electronics

Article

Application of Improved Butterfly Optimization Algorithm Combined with Black Widow Optimization in Feature Selection of Network Intrusion Detection

Hui Xu *, Yanping Lu and Qingqing Guo

School of Computer Science, Hubei University of Technology, Wuhan 430068, China
* Correspondence: xuhui@hbut.edu.cn

Abstract: Feature selection is a very important direction for network intrusion detection. However, current feature selection technology of network intrusion detection has the problems of low detection rate and low accuracy due to feature redundancy. An improved Butterfly Optimization Algorithm combined with Black Widow Optimization (BWO-BOA) is proposed in this paper, which introduces a dynamic adaptive search strategy in the global search phase of the Butterfly Optimization Algorithm (BOA), uses the movement search process of Black Widow Optimization (BWO) algorithm as the local search, and at the same time, in order to overcome the improved butterfly optimization algorithm easily falling into a local optimum in local search phase, takes advantage of the small probability mutation strategy to filter out the redundant features. This paper then tries to apply the proposed BWO-BOA algorithm to feature selection of network intrusion detection. In order to verify the performance of the proposed BWO-BOA algorithm, the UNSW-NB15 dataset is selected for binary classification and multi-classification simulation experiments, and the feature selection models of BWO-BOA algorithm, BOA algorithm, BWO algorithm, Particle Swarm Optimization, Salp Swarm Algorithm, Whale Optimization Algorithm and improved Butterfly Optimization Algorithm are compared for validation. The experimental results show that the proposed BWO-BOA algorithm can enhance the performance of the feature selection model in network intrusion detection and significantly boost the reduction of feature dimensions.

Keywords: network security; intrusion detection; feature selection; butterfly optimization algorithm; black widow optimization

check for updates

Citation: Xu, H.; Lu, Y.; Guo, Q. Application of Improved Butterfly Optimization Algorithm Combined with Black Widow Optimization in Feature Selection of Network Intrusion Detection. *Electronics* **2022**, *11*, 3531. https://doi.org/10.3390/electronics11213531

Academic Editor: Martin Reisslein

Received: 17 September 2022
Accepted: 19 October 2022
Published: 29 October 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

In recent years, the frequent occurrence of network security incidents, the diversification of intrusion methods and the increasing frequency of intrusions led to the challenges of low accuracy of abnormal traffic detection and identification and classification of existing network intrusion detection technologies. Network intrusion detection models are based on intrusion detection algorithms and network datasets [1]. Network datasets usually include false positives, irrelevant and redundant features, which not only slow down the detection speed, but also consume a lot of computing resources. Feature selection is the process of selecting the most relevant features that help build robust models [2–4]. It is not only used in breast cancer detection and coronary heart disease detection, but is also an important step for data preprocessing in intrusion detection [5–8]. Although it reduces processing costs and minimizes storage space, it faces significant challenges in terms of dimension disasters and classification accuracy [9].

Swarm intelligence optimization is widely used for feature selection in intrusion detection systems because of its high accuracy [10]. In this context, swarm intelligence is an important technology for implementation and classification. A variety of intelligent optimization algorithms are used in feature selection of network intrusion detection, such as

the Moth Flame Optimization (MFO) algorithm [11], Ant Colony Optimization (ACO) algorithm [12], Cuckoo Search (CS) [13], Butterfly Optimization Algorithm (BOA) [14,15], Firefly Algorithm (FA) [16–19], Krill Herd Algorithm [20], Sparrow Search Algorithm (SSA) [21], Artificial Bee Colony (ABC) algorithm [22], Salp Swarm Algorithm (SSA) [23] and Gray Wolf Optimization (GWO) algorithm [24], etc. However, due to the limitations of a single swarm intelligence algorithm, there are many measures to improve the swarm intelligence algorithm, such as changing the initialization mode to increase randomness, using mutations to accelerate the convergence speed or making a fine-grained search strategy [25–28]. At the same time, an increasing number of scholars choose a hybrid algorithm, which synergizes the characteristics of different algorithms; for example, Mojtahedi proposed that the application of a combined genetic algorithm and whale optimization algorithm in feature selection to enhance the accuracy of network intrusion detection [29]. Yuan et al. put forward the combination of a genetic algorithm and improved ant colony algorithm for feature extraction [30]. Xu et al. fused cuckoo search and the gray wolf optimization algorithm to increase the global search ability of feature selection for network intrusion detection [31]. Kang et al. proposed the application and feature selection problems of a hybrid improved flower pollination algorithm and gray wolf algorithm [32]. Therefore, in the field of feature selection, combination algorithms perform better than single algorithms.

Based on the idea of applying these algorithms to improve the network intrusion detection system, this paper selects the Butterfly Optimization Algorithm (BOA) [33] and Black Widow Optimization (BWO) algorithm [34], which have complementary advantages. The BOA has a simple structure and easy realization, but it also has the problems of slow convergence and easily falling into the local optimum. The BWO algorithm is linear and spiral in the way of motion, which can carry out a fine-grained search to prevent the algorithm from falling into the local optimum. This paper proposes a new algorithm to improve the Butterfly Optimization Algorithm combined with Black Widow Optimization (BWO-BOA), according to the characteristics of the BOA algorithm and the BWO algorithm. The new algorithm can improve the problems of easily to falling into local optima and slow convergence, and thus has some advantages in the feature selection model.

The main contributions of this paper are as follows: The BWO-BOA algorithm is proposed. Based on the BOA framework, the BWO-BOA algorithm makes the following improvements. Firstly, the BOA is improved by using the dynamic adaptive search strategy, the search strategy of the BWO algorithm and small probability mutation strategy to enhance the convergence speed and the global optimization ability of the algorithm. Then, the intrusion detection model of the improved algorithm is tested with the UNSW-NB15 dataset. Simulation results and experimental analysis verify the effectiveness of the proposed model.

2. Basic Algorithms

2.1. Butterfly Optimization Algorithm (BOA)

The BOA is derived by simulating a butterfly to analyze the odors in the air to locate a food source. Each butterfly has a different fragrance. Butterflies can smell and analyze the fragrance of other butterflies in the air to determine the direction of movement to the global optimal position. The intensity of the fragrance depends on the fitness of the butterfly itself, and the fitness of the butterfly constantly changes during its movement. The expression for fragrance concentration is shown in Equation (1).

$$f_i = cI^a, \tag{1}$$

where f_i is the fragrance concentration and c is the perceptual form. I is the stimulus intensity and the I value is determined by the fitness of the current butterfly individual. a represents the power index of the dependent perception form.

Due to the mutual attraction of the fragrances between the butterflies, the location of the best butterfly is the one with the largest fragrance. If the fragrance concentration is strong, the butterflies move towards the best location. This phase is called global search.

Conversely, when the butterfly is not sensitive to the smell, the direction of movement is not specified, this phase is called local search. In the process of population movement, the conversion probability p determines the search phase. In each iteration, *rand* is generated and compared with the conversion probability p. If the *rand* is lower than p, a global search is performed; otherwise, a local search is performed.

In the global search phase, the butterfly moves towards the global optimal position, and the updated position is shown in Equation (2).

$$x_i^{t+1} = x_i^t + \left(r^2 \times g^* - x_i^t\right) \times f_i, \tag{2}$$

where x_i^t is the position of the i-th butterfly in the t-th iteration, the range of r is [0, 1], x_i^{t+1} is the updated butterfly position, g^* indicates that the position of the butterfly is the global optimum and f_i is the intensity of fragrance emitted by the i-th butterfly.

During the local search phase, the butterfly randomly moves its position, and the position update formula is as follows.

$$x_i^{t+1} = x_i^t + (r^2 \times x_j^t - x_k^t) \times f_i, \tag{3}$$

where x_j^t and x_k^t stand for two butterflies randomly selected from the same population. The r is in the range of [0, 1].

