Computers, Materials & Continua
DOI: 10.32604/cmc.2023.033513
Article

Tech Science Press

check for updates

Voting Classifier and Metaheuristic Optimization for Network Intrusion Detection

Doaa Sami Khafaga[1], Faten Khalid Karim[1,*], Abdelaziz A. Abdelhamid[2,3], El-Sayed M. El-kenawy[4], Hend K. Alkahtani[1], Nima Khodadadi[5], Mohammed Hadwan[6] and Abdelhameed Ibrahim[7]

[1]Department of Computer Sciences, College of Computer and Information Sciences, Princess Nourah bint Abdulrahman University, P.O. Box 84428, Riyadh, 11671, Saudi Arabia
[2]Department of Computer Science, Faculty of Computer and Information Sciences, Ain Shams University, Cairo, 11566, Egypt
[3]Department of Computer Science, College of Computing and Information Technology, Shaqra University, 11961, Saudi Arabia
[4]Department of Communications and Electronics, Delta Higher Institute of Engineering and Technology, Mansoura, 35111, Egypt
[5]The Department of Civil and Environmental Engineering, Florida International University, Miami, FL, USA
[6]Department of Information Technology, College of Computer, Qassim University, Buraydah, 51452, Saudi Arabia
[7]Computer Engineering and Control Systems Department, Faculty of Engineering, Mansoura University, Mansoura, 35516, Egypt
*Corresponding Author: Faten Khalid Karim. Email: fkdiaaldin@pnu.edu.sa
Received: 19 June 2022; Accepted: 11 August 2022

Abstract: Managing physical objects in the network's periphery is made possible by the Internet of Things (IoT), revolutionizing human life. Open attacks and unauthorized access are possible with these IoT devices, which exchange data to enable remote access. These attacks are often detected using intrusion detection methodologies, although these systems' effectiveness and accuracy are subpar. This paper proposes a new voting classifier composed of an ensemble of machine learning models trained and optimized using metaheuristic optimization. The employed metaheuristic optimizer is a new version of the whale optimization algorithm (WOA), which is guided by the dipper throated optimizer (DTO) to improve the exploration process of the traditional WOA optimizer. The proposed voting classifier categorizes the network intrusions robustly and efficiently. To assess the proposed approach, a dataset created from IoT devices is employed to record the efficiency of the proposed algorithm for binary attack categorization. The dataset records are balanced using the locality-sensitive hashing (LSH) and Synthetic Minority Oversampling Technique (SMOTE). The evaluation of the achieved results is performed in terms of statistical analysis and visual plots to prove the proposed approach's effectiveness, stability, and significance. The achieved results confirmed the superiority of the proposed algorithm for the task of network intrusion detection.

Keywords: Voting classifier; whale optimization algorithm; dipper throated optimization; intrusion detection; internet-of-things

1 Introduction

Growth and flexibility in every area have been seen in the Internet of Things (IoT) during the last several years [1–5]. The IoT in real-life is depicted in Fig. 1. An increasing number of intelligent systems are based on IoT, and securing these systems is a significant challenge [6–10]. In the current literature, cyber attack detection strategies for smart systems have been shown to be of great value. In the past, an IoT device breached by an attacker led to a power outage, affecting 225000 people [11] because the security mechanism was of insufficient quality. Interdependence among devices, limited variety, and more are only some of the characteristics of IoT technology [12]. We can better protect our smart systems by fully comprehending their functionalities. Since IoT devices communicate data with one another and with one other, the interdependence of these gadgets necessitates fewer human decisions and requires less human engagement. For example, an imaginative home scenario where the thermostat measures the temperature in the house and compares it to a preset threshold. When the thermostat detects a deviation from the preset temperature range, it attempts to restore equilibrium to the surrounding air. When it comes to maintaining a comfortable temperature, the smart plug is checked. When the AC is disconnected, the windows are automatically opened to maintain a stable temperature and allow for ventilation. Intruders can access a building by hacking into an IoT device and opening a door or window if the system is not adequately secured.

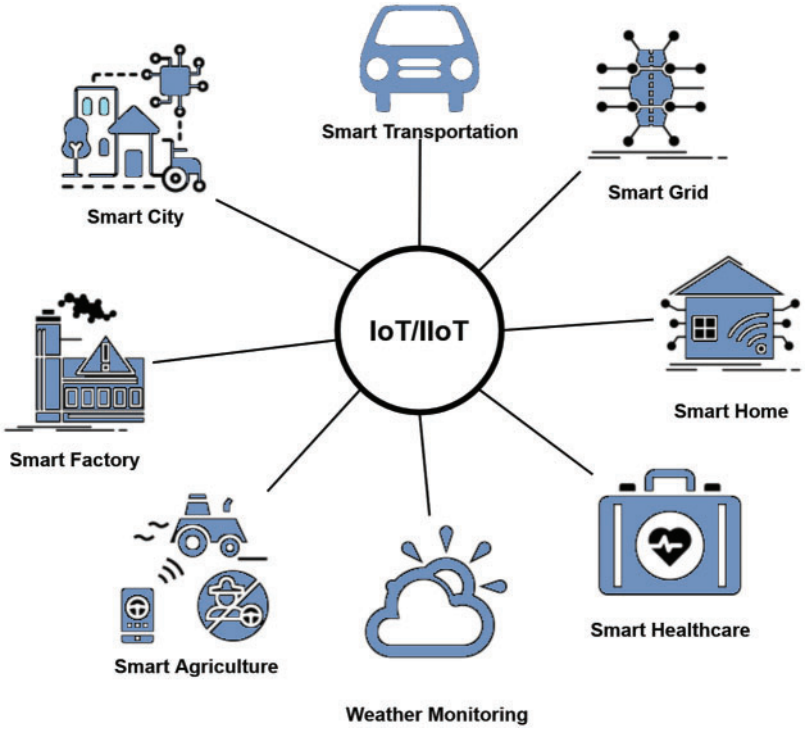

Figure 1: The domain of the internet of things (IoT) in real life

Code injection and Man in the Middle (MitM) attacks [13] are two examples of attacks that might exploit an IoT device. Adversaries use code injection to modify data on IoT devices. These attacks employ techniques similar to listening for and intercepting communications between two nodes. An attacker can control a node in this scenario. Using an ensemble-based voting classifier for intrusion detection, this article examines how well it protects individual IoT devices while securing the network's dependency. The typical machine learning (ML) method was merged with the ensemble-based voting classifier, which then cast votes on each prediction to arrive at a final prediction. Soft voting and hard voting are two examples of voting. Section 3 contains a complete mathematical description of the suggested algorithm [14]. A real-world IoT network dataset named Ton-IoT is used to evaluate the performance of the technique proposed in this paper [15].

