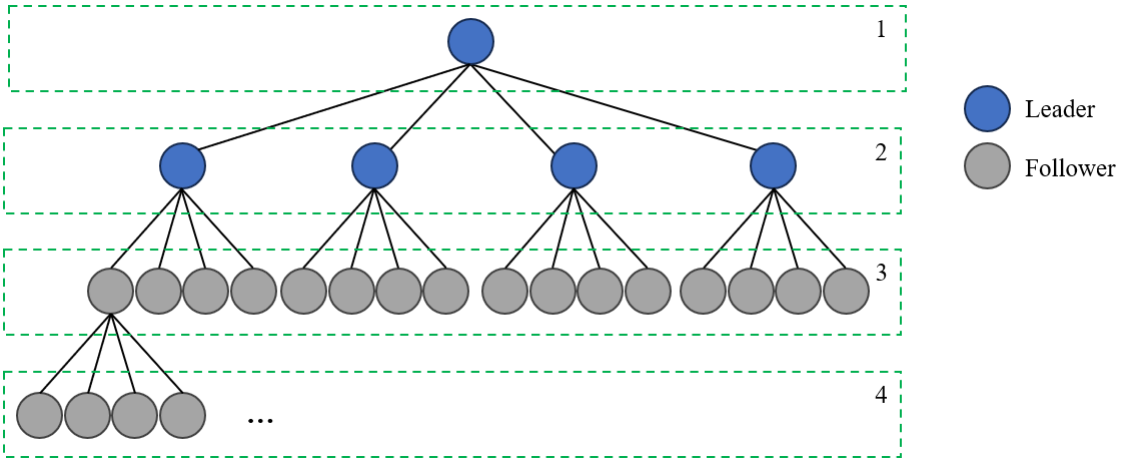

Fig. 4. Adaptive hierarchical structure of leaders and followers in HSMAOA.

E. Step-by-Step Implementation

As illustrated in Fig. 5, the architecture employs a systematic methodology comprising three primary phases: dataset preparation, feature extraction, and intrusion detection. At the heart of the method is HSMAOA, which enhances the feature extraction process to enhance the global search capability and the accuracy of detection of malicious behavior in IoT settings.

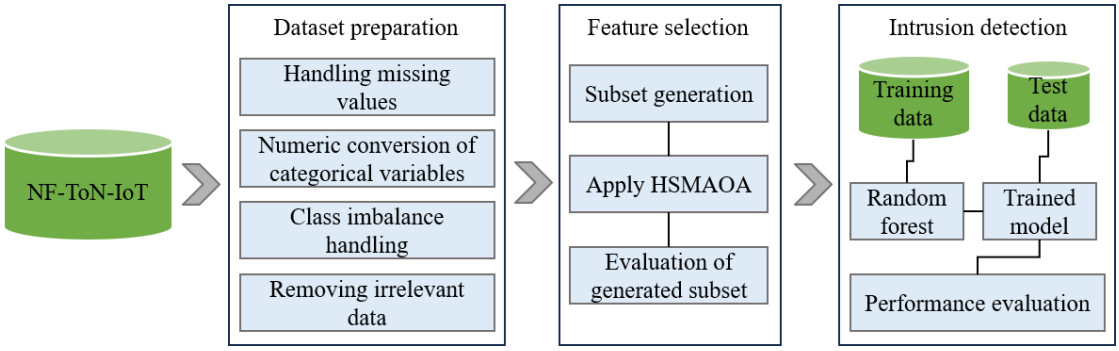

Fig. 5. Overview of the proposed IDS framework utilizing HSMAOA for feature selection.

HSMAOA leverages the advantages of a hierarchical structure and self-adaptive elements to mitigate the computational expense of selecting attributes from large datasets, such as NF-ToN-IoT. The IDS system becomes more accurate, reduces the dimensions of the data, and exhibits improved overall performance due to the altered optimization strategy. NF-ToN-IoT serves as training and test datasets, with pre-processed data used before attribute selection and classification.