2.2. Black Widow Optimization (BWO) Algorithm

The BWO algorithm is a swarm intelligence algorithm proposed by Peña-Delgado et al. in 2020, which is inspired by mating behavior of black widow spiders. The western black widow spider is a poisonous spider found from western Canada to southern Mexico. The body of female black widows contains a powerful neurotoxin. In addition, the neurotoxin is one of the most dangerous to humans, as a single bite can lead to death. Black widow spiders feed on insects such as cockroaches, beetles and butterflies; they weave webs in trees and inhabit forests and swamps. Males use sex pheromones to identify mating patterns of females, and they are not interested in hungry or malnourished females, because females exhibit cannibalism. Black widow spiders have two strategies: the movement strategy and the pheromone strategy.

Movement strategy: the movement of a black widow spider on a web is simulated as linear and spiral, using the optimal position of the current black widow spider to move in a linear and spiral manner, with a conversion probability of 0.3.

Pheromone strategy: after normalized pheromone processing, if the pheromone is below or equal to 0.3, the spider is replaced by other spiders.

3. Proposed BWO-BOA Algorithm

The BWO-BOA algorithm improves three areas based on the BOA framework. Firstly, a dynamic adaptive search strategy is introduced to improve the global search ability of the BOA algorithm, which balances the whole search ability of the global search. Then, the local search phase is improved by fusing the movement strategy of the BWO algorithm to search more precisely. Finally, the small probability mutation strategy and the pheromone strategy of the BWO algorithm are used to update the location of the algorithm to avoid the algorithm falling into the local optimum, which improves the global search ability of the algorithm.

3.1. Dynamic Adaptive Search Strategy

Due to the slow convergence speed and poor precision of the algorithm when the butterfly searches for the optimal solution, this paper randomly distributes the population every time and balances the global and local search ability of the algorithm by improving the fragrance formula and introducing the inertia weight. Since the improved fragrance formula

and the inertia weight formula dynamically adapt the search ability of the algorithm, this paper is collectively called the dynamic adaptive search strategy.

In the BOA, the intensity of fragrance directly affects the range of the butterfly search, and a change of I in the fragrance intensity calculation formula is determined by the value of fitness; the curve slope of the fitness affects the convergence speed and the solution precision of the algorithm. In order to make the global search more efficient, the formula for fragrance concentration is improved by using an improved fragrance formula search strategy, such as (4).

$$f = rand \times \left(4 - \frac{4t}{N_{iter}}\right) + \frac{2t}{N_{iter}} - 2, \tag{4}$$

where $rand$ is in the range of [0, 1]. In the iterative process, f can dynamically adapt the search ability of the early and late phases of the algorithm, help the algorithm to find the global optimum quickly and improve the solving ability of the algorithm.

Although the overall convergence ability of the algorithm is enhanced after improving the fragrance concentration formula, the accuracy and convergence ability of the algorithm still cannot reach the expectation in the optimization process. Hence, this paper introduces the inertia weight search strategy to balance the global and local search ability of the algorithm and strengthen the optimization and convergence ability of the algorithm. Adaptive hybrid inertia weights were introduced in [35,36] to improve the search capability of the algorithm in the early and late phases, and nonlinear inertia weights were introduced in [37,38]. This paper introduces the inertia weight search strategy to dynamically adapt the algorithm to the overall search ability, using the inertia weight ω calculation as shown in Equation (5).

$$\omega(t) = 2 \times exp\left(-\left(4 \times \frac{t}{N_{iter}}\right)^2\right), \tag{5}$$

where exp is an exponential function. The ω has a wide change range and big change of slope; a search in a wide range in the initial phase and the range is smaller and smaller with the number of iterations changing after introducing ω, which is helpful to enhance the convergence ability of the algorithm. The global search position update formula after introducing ω is as Equation (6).

$$x_i^{t+1} = \omega(t) \times x_i^t + (r^2 \times g^* - x_i^t) \times f \tag{6}$$

3.2. Black Widow Search Strategy

This paper uses the movement strategy of the BWO algorithm to improve the local search formula of the BOA and uses a 0.3 conversion probability to decide whether the local search should be a linear search or spiral search. The conversion probability p is changed to 0.5, which makes the local search process more precise, thus greatly improving the local development ability and convergence ability of the algorithm. The local position update formula after improving is as follows.

$$x_i^{t+1} = \begin{cases} g^* - \mu x_l^t, & rand > p1 \\ g^* - cos(2\pi\theta) x_i^t, & others \end{cases}, \tag{7}$$

where $p1$ is 0.3, the variable μ is a floating-point number randomly generated in [0.4, 0.9], l is an individual randomly selected from the population (l not equal i) and θ is a random floating-point number from [−1.0, 1.0].

Although this paper uses the movement search strategy of the BWO algorithm to improve the local search ability of the BOA, only the movement strategy makes the algorithm easily fall into the local optimal. Genetic algorithms have individual variation; the probability of a small mutation can prevent the phenomenon of convergence of the algorithm too early [39]. In this paper, $p2$ is used to represent the probability of variation, the value of $p2$ is 0.1, so there is a 90% probability of no mutation in the algorithm, and the

remaining 10% is the probability of mutation; this strategy is called the small probability mutation strategy.

This paper introduces a small probability mutation strategy, which helps move out of the local optimal to the global optimal solution and improves the ability of local convergence. The exploration mode of the local search phase is determined by $p2$, using the low pheromone substitution strategy of the pheromone strategy in the BWO algorithm to update the butterfly position. The position after mutation is as follows.

$$x_i^{t+1} = g^* + \frac{1}{2}\left[x_j^t - (-1)^\sigma \times x_k^t\right], \tag{8}$$

where σ is a random binary number randomly generated, $\sigma \in \{0, 1\}$.

3.3. Pseudo-Code of BWO-BOA Algorithm

The pseudocode for BWO-BOA is shown in Algorithm 1.

Algorithm 1: BWO-BOA pseudocode.

Initialize population N, iteration times T, Dimension D, upper and lower bounds and parameters.
Calculate the fitness of each butterfly and record the optimal position.
while (t<T)
  for i<N
    Calculate f and $\omega(t)$, respectively, using Equations (4) and (5).
    if rand<p
      Update the global position with Equation (6).
    else
      if rand<$p2$
        Update local position with Equation (7).
      else
        The position of the mutation is updated with Equation (8).
      end if
    end if
  end for
end while
Output the optimal value.

The BWO-BOA is based on improving the BOA framework, which is mainly divided into three steps. In the first step, the dynamic adaptive search strategy is introduced into the BOA, which utilizes the improved fragrance formula and inertia weight to dynamically adapt the search ability of the BOA in the early and late phases, causing the algorithm to achieve a balanced and coordinated overall search ability. The second step is to update the position of the butterfly in the local search process of the improved BOA by incorporating a movement strategy of the BWO, which makes the local search process more precise and changes the conversion probability p from 0.8 to 0.5, to increase the probability of the local search process, so as to increase the ability of local search process. Lastly, in order to avoid the algorithm falling into the local optimum and improve the local search ability of algorithm, this paper uses a small probability mutation probability and utilizes the low pheromone substitution strategy of the BWO to update the position of the mutated butterfly. The flow chart of the BWO-BOA algorithm is shown in Figure 1.

In conclusion, after three steps of improvement, the BWO-BOA algorithm effectively enhances the search ability, improves the ease with which the original BOA fall into the local optimal and balances the search process of the earlier and later phases; the overall performance of the algorithm has been greatly improved.