This paper includes the following: 1) A novel attack categorization model based on an ensemble has been suggested. Assessment of the presented method for IoT datasets 3) A comparison between the proposed and existing approaches is examined using various criteria. The rest of the document is structured as follows: Section 2 provides background information and an in-depth examination of IoT-related activities and dangers. Section 3 explores the solution under consideration. Data selection and pre-processing are discussed in Section 4. Section 5 wraps up the project and suggests some possible future avenues of exploration.

2 Literature Review

By automating the working environment to decrease human participation and increase system efficiency, the Internet of Things (IoT) is a new technology [16]. As IoT technology continues to advance, new forms of cyberattacks are being developed daily. These networks are easy targets for cybercriminals because they lack adequate security measures. When an IoT device is hacked, attackers can control all of the other devices connected to that device [17]. Detecting an attacker's infiltration or malicious behavior on the network can help prevent these attacks from happening in the first place. Intrusion detection methods are crucial in spotting unwanted activity on networks like this. Tab. 1 summarizes the content of this chapter. Ensemble approaches, including Boosted Trees, Bagged Trees, Subspace Discriminant, and Boosted Trees, have been applied in [18] to offer routing protocol for low-power and lossy networks (RPL)-based network intrusion detection for IoT networks. Network intrusion detection systems (NIDS) are tested using the RPL-NIDS17 dataset. Authors in [19] presented an ensemble-based intrusion detection system to avoid harmful events in IoT networks, especially the botnet attack against hypertext transfer protocol (HTTP), message queuing telemetry transport (MQTT), and domain name space (DNS) protocols. An adaptive boost ensemble technique for attack detection was developed by combining three machine learning algorithms: decision tree (DT), naïve Bayes (NB), and artificial neural network (ANN). An evaluation of the proposed technique was performed using the Network Information Management System (NIMS) botnet data sets.

A two-stage intrusion detection system was proposed in [20]. It is hypothesized that a voting ensemble classifier may be developed by selecting optimum features and combining C4.5 and RF and Random Forest by Penalizing Attribute (RF-PA) machine learning methods. Evaluation measures Accuracy, Precision, Detection rate, and F-measure were used to assess the presented techniques. Authors in [21] proposed employing the XGBoost model for intrusion detection. Evaluation measures included accuracy, precision, recall, and the F1 measure for the test dataset. As an assessment metric for an intrusion detection system based on Extreme Gradient Boosting (XGBoost), the proposed method was put to a test dataset [22]. To guard against attacks on wifi, authors in [23] suggested

an ensemble-based intrusion detection system based on a dataset and assessed using the evaluation metrics accuracy, precision, recall, and F1-measure. This system is essential to enable narrowband and broadband IoT applications. To detect infiltration in IoT networks, Authors in [24] evaluated eleven methods, including seven supervised and three unsupervised ones. Unsupervised algorithms that performed best were found to use XGBOOST and Expectation-Maximization (EM). The accuracy, area under the curve (AUC), and Matthews correlation coefficient (MCC), of 11 algorithms were evaluated.

Table 1: The models used in the literature for intrusion detection

Paper	Model	Evaluation metrics
[15]	Lineat Regression (LR), Linear Discriminant Analysis (LDA), k-Nearest Neighbors (KNN), Random Forest (RF), Classification and Regression Trees (CART)	Accuracy, Precision Recall, F-measure
[18]	Ensemble learning	Accuracy, Receiver Operating Characteristic (ROC)
[19]	Adaptive boost	ROC, Detection rate, Accuracy, False positive rate, ROC
[20]	Ensemble voting	F-measure, Detection rate, Accuracy, Precision
[21]	XGBoost	Precision, Accuracy ROC, F-measure, Recall
[22]	XGBoost	Accuracy
[23]	Ensemble learning	Precision, Accuracy, F1-Measure, Recall

As with intrusion detection, however, the assessment of datasets [25] plays a significant role. IoT 4.0 telemetry datasets were proposed in [15] by writers who used a variety of attack scenarios to create a whole new generation of data. An IoT device data collection containing actual sensor readings from seven IoT sensors. DT and RF beat ML and deep learning (DL) algorithms in evaluation criteria like accuracy, precision, recall, F-measure, etc. Results reveal that a single machine learning algorithm's performance changes when the data from sensors vary. The best attack detection performance on every sensor, thus, requires ensemble-based learning. As previously stated, designing an optimum intrusion detection system requires a realistic dataset near real-time scenarios [26,27]. For this study, many openly accessible datasets were compared and analyzed using various criteria such as different attack scenarios, data from IoT telemetry, and independent datasets for each type of IoT item. Several publicly accessible datasets may be used to construct and analyze an intrusion detection system (IDS), such as the Labeled Wireless Sensor Network Data Repository (LWSNDR). Data created for evaluating IDSs in IoT and Industrial internet of things (IIoT) networks is now publicly available through a new dataset [28–31].

Using Telemetry data from several IoT/IIoT services, this dataset includes information on a wide range of attacks. Fridge sensor, Garage door, Global Positioning Sensor (GPS), Weather, Motion light sensor, and Thermostat are all included in the dataset's 7 IoT devices. The data recorded in these datasets differ; hence, the retrieved dataset is derived from various sources. Garage door IoT devices, for example, only deal with 'ON' or 'OFF', signifying the door's status, as not all IoT devices deal

with the same kind of data. Some devices also deal with real-valued numeric data in the same way. The typical ML algorithm's performance does not remain constant as the kind of data changes. For this reason, we proposed an accurate classifier that could handle data and operate optimally on most devices in IoT networks by using DT, KNN, RF, and Naïve Bayes (NB) algorithms, which we have integrated. An IDS may be analyzed using the dataset listed in Tab. 2. An IDS for IoT devices can use the Ton-IoT dataset, which is based on data from various/separate IoT devices, as shown in the table above.