1) Dataset preparation. The data preparation gives the dataset for the subsequent machine learning process and future analysis. In it, different inconsistencies within the data are corrected, including how the gap within the data, imbalance within the classes, and inconsistency can impact the outcome within the model negatively. The NF-ToN-IoT dataset goes through the preprocessing wherein the missing values are addressed, categorical variables are converted into numeric, the problem of class imbalance is addressed, and the irrelevant data is deleted.

Missing values in the dataset are imputed before passing the data to the HSMAOA. Imputation of the data is done in two phases: firstly, the most common value for each attribute is employed to replace the missing entries. Secondly, when required, imputed statistical values such as the mean or median value are used based on the attribute type so that the dataset is ready to be optimized.

NF-ToN-IoT contains categorical variables, and hence, the variables must be converted into a numeric form before being input into machine learning models. The categorical features are converted into a machine-learning input-compatible format using the technique of one-hot encoding.

A common challenge in IoT network datasets, such as NF-ToN-IoT, is the imbalance between classes (e.g., regular traffic vs. attack traffic). To address this issue, SMOTE (Synthetic Minority Over-sampling Technique) is applied, which generates synthetic data points for the minority class, thereby improving the classifier's ability to identify less frequent attack types.

Irrelevant attributes, such as timestamps, ports, and IP addresses, are removed from the dataset. These are generally not useful for identifying network intrusions within the context of an IDS and may result in overfitting or added complexity to the system. Eliminating irrelevant data aids in system performance and generalization.

2) Feature selection using HSMAOA. The second step in the process is the selection of the features, a step essential for the enhancement of the IDS' performance. The HSMAOA is utilized for the optimization of the subset of the best features for the task of intrusion detection. The hierarchical and adaptive approach of the algorithm ensures the selection of the best possible optimum features for training the model, striking a balance between exploration and exploitation on each optimization step. The choice of the features is made by optimizing the subset best describing the data using the HSMAOA.

HSMAOA employs the initial random subset generation method for generating various probable subsets. The subsets are then evaluated for the best subset of features that yields higher accuracy without an increase in the number of features. The goal is the identification of the best subset of features with the greatest accuracy and least computational cost.

Upon generation of the subset, it is assessed using the fitness function that balances the accuracy of the classifications and the number of selected features. The fitness function aims to make the selected features informative to detect intrusions without overfitting because there are too many features. Eq. (13) is the fitness evaluation of the feature subsets by HSMAOA.

$$Objective\ function = C_1 \times A + \frac{C_2}{NF} \tag{13}$$

where, A is the classification accuracy achieved by the subset, NF is the number of features in the subset, C_1 and C_2 are constants that balance the priority between accuracy and the number of features.

3) Intrusion detection using random forest. Once the feature selection process is completed, it is followed by

utilizing the chosen features at an intrusion detection level. The latter utilizes the RF algorithm, an efficient ensemble learning algorithm. The algorithm comprises several random subsets of data, on which decision trees have been generated. Each tree contributes one vote to the final classification outcome. The algorithm is best suited for classifying intrusions within networks because it can handle both the randomness of the data and the heterogeneity of the IoT dataset.

The chosen features are divided into the training data and the test data. The train data, which includes the selected characteristics and labels, is used for training the RF algorithm. The test data is utilized for testing the RF algorithm for determining the accuracy of the classifier. The labeled data trains the model throughout the process based on the construction of decision trees for determining the type of network traffic as benign or malicious based on the selection of the feature. Bootstrapping is invoked by the RF model, where different models are applied and fitted on various random data sets, resulting in a superior model. In the construction of a model, the performance can be estimated based on the calculation of the result on the test dataset. The true labels are compared with the labels predicted based on the model to determine the accuracy, precision, etc.

Several mitigation techniques have been implemented to stabilize the model, preventing it from overfitting. The process is validated by K-fold cross-validation, where the model is tested across data partitions. The algorithm minimizes bias during training. Regularization techniques, such as pruning decision trees and limiting tree depth, prevent overfitting by reducing complexity and simplifying the model. The ensemble method, which utilizes the RF algorithm, combines the outputs of multiple decision trees to enhance the stability and consistency of the model, thereby avoiding overfitting. The gains achieved by these countermeasure methods are measured through comparative analysis, with assistance from metrics such as accuracy, precision, and recall, to determine the value of the model.