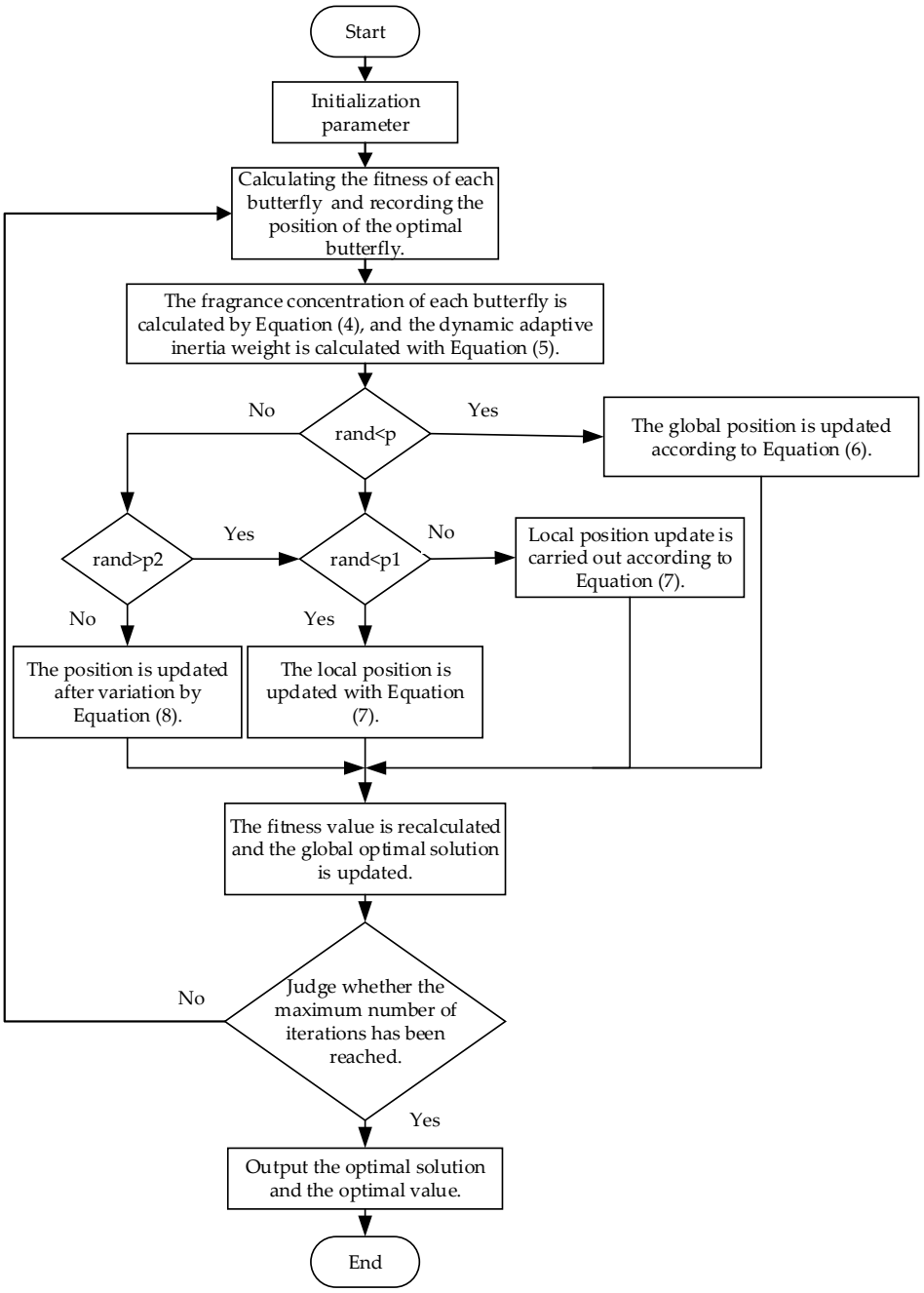


Figure 1. Flow chart of proposed BWO-BOA algorithm.

4. Feature Selection Model Based on BWO-BOA Algorithm for Network Intrusion Detection

Feature selection is a kind of preprocessing technology to remove the irrelevant, noisy and redundant feature data, mainly by finding the best feature subset from the original feature set to reduce the dimensionality of data processing, decrease the computational pressure of storage and enhance the classification performance of the model.

4.1. Fitness Function

Since the number of selected features is not as small as possible, to ensure the best effect of the feature selection model, the fitness function should effectively integrate the classification accuracy and the number of selected features. To balance the highest possible classification accuracy with the lowest possible number of features, the fitness function is set as follows.

$$fitness = \alpha \times error + \beta \times \frac{num_feat}{max_feat}, \tag{9}$$

$$error = 1 - Accuracy, \tag{10}$$

where the *error* represents the classification error rate and α and β denote the weight of the error rate and the feature subset, respectively. In this paper, α is taken as 0.99 and $\beta = 1 - \alpha$. num_feat is the length of the selected feature and max_feat is the total length of the dataset features.

4.2. Evaluation Indicators

In this paper, the accuracy rate (*Acc*), precision rate (*Pre*), recall rate (*Rec*) and F1 score ($F1$) are used as the indicators of feature selection model evaluation. *Acc* refers to the prediction accuracy of the classification of sample data, *Pre* refers to the ratio of the data with positive results to the data predicted to be positive, *Rec* refers to the ratio of the predicted positive samples to the actual positive results and the $F1$ refers to both the precision rate and the recall rate so that a balance between the two can obtain an optimal result. The evaluation indicators are calculated as follows.

$$Acc = \frac{TP + TN}{TP + TN + FP + FN}, \tag{11}$$

$$Pre = \frac{TP}{TP + FP}, \tag{12}$$

$$Rec = \frac{TP}{TP + FN}, \tag{13}$$

$$F1 = \frac{2 \times Pre \times Rec}{Pre + Rec}, \tag{14}$$

where TP means that the positive class of the sample is predicted to be a positive class (True Positive), TN means that the negative class of the sample is predicted to be a negative class (True Negative), FP means that the negative class of the sample is predicted to be a positive class (False Positive) and FN means that the positive class of the sample is predicted to be a negative class (False Negative).

4.3. Proposed Feature Selection Model Based on BWO-BOA Algorithm

As for network intrusion detection, the proposed BWO-BOA algorithm is utilized for feature selection, which not only ensures the accuracy of feature selection and extraction of key features, but also reduces the interference of redundant features. The feature selection model based on the BWO-BOA algorithm is shown in Figure 2.

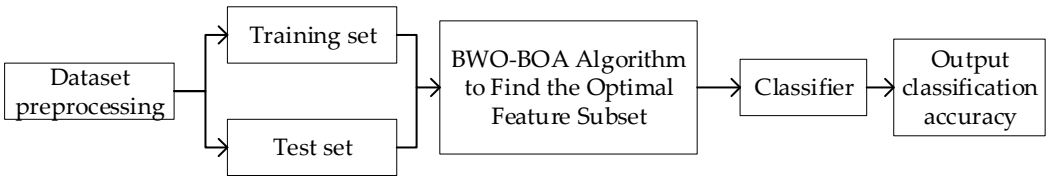


Figure 2. Feature selection model based on BWO-BOA algorithm.

As is indicated in Figure 2, the steps of feature selection based on the BWO-BOA algorithm are as follows.

Step1: Data preprocessing is performed on the test and training sets of the dataset.

Step2: Initialize the population and related parameters of the BWO-BOA. Use the evaluation function fitness to find out the feature subset selected by the optimal individual in the BWO-BOA algorithm.

Step3: Input the optimal feature subset into the classifier for classification.

Step4: Output the *Acc*, *Pre*, *Rec*, *F*1 and selected feature subset of the model.

5. Experimental Results

The simulation experiment environment in this paper is a 64-bit Windows 10 operating system, the main frequency of the machine is 3.30 GHz, the memory is 16 GB, and the algorithm is implemented by using the Sklearn library in Python 3.10.

For the sake of verifying the effectiveness of the feature selection model based on the BWO-BOA algorithm proposed in this paper, this paper uses the UNSW-NB15 dataset to conduct simulation experiments on the KNN classifier [25,40–42]. The average classification accuracy, average classification precision, average recall rate, average F1-score and average optimal feature subset are evaluated. In order to fully verify the effectiveness of the algorithm in this paper, the BWO-BOA algorithm is compared with the original BOA, BWO algorithm, Particle Swarm Optimization (PSO) [43], Salp Swarm Algorithm (SSA) [44], Whale Optimization Algorithm (WOA) [45] and the improved butterfly optimization algorithm using Gaussian mutation and dynamic variance (IBOA) [27]. In the simulation experiments, the population size was set to 20, the maximum number of iterations was 50, and each feature selection model was independently run 20 times on the dataset. The rest of the algorithm parameters are shown in Table 1.

Table 1. Parameter settings of seven algorithms.

