IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received January 20, 2020, accepted February 10, 2020, date of publication March 3, 2020, date of current version March 31, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2978035

A Novel Feature Selection Method Using Whale Optimization Algorithm and Genetic Operators for Intrusion Detection System in Wireless Mesh Network

R. VIJAYANAND[1] AND D. DEVARAJ[2], (Senior Member, IEEE)
[1]J.B. Institute of Engineering and Technology, Hyderabad 500075, India
[2]Kalasalingam Academy of Education and Research, Krishnankoil 626128, India

Corresponding author: R. Vijayanand (rkvijayanand@gmail.com)

ABSTRACT Machine learning–based intrusion detection system (IDS) is an important requirement for securing data traffic in wireless mesh networks. The noisy and redundant features of network data tend to degrade the performance of the attack detection classifiers. *Therefore*, the selection of informative features plays a vital role in the enhancement to the IDS. In this paper, we propose a wrapper-based approach using the modified whale optimization algorithm (WOA). One drawback of WOA is that premature convergence results in a local optimal solution. To overcome this limitation, we proposed a method in which the genetic algorithm operators were combined with the WOA. The crossover operator was used to further improve the search space of whales, and the mutation operator helped to avoid being stuck in the local optimum. The proposed method selects the informative features in the network data, which helps to accurately detect intrusions. Using a support vector machine (SVM), we identified the types of intrusions based on the selected features. The performance of the improved method was analyzed by using the CICIDS2017 and ADFA-LD standard datasets. Our proposed method had better attack detection rate than the standard WOA and other evolutionary algorithms; it also had good accuracy and was suitable for IDS in the wireless mesh networks. The performance of the IDS was increased by selecting the informative features with the improved whale optimization algorithm. The attack detection ratio was higher than that of the standard WOA.

INDEX TERMS Crossover and mutation operator, IDS for WMN, improved whale optimization algorithm, WOA-based feature selection method, WOA + genetic operators.

I. INTRODUCTION

A wireless mesh network (WMN) is a communication technology suitable for cyber physical applications, such as healthcare devices, smart grids, and the Internet of Things. The low-cost, dynamic, and self-healing nature of the mesh architecture provides reliable communication to future-generation devices. It uses hop-by-hop forwarding technique to transmit the data from the sender to the receiver [1]. This multi-hop data transmission increases the vulnerability to various kinds of network attacks such as false data injection, denial of service (DoS), and worms. Various security mechanisms are available to protect the network from attacks [2].

Intrusion detection system (IDS) is a notable security mechanism that helps to detect the attacks on the network traffic data. For detection, the IDS uses the classifier, which separates the normal data from the network attack data. In recent years, machine learning algorithms such as naïve Bayes, neural networks, and support vector machines (SVMs), have been used as IDS classifiers. The presence of irrelevant features will deteriorate the performance of the classifier. Identifying and removing the features will increase the network attack detection ratio of the classifier. Hence, it is essential to select informative feature to improve the performance of IDS classifiers. However, this is not an easy task. The classifier has the complexity of non-deterministic polynomial-time hardness because the selection of the informative feature subset will include the analysis of all the available features [3].

The aim of the informative feature selection methods is to reduce the dimensionality of the input data and provide

The associate editor coordinating the review of this manuscript and approving it for publication was Mohammad S. Khan.

discriminative features to the classifier algorithms. The available methods are categorized into the filter, wrapper, and embedded methods. Filter methods select the informative features by computing the relationship between the input and output data using evaluation criteria such as the correlation [4] and probability [5]. Then, a threshold value is defined and the features within or above that range are selected as the informative features. However, in the wrapper method, a greedy search algorithm for selecting the optimal features by using randomly generated features is introduced iteratively to the classifier.

Currently, bio-inspired population-based algorithms are being used to select the informative features. It has good accuracy as compared with the filter method, and overcomes the local optimal problem of the greedy-based wrapper methods [6]. In [7], the genetic algorithm (GA)–based wrapper method is used to detect the informative features for improving the performance of the SVM classifier. The performance can be further improved by updating the available methods or replacing them with the techniques of other algorithms. For instance, the author in [8] used the GA-based wrapper method for the selection of semi-informative features and mutual information–based filter method for finding informative features from the selected features. In [9], the differential evolution algorithms were improved by efficiently designing the mutation rule. Similarly, the multiple SVMs were arranged in a hierarchical order for the detection of attacks in WMN [10].

Although, large number of research papers have been published in the area of bio-inspired–based feature selection method for the IDS, there is a large potential to improve the solutions of the optimization algorithms. Whale optimization algorithm (WOA) is a population-based meta-heuristic algorithm that has better performance (i.e., selection of a few parameters, overcoming the local optimum entrapment, and fast convergence to the best solution) than algorithms such as GA, particle swarm optimization, and binary bat algorithm. Very few researchers have used WOA to avoid slow convergence and the local optimal trap during the development of classifiers, such as artificial neural networks, in the intrusion detection model.