IV. RESULTS

A performance analysis of the proposed HSAOA for feature selection in IDS is summarized in this section. The performance analysis is conducted using the NF-ToN-IoT dataset, where the data is divided into 80% for training and 20% for testing, as shown in Fig. 6. The simulation parameters used to evaluate the model are presented in Table II. The execution used Python on the Jupyter Notebook platform and the Pandas library. The hardware-software setup for experimentation consisted of an Intel Core i3-1115G4 processor at 2.95 GHz, 16 GB of RAM, 64-bit Windows 10, and a 256 GB SSD.

HSAOA effectively reduced the dimensionality of the NF-ToN-IoT dataset by selecting the 22 best features while preserving the accuracy of the classifier. The selected features are listed in Table III. The convergence behavior of HSAOA was compared with that of the baseline AOA in terms of fitness minimization over 20 iterations. As depicted in Fig. 7, it can be seen that HSAOA exhibits better convergence than AOA, with the significant change occurring at the 10th iteration. The objective function reaches an optimal and stable value, demonstrating improved global exploration capability and feature discrimination ability of HSAOA.

HSAOA-based feature engineering was designed for extracting discriminative features most informative for IDS classification performance. Features from have been chosen as they have high potential for characterizing abnormal activity. The feature set was extended for more discrimination between legitimate and malicious traffic, and hence reduced false positives. The effect of fitness value on classification accuracy is evident from Fig. 8. As the algorithm optimizes with a diminishing objective function, accuracy increases. The maximum accuracy obtained by the IDS system employing feature selection with HSAOA-converged was 99.94%.

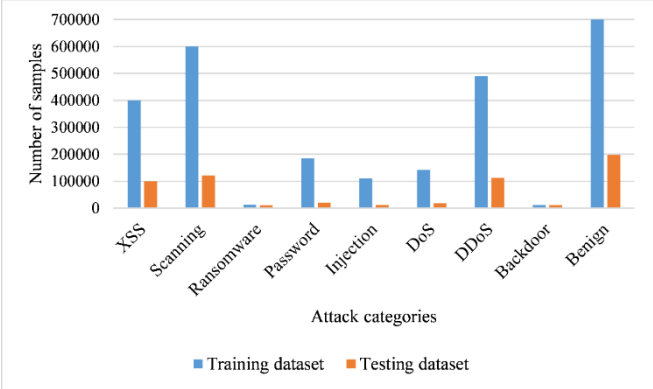

Fig. 6. Distribution of attack categories across the training and testing sets.

TABLE II. CONFIGURATION SETTINGS USED IN THE SIMULATION EXPERIMENTS

Parameter	Assigned value
Maximum iteration count	20
Population size	12
Total execution cycles	20
Weight factor C_1	0.6
Weight factor C_2	0.4

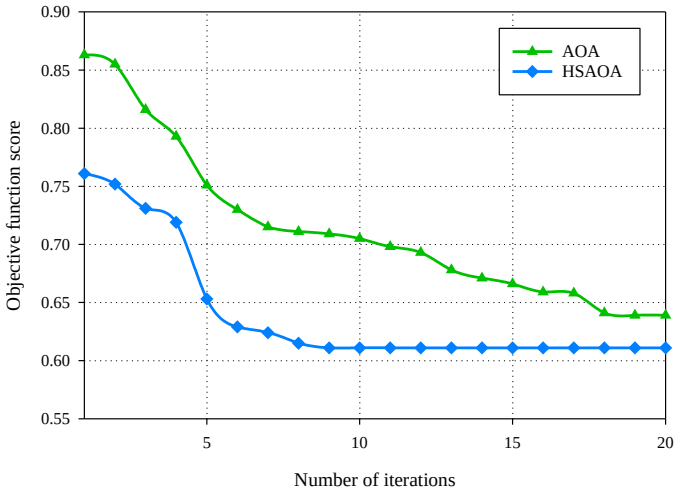

Fig. 7. Comparison of convergence behavior between the conventional AOA and the proposed HSAOA.