Table 2: Dataset preprocessing using locality sensitive hashing (LSH)-synthetic minority oversampling technique (SMOTE) balancing

Category	Total instances in dataset	Utilized instances	Using LSH-SMOTE
Attack	33,337	33,337	1,33,348
Normal	4,31,981	1,33,348	1,33,348

3 Methodology

To detect attacks on RPL-based IoT networks, we proposed a new algorithm based on dipper throated optimization (DTO) and whale optimization algorithm (WOA) in this paper. We referred to as (DTO-Guided WOA). This algorithm is employed to optimize the parameters of the voting classifier based on three classifiers, namely, neural networks (NN), RF, and KNN. Data collection, processing, and detection are all covered in the design of the suggested technique. Sniffer and sensor systems comprise the data collecting system. A sniffer may be necessary to access IPv6 over low-power wireless personal area networks (6LoWPAN). A database of sensor events and packets that have been intercepted and routed may be accessed. The dataset's most essential properties are then identified using a feature selection technique. The detecting system includes an alarm/attack notification module. Regularly, it does traffic analysis to offer user interfaces with log data. The architecture of the proposed approach is depicted in Fig. 2.

3.1 Dataset Collection

The proposed method is trained using the RPL-NIDS17 dataset [32]. This data set was generated with the help of the NetSim application. Simulating many sorts of network infrastructures is easy using NetSim. The Internet of Things network includes a gateway, sensor nodes, a wired node, and a router. Each attack is documented in great detail in a comma-separated values (CSV) file. It's possible to merge all of the CSV files into a single dataset 20 features may be tagged using this dataset's time, essential, and flow properties. Aside from the primary traffic patterns, hello flooding, and selective forwarding that may be used in routing attacks, include Sybil (blackhole), sinkhole, and clone intrusion detections. This dataset only contains 33,337 routing attacks and 431,981 pieces of regular traffic. Because of the imbalance, the data is skewed.

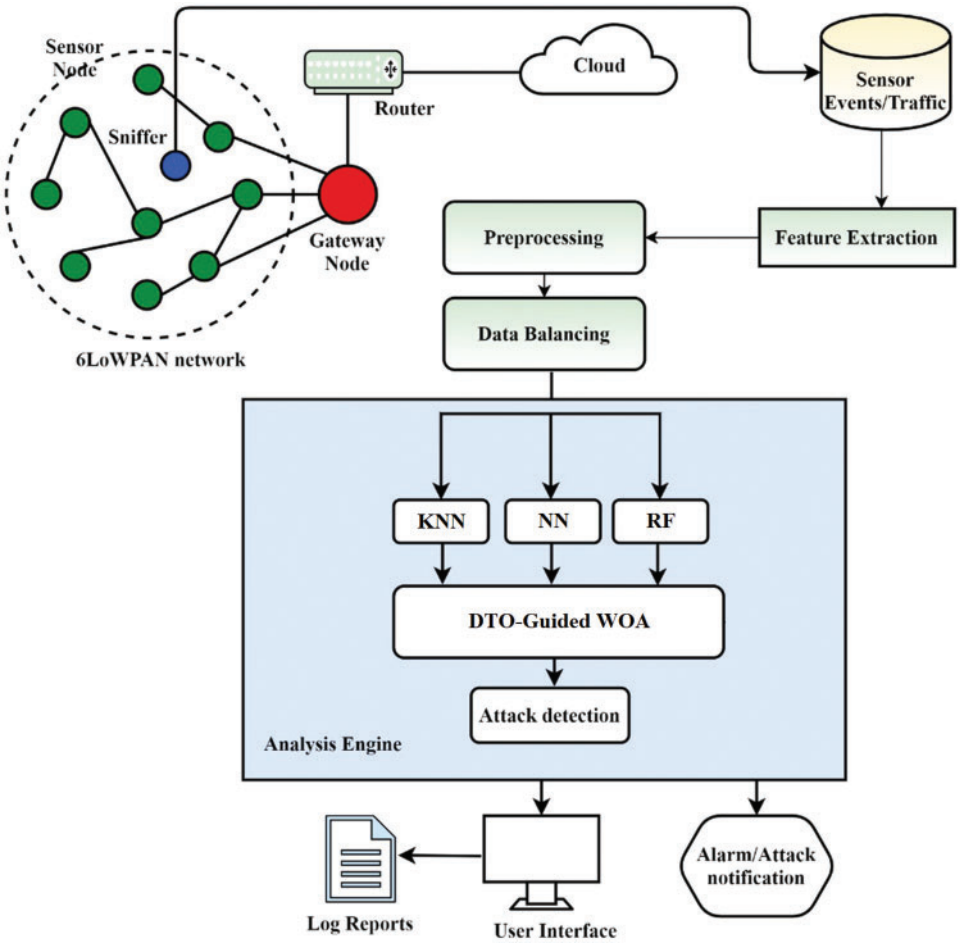

Figure 2: The architecture of the proposed network intrusion detection system

3.2 Dataset Preprocessing

The first step in getting ready is to clean up your data. Encrypting and de-duplicating the data is also part of this process. Even though computers can only read numeric data, the dataset comprises numeric and nominal data. As a result, the dataset's characters have been converted to numeric values for storage. After all of this, data scaling is employed to speed things up. The dataset contains a wide range of characteristics, both in size and unit. Data integrity may be maintained throughout the time when scaled [33–40]. As part of data preparation, the amount of samples in the dataset is balanced such that each class has equal numbers of samples [41–47]. The locality-sensitive hashing and synthetic minority oversampling techniques were used here to accomplish this aim. Before and after balancing, the number of samples in the dataset is shown in Tab. 2.