In this research, we used the WOA to improve the selection of informative features for IDS. It was enhanced by the crossover and mutation operators of the GA. The SVM was used as a classifier that helps to select the informative features of IDS in a WMN. The performance of the improved algorithm was evaluated by using standard network-intrusion datasets such as CICIDS2017 and ADFA-LD. Our simulation results proved that the proposed algorithm had good accuracy, and the improved results were comparable with the standard WOA.

The contributions of this work are summarized as follows,

1. The performance of intrusion detection system was increased by selecting the informative features using improved whale optimization algorithm. It has high attack detection ratio compared to standard WOA.
2. The standard WOA based feature selection method was enhanced by the crossover and mutation operators of genetic algorithm.
3. Finally, the IDS with improved WOA based feature selection method was evaluated with the standard datasets. The proposed method is compared with traditional WOA and other evolutionary algorithms. The comparison results proved that the proposed method has good accuracy and is suitable for IDS in WMN.

The remaining sections of this paper are organized as follows. Section 2 describes the requirement of IDS for WMN. The implementation of the proposed method is elaborately discussed in Section 3. Sections 4, 5, and 6 describe the detailed architecture of the WOA, genetic operators and SVMs, respectively. The simulation of the proposed method is demonstrated in Section 7. Finally, we conclude the paper in Section 8.

II. IDS FOR WMN

The mesh technology provides reliable routing mechanism to all kinds of wireless networks. It inherits the protocol of the corresponding networks for communication, which makes the network vulnerable to various routing-based attacks. The attackers frequently use attacks, such as black hole and worm hole, to damage the network [11]. IDS is a widely used tool that detects the attacks by analyzing the WMN traffic. Based on the architectural model, we can classify IDSs into various types such as monitored, cooperative, and traffic aware [2]. In a monitored IDS, some nodes dedicatedly monitor a specific part of the network. This mechanism increases the implementation cost of the network [12]. Cooperative IDS does not need extra monitoring nodes in which each node has an intrusion detection engine to detect attacks from the network data. It updates automatically by periodically exchanging the audit data with the neighboring nodes [13]. This overhead message exchange decreases the IDS performance. In [14], the author proposed an alternative method that uses a traffic-aware–based IDS instead of a domain-based IDS. This method places the IDS along the routing path for detecting the routing-based attacks. The traffic-based IDS detects the attacks more effectively than other methods. A major drawback of this method is the lack of a common vantage point and the requirement of complete routing knowledge of the network [8].

The classifier plays a dominant role in the IDS that separates the normal data and network attack data. The physical properties of WMNs, such as the noisy environment, quality of nodes, and network quality, reduce the detection ratio of classifiers. In recent times, machine learning algorithms, such as SVMs, extreme learning machines, and artificial neural networks, are being used as classifiers. It has a good intrusion detection rate with a low false positive rate. The performance of each classifier varies with different applications, and the selection of suitable classifiers for a particular application is a complex task.

In [15], an SVM classifier ensemble with the feature augmentation technique was used to improve the performance of the attack detection model. The deep reinforcement learning algorithm was proposed to detect attacks in supervised problems in recent times [16]. The presence of non-informative and redundant features in the collected traffic data reduces the attack detection rate of the classifier. Thus, feature selection algorithms are used to avoid performance degradation of the IDS classifier. In this study, we used the improved WOA to select the informative features for enhancing the performance of the SVM classifier–based IDS.

III. PROPOSED FEATURE SELECTION METHOD FOR IDS

An SVM classifier has a good convergence rate and detection ratio in multiclass classification problems [17]. The classifier has been further enhanced by the improved WOA that selects the most informative features as inputs to the classifier. The conventional WOA has the problem of being confined to the local optimum [18]. Therefore, in the proposed method, the crossover and mutation operators of the GA were added with the traditional WOA algorithm. The crossover operator helps to generate the new population, and the mutation operator is used to avoid the problem of being stuck in the local optimum. The proposed feature selection method is shown in Fig. 1.

In our proposed method, the initial positions of the whales were generated by using randomly selected features of the dataset, which acted as the initial population. We evaluated the fitness of each whale's position in the population by using the SVM algorithm to determine the search agent, that is, the whale position nearest to the prey. The positions of the other whales were updated based on the best solution.

Then, the whales enhanced their positions by using the crossover and mutation operators of the GA. This helped to avoid the search from being stuck in the local optimum and to increase the diversity of the solutions. For effective implementation of the proposed method, the best position of the whale was selected using the elitism method, and the remaining positions were improved by the crossover operator for extending the depth search space. Finally, the mutation operator was applied to explore the solution in a breadth search and avoid the problem of being stuck in the local optimum. The output of the current iteration was used as the input position of the next iteration. The steps were iteratively used until the last iteration to find the informative features set from the available features.