Algorithm	Parameter
BWO-BOA	$p = 0.5$, $p1 = 0.1$, $p2 = 0.3$
BOA	$p = 0.8$, $a = 0.1$, $c = 0.01$
BWO	$m = 0.4 \sim 0.9$, $\beta = -1 \sim 1$
PSO	$W = 0.9$, $c1 = c2 = 3$
SSA	$r1$, $r2$, $r3 = 0 \sim 1$
WOA	$b = 1$, $C = 0 \sim 2$, $l = -1 \sim 1$
IBOA	$p_{max} = 0.8$, $p_{min\ =\ 0.3}$, $a = 0.1$, $c = 0.01$, $\sigma_{max} = 1.5$, $\sigma_{min} = 0.4$, $pg = 0.5$

5.1. Dataset

The dataset for this experiment is the UNSW-NB15 dataset, the original data of which were created by the Ixia PerfectStorm tool, and the file type is saved in CSV format. The dataset contains 10 behavior categories, one of which is normal behavior and the other nine are attack behavior; the nine types of attacks are Fuzzers, Analysis, Backdoors, Dos, Exploits, Generic, Reconnaissance, Shellcode and Worms [46]. The dataset was divided into the training set and test set. There are 175,341 data in the training set and 82,332 in the test set. Since it lacks Shellcode attack, there are only eight attack types, and each data point has 45 features, including id and its label. The specific feature names are shown in Table 2.

Table 2. Feature classification table.

Type	Feature
Object	proto, service, state, attack_cat
Integer	spkts, dpkts, sbytes, dbytes, sttl, dttl, sload, dload, swin, stcpb, dtcpb, dwin, smean, dmean, trans_depth, response_body_len, ct_srv_src, ct_state_ttl, ct_src_dport_ltm, ct_dst_ltm, ct_ftp_cmd, ct_flw_http_mthd, ct_src_ltm, ct_dst_sport_ltm, ct_dst_src_ltm, ct_srv_dst
Float	dur, rate, sloss, dloss, sinpkt, dinpkt, sjit, djit, tcprtt, synack, ackdat
Binary	is_sm_ips_ports, is_ftp_login, label

The original dataset contains data that affect the effect of simulation experiments, so data preprocessing is required. Data preprocessing includes the following three steps.

(1) Data cleaning

The 'service' column in the dataset represents the type of communication service, including 'HTTP', 'FTP', 'SMTP', 'SSH', 'DNS', '-'. Where '-' represents a protocol that the

model cannot recognize, it was replaced by a null value during processing in this paper. Rows with null values are deleted when cleaning data so they do not affect results. There are 94,168 rows with null values in the training set, and 81,173 records after processing, 47,153 rows with null values in the test set, and 35,179 records after processing.

(2) Data mapping

The 'proto', 'state' and 'service' features in the UNSW-NB15 dataset are strings, which cannot be recognized by the detection model, so feature mapping is implemented with one-hot encoding pairs. The values of 'proto' are 'TCP' and 'UDP'. After one-hot encoding mapping, they are 01 and 10, respectively. The values of 'state' are 'CON', 'FIN', 'INT', 'REQ' and 'RST'. After mapping, they are 10000, 01000, 00100, 00010 and 00001 respectively. In addition, values of 'service' are 'snmp', 'smtp', 'ftp', 'irc', 'pop3', 'ssh', 'http', 'radius', 'ftp-data', 'ssl', 'dhcp' and 'dns'; the corresponding codes after mapping are no longer listed. The three feature subsets become 19 columns after one-hot encoding; hence, removed the id column, there are 60 columns per row of the dataset.

(3) Normalization

The values between the features in the dataset are in different ranges, which affects the accuracy of the model. In order to ensure that the features are in the same index range, this paper adopts data normalization processing to map the data in the range of [0, 1]. The processing formula is as follows.

$$Z^* = \frac{Z - Z_{min}}{Z_{max} - Z_{min}}, \tag{15}$$

where Z^* is the normalized data, Z is the original data and Z_{max} and Z_{min} are the maximum and minimum values of the data, respectively.

5.2. Comparative Analysis of Feature Selection Models with Different Improvements in BWO-BOA Algorithm

In order to test the effectiveness of different strategies and their influence on the classification accuracy of the feature selection model, the original BOA is compared with the BOA with a dynamic adaptive search strategy (BOA1), the BOA1 and the movement strategy of BWO algorithm are combined (BOA2) and the small probability mutation strategy and updating the position by using the pheromone strategy in BWO algorithm are introduced to the BOA2 (BWO-BOA). *Acc*, *Pre*, *Rec* and *F*1 in the following tables are the averages of 20 times, and n is the feature mean selected for each model.

5.2.1. Comparative Analysis of Different Improvements in Binary Classification and Multi-Classification

The results of binary classification and multi-classification tests with 5% of the training set and 5% of the test set are shown in the following table. Table 3 shows the results of binary classification tests with different improvement strategies. Table 4 shows the results of different improvement strategies on multi-classification.

Through the comparison of several indicators—*Acc*, *Pre*, *Rec* and *F*1—in Table 3, it can be seen that the improvement of different strategies affects the feature selection model. The BOA1 that introduces the dynamic adaptive search strategy in the binary classification has different degrees of improvement in *Pre* and *Rec* than the original BOA, and the number of selected features is also reduced. At the same time, BOA2, after integrating the movement strategy of the BWO algorithm, has a great improvement in *Acc*, *Pre*, *Rec* and *F*1 compared with the BOA and BOA1. The classification accuracy of the model after integrating the movement strategy of BWO algorithm is improved from 94.37% to 96.23%, which indicates the classification accuracy and classification efficiency of the feature selection model can be effectively improved. The feature subset selected by the BWO-BOA is eight, which is the best one among the four algorithms. It shows that the introduction of the small probability mutation strategy and the use of the BWO algorithm pheromone strategy to update the position have obvious effects on reducing the redundancy of the feature selection subset.

Table 3. Binary classification results with different improvement strategies.

Algorithm	*Acc* (%)	*Pre* (%)	*Rec* (%)	*F*1 (%)	*n*
BWO-BOA	96.28	97.15	93.57	95.15	8
BOA	94.49	95.69	90.56	92.68	24
BOA1	94.37	95.74	90.22	92.50	23
BOA2	96.23	96.90	93.67	95.10	23

Table 4. Multi-classification results for different strategies.

Algorithm	Indicator	Analysis	Backdoor	DoS	Exploits	Fuzzers	Generic	Normal	Reconnaissance	Worms
BWO-BOA	*Pre*(%)	0	5.00	32.40	76.34	44.76	99.88	97.13	45.40	10.71
	Rec(%)	0	10.00	18.30	91.14	63.25	98.33	88.27	38.13	10.71
	*F*1(%)	0	6.66	22.66	83.04	51.86	99.10	92.45	39.94	9.52
	Acc(%)					91.20				
	n					8				
BOA	*Pre*(%)	0	0	22.90	67.81	35.36	99.73	96.90	46.05	2.22
	Rec(%)	0	0	12.98	89.51	46.59	98.04	80.26	32.26	6.66
	*F*1(%)	0	0	15.69	76.98	39.20	98.88	87.73	35.74	3.33
	Acc(%)					88.21				
	n					23				
BOA1	*Pre*(%)	0	0	29.03	67.91	35.88	99.78	97.48	36.35	4.44
	Rec(%)	0	0	13.67	90.73	51.39	98.13	79.44	28.30	6.66
	*F*1(%)	0	0	17.41	77.55	41.44	98.94	87.47	29.33	5.33
	Acc(%)					88.26				
	n					28				
BOA2	*Pre*(%)	0	5.55	32.18	75.36	44.52	99.83	97.36	47.51	9.33
	Rec(%)	0	11.11	17.62	91.16	62.77	98.25	87.14	41.62	10.00
	*F*1(%)	0	7.40	21.96	82.46	51.46	99.03	91.94	42.56	8.84
	Acc(%)					91.01				
	n					28				

From the analysis in Table 4, the BWO-BOA has the highest overall classification accuracy, followed by the BOA2, BOA1 and BOA, which indicates that the integration of different improvement strategies can improve the classification accuracy of the model. In addition to 'Backdoor' and 'Reconnaissance' attacks, compared with the BOA, the BOA1 improves the classification accuracy of other types of attacks, which shows the effectiveness of the dynamic adaptive search strategy. The BOA2 not only improves the classification accuracy of attack types other than 'Normal', but also detects 'Backdoor' attacks that the BOA and BOA1 cannot detect, which demonstrates the integration of the movement strategy of the BWO algorithm can effectively improve the detection efficiency of multi-classification. The BWO-BOA improves the classification accuracy of 'Dos', 'Fuzzers', 'Generic' and 'Worms' attacks, which illustrates the introduction of a small probability mutation strategy and the use of the pheromone strategy of the BWO algorithm to update the location can also improve the classification efficiency of multi-classification.