3.3 Dipper Throated Optimization

Dipper throated optimization (DTO) is based on tracking the locations and speeds of swimming and flying birds to simulate the genuine process of seeking food. Swimming birds' positions and speeds are updated using these equations.

$$BL_{nd}(t+1) = BL_{best}(t) - C_1.|C_2.BL_{best}(t) - BL_{nd}(t)| \quad (1)$$

where $BL_{nd}(t)$ and $BL_{best}(t)$ are the normal location and best location of the bird at iteration t, and C_1 and C_2 are adaptive values whose values are changed during the optimization process based on the iteration number and random values. The flying bird's location is updated using the following equation.

$$BS(t+1) = C_3 BS(t) + C_4 r_1 (BL_{best}(t) - BL_{nd}(t)) + C_5 r_1 (BL_{Gbest} - BL_{nd}(t)) \quad (2)$$

$$BL_{nd}(t+1) = BL_{nd}(t) + BS(t+1) \quad (3)$$

where $BS(t+1)$ is the updated speed of each bird, r_1 is a random number in [0; 1], BL_{Gbest} is the global best location, and C_3 is a weight value, C_4 and C_5 are constants.

3.4 Whale Optimization

Humpback whales forage for food using WOA, a novel metaheuristic algorithm described in [29]. Whales searching for tiny fish near the surface swim in a circle and make bubbles along a route that looks like a "9," as illustrated in Fig. 3. In the first phase of the algorithm, the encircling prey and the spiral bubble-net attack approach were depicted; in the second phase, the system searched for prey randomly (exploration phase). In the following sections, we will go through some details of the mathematical model of each phase. Random numbers will be generated using a uniform distribution, as shown in the formulae.

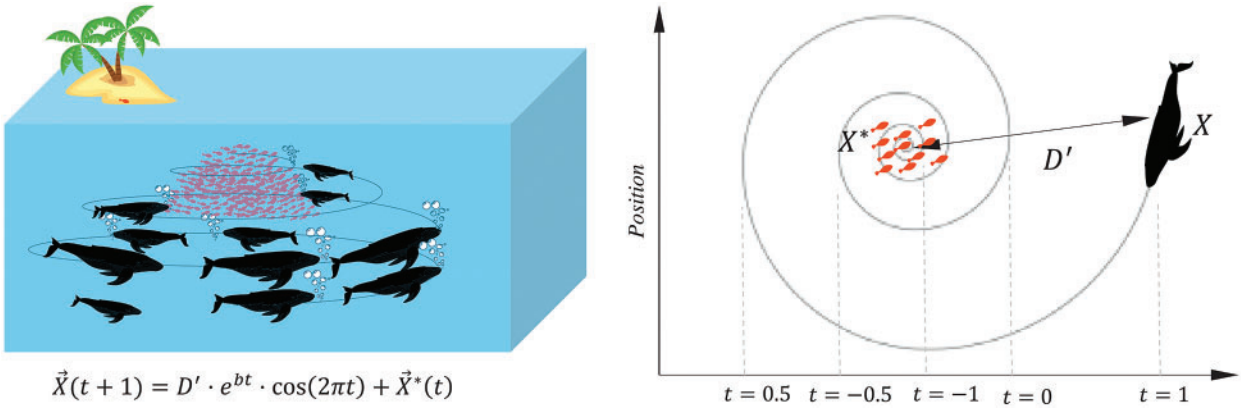

Figure 3: The hunting process of grey wolf optimization

3.4.1 Exploitation Phase

Humpback whales initially encircle their prey to catch it. Eqs. (4) and (5) can be used to represent this behavior quantitatively.

$$D = \left| C \cdot \vec{X^*}(t) - \vec{X}(t) \right| \quad (4)$$

$$\vec{X}(t+1) = \overrightarrow{X^*}(t) - \vec{A} \,.\, D \tag{5}$$

In this case, t is the current iteration, $\overrightarrow{X^*}$ represents the best solution found thus far, $\vec{X}$ is the position vector, and |. | denotes the absolute value. The coefficient vectors A and C are also determined as in Eqs. (3) and (4):

$$\vec{A} = 2\vec{a}.\vec{r} - \vec{a} \tag{6}$$

$$\vec{C} = 2.\vec{r} \tag{7}$$

In the exploration and exploitation stages, 'a' declines linearly from 2 to 0, and r is produced randomly with uniform distribution in the interval [0,1]. Search agents (whales) update their locations based on the best-known solution's position (prey), according to Eq. (5). Predators can only be located in the vicinity of a whale by altering the values of A and C vectors. According to Eq. (8), t is the iteration number, and the value in Eq. (6) is decreased to produce the shrinking encircling behavior.

$$a = 2 - t\frac{2}{Max\_Iterations} \tag{8}$$

Search agent (X) and best known search agent (X*) distances are obtained as shown in Fig. 3; then the spiral in Eq. (9) is used to produce the neighbour search agent's position.

$$\vec{X}(t+1) = \grave{D}.e^{bl}.Cos(2\pi l) + \overrightarrow{X^*}(t) \tag{9}$$

$$\grave{D} = \left|\overrightarrow{X^*}(t) - \vec{X}(t)\right| \tag{10}$$

where l is a random number within the range [−1, 1], the logarithmic spiral's constant is denoted by b. In Eq. (11), we suppose there is a 50% chance that the optimization process will pick between the spiral-shaped approach and shrinking encircling when p is some random value in the range [0,1].

$$\vec{X}(t+1) = \begin{cases} \overrightarrow{X^*}(t) - \vec{A}.D, & if\ (p < 0.5) \\ \grave{D}.e^{bl}.Cos(2\pi l) + \overrightarrow{X^*}(t), & if\ (p \geq 0.5) \end{cases} \tag{11}$$

3.4.2 Exploration Phase

Instead of adjusting the search agents' placement based on the position of the best one found thus far, a random search agent is used to direct the search in WOA. Since A is used to compel the search agent to wander far away from the best-known search agent, random values larger than 1 or less than −1 are employed. $\vec{X}_{rand}$ is a randomly selected whale from the present population in Eq. (12), which mathematically models this procedure.

$$\vec{D} = \vec{X}_{rand} - \vec{A}.\left|\vec{C}.\vec{X}_{rand} - \vec{X}\right| \tag{12}$$

Algorithm 1: The proposed DTO-Guided WOA algorithm

```
Generate Initial Population X i (i = 1, 2, …, n)
Calculate the fitness of each solution
X* = the best search agent
while (t < Max_Iteration)
    for each solution
        Update a, A, C, l, and p
        if (p < 0.5)
            if (|A| < + 1)
                Update the position of the current solution by Eq. (5)
            else if (|A| > + 1)
                Select a random search agent ()
                Update the position of the current search agent by the Eq. (12)
            end if
        else if (p ≥ 0.5)
            Update the position of the current search using DTO
        end if
    end for
    Check if any solution goes beyond the search space and amend it
    Calculate the fitness of each solution
    Update X ∗ if there is a better solution t = t + 1
end while
return X ∗
```

3.5 The Proposed Optimization Algorithm

The proposed DTO-Guided WOA employed in optimizing the voting classifiers parameter is listed in Algorithm 1. This algorithm is used to optimize the parameters of the classifiers and the voting ensemble to boost the overall classification accuracy of the network attacks.