TABLE III. FINALIZED FEATURE SUBSET SELECTED USING THE PROPOSED HSAOA

Feature No.	Feature name	Feature No.	Feature name
F_1	Protocol type	F_{12}	Average throughput (Src to Dst)
F_2	Connection protocol	F_{13}	Average throughput (Dst to Src)
F_3	Packets received (inbound)	F_{14}	Packet Count ≤128 bytes
F_4	Packets sent (outbound)	F_{15}	Packets 128–256 bytes
F_5	TCP flag summary	F_{16}	Packets 256–512 bytes
F_6	Millisecond flow duration	F_{17}	Packets 512–1024 bytes
F_7	Outbound session duration	F_{18}	Packets 1024–1514 bytes
F_8	Minimum TTL	F_{19}	FTP command response code
F_9	Maximum TTL	F_{20}	DNS query category
F_{10}	Maximum IP packet length	F_{21}	DNS query identifier
F_{11}	Retransmitted byte count	F_{22}	IPv4 ICMP type

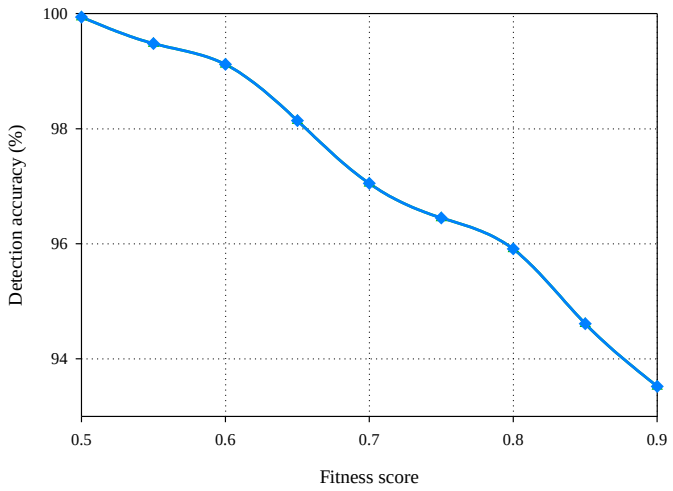

Fig. 8. Correlation between detection accuracy and fitness values during feature selection using the HSAOA.

After feature selection, the RF classification model was used for detection. The model's parameters are listed in Table IV. Performance metrics, detection rate, positive predictive value, sensitivity, and F1 score were determined using True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN) values, as outlined in Eq. (14) to Eq. (17). Table V reports comparative results with and without mitigation techniques (i.e., regularization, bootstrapping, and class balancing). Performance parameters improved substantially with mitigation to 99.94%, 99.87%, 99.78%, and 99.80%, respectively.

$$Detection\ rate = \frac{TN + TP}{FN + FP + TN + TP} \quad (14)$$

$$Positive\ predictive\ value = \frac{TP}{FP + TP} \quad (15)$$

$$Sensitivity = \frac{TP}{FN + TP} \quad (16)$$

$$F1\ score = \frac{2 \times Positive\ predictive\ value \times Sensitivity}{Positive\ predictive\ value \times Sensitivity} \quad (17)$$

TABLE IV. CONFIGURATION SETTINGS FOR RANDOM FOREST USED IN THE CLASSIFICATION TASK

Hyperparameter	Assigned value
Number of trees	100
Maximum tree depth	10
Minimum leaf samples	2
Minimum split samples	5
Class weighting strategy	Balanced
Randomization seed	42
Bootstrap sampling	Enabled

TABLE V. EVALUATION OUTCOMES OF CLASSIFICATION METRICS BEFORE AND AFTER MITIGATION ENHANCEMENTS

Evaluation measure	Pre-mitigation (%)	Post-mitigation (%)
F1 score	96.21	99.88
Sensitivity	95.52	99.83
Positive predictive value	96.87	99.89
Detection rate	97.15	99.95

The IDS was tested on various classes of attacks, as outlined in Table VI. The accuracy across the classes remained high throughout, with DDoS and scanning attacks achieving precision and recall of over 99%. These results provide evidence of the HSAOA-based IDS model's generalizability and the system's resistance to attacks. A comparison with other feature selection techniques on the given data is provided in Table VII. Compared to other state-of-the-art methods, the HSAOA method presents higher accuracy using fewer selected features without compromising detection power, proving efficient and practical.