5.2.2. Comparison of Fitness for Different Strategies

For the purpose of observing the effect of different improvement strategies more intuitively, this paper makes a comparison of the fitness for different strategies in 50 iterations. Figure 3a shows the fitness of different strategies in binary classification, and Figure 3b shows the fitness of different strategies in multi-classification.

From Figure 3, the order of fitness is BWO-BOA, BOA2, BOA1, BOA; it is shown that fusing each strategy into the BOA algorithm can improve the classification accuracy of the model and reduce the number of features selected.

Based on the above analysis, this paper introduces a dynamic adaptive search strategy to balance the ability of early and late searches in the global search and enhance the classification accuracy of the model. At the same time, the movement strategy of BWO is integrated to improve the local search process and further improve the classification accuracy and efficiency of the model; furthermore, the location updating strategy of small

probability mutation and pheromone strategy of BWO improved the global search ability of the algorithm and then reduced the redundancy of feature subset selected by the model.

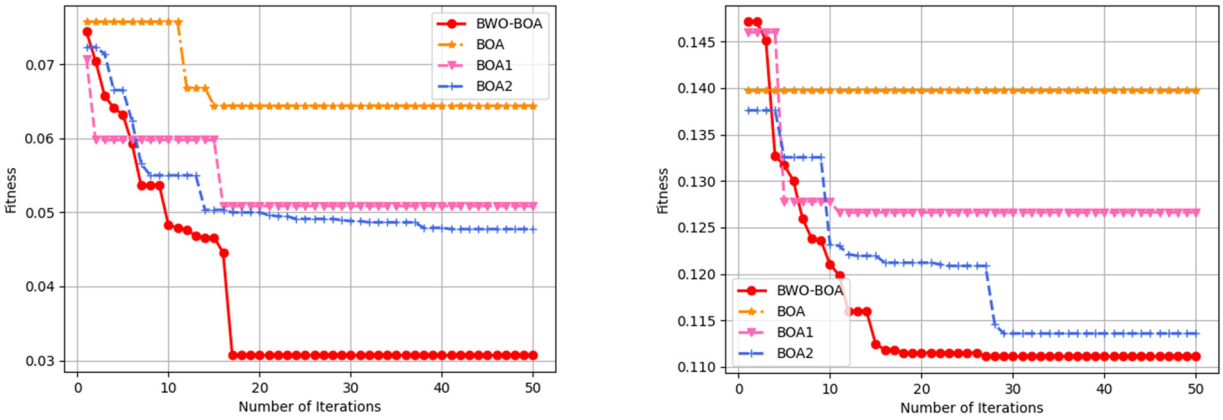


Figure 3. (**a**) The fitness of different strategies in binary classification; (**b**) The fitness of different strategies in multi-classification.

5.3. Comparative Experimental Results of Different Algorithms

To verify the effectiveness of the model proposed in this paper, the model is compared and analyzed with the six algorithms—BOA, BWO, PSO, SSA, WOA and IBOA—used in feature selection. Since the dataset is too large, this paper takes 5%, 10%, 20% and 30% of the training set and test set, respectively, for simulation experiments. *Acc*, *Pre*, *Rec* and *F*1 in the following tables are the averages of 20 times, and *n* is the average number of features selected by each model.

5.3.1. Comparative Analysis of Binary Classification Results

Tables 5–8 are the binary classification results of the seven algorithms on 5%, 10%, 20% and 30% of the datasets, respectively. It can be seen from the observation that the feature selection model of the proposed algorithm is better than other algorithms in terms of *Acc*, *Pre*, *Rec* and *F*1, and compared with the original BOA and BWO algorithms, the result of the feature selection model is greatly improved. For example, with 30% of the dataset, compared with BWO, the BWO-BOA improved *Acc*, *Pre*, *Rec* and *F*1 by 0.18%, 0.24%, 0.23% and 0.23%, respectively. This shows that the BWO-BOA algorithm with the Black Widow search strategy is better than the original BOA and BWO algorithms. Compared with the IBOA, the number of features selected by the BWO-BOA is similar, but the *Acc*, *Pre*, *Rec* and *F*1 of the BWO-BOA are higher than those of the IBOA algorithm. It also further shows that the application of the BWO-BOA in the feature selection model can improve its classification accuracy and precision and has a good classification effect in data preprocessing.

Table 5. Binary classification results of seven algorithms in 5% of the dataset.

Algorithm	*Acc* (%)	*Pre* (%)	*Rec* (%)	*F*1 (%)	*n*
BWO-BOA	96.14	96.88	93.49	94.98	8
BOA	94.43	95.64	90.14	92.41	24
BWO	96.10	96.71	93.52	94.94	7
PSO	96.08	96.76	93.24	94.79	20
SSA	94.58	95.72	90.46	92.66	26
WOA	95.11	96.09	91.43	93.41	16
IBOA	94.72	95.45	91.36	93.08	6

Table 6. Binary classification results of seven algorithms in 10% of the dataset.

Algorithm	*Acc* (%)	*Pre* (%)	*Rec* (%)	*F*1 (%)	*n*
BWO-BOA	95.93	96.85	92.96	94.66	7
BOA	94.44	95.75	90.30	92.57	24
BWO	95.90	96.76	92.96	94.64	7
PSO	95.84	96.66	92.88	94.54	23
SSA	94.71	95.86	90.87	92.96	27
WOA	94.95	95.88	91.44	93.32	21
IBOA	95.27	96.06	92.12	93.80	6

Table 7. Binary classification results of seven algorithms in 20% of the dataset.

Algorithm	*Acc* (%)	*Pre* (%)	*Rec* (%)	*F*1 (%)	*n*
BWO-BOA	95.92	96.66	93.09	94.66	7
BOA	94.55	95.78	90.61	92.75	24
BWO	95.84	96.55	93.01	94.57	7
PSO	95.91	96.51	93.23	94.68	22
SSA	95.34	96.22	92.10	93.88	27
WOA	95.39	96.19	92.25	93.96	19
IBOA	94.95	95.94	91.41	93.31	7

Table 8. Binary classification results of seven algorithms in 30% of the dataset.

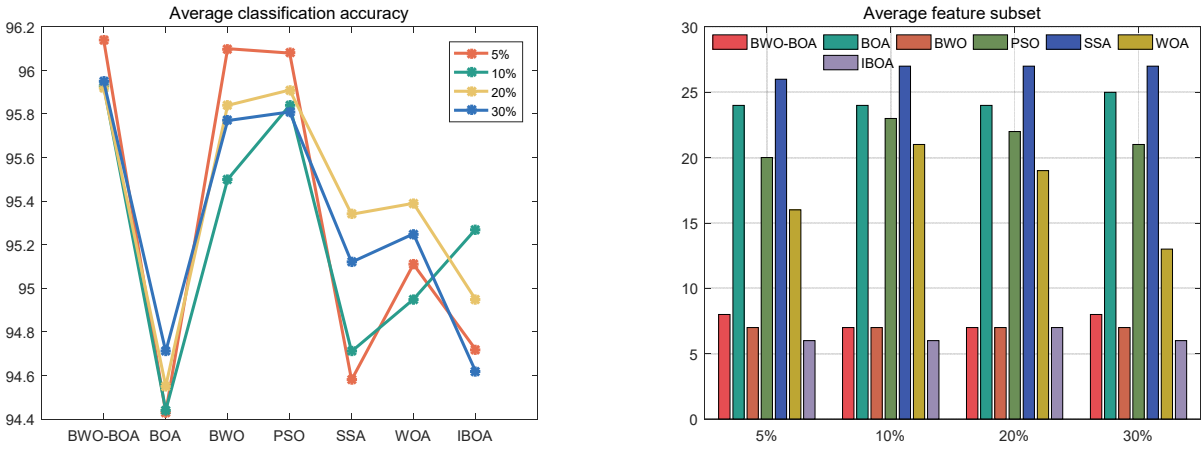
Algorithm	*Acc* (%)	*Pre* (%)	*Rec* (%)	*F*1 (%)	*n*
BWO-BOA	95.95	96.64	93.15	94.69	8
BOA	94.71	95.88	90.86	92.97	25
BWO	95.77	96.40	92.92	94.46	7
PSO	95.81	96.68	92.80	94.50	21
SSA	95.12	96.28	91.51	93.55	27
WOA	95.25	96.31	91.79	93.74	13
IBOA	94.62	95.66	90.86	92.89	6

To intuitively observe the effect of the BWO-BOA in the feature selection model, Figure 4a is a line chart comparing the average classification accuracy of the seven algorithms in the binary classification. The average accuracy of the BWO-BOA on the feature selection model is higher than that of other algorithms on four datasets, which shows that the proposed algorithm has some advantages over other algorithms in the feature selection model.