4 Experimental Results

The tests are conducted on a Windows 11 laptop with a 2.33 GHz Intel Core i5 and 16 GB of random access memory (RAM). MATLAB R2020a was used to build and evaluate the suggested framework. Text Analytics Toolbox MATLAB is used for preparing the dataset. An evaluation and comparison of the proposed approach's performance are carried out here.

The achieved results based on the proposed feature selection approach are presented in Tab. 3, with 95.1% accuracy. In addition, a comparison between the performance of the majority voting and the proposed voting algorithm is presented in Tabs. 4–6. In these tables, the performance of the proposed voting algorithm is much better than the traditional majority voting. The achieved AUC is (0.99) using the proposed approach, whereas the AUC value using the traditional voting is (0.974). In addition, the proposed approach's mean square error (MSE) is (2.50E-08), which reflects the superiority of the proposed approach when compared to the traditional voting algorithm.

Table 3: Evaluation results of the results achieved with/without data preprocessing

Metric	NN	KNN	Random Forest
AUC without SMOTE	0.813	0.861	**0.893**
MSE without SMOTE	0.052373	0.04932	**0.032853**
AUC with SMOTE	0.861	0.917	**0.931**
MSE with SMOTE	0.006708	0.005852	**0.0035723**
AUC with LSH-SMOTE	0.897	0.936	**0.951**
MSE with LSH-SMOTE	0.000574	0.000395	**0.0001012**

Table 4: Comparison between the results achieved by the majority voting and the proposed approach

Metric	Majority voting	Voting (DTO _Guided WOA)
AUC with LSH-SMOTE	0.974	**0.999975**
MSE with LSH-SMOTE	0.000005931	**2.50E-08**

Table 5: Assessment of the voting approach using the proposed optimization algorithm and other algorithms

	AUC with LSH-SMOTE	MSE with LSH-SMOTE
Voting (DTO _Guided WOA)	**0.999975**	**2.50E-08**
Voting particle swarm optimimizatiion (PSO)	0.9914	0.00000151
Voting genetic algorithm (GA)	0.989	0.000002721
Voting grey wolf optimization (GWO)	0.986	0.000003025
Voting whale optimization algorithm (WOA)	0.981	0.000004084

Table 6: Statistical analysis of the results achieved by the proposed optimization algorithm and other algorithms

	Voting (DTO _Guided WOA)	Voting WOA	Voting GWO	Voting GA	Voting PSO
Number of values	14	**14**	14	14	14
Minimum	**1**	0.971	0.976	0.979	0.9714
25% Percentile	**1**	0.981	0.986	0.989	0.9914
Median	**1**	0.981	0.986	0.989	0.9914
75% Percentile	**1**	0.981	0.986	0.989	0.9914
Maximum	**1**	0.991	0.996	0.999	0.9914
Range	**0**	0.02	0.02	0.02	0.02
Mean	**1**	0.981	0.986	0.989	0.9893

(Continued)

Table 6: Continued

	Voting (DTO _Guided WOA)	Voting WOA	Voting GWO	Voting GA	Voting PSO
Std. Error of mean	**0**	0.001048	0.001048	0.001048	0.001547
Std. Deviation	**0**	0.003922	0.003922	0.003922	0.005789
Sum	**14**	13.73	13.8	13.85	13.85

Tabs. 7 and 8 show the findings of the analysis of variance (ANOVA) and Wilcoxon signed-rank tests, on the other hand. As can be seen from the tables, the proposed strategy is statistically significant, just like the different strategies. Therefore, the suggested method is suited to the task of selecting features. Tab. 7 shows the ANOVA test results for validating the proposed approach's stability and effectiveness. These tests stress the statistical significance and efficacy of the suggested method based on the hypotheses of these tests.

Table 7: One-way analysis of variance test

Metric	SS	DF	MS	F (DFn, DFd)	*P* value
Treatment	0.002709	4	0.000677	F (4, 65) = 42.50	$P < 0.0001$
Residual	0.001036	65	1.59E-05		
Total	0.003745	69			

Table 8: Wilcoxon signed rank test

	Voting (DTO _Guided WOA)	Voting WOA	Voting GWO	Voting GA	Voting PSO
Theoretical median	**0**	0	0	0	0
Number of values	**14**	14	14	14	14
Actual median	**1**	0.981	0.986	0.989	0.9914
Discrepancy	**1**	0.981	0.986	0.989	0.9914
Significant (alpha = 0.05)?	**Yes**	Yes	Yes	Yes	Yes
Exact or estimate?	**Exact**	Exact	Exact	Exact	Exact
P value (two tailed)	**0.0001**	0.0001	0.0001	0.0001	0.0001
Sum of negative ranks	**0**	0	0	0	0
Sum of positive ranks	**105**	105	105	105	105
Sum of signed ranks (W)	**105**	105	105	105	105

The attained outcomes are shown in Fig. 4 to demonstrate the approach's efficacy and superiority. It's easy to see that the proposed technique is highly accurate based on the data in this image because the residual error is so little. Quantile-by-quantile (QQ), heatmaps, ROCs, and histogram plots are utilized to demonstrate the suggested method's efficiency. Plots like this demonstrate how superior the recommended strategy is to the alternatives.