TABLE VI. DETECTION PERFORMANCE PER INTRUSION TYPE USING HSAOA-BASED FEATURE SELECTION AND RANDOM FOREST CLASSIFIER

Type of attack	Detection rate (%)	Positive predictive value (%)	Sensitivity (%)	F1 score (%)
XSS	99.91	99.93	99.72	99.89
Scanning	98.78	98.99	98.12	98.95
Ransomware	99.95	99.53	99.54	96.82
Password	98.96	98.95	97.71	97.96
Injection	98.98	98.48	98.57	98.48
DoS	99.94	99.88	99.88	99.93
DDoS	99.94	99.91	99.89	99.92
Backdoor	99.49	99.51	99.42	99.57

TABLE VII. ACCURACY-BASED COMPARISON OF RECENT IDS APPROACHES USING THE TON-IOT DATASET

Reference	Classification model	Feature selection method	No. of selected attributes	Reported detection rate (%)
[17]	XG-Boost (v20)	Chi-squared + SMOTE	20	99.10
[18]	XG-Boost (v19)	Secure privacy-preserving framework	19	98.84
[19]	Random forest	Tabu search	16	83.12
[20]	Neural network	ReliefF	20	98.39
[21]	SVM	Filter + NSGA-II	13	99.48
[21]	SVM	Basic NSGA-II	18	98.86
[22]	CNN	Firefly-based optimization	20	96.66
[23]	CNN	Regularization strategy	8	97.95
[24]	CNN	Pearson correlation-based selection	85	98.75
Present study	Random forest	HSAOA	22	99.94

Security compliance was ensured through both formal and informal security analyses. Through ProVerif verification, formal analysis assures data confidentiality and authentication, as given in Table VIII. Queries 1 and 3 assessed the protocol's resistance to eavesdropping and impersonation, and query 2 raised an issue concerning data confidentiality under given conditions. Informal analysis, as depicted in Table IX, identified the system's response to specific attacks, such as DoS and Replay. Although the system was highly resilient to most attacks, some precautions, such as restricting repetitive requests, are proposed to provide enhanced protection against replay attacks.

TABLE VIII. FORMAL SECURITY VALIDATION OUTCOMES USING THE PROVERIF FRAMEWORK

Queries	Security goal assessed	Outcome	Clarification
Q1	Key/data confidentiality against attacker interception	Yes	Key is protected from exposure; communication is secure against interception.
Q2	Data confidentiality under attacker surveillance	No	Sensitive data may be accessed through insecure communication channels.
Q3	Authentication via request-response correspondence	Yes	Each response is verified to match a valid request, confirming authenticity.

TABLE IX. INFORMAL ASSESSMENT OF SYSTEM VULNERABILITIES AND CORRESPONDING RECOMMENDATIONS

Threat category	Vulnerability identified	Root cause explanation	Proposed mitigation
Replay attack	Present	System allows duplicate message transmissions	Implement time-based tokens or session tags
Injections	None	Alteration of data inputs is successfully detected	No intervention needed
DoS attack	Partial	System handles low volume but weak under heavy load	Enforce connection rate limiting
Backdoor	None	Authentication mechanism verified successfully	No further action necessary

V. DISCUSSION

The optimization gains made with the proposed HSAOA are a result of its ability to solve the significant problems outlined for IoT-based IDS, especially the high dimensionality of the data, computational intensity, and the low detection effectiveness resulting from the poor selection of features.

The integration of a spiral-oriented stochastic search process and adaptive hierarchical leader-follower structure into HSAOA also enhances the capability of the algorithm to exploit and explore the space of features more effectively. Unlike conventional AOA, which gets stuck depending on the fixedness of the arithmetic operator, HSAOA offers enhanced randomness and directional diversity across initial iterations. It facilitates greater discrimination across features and the ability of the model to select the most relevant attributes and delete noise and redundancy.

Another drawback of traditional metaheuristics is that they prematurely converge into local optima for highly complex and high-dimensional data like the kind used in IoT contexts. The rank-based differential mutation strategy suggested for HSAOA maintains population diversity throughout the optimization process, increasing global optima discovery potential. The outcome includes improved IDS accuracy and robustness when attack patterns evolve.