Figure 4b shows the average number of feature subsets selected by the seven algorithms in the four datasets. It can be seen that, although the average number of features selected by the feature selection model of the BWO-BOA in the 10% and 20% datasets is the same as that of the BWO, and only one difference from the BWO algorithm in the 5% and 30% datasets, the average number of selected features is also similar to the IBOA, but they are largely lower than the average number of features selected by other algorithms, including the BOA, which proves that the feature model of the BWO-BOA is effective for reducing the number of features selected by the feature model of the original BOA. This indicates that the model in this paper has a certain role in decreasing the redundancy of features and reducing dimensionality.

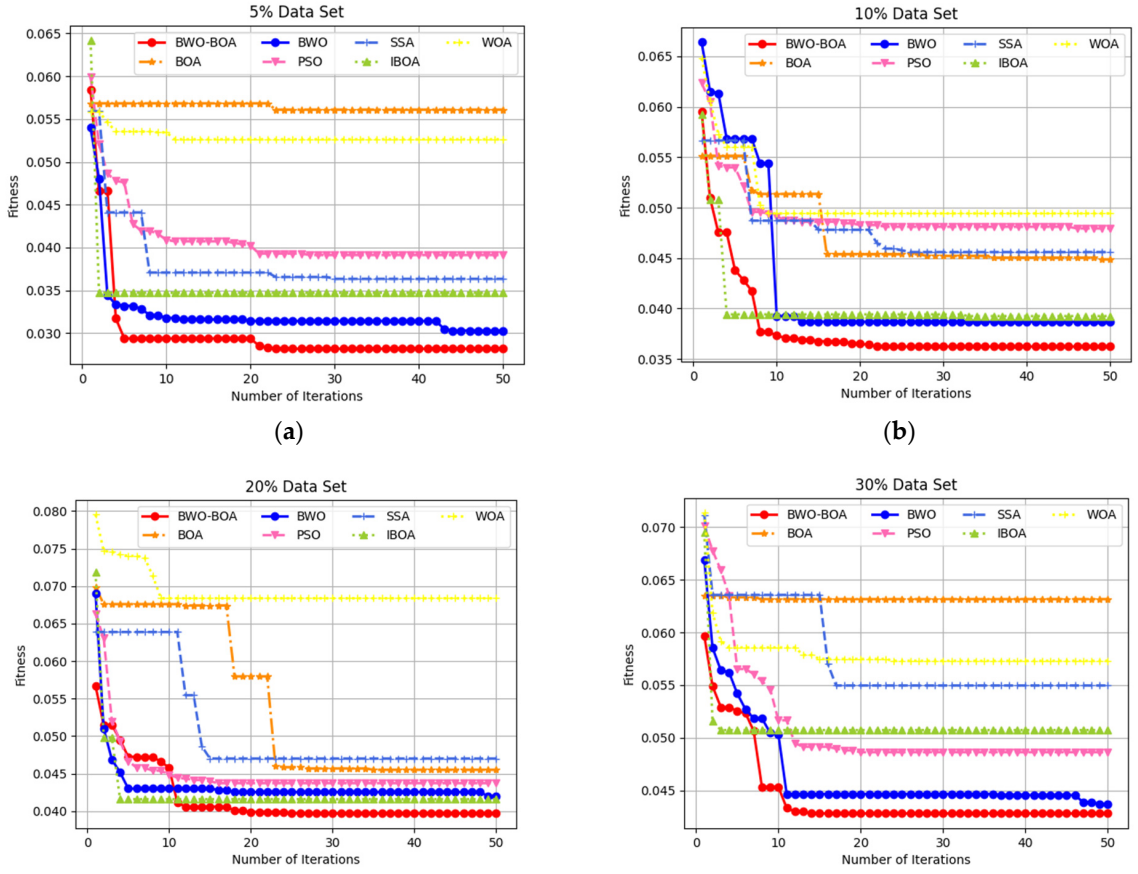
For the sake of judging the convergence speed and accuracy of the BWO-BOA in the binary classification, Figure 5 shows the fitness changes of different algorithms in the 50 iterations of the four datasets. It can be seen in Figure 5 that the fitness of the BWO-BOA is not high in the first-generation iteration process on the 20% and 30% datasets, which indicates that the BWO-BOA has a strong search ability from the beginning. In the four datasets, the fitness of the BWO-BOA is lower than the other algorithms, and it is the largest difference from the original BOA, which shows that the BOA combined

with the movement strategy of BWO can reduce feature redundancy, while ensuring high classification accuracy. It proves that the BWO-BOA can well integrate the characteristics of the BOA and BWO and play a good role in binary classification.



(**a**) (**b**)

Figure 4. (**a**) The average accuracy of the seven algorithms in binary classification; (**b**) The average number of features selection for seven algorithms in binary classification.



(**a**) (**b**)

(**c**) (**d**)

Figure 5. (**a**) Fitness curves of binary classification for each algorithm on 5% datasets; (**b**) Fitness curves of binary classification for each algorithm on 10% datasets; (**c**) Fitness curves of binary classification for each algorithm on 20% datasets; (**d**) Fitness curves of binary classification for each algorithm on 30% datasets.

5.3.2. Comparative Analysis of Multi-classification Results

Tables 9–12 are the multi-classification results of the seven algorithms for the four datasets of 5%, 10%, 20% and 30%, respectively, which includes the *Acc*, *Pre*, *Rec* and *F*1 results for the nine types of attacks. Overall, except for BWO and IBOA algorithms, the average number of features selected by the BWO-BOA is lower than that selected by other algorithms; especially compared with the original BOA, the redundancy of features is reduced to a great extent. Although it is not much different from the average number of features selected by BWO, the average classification accuracy of the BWO-BOA in the feature selection model is higher than all algorithms. From the identification of each attack category, perhaps because the number of attacks in the dataset is too small, the 'Analysis' attack is the most difficult type to identify, and the average classification accuracy of the seven algorithms in the four datasets is 0.

Table 9. Multi-classification results of seven algorithms in 5% of the dataset.