IV. WOA-BASED FEATURE SELECTION ALGORITHM

WOA is a recently developed optimization algorithm based on the hunting behavior of whales. It is practiced using a special method for the successful hunting of prey [19]. This method contains the following three stages:

- Circling hunting
- Bubble-net attacking
- Prey hunting

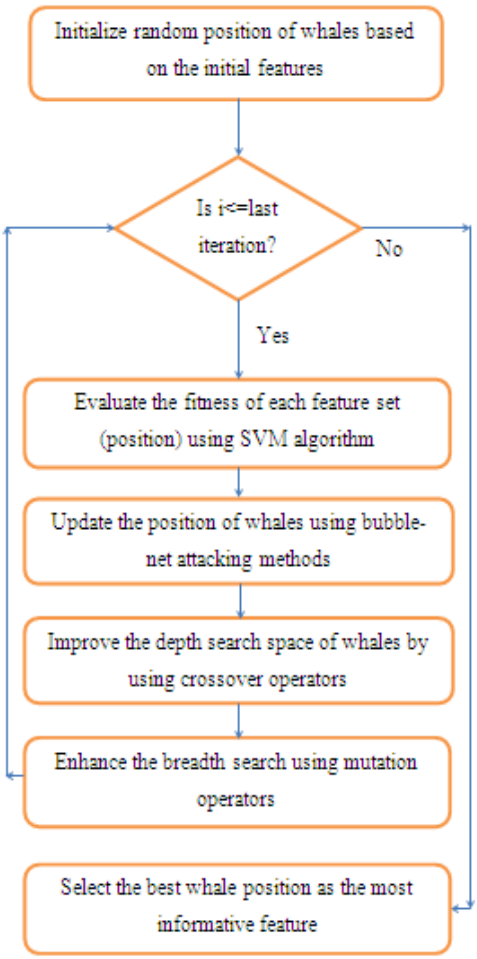

FIGURE 1. Proposed WOA-based feature selection method.

In the circling method, the whales initially set the trap by circling the prey. Then, it selects a search agent based on the individual whale's distance from the prey. After identifying the search agent, all the whales in the group update their positions based on that the search agent. Mathematically, this can be represented as follows:

$$\vec{D} = \left[\vec{C} \cdot \vec{X} * (t) - \vec{X}(t)\right] \tag{1}$$

Here, $\vec{C} = 2 * \vec{r}$, r is a random number between 0 and 1; $\vec{X} - *$ is the local optimal position; $\vec{X}$- Current is the current position; *is* the number of iterations; and is the distance between each whale and the search agent.

Then, the whales engage in the bubble-net attacking method by using the spiral around and spiral update methods [20]. The whales move toward the prey in a spiral manner based on the search agent. The updated position of the other search agents moving toward the best agent can be found by using Eq. (2).

$$\vec{X}(t+1) = \vec{X} * (t) - \vec{A} \cdot \vec{D} \tag{2}$$

$$\vec{A} = 2 \cdot \vec{a} \cdot \vec{r} - \vec{a} \tag{3}$$

Here, a linearly decreases in (2, 0); A is the value between $[-2, +2]$. A is an important parameter used for both the exploration and development phases and is formulated as Eq. (3) [21]. Exploration means finding the search agent. If $A > 1$ or $A == 1$, the search agent position is updated randomly instead of moving toward the current best-solution whale. Now, the random position of the search agent is calculated by using Eq. (4) and Eq. (5).

$$\vec{D} = \left[\vec{C} \cdot \vec{X}_{rand} - \vec{X}\right] \tag{4}$$

$$\vec{X}(t+1) = \vec{X}_{rand} - \vec{A} \cdot \vec{D} \tag{5}$$

Otherwise, the whale starts to change its position toward the search agent in the developmental phase. If $A < 1$, the whale updates the shrink position in the helix-shaped movement toward the prey, which is represented in Eq. (6):

$$\vec{X}(t+1) = \vec{D} \cdot e^{bL} \cdot \cos(2\pi L) + \vec{X} * t \tag{6}$$

Here, $\vec{X}(t+1)$ = Updated is the updated position of the whales; b is a constant that represents the shape of the logarithmic spiral; l is the distance between the whale and the food. $L = -1$ is the shortest distance to the food, and $L = +1$ is the longest distance to the food. Thus, the whales used the shrink-around and shrink-update methods simultaneously to approach the prey. Similarly, in the Zheng *et al.* method [21], the probability of choosing a method for particular situation was assumed to be 50%,and the chance of selecting the path is given by Eq. (7).

$$\vec{X}(t+1) = \begin{cases} \vec{X} * (t) - \vec{A} \cdot \vec{D} & p < 0.5 \\ \vec{D} \cdot e^{bL} \cdot \cos(2\pi L) + \vec{X} * t & P > 0.5 \end{cases} \tag{7}$$

Here, p is a number that is randomly selected between 0 and 1. After the whale reaches the prey, hunting will start.

In the WOA-based feature selection, we used the randomly selected features as the positions of whales, and the fitness of an individual feature subset was found by using learning algorithms. The feature subset having the best solution was considered to be a search agent. Then, the other feature subsets, that is, the whales' positions were updated based on the best feature subset using the bubble-net attacking method. The updated positions were used as the whales' positions for the next iteration; this was repeated until the last iteration to find the most informative features subset. The selected features were used as the input to the classifier that effectively detects the attacks in WMN.