The IDS, after feature reduction via HSAOA, demonstrated consistent high classification accuracy and generalizability across different attack types in the NF-ToN-IoT dataset. This shows that the selected features retain sufficient discriminatory power for real-world use cases while reducing computational load. The use of RF as a robust classifier further contributed to detection stability, mainly when supported by mitigation strategies such as bootstrapping, SMOTE, and regularization.

Although the NF-ToN-IoT dataset provides a rich and diverse set of traffic and attacks, the result could possibly fail to generalize across the entire IoT application spaces (e.g., edge computing or industrial IoT applications). Testing using more benchmark sets should be contemplated to measure broader application. Even though HSAOA greatly reduces dimensionality, the initial search and mutation stages continue to be costly computationally. In extremely dynamic or real-time IoT applications, the overhead might limit performance unless reduced further or parallelized.

From the informal analysis output, the system was susceptible to the type of threats based on replays. Though a non-transparent outcome, it nonetheless implies that the adoption of time-based attributes or time-aware models can be beneficial for enhancing robustness against the same. The algorithm's behavior depends on some parameters that have been empirically set. Though optimally set using experimentation, there could also be automated or adaptive tuning procedures that could raise consistency within the use cases.

VI. CONCLUSION

This research introduced HSAOA to address the major challenge of feature selection for IDS in IoT settings. By employing an adaptive multi-layer hierarchy, spiral-guided random walk, and differential mutation operators, HSAOA improved the global search capability and local exploitation of the traditional AOA algorithm. The proposed approach was tested on the NF-ToN-IoT dataset, yielding a better convergence rate, feature reduction, and classification accuracy. Compared to traditional techniques, HSAOA significantly selected an optimal but highly informative feature subset, facilitating the RF classifier to achieve 99.94% accuracy while sustaining high precision, recall, and F1 measure over diverse types of attacks.

Formal and informal security analyses validated the resistance of the proposed IDS system to various attacks, verifying applicability to realistic IoT environments. Further work can be explored, where HSAOA is combined with deep learning techniques, running the model into actual IoT environments and exploring scalability over large-scale heterogeneous data sets. Additional optimization by adaptive parameter management and other optimization techniques are also ready to propel smart IoT security to new heights.

REFERENCES

[1] B. Pourghebleh, N. Hekmati, Z. Davoudnia, and M. Sadeghi, "A roadmap towards energy - efficient data fusion methods in the Internet of Things," Concurrency and Computation: Practice and Experience, vol. 34, no. 15, p. e6959, 2022.

[2] E. Geo Francis, S. Sheeja, E. Antony John, and J. Joseph, "IoT and Smart Device Security: Emerging Threats and Countermeasures," Securing the Digital Frontier: Threats and Advanced Techniques in Security and Forensics, pp. 217-241, 2025.

[3] R. Saadouni, C. Gherbi, Z. Aliouat, Y. Harbi, and A. Khacha, "Intrusion detection systems for IoT based on bio-inspired and machine learning techniques: a systematic review of the literature," Cluster Computing, vol. 27, no. 7, pp. 8655-8681, 2024.

[4] M. M. Rahman, S. Al Shakil, and M. R. Mustakim, "A survey on intrusion detection system in IoT networks," Cyber Security and Applications, vol. 3, p. 100082, 2025.

[5] B. Pourghebleh, K. Wakil, and N. J. Navimipour, "A comprehensive study on the trust management techniques in the Internet of Things," IEEE Internet of Things Journal, vol. 6, no. 6, pp. 9326-9337, 2019.

[6] F. Cerasuolo, G. Bovenzi, D. Ciuonzo, and A. Pescapè, "Adaptable, incremental, and explainable network intrusion detection systems for internet of things," Engineering Applications of Artificial Intelligence, vol. 144, p. 110143, 2025.

[7] A. Davahli, M. Shamsi, and G. Abaei, "Hybridizing genetic algorithm and grey wolf optimizer to advance an intelligent and lightweight intrusion detection system for IoT wireless networks," Journal of Ambient Intelligence and Humanized Computing, vol. 11, no. 11, pp. 5581-5609, 2020.