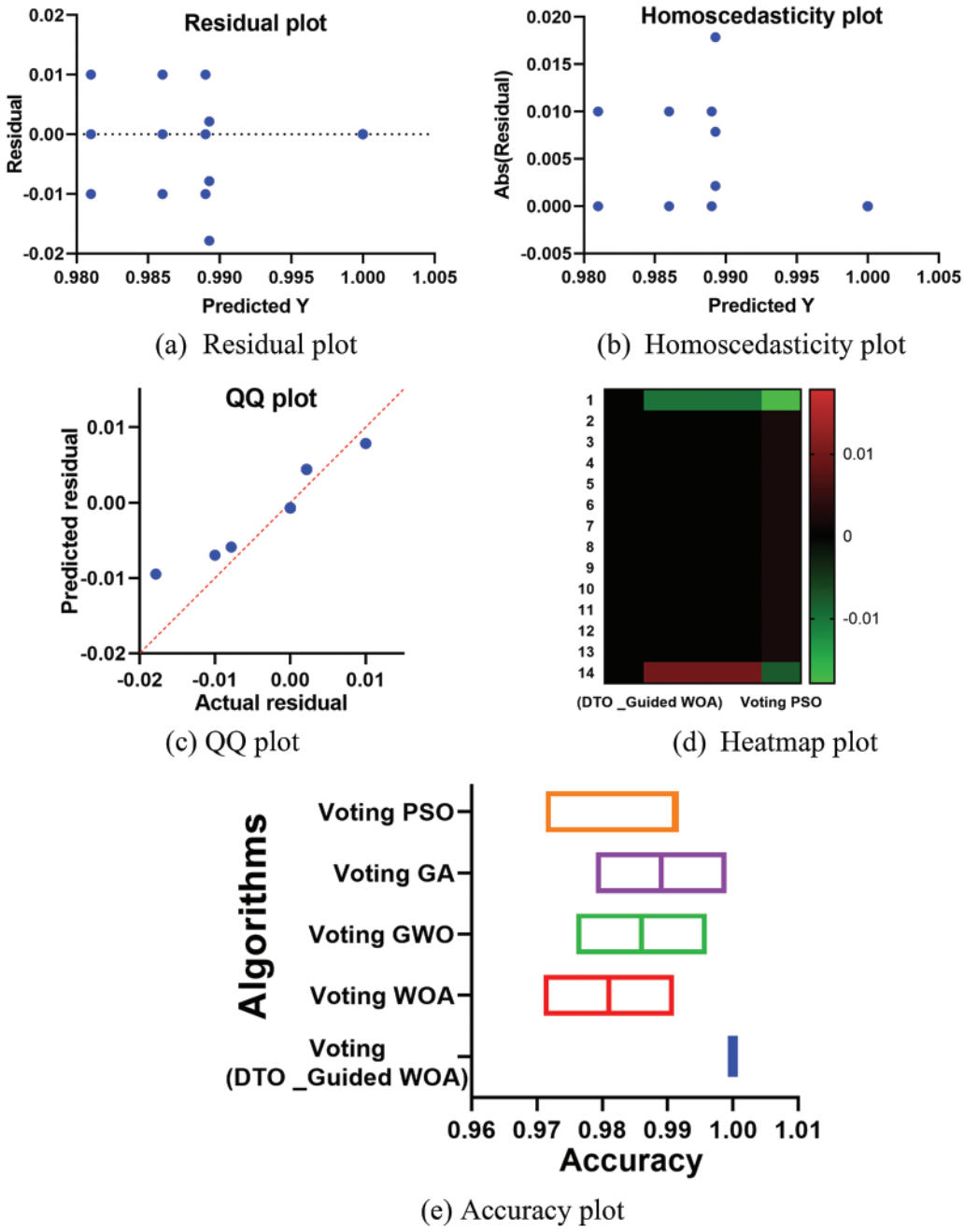

Figure 4: Visualization of the achieved results using the proposed methodology

The accuracy of the achieved results using the proposed approach is presented in a histogram in Fig. 5. As shown in this figure, the proposed approach achieves the best results compared to the other voting classifier approaches. These results prove the proposed approach's superiority in accurately detecting network attacks.

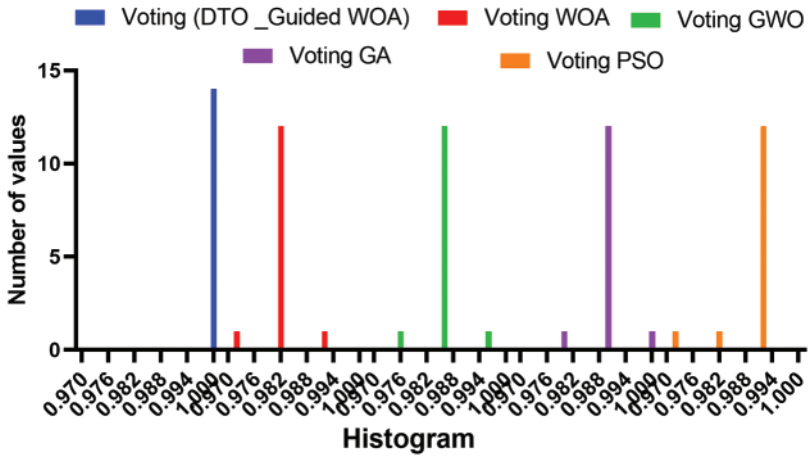

Figure 5: Histogram of the accuracy achieved by the proposed approach and other approaches

5 Conclusions

This paper proposes a new optimization algorithm for optimizing the classifiers used in intrusion detection systems. The proposed algorithm is based on the dipper throated and whale optimization algorithms. The proposed algorithm is used in an architecture designed to detect network attacks in IoT environments. To validate the effectiveness of the proposed approach, several experiments were conducted to evaluate the stages of the proposed framework. Evaluation results showed the effectiveness of the proposed method. On the other hand, the comparison is conducted to show the superiority of the proposed approach. In addition, a statistical analysis is performed to prove the stability and significance of the proposed method for intrusion detection tasks. The recorded results confirm the findings and emphasize the significance of the proposed approach.

Acknowledgement: Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2022R300), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia.

Funding Statement: Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2022R300), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

[1] M. M. Islam, A. Rahaman and M. R. Islam, "Development of smart healthcare monitoring system in IoT environment," *SN Computer Science*, vol. 1, no. 3, pp. 1–11, 2020.