V. GENETIC OPERATORS

GA is a heuristic method used to optimize the outputs in various applications. The input populations are represented in the form of chromosomes; they use three operations (selection, crossover, and mutation) to optimize the parent population of each generation [22]. The crossover operation is used to combine the characters of two parents by selecting the crossover point in the parents. Then, the latter part of first parent chromosome after that point is swapped to the second parent and vice versa; this helps to generate the new population.

A									
1	0	1	0	1	1	0	0	1	1
B									
1	1	1	0	0	1	1	1	0	0

FIGURE 2. Parent chromosome.

Consider two individuals A and B as shown in Fig. 2. Assume that the crossover point is at the 6th position, then both the individuals after the exchange are shown in Fig. 3. The generated child individuals have the characteristics of both parents. The child individual may give better or worst results as compared with the parents' chromosomes.

A									
1	0	1	0	1	1	1	1	1	1
B									
1	1	1	0	0	1	0	0	0	0

FIGURE 3. Child chromosome after the crossover operation.

The mutation operation was used to avoid the problem of being stuck in the local optimum [23]. This operation generated new individuals by altering the nature of some random features. It was applied by flipping the 1s and 0s at random positions into 0s and 1s, respectively. For instance, if a mutation point was applied at position 8 of the A individual, then the mutated chromosome will appear as shown in Fig. 4.

A: Before Mutation									
1		1	1	0	0	1	0	**0**	1	1
A: After Mutation									
1	1	1	0	0	1	0	**1**	1	1

FIGURE 4. Implementation of the mutation operator.

In the GA-based feature selection method, the index of the features is represented as individuals in the form of 0s and 1s. If the feature position is 0, the feature is not selected; otherwise, the feature is included in the feature subset. The initial feature subset is evaluated by using the machine learning algorithms. Then, the crossover and mutation operator are applied to the generated offspring that selects the parents forth next generation. This process is repeated until the condition is satisfied or until last iteration for obtaining the most informative features is reached.

VI. SUPPORT VECTOR MACHINE

An SVM is a supervised machine learning algorithm that classifies the data using a hyperplane hyperplane is generally represented as a linear line in two-dimensional vectors. It is mathematically [24] represented as follows:

$$f(x) = W \cdot x + b$$
$$W \cdot x + b = 0 \tag{8}$$

Here, W is the weight vector, and b is the bias. Many hyperplanes are available between the data, but the selection of a suitable hyperplane will give the maximum accuracy rate. The linear line that classifies the data with the largest margin is selected as a suitable hyperplane. The margins are derived with the help of support vectors and the maximal neighborhood data points. Most of the applications have used the calculation of (2/||W||) as a standard for the global margin.

In case of real-world applications, the data is overlapped and complex. It requires a curve instead of a line for classification [25]. Thus, the data is initially mapped into a multiple feature space and is then classified using a linear line. This mapping and classifying operation was performed using kernel functions. Linear, quadratic, polynomial multilayer perceptron and radial basis function (RBF) are some of the kernel functions widely used in various applications. The RBF kernel has good convergence rate and uses the exponential decay function to determine the margin of the classifier [26]. The mathematical representation of the RBF function to transform the problem into a linear separable form is given as,

$$f(x, y) = e^{(-(||x-y||)^2)/2\sigma^2} \tag{9}$$

where e is the exponential operation; x and y are the data points; ||x − y|| is the Euclidean distance; and σ is the influence distance of each data. Eq. (9) helps to find the support vectors and is then decayed in multiple directions to obtain a suitable hyperplane.

VII. RESULTS AND DISCUSSION

To evaluate the performance of the proposed method, we performed a simulation using Matlab 2014b in a system with a 4-GB RAM and an i3 processor.

A. DATASET

The improved WOA-based feature selection method was analyzed by the standard ADFA-LD and CICIDS2017 datasets. These datasets were being generated based on the properties of the wireless networks. The parameters of the standard datasets were similar to those of WMN, and they were used to evaluate the performance of the mesh network IDS. The hold-out cross-validation method was used to select the training and testing dataset for analyzing the proposed IDS with the SVM classifier.

The ADFA-LD dataset was found in the University of New South Wales repository, and the details of the generated data record are given in [27]. The records in the dataset were categorized into nine network attacks data and a normal class data. In this research, we considered only the general features and the attacks relevant to WMN, that is, 44 features were grouped under six attacks and the normal data class was used for evaluation. The test dataset had normal, exploits, DoS, reconnaissance, generic, and worm data of sizes 795, 116, 21, 156, 911, and 1, respectively. Similarly, the features of the CICIDS2017 dataset had the characteristics of WMN attacks and were represented using 77 features [28]. In this experiment, we analyzed the benign, DoS, portscan, web attack, bot, ftp_parator, and ssh_parator class data were analyzed with the sizes 1788, 838, 779, 22, 996, 160, and 337, respectively.

TABLE 1. Features selected as most informative from the standard datasets.