Algorithm	Indicator	Analysis	Backdoor	DoS	Exploits	Fuzzers	Generic	Normal	Reconnaissance	Worms
BWO-BOA	*Pre*(%)	0	0	27.64	74.79	46.50	99.84	97.89	43.23	23.14
	Rec(%)	0	0	17.67	91.29	61.22	98.33	87.59	42.26	22.22
	*F*1(%)	0	0	20.99	82.12	51.61	99.07	92.41	41.26	22.03
	Acc(%)					91.28				
	n					8				
BOA	*Pre*(%)	0	0	23.51	66.23	40.38	99.72	97.16	34.27	12.50
	Rec(%)	0	0	14.94	90.04	50.63	98.19	78.76	27.80	20.83
	*F*1(%)	0	0	17.23	76.13	43.31	98.95	86.88	28.64	15.55
	Acc(%)					88.20				
	n					24				
BWO	*Pre*(%)	0	0	30.63	73.49	47.10	99.71	97.48	44.41	8.33
	Rec(%)	0	0	13.54	90.86	66.32	98.08	85.42	41.89	4.62
	*F*1(%)	0	0	17.89	81.19	54.42	98.89	90.99	41.48	5.55
	Acc(%)					90.30				
	n					7				
PSO	*Pre*(%)	0	12.50	29.66	75.08	45.15	99.80	97.55	46.66	21.05
	Rec(%)	0	10.00	21.22	91.09	62.48	98.34	85.59	39.40	17.10
	*F*1(%)	0	10.00	24.25	82.23	51.63	99.07	91.11	41.54	17.89
	Acc(%)					90.76				
	n					23				
SSA	*Pre*(%)	0	1.92	29.81	65.17	42.30	95.80	97.76	38.03	16.66
	Rec(%)	0	3.84	16.62	81.21	62.70	94.77	78.65	31.38	11.11
	*F*1(%)	0	2.56	20.50	70.55	49.79	95.22	87.11	32.72	12.96
	Acc(%)					88.36				
	n					28				
WOA	*Pre*(%)	0	3.12	29.57	69.95	51.05	91.49	91.68	40.48	22.64
	Rec(%)	0	6.25	22.82	89.40	62.15	89.19	79.46	33.31	23.52
	*F*1(%)	0	4.16	24.92	77.94	54.50	89.99	84.82	35.05	22.54
	Acc(%)					89.33				
	n					17				
IBOA	*Pre*(%)	0	0	25.05	68.68	37.74	99.74	96.26	31.00	1.42
	Rec(%)	0	0	12.17	91.24	46.99	98.03	82.68	20.10	3.57
	*F*1(%)	0	0	15.61	78.21	39.29	98.88	88.85	21.99	2.04
	Acc(%)					89.23				
	n					6				

As shown in Table 9, the BWO-BOA, BOA, BWO and IBOA algorithms failed to identify the 'Backdoor' attack, but compared with the original BOA and the improved IBOA, the average classification precision of the improved BWO-BOA for the remaining seven attacks is improved. In addition, compared with BWO, the classification precision of 'Exploits', 'Generic' and 'Worms' attacks are greatly improved. As shown in Table 10, due to the increase in the number of datasets, the BWO-BOA can identify 'Backdoor' attack, while BWO and the IBOA fails to identify 'Backdoor' attack. The average precision of the BWO-BOA in 'Fuzzers', 'Generic' and 'Worms' attacks is also the highest among the seven algorithms; compared with the BOA, the average classification precision of the other attacks except the 'Backdoor' attack has been greatly improved. The average classification precision of each attack type is higher than that of the IBOA. As shown in Table 11, although

the precision of the BWO-BOA in each attack category is not the highest, compared with the original BOA and BWO, the classification effect of each attack is greatly improved. Except for the 'Backdoor' and 'Worms' attacks, the average classification precision is higher than that of the IBOA. As shown in Table 12, for the hard-to-recognize 'Backdoor' attacks, the recognition effect of the BWO-BOA is higher than other algorithms; the recognition effect of 'DoS', 'Exploits', 'Fuzzers', 'Reconnaissance' and 'Worms' attacks are also better than other algorithms, and the classification effect of the BWO-BOA on the attack is much higher than the BOA, BWO and the improved IBOA algorithms.

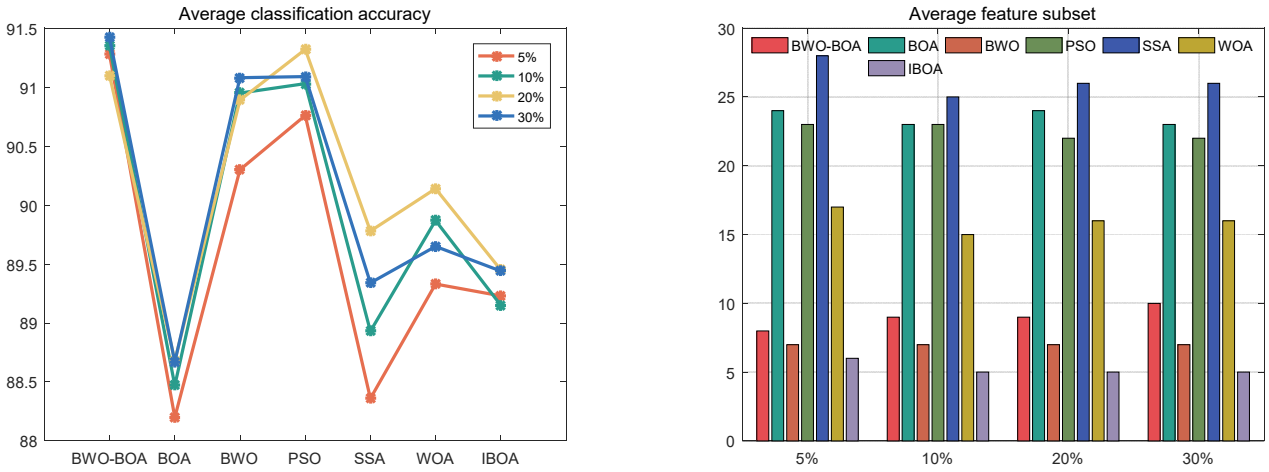
Table 10. Multi-classification results of seven algorithms in 10% of the dataset.

Algorithm	Indicator	Analysis	Backdoor	DoS	Exploits	Fuzzers	Generic	Normal	Reconnaissance	Worms
BWO-BOA	*Pre*(%)	0	6.14	32.07	76.39	46.78	99.84	97.71	59.24	35.96
	Rec(%)	0	7.89	19.19	91.52	65.74	98.43	86.70	50.36	32.42
	*F*1(%)	0	6.84	23.49	83.24	54.00	99.13	91.85	53.08	31.85
	Acc(%)					91.35				
	n					9				
BOA	*Pre*(%)	0	8.33	19.72	68.27	39.83	97.25	95.05	33.03	18.13
	(%)	0	5.55	12.42	88.71	53.37	95.91	79.01	24.62	9.21
	*F*1(%)	0	6.48	14.20	76.89	45.00	96.57	86.19	27.43	11.84
	Acc(%)					88.47				
	n					23				
BWO	*Pre*(%)	0	0	35.50	75.93	44.25	99.82	97.13	49.11	28.38
	Rec(%)	0	0	20.22	91.03	58.89	98.37	86.46	43.56	18.09
	*F*1(%)	0	0	25.44	82.76	47.23	99.09	91.43	43.67	21.57
	Acc(%)					90.95				
	n					7				
PSO	*Pre*(%)	0	15.74	35.48	76.48	45.86	99.83	97.92	52.06	27.31
	Rec(%)	0	12.96	23.29	91.51	64.76	98.45	85.72	42.52	12.18
	*F*1(%)	0	13.70	27.53	83.28	53.38	99.13	91.38	45.97	15.67
	Acc(%)					91.03				
	n					23				
SSA	*Pre*(%)	0	9.72	29.18	69.92	40.39	99.82	97.45	41.81	15.88
	Rec(%)	0	8.33	15.33	90.25	60.10	98.30	80.78	33.14	9.36
	*F*1(%)	0	8.33	19.52	78.68	47.91	99.05	88.27	35.88	11.03
	Acc(%)					88.93				
	n					25				
WOA	*Pre*(%)	0	5.20	28.78	72.10	38.49	99.75	97.42	40.73	19.44
	Rec(%)	0	5.20	15.79	90.03	55.15	98.34	84.09	33.22	11.62
	(%)	0	5.20	19.15	79.99	44.62	99.04	90.16	35.99	13.24
	Acc(%)					89.87				
	n					15				
IBOA	*Pre*(%)	0	0	21.40	69.12	37.09	99.70	97.11	27.86	13.85
	Rec(%)	0	0	11.02	91.32	40.16	98.23	82.60	22.82	12.69
	*F*1(%)	0	0	13.67	78.48	36.27	98.96	89.23	22.42	11.52
	Acc(%)					89.15				
	n					5				

In summary, the BWO-BOA in this paper has some advantages over the other six algorithms in multi-classification of feature selection models; especially compared with the original BOA, BWO and the improved IBOA algorithm, the classification of each attack is greatly improved, the validity of the proposed model is verified.