[2] G. Mois, S. Folea and T. Sanislav, "Analysis of three IoT-based wireless sensors for environmental monitoring," *IEEE Transactions on Instrumentation and Measurement*, vol. 66, no. 8, pp. 2056–2064, 2017.

[3] A. Salam and S. Shah, "Urban underground infrastructure monitoring IoT: The path loss analysis," in *2019 IEEE 5th World Forum on Internet of Things (WF-IoT)*, Limerick, Ireland, pp. 398–401, 2019.

[4] Y. Hajjaji, W. Boulila, I. R. Farah, I. Romdhani and A. Hussain, "Big data and IoT-based applications in smart environments: A systematic review," *Computer Science Review*, vol. 39, no. 7, pp. 1–13, 2021.

[5] A. -R. Al-Ali, I. A. Zualkernan, M. Rashid, R. Gupta and M. Alikarar, "A smart home energy management system using IoT and big data analytics approach," *IEEE Transactions on Consumer Electronics*, vol. 63, no. 4, pp. 426–434, 2017.
[6] A. Churcher, R. Ullah, J. Ahmad, F. Masood, M. Gogate *et al.,* "An experimental analysis of attack classification using machine learning in IoT networks," *Sensors*, vol. 21, no. 2, pp. 446–462, 2021.
[7] A. Shafique, J. Ahmed, W. Boulila, H. Ghandorh, J. Ahmad *et al.,* "Detecting the security level of various cryptosystems using machine learning models," *Algorithms*, vol. 1, pp. 1–5, 2021.
[8] A. Ali, S. Shaukat, M. Tayyab, M. A. Khan, J. S. Khan *et al.,* "Network intrusion detection leveraging machine learning and feature selection," in *IEEE 17th Int. Conf. on Smart Communities: Improving Quality of Life Using ICT*, Charlotte, NC, USA, pp. 49–53, 2020.
[9] A. Qayyum, J. Ahmad, W. Boulila, S. Rubaiee, F. Masood *et al.,* "Chaos-based confusion and diffusion of image pixels using dynamic substitution," *IEEE Access*, vol. 8, pp. 140 876–140 895, 2020.
[10] F. Masood, W. Boulila, J. Ahmad, S. Sankar, S. Rubaiee *et al.,* "A novel privacy approach of digital aerial images based on mersenne twister method with DNA genetic encoding and chaos," *Remote Sensing*, vol. 12, no. 11, pp. 1893–1910, 2020.
[11] L. Robert, J. Michael and C. Tim, "Analysis of the cyber attack on the Ukrainian power grid," *USA: Electricity Information Sharing and Analysis Centre*, vol. 1, no. 1, pp. 1–20, 2016.
[12] W. Zhou, Y. Jia, A. Peng, Y. Zhang and P. Liu, "The effect of IoT new features on security and privacy: New threats, existing solutions, and challenges yet to be solved," *IEEE Internet of Things Journal*, vol. 6, no. 2, pp. 1606–1616, 2018.
[13] M. Zolanvari, M. A. Teixeira, L. Gupta, K. M. Khan and R. Jain, "Machine learning-based network vulnerability analysis of industrial internet of things," *IEEE Internet of Things Journal*, vol. 6, no. 4, pp. 6822–6834, 2019.
[14] N. Mustafa, "Ton-IoT dataset," 2019. [Online]. Available: https://cloudstor.aarnet.edu.au/plus/s/ds5zW91vdgjEj9i, (accessed: 29-May-2022).
[15] A. Alsaedi, N. Moustafa, Z. Tari, A. Mahmood and A. Anwar, "Ton IoT telemetry dataset: A new generation dataset of IoT and IIoT for data-driven intrusion detection systems," *IEEE Access*, vol. 8, pp. 130–165, 2020.
[16] S. Batool, N. A. Saqib, M. K. Khattack and A. Hassan, "Identification of remote IoT users using sensor data analytics," in *Future of Information and Communication Conf.*, San Francisco, CA, USA, pp. 328–337, 2019.
[17] G. Falco, C. Caldera and H. Shrobe, "IIoT cybersecurity risk modeling for SCADA systems," *IEEE Internet of Things Journal*, vol. 5, no. 6, pp. 4486–4495, 2018.
[18] A. Verma and V. Ranga, "Elnids: Ensemble learning based network intrusion detection system for RPL based internet of things," in *4th Int. Conf. on Internet of Things: Smart Innovation and Usages (IoT-SIU)*, Ghaziabad, India, pp. 1–6, 2019.
[19] N. Moustafa, B. Turnbull and K. -K. R. Choo, "An ensemble intrusion detection technique based on proposed statistical flow features for protecting network traffic of internet of things," *IEEE Internet of Things Journal*, vol. 6, no. 3, pp. 4815–4830, 2018.
[20] Y. Zhou, G. Cheng, S. Jiang and M. Dai, "Building an efficient intrusion detection system based on feature selection and ensemble classifier," *Computer Networks*, vol. 174, no. 8, pp. 107247–107266, 2020.
[21] S. S. Dhaliwal, A. -A. Nahid and R. Abbas, "Effective intrusion detection system using XGBoost," *Information*, vol. 9, no. 7, pp. 149–155, 2018.
[22] A. Bansal and S. Kaur, "Extreme gradient boosting based tuning for classification in intrusion detection systems," in *Int. Conf. on Advances in Computing and Data Sciences*, Dehradun, India, pp. 372–380, 2018.
[23] F. D. Vaca and Q. Niyaz, "An ensemble learning based wi-fi network intrusion detection system (WNIDS)," in *IEEE 17th Int. Symp. on Network Computing and Applications (NCA)*, Cambridge, MA, USA, pp. 1–5, 2018.
[24] R. Panigrahi and S. Borah, "A detailed analysis of cicids2017 dataset for designing intrusion detection systems," *International Journal of Engineering & Technology*, vol. 7, no. 3.24, pp. 479–482, 2018.