Dataset	Number of available features	Number of selected features	Index of informative features
CICIDS2017	77	35	3, 5, 7, 12, 13, 14, 17, 18, 19, 21, 23, 27, 29, 30, 31, 32, 34, 35, 38, 39, 41, 46, 51, 52, 53, 56, 57, 58, 63, 64, 67, 69, 71, 73, 76
ADFA-LD	44	25	2, 3, 5, 6, 7, 8, 9, 10, 16, 18, 19, 21, 24, 25, 26, 27, 29, 30, 32, 33, 36, 37, 39, 40, 41

B. INFORMATIVE FEATURE SELECTION

The purpose of the feature selection methods is to select the informative features that improve the intrusion detection rate and reduce the computational complexity of the classifier. The proposed wrapper-based method selects the most informative features for SVM-based IDS. The features found as informative for the classification of normal and intrusion data from the CICIDS2017 and ADFA-LD dataset are given in Table 1.

C. PERFORMANCE ANALYSIS

In this research, RBF was used as a kernel function for SVM-based IDS. The SVM classifier was trained using the training data, which was generated using the informative features selected by the improved WOA. The trained SVM was then assessed with the testing data. The parameters used to assess the performance of classifiers in IDS were the true positive (TP), false positive (FP), false negative (FN), and true negative (TN) rates [29]. The accuracy was used to evaluate the fitness of each individual in the population. The mathematical formula used to calculate the accuracy of the classifier is given as follows:

$$\text{Accuracy} = \frac{(\text{TP} + \text{TN})}{(\text{TP} + \text{FP} + \text{FN} + \text{TN})} \tag{10}$$

Here, TP is the correctly detected real positive data;

TN is the correctly detected real negative data;

FP is the data for positives wrongly detected as negatives; and

FN is the data for negatives wrongly detected as positives.

In this paper, the evaluation metrics of sensitivity, specificity and precision derived from the false positive rates (FPRs) and false negative rates (FNRs) were used to analyze the effectiveness of the proposed IDS; these three metrics are defined as follows:

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{11}$$

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}} \tag{12}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{13}$$

TABLE 2. Simulation results of WOA + genetic operators–based feature selection on the CICIDS2017 dataset.

Attacks	Test Data	True Positive	False Positive	False Negative	Sensitivity	Specificity	Precision
Benign	1788	1693	94	95	0.9469	0.9702	0.9474
DoS	838	833	4	5	0.9940	0.9990	0.9952
Portscan	779	775	4	4	0.9949	0.9990	0.9949
Web attack	22	22	0	0	1.0000	1.0000	1.0000
Bot	996	922	62	74	0.9257	0.9841	0.9370
Ftp-Parator	160	154	6	6	0.9625	0.9987	0.9625
SSh-Parator	377	358	19	19	0.9496	0.9957	0.9496

TABLE 3. Simulation results of WOA + genetic operators–based feature selection on the ADFA-LD dataset.

Attacks	Test Data	True Positive	False Positive	False Negative	Sensitivity	Specificity	Precision
Normal	795	790	5	0	1	0.9955	0.9937
Exploits	116	97	27	67	0.5914	0.9851	0.7822
DoS	21	19	0	0	1	1	1
Reconnaissance	156	90	17	60	0.6	0.9906	0.8411
Generic	911	891	1	12	0.9867	0.999	0.9988
Worms	1	1	38	0	1	0.9803	0.0256

The WOA parameters considered for the evaluation was population size = 20 and search agents = 20. The crossover and mutation values used to improve the positions of the whales were 0.6 and 0.05, respectively. We evaluated the proposed method using various values for the training and testing dataset; these values were collected from multiple datasets. This evaluation mechanism helps to determine the suitability of the proposed method in different environments [30].

The validation results of the proposed system using the CICIDS2017 and ADFA-LD standard datasets are given in Table 2 and Table 3, respectively. The results shown in Table 2 prove the efficiency of the proposed WOA-based feature selection method using genetic operators for detecting network attacks in WMNs. This method detects the normal data and the data on network attacks with low FPRs and FNRs. In Table 3, the IDS detected the normal, DoS, and generic data with low FPRs. However, for exploits and reconnaissance attacks, the classifier had high FPRs and FNRs; this may be due to similarities in both the FPR and FNR attacks on data. A small number of training data may be the reason for high FPRs in the worm data. Therefore, these results show that the improved WOA-based analysis of WMN with parameters of the CICIDS2017 dataset had better detection rate than the ADFA-LD parameters.

Table 4 and 5 show the comparison results of the overall network attack detection ratio of the proposed method using the conventional WOA and the WOA along with the feature selection method based on the mutation operator. The results prove that the features selected by the proposed WOA with the crossover and mutation operators have improved the classification rate as compared with the traditional WOA method.

TABLE 4. Comparison of the proposed method with the conventional WOA method on the CICIDS2017 dataset.

Training data = 4959	Test data =4960
Techniques	Detection Rate (%)
Standard GA	88.85
Standard WOA	93.38
WoA +Mutation	95.54
Proposed method	95.91

TABLE 5. Comparison of the proposed method with the conventional WOA method on the ADFA-LD dataset.