[8] N. H. Shahapure and M. Punitha, "Water moth search algorithm-based deep training for intrusion detection in IoT," Journal of Web Engineering, vol. 20, no. 6, pp. 1781-1812, 2021.

[9] R. Gangula and M. M. V, "Network intrusion detection system for Internet of Things based on enhanced flower pollination algorithm and ensemble classifier," Concurrency and Computation: Practice and Experience, vol. 34, no. 21, p. e7103, 2022.

[10] O. A. Alghanam, W. Almobaideen, M. Saadeh, and O. Adwan, "An improved PIO feature selection algorithm for IoT network intrusion detection system based on ensemble learning," Expert Systems with Applications, vol. 213, p. 118745, 2023.

[11] Y. K. Saheed, O. H. Abdulganiyu, and T. Ait Tchakoucht, "Modified genetic algorithm and fine-tuned long short-term memory network for intrusion detection in the internet of things networks with edge capabilities," Applied Soft Computing, vol. 155, p. 111434, 2024.

[12] S. Sharma, V. Kumar, and K. Dutta, "Multi - objective prairie dog optimization algorithm for IoT - based intrusion detection," Internet Technology Letters, vol. 7, no. 6, p. e516, 2024.

[13] T. B. Ogunseyi and G. Thiyagarajan, "An Explainable LSTM-Based Intrusion Detection System Optimized by Firefly Algorithm for IoT Networks," Sensors, vol. 25, no. 7, p. 2288, 2025.

[14] J. Li, H. Chen, M. S. Othman, N. Salim, L. M. Yusuf, and S. R. Kumaran, "NFIoT-GATE-DTL IDS: Genetic algorithm-tuned ensemble of deep transfer learning for NetFlow-based intrusion detection system for internet of things," Engineering Applications of Artificial Intelligence, vol. 143, p. 110046, 2025.

[15] A. Grandhi and S. K. Singh, "Interrelated dynamic biased feature selection and classification model using enhanced gorilla troops optimizer for intrusion detection," Alexandria Engineering Journal, vol. 114, pp. 312-330, 2025.

[16] L. Abualigah, A. Diabat, S. Mirjalili, M. Abd Elaziz, and A. H. Gandomi, "The arithmetic optimization algorithm," Computer methods in applied mechanics and engineering, vol. 376, p. 113609, 2021.

[17] A. R. Gad, A. A. Nashat, and T. M. Barkat, "Intrusion detection system using machine learning for vehicular ad hoc networks based on ToN-IoT dataset," IEEE access, vol. 9, pp. 142206-142217, 2021.

[18] P. Kumar, G. P. Gupta, and R. Tripathi, "TP2SF: A Trustworthy Privacy-Preserving Secured Framework for sustainable smart cities by leveraging blockchain and machine learning," Journal of Systems Architecture, vol. 115, p. 101954, 2021.

[19] A. Nazir and R. A. Khan, "A novel combinatorial optimization based feature selection method for network intrusion detection," Computers & Security, vol. 102, p. 102164, 2021.

[20] R. H. Mohamed, F. A. Mosa, and R. A. Sadek, "Efficient intrusion detection system for IoT environment," International Journal of Advanced Computer Science and Applications, vol. 13, no. 4, 2022.

[21] A. K. Dey, G. P. Gupta, and S. P. Sahu, "Hybrid meta-heuristic based feature selection mechanism for cyber-attack detection in IoT-enabled networks," Procedia Computer Science, vol. 218, pp. 318-327, 2023.

[22] M. Dobrojevic, M. Zivkovic, A. Chhabra, N. S. Sani, N. Bacanin, and M. M. Amin, "Addressing internet of things security by enhanced sine cosine metaheuristics tuned hybrid machine learning model and results interpretation based on shap approach," PeerJ Computer Science, vol. 9, p. e1405, 2023.

[23] B. Ibrahim Hairab, H. K. Aslan, M. S. Elsayed, A. D. Jurcut, and M. A. Azer, "Anomaly detection of zero-day attacks based on CNN and regularization techniques," Electronics, vol. 12, no. 3, p. 573, 2023.

[24] S. Yaras and M. Dener, "IoT-based intrusion detection system using new hybrid deep learning algorithm," Electronics, vol. 13, no. 6, p. 1053, 2024.