[25] M. S. Amin, L. Hassan, A. A. Shah, U. Akbar and H. A. Niaz, "Accelerated GPU based protein sequence alignment-An optimized database sequences approach," *IJCSNS*, vol. 17, no. 10, pp. 231–245, 2017.
[26] J. Asharf, N. Moustafa, H. Khurshid, E. Debie, W. Haider *et al.,* "A review of intrusion detection systems using machine and deep learning in internet of things: Challenges, solutions and future directions," *Electronics*, vol. 9, no. 7, pp. 1177–1184, 2020.
[27] S. Latif, Z. Zou, Z. Idrees and J. Ahmad, "A novel attack detection scheme for the industrial internet of things using a lightweight random neural network," *IEEE Access*, vol. 8, pp. 89 337–89 350, 2020.
[28] A. Dogan and D. Birant, "A weighted majority voting ensemble approach for classification," in *4th Int. Conf. on Computer Science and Engineering*, Samsun, Turkey, pp. 1–6, 2019.
[29] B. Zhang, H. Ye, W. Lu, W. Huang, B. Wu *et al.,* "A spatiotemporal change detection method for monitoring pine wilt disease in a complex landscape using high-resolution remote sensing imagery," *Remote Sensing*, vol. 13, no. 11, pp. 1–17, 2021.
[30] S. Raschka, "Python Machine Learning: Unlock deeper insights into Machine Leaning with this vital guide to cutting-edge predictive analytics," *Packt Publishing*, 2015. [Online]. Available: https://www.amazon.com/Python-Machine-Learning-Sebastian-Raschka-ebook/dp/B00YSILNL0, (accessed: 26-05-2022).
[31] A. A. Aburomman and M. B. I. Reaz, "A novel SVM-KNN-PSO ensemble method for intrusion detection system," *Applied Soft Computing*, vol. 38, no. 1, pp. 360–372, 2016.
[32] A. Geron, *Hands-on Machine Learning with Scikit-Learn and Tensor-Flow: Concepts, Tools, and Techniques to Build Intelligent Systems*. 1st ed., USA: O'Reilly Media, 2017.
[33] A. Verma and V. Ranga, "RPL-NIDDS17-A data set for intrusion detection in RPL based 6LoWPAN networks," *Internet of Things*, vol. 5, no. 1, pp. 1–20, 2018.
[34] E. -S. M. El-Kenawy, S. Mirjalili, A. Ibrahim, M. Alrahmawy, M. El-Said *et al.,* "Advanced meta-heuristics, convolutional neural networks, and feature selectors for efficient COVID-19 X-ray chest image classification," *IEEE Access*, vol. 9, no. 1, pp. 36019–36037, 2021.
[35] A. Abdelhamid and S. Alotaibi, "Optimized two-level ensemble model for predicting the parameters of metamaterial antenna," *Computers, Materials & Continua*, vol. 73, no. 1, pp. 917–933, 2022.
[36] A. Abdelhamid and S. R. Alotaibi, "Robust prediction of the bandwidth of metamaterial antenna using deep learning," *Computers, Materials & Continua*, vol. 72, no. 2, pp. 2305–2321, 2022.
[37] D. Sami Khafaga, A. Ali Alhussan, E. M. El-kenawy, A. E. Takieldeen, T. M. Hassan *et al.,* "Meta-heuristics for feature selection and classification in diagnostic breast cancer," *Computers, Materials & Continua*, vol. 73, no. 1, pp. 749–765, 2022.
[38] D. Sami Khafaga, A. Ali Alhussan, E. M. El-kenawy, A. Ibrahim, S. H. Elkhalik *et al.,* "Improved prediction of metamaterial antenna bandwidth using adaptive optimization of LSTM," *Computers, Materials & Continua*, vol. 73, no. 1, pp. 865–881, 2022.
[39] E. -S. M. El-Kenawy, S. Mirjalili, F. Alassery, Y. Zhang, M. Eid *et al.,* "Novel meta-heuristic algorithm for feature selection, unconstrained functions and engineering problems," *IEEE Access*, vol. 10, pp. 40536–40555, 2022.
[40] A. Abdelhamid, E. -S. M. El-kenawy, B. Alotaibi, M. Abdelkader, A. Ibrahim *et al.,* "Robust speech emotion recognition using CNN+LSTM based on stochastic fractal search optimization algorithm," *IEEE Access*, vol. 10, pp. 49265–49284, 2022.
[41] M. M. Eid, E. -S. M. El-Kenawy and A. Ibrahim, "A binary sine cosine-modified whale optimization algorithm for feature selection," in *4th National Computing Colleges Conf. (NCCC 2021)*, IEEE, Taif, Saudi Arabia, pp. 1–6, 2021.
[42] E. -S. M. El-Kenawy, S. Mirjalili, S. S. M. Ghoneim, M. M. Eid, M. El-Said *et al.,* "Advanced ensemble model for solar radiation forecasting using sine cosine algorithm and newton's laws," *IEEE Access*, vol. 9, pp. 115750–115765, 2021.
[43] A. Salamai, E. -S. M. El-kenawy and A. Ibrahim, "Dynamic voting classifier for risk identification in supply chain 4.0," *CMC-Computers, Materials & Continua*, vol. 69, no. 3, pp. 3749–3766, 2021.

[44] A. Ibrahim, S. Mirjalili, M. El-Said, S. S. M. Ghoneim, M. Al-Harthi *et al.,* "Wind speed ensemble forecasting based on deep learning using adaptive dynamic optimization algorithm," *IEEE Access*, vol. 9, pp. 125787–125804, 2021.

[45] E. -S. M. El-kenawy, H. F. Abutarboush, A. W. Mohamed and A. Ibrahim, "Advance artificial intelligence technique for designing double T-shaped monopole antenna," *CMC-Computers, Materials & Continua*, vol. 69, no. 3, pp. 2983–2995, 2021.

[46] E. -S. M. El-kenawy, A. Ibrahim, N. Bailek, B. Kada, M. Hassan *et al.,* "Sunshine duration measurements and predictions in Saharan Algeria region: An improved ensemble learning approach," *Theoretical and Applied Climatology*, vol. 147, no. 3–4, pp. 1015–1031, 2022.

[47] E. -S. M. El-Kenawy, S. Mirjalili, F. Alassery, Y. Zhang, M. Eid *et al.,* "Novel meta-heuristic algorithm for feature selection, unconstrained functions and engineering problems," *IEEE Access*, vol. 10, pp. 40536–40555, 2022.