Training data = 2000	Test data = 2000
Techniques	Detection Rate (%)
Standard GA	84.95
WOA	92.65
WOA +Mutation*[r32]*	93.80
Proposed method	94.44

TABLE 6. Comparison of the proposed method with the conventional WOA and WOA + mutation method on the CICIDS2017 dataset based on individual attack detection.

Attacks	Test Data	WoA	WoA + Mutation	Proposed Method
Benign	1788	1714	1693	1693
DoS	838	833	833	833
Portscan	779	775	775	775
Web attack	22	22	22	22
Bot	996	856	903	922
FTP-Parator	160	100	154	154
SSH- Parator	377	356	359	358

TABLE 7. Comparison of the proposed method with the conventional WOA method on the ADFA-LD dataset on individual attack detection.

Attacks	Test Data	WoA	WoA + Mutation	Proposed Method
Normal	795	790	793	793
Exploits	116	81	97	97
DoS	21	5	9	13
Reconnaissance	156	89	94	90
Generic	911	882	876	891
Worms	1	1	1	1

The individual attack detection rate of the proposed method and the traditional WOA based feature selection method on the CICIDS2017 and ADFA-LD datasets are given in Tables 6 and 7, respectively. The results show that the proposed method has good detection rate on the normal data in all the datasets. It also has good accuracy for the detection of most attacks in both the datasets except for attacks such as DoS and reconnaissance of the ADFA-LD dataset. This may be due to the dominance of the informative features of classes having a large number of datasets over the classes having a small number of datasets.

The proposed method was further validated using a statistical technique called K-fold cross validation for estimating the prediction error. The result in Table 8 proves that the modified

TABLE 8. Validation of the proposed method with the K-fold cross-validation technique.

Dataset	Detection Rate (%)	
	GA	Proposed Method
CICIDS2017	88.85±1.93	95.91±1.82
ADFA-LD	84.95±2.56	94.44±1.26

TABLE 9. Training and testing time of the classifier on CICIDS2017 datasets.

IDS technique	CICIDS2017 Dataset		ADFA-LD Dataset	
	Training Time	Testing Time	Training Time	Testing Time
GA	65.76–83.24	0.705–1.236	42.56–58.43	0.0156–0.078
WOA	25.38–31.76		6.474–7.7688	
WOA + Mutation	25.11–31.03		6.7392–8.19	
WOA + Genetic operators	19.09–27.19		6.3648–7.432	

WOA algorithm has good detection rate and is suitable for the detection of attacks in wireless networks.

The execution time of the classifier is another parameter used to analyze the performance of the classifier. In IDS, the time taken by the classifier for training and testing with the selected features is considered. The training of the classifier usually requires more time than testing, and the training is more importance than the testing [31]. Table 9 shows the training and testing time of the classifier using the proposed feature selection method. The result shows that the proposed method takes more time for training because of the selection of more informative features; however, this is not a major drawback in the offline processing system.

FIGURE 5. Performance analysis of the proposed IDS over the ANN and RF classifiers using the CICIDS2017 dataset.

The suitability of the SVM classifier for the proposed feature selection method was evaluated by comparing with the artificial neural networks (ANNs) and Random forest (RF) learning algorithms. ANN is a standard classifier used in various classification problems for more than a decade, and RF classifier was recently used in multiple applications. Figs. 5 and 6 show the performance analysis of the classifier using the CICIDS2017 and ADFA-LD dataset, respectively. The figures clearly indicate that the performance of the SVM

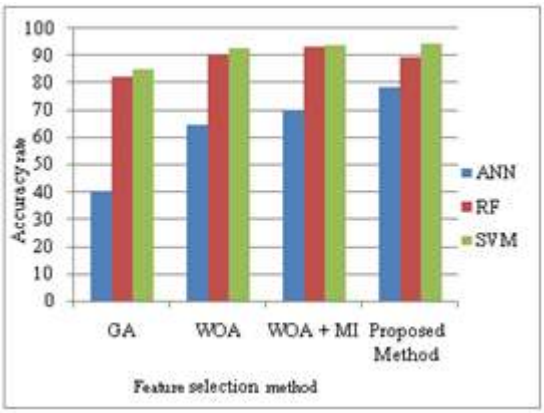

FIGURE 6. Performance analysis of the proposed IDS over the ANN and RF classifiers using the ADFA-LD dataset.

TABLE 10. Computational complexity analysis of the proposed and conventional feature selection methods.

IDS technique	Time complexity
IDS with conventional WOA	O(P×N) P: Position of the whales N: Size of the whale group
IDS with WOA + Mutation	O(P×N) + O(MT) MT: Time taken for mutation operation
IDS with WOA + Mutation + Crossover	O(P×N) + O(CT) +O(MT) CT: Time taken for the crossover operation

classifier was much better than that of ANN and a little superior to that of the RF classifier.

Computational complexity helps to analyze the processing overhead of the proposed method [32]. Table 10 shows that the computational complexity of the IDS with the proposed WOA-based feature selection method is greater than that of the conventional WOA because of using genetic operators such as crossover and mutation. The mutation time includes the random mutation point selection time and bit conversion time. However, the crossover time includes the crossover point selection and the swapping operation time. This time delay has not influenced the performance of the proposed IDS in high processing systems.

Thus, from the above-mentioned results, we verified that the proposed feature selection with the SVM classifier detects network attacks with high accuracy. As compared with the ANN and RF classifiers, the SVM classifier is more suitable for the modified WOA algorithm. An analysis of the proposed WOA with the training time demonstrates its efficiency in the IDS. Even though, the attack detection rate was improved by using the proposed method, the training time still needs to be reduced for the WMN-based highly sensitive networks.

VIII. CONCLUSION

In this paper, we proposed a novel feature selection method using WOA and genetic operators for improving the performance of IDS in WMNs. The proposed method selects the most informative features from the network data. The selected informative features help to improve the accuracy of the SVM-based IDS. We evaluated the performance of the

proposed method by using standard network-intrusion datasets such asCICIDS2017 and ADFA-LD. A comparison of the proposed method with the conventional WOA and WOA along with the mutation operator on the basis of detection rate, execution time, and computational complexity proved the efficiency of the proposed method. The results clearly indicated that the inclusion of the crossover and mutation operators enhances the global search space of the whales and overcomes the problem of being stuck in the local optimum. The simulation results proved the suitability of the improved WOA-based feature selection method for the IDS in WMN.

In the future, we planned to further optimize the modified WOA using filter-based feature selection methods such as information gain. As an extension, the performance of the developed method was analyzed in highly sensitive real-time networks such as smart meter communications.

REFERENCES

[1] R. Hou, K.-S. Lui, F. Baker, and J. Li, "Hop-by-Hop routing in wireless mesh networks with bandwidth guarantees," *IEEE Trans. Mobile Comput.*, vol. 11, no. 2, pp. 264–277, Feb. 2012.
[2] A. Hassanzadeh, Z. Xu, R. Stoleru, G. Gu, and M. Polychronakis, "PRIDE: A practical intrusion detection system for resource constrained wireless mesh networks," *Comput. Secur.*, vol. 62, pp. 114–132, Sep. 2016.
[3] M. Montazeri, M. Montazeri, H. R. Naji, and A. Faraahi, "A novel memetic feature selection algorithm," in *Proc. 5th Conf. Inf. Knowl. Technol.*, May 2013, pp. 295–300.
[4] M. A. Hall, "Correlation-based feature selection for machine learning," Ph.D. dissertation, Dept. Comput. Sci., Univ. Waikato, Hamilton, New Zealand, 1999.
[5] S. Sharmin, M. Shoyaib, A. A. Ali, M. A. H. Khan, and O. Chea, "Simultaneous feature selection and discretization based on mutual information," *Pattern Recognit.*, vol. 91, pp. 162–174, Jul. 2019.
[6] L. Cao, L. Xu, and E. D. Goodman, "A guiding evolutionary algorithm with greedy strategy for global optimization problems," *Comput. Intell. Neurosci.*, vol. 2016, pp. 1–10, May 2016.
[7] C. Huang and C. Wang, "A GA-based feature selection and parameters optimizationfor support vector machines," *Expert Syst.*, vol. 31, pp. 231–240, Aug. 2006.
[8] R. Vijayanand, D. Devaraj, and B. Kannapiran, "A novel intrusion detection system for wireless mesh network with hybrid feature selection technique based on GA and MI," *J. Intell. Fuzzy Syst.*, vol. 34, no. 3, pp. 1243–1250, Mar. 2018.
[9] Y. Gao and J. Liu, "A new differential evolution algorithm with random mutation," in *Emerging Intelligent Computing Technology and Applications. With Aspects of Artificial Intelligence* (Lecture Notes in Computer Science), vol. 5755. Berlin, Germany: Springer, 2009, pp. 209–214.
[10] R. Vijayanand, D. Devaraj, and B. Kannapiran, "Intrusion detection system for wireless mesh network using multiple support vector machine classifiers with genetic-algorithm-based feature selection," *Comput. Secur.*, vol. 77, pp. 304–314, Aug. 2018.
[11] S. Alanazi, K. Saleem, J. Al-Muhtadi, and A. Derhab, "Analysis of denial of service impact on data routing in mobile eHealth wireless mesh network," *Mobile Inf. Syst.*, vol. 2016, pp. 1–19, Jul. 2016.
[12] D. Subhadrabandhu, S. Sarkar, and F. Anjum, "A framework for misuse detection in ad hoc networks-part i," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 2, pp. 274–289, Feb. 2006.
[13] A. Hassanzadeh and R. Stoleru, "Towards optimal monitoring in cooperative IDS for resource constrained wireless networks," in *Proc. 20th Int. Conf. Comput. Commun. Netw. (ICCCN)*, Jul./Aug. 2011, pp. 1–8.
[14] V. Sekar, R. Krishnaswamy, A. Gupta, and M. K. Reiter, "Network-wide deployment of intrusion detection and prevention systems," in *Proc. 6th Int. Conf. (Co-NEXT)*, 2010, pp. 1–12.
[15] J. Gu, L. Wang, H. Wang, and S. Wang, "A novel approach to intrusion detection using SVM ensemble with feature augmentation," *Comput. Secur.*, vol. 86, pp. 53–62, Sep. 2019.
[16] M. Lopez-Martin, B. Carro, and A. Sanchez-Esguevillas, "Application of deep reinforcement learning to intrusion detection for supervised problems," *Expert Syst. Appl.*, vol. 141, Mar. 2020, Art. no. 112963.
[17] Y. Zhang, L. Wang, W. Sun, R. C. Green II, and M. Alam, "Distributed intrusion detection system in a multi-layer network architecture of smart grids," *IEEE Trans. Smart Grid*, vol. 2, no. 4, pp. 796–808, Dec. 2011.
[18] Y. Zheng, Y. Li, G. Wang, Y. Chen, Q. Xu, J. Fan, and X. Cui, "A novel hybrid algorithm for feature selection based on whale optimization algorithm," *IEEE Access*, vol. 7, pp. 14908–14923, 2019.
[19] S. Mirjalili and A. Lewis, "The whale optimization algorithm," *Adv. Eng. Softw.*, vol. 95, pp. 51–67, May 2016.
[20] M. Azizi, R. G. Ejlali, S. A. Mousavi Ghasemi, and S. Talatahari, "Upgraded whale optimization algorithm for fuzzy logic based vibration control of nonlinear steel structure," *Eng. Struct.*, vol. 192, pp. 53–70, Aug. 2019.
[21] H. Zheng, L. Tang, C. Yang, and S. Lan, "Locating electric vehicle charging stations with service capacity using the improved whale optimization algorithm," *AdvancedEng. Informat.*, vol. 42, Aug. 2019, Art. no. 100901.
[22] D. E. Goldberg, *Genetic algorithms in Search, Optimization, and Machine Learning*. Reading, MA, USA: Addison-Wesley, 1989.
[23] P. S. Oliveto, T. Paixão, J. Pérez Heredia, D. Sudholt, and B. Trubenová, "How to escape local optima in black box optimisation: When non-elitism outperforms elitism," *Algorithmica*, vol. 80, no. 5, pp. 1604–1633, May 2018.
[24] C. Cortes and V. Vapnik, "Support-vector networks," *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, 1995.
[25] S. K. Biswas and M. M. A. Mia, "Image reconstruction using multilayer perceptron and support vector machine classifier and study of classification accuracy," *Int. J. Sci. Technol. Res.* , vol. 4, no. 2, pp. 226–231, 2015.
[26] A. Rahimi and B. Recht, "Random features for large-scale kernel machines," in *Proc. Neural Inf. Process. Syst., Int. Conf. ACM*, 2007, pp. 1177–1184.
[27] G. Creech and J. Hu, "Generation of a new IDS test dataset: Time to retire the KDD collection," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Apr. 2013, pp. 4487–4492.
[28] I. Sharafaldin, A. H. Lashkari, and A. A. Ghorbani, "Toward generating a new intrusion detection dataset and intrusion traffic characterization," in *Proc. 4th Int. Conf. Inf. Syst. Secur. Privacy*, 2018, pp. 108–116.
[29] A. Shenfield, D. Day, and A. Ayesh, "Intelligent intrusion detection systems using artificial neural network," *ICT Express*, vol. 4, no. 2, pp. 95–99, 2018.
[30] B. Mukherjee, L. T. Heberlein, and K. N. Levitt, "Network intrusion detection," *IEEE Netw.*, vol. 8, no. 3, pp. 26–41, May 1994.
[31] X.-S. Gan, J.-S. Duanmu, J.-F. Wang, and W. Cong, "Anomaly intrusion detection based on PLS feature extraction and core vector machine," *Knowl.-Based Syst.*, vol. 40, pp. 1–6, Mar. 2013.
[32] H. H. Soliman, N. A. Hikal, and N. A. Sakr, "A comparative performance evaluation of intrusion detection techniques for hierarchical wireless sensor networks," *Egyptian Inf. J.*, vol. 13, no. 3, pp. 225–238, Nov. 2012.

R. VIJAYANAND received the B.E. degree in computer science and engineering from the PTR College of Engineering and Technology, Madurai, the M.Tech. degree in computer science and engineering from Kalasalingam University, and the Ph.D. degree in network and smart meter communication, in 2018. He is currently an Assistant Professor with the Department of Computer Science and Engineering, J.B. Institute of Engineering and Technology, Hyderabad, India.

D. DEVARAJ (Senior Member, IEEE) received the B.E. degree in electrical and electronics engineering and the M.E. degree in power system engineering from the Thiagarajar College of Engineering, Madurai, in 1992 and 1994, respectively, and the Ph.D. degree from IIT Madras, in 2001. He guided more than 22 Ph.D. scholars. He is currently a Professor with the Department of Electrical and Electronics Engineering, Kalasalingam University, India. His research interests include power system security, voltage stability, smart grid, and evolutionary algorithm.