For the purpose of observing the effect of the proposed model more directly, Figure 6a shows the average classification accuracy of the seven algorithms for multi-classification on four datasets. Through the analysis of Figure 6a, this shows that the accuracy of the BWO-BOA is higher than other algorithms, which can prove that the BWO-BOA has some advantages in the feature selection model. Figure 6b shows the average number of features selected by the seven algorithms for multi-classification in the four datasets. It can be seen in Figure 6b that in the feature selection model of the seven algorithms, except BWO and IBOA algorithms, the BWO-BOA algorithm has the lowest average number of features selected; especially in the 5% dataset, the minimum average number of features is 8, which is 16 fewer than the original BOA, and 20 fewer relative to the SSA of the largest number for the seven algorithms. Furthermore, it is a small difference from the average number of

features selected by the BWO and IBOA algorithms. This shows that the proposed model is effective in both the average classification accuracy and the average number of selected features, which verifies the effectiveness of the proposed model.



(**a**) (**b**)

Figure 6. (**a**) The average accuracy of the seven algorithms in multi-classification; (**b**) The average number of features selected for seven algorithms in multi-classification.

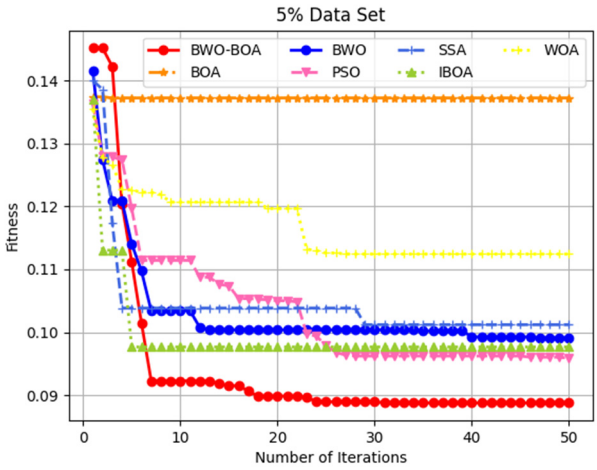
Table 11. Multi-classification results of seven algorithms in 20% of the dataset.

Algorithm	Indicator	Analysis	Backdoor	DoS	Exploits	Fuzzers	Generic	Normal	Reconnaissance	Worms
BWO-BOA	*Pre*(%)	0	8.47	37.64	75.56	46.23	99.87	97.49	48.61	33.01
	Rec(%)	0	9.33	21.73	91.65	66.22	98.37	86.22	40.62	29.94
	F1(%)	0	7.93	26.75	82.77	54.06	99.11	91.49	43.14	30.16
	Acc(%)					91.10				
	n					9				
BOA	*Pre*(%)	0	5.55	26.78	68.50	41.56	99.82	97.78	39.25	23.33
	Rec(%)	0	2.66	14.30	90.19	53.92	98.05	81.57	30.37	8.33
	F1(%)	0	3.21	17.73	77.74	45.82	98.93	88.91	32.67	11.43
	Acc(%)					88.68				
	n					24				
BWO	(%)	0	8.33	34.85	75.55	43.60	99.83	97.48	40.46	25.43
	Rec(%)	0	5.41	20.66	92.01	62.66	98.37	86.07	34.38	26.80
	F1(%)	0	5.61	25.48	82.93	51.30	99.10	91.41	36.15	24.63
	Acc(%)					90.89				
	n					7				
PSO	*Pre*(%)	0	5.25	38.83	77.11	46.80	99.86	97.47	51.28	37.43
	Rec(%)	0	12.50	24.47	91.21	65.11	98.39	87.12	42.63	29.97
	F1(%)	0	7.30	29.66	83.55	54.13	99.12	91.98	45.71	31.57
	Acc(%)					91.32				
	n					22				
SSA	*Pre*(%)	0	9.76	34.69	72.22	41.32	99.78	97.52	46.11	36.72
	Rec(%)	0	12.08	19.04	90.02	58.87	98.22	83.73	29.98	20.94
	F1(%)	0	9.66	24.04	80.01	48.22	98.99	90.04	34.64	25.22
	Acc(%)					89.78				
	n					26				
WOA	*Pre*(%)	0	4.30	33.79	79.21	41.33	99.80	96.25	41.74	30.31
	Rec(%)	0	5.41	17.29	90.19	56.56	98.21	85.51	36.00	21.24
	F1(%)	0	4.16	21.68	80.54	46.57	99.00	90.40	37.76	22.84
	Acc(%)					90.14				
	n					16				
IBOA	*Pre*(%)	0	10.83	24.93	70.25	40.21	99.73	97.04	31.46	45.11
	Rec(%)	0	15.41	15.96	91.09	53.09	98.30	82.63	25.68	23.75
	F1(%)	0	12.20	19.20	79.24	45.10	99.01	89.23	27.26	27.56
	Acc(%)					89.45				
	n					5				

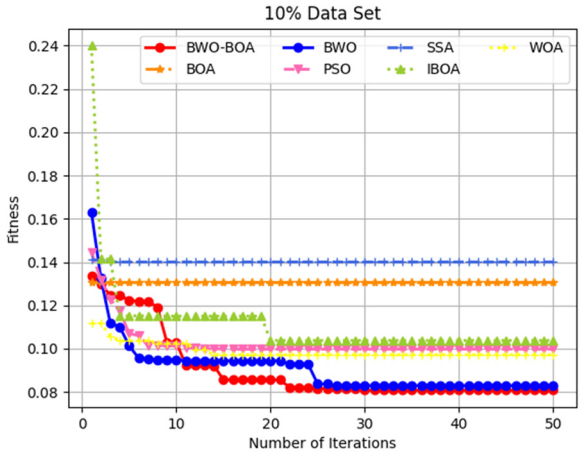
Table 12. Multi-classification results of seven algorithms in 30% of the dataset.

Algorithm	Indicator	Analysis	Backdoor	DoS	Exploits	Fuzzers	Generic	Normal	Reconnaissance	Worms
BWO-BOA	*Pre*(%)	0	8.33	40.22	77.69	46.93	99.85	97.55	52.06	47.28
	Rec(%)	0	13.42	25.58	91.13	69.03	98.61	85.79	45.18	40.10
	*F*1(%)	0	9.79	30.62	83.86	55.66	99.23	91.28	47.59	42.41
	Acc(%)					91.42				
	n					10				
BOA	*Pre*(%)	0	2.50	28.51	69.05	38.47	99.81	97.97	35.75	21.68
	Rec(%)	0	3.75	16.42	90.61	56.33	98.05	79.05	26.21	11.81
	*F*1(%)	0	2.91	20.24	78.24	45.42	98.92	87.43	29.75	14.12
	Acc(%)					88.67				
	n					23				
BWO	*Pre*(%)	0	3.08	35.08	75.60	45.79	99.84	96.74	39.66	34.31
	Rec(%)	0	4.16	21.22	91.56	60.94	98.40	86.70	34.83	27.26
	*F*1(%)	0	3.47	25.77	82.73	51.66	99.12	91.38	35.62	28.11
	Acc(%)					91.08				
	n					7				
PSO	*Pre*(%)	0	7.34	38.29	77.60	45.79	99.86	97.90	48.88	42.84
	Rec(%)	0	18.41	25.43	90.74	62.00	98.54	85.76	44.16	31.57
	*F*1(%)	0	9.85	30.31	83.61	52.20	99.20	91.42	45.79	34.51
	Acc(%)					91.09				
	n					22				
SSA	*Pre*(%)	0	4.55	27.80	72.12	40.85	99.84	96.86	39.93	35.19
	Rec(%)	0	3.91	18.30	89.54	57.79	98.43	82.05	32.29	14.85
	*F*1(%)	0	3.54	21.78	79.82	47.35	99.13	88.78	35.26	18.47
	Acc(%)					89.34				
	n					26				
WOA	*Pre*(%)	0	3.62	31.17	73.39	41.43	99.80	97.28	44.32	37.10
	Rec(%)	0	7.00	18.75	90.87	61.34	98.40	82.26	33.77	23.84
	*F*1(%)	0	4.05	23.17	81.11	49.35	99.09	89.09	36.08	26.73
	Acc(%)					89.65				
	n					16				
IBOA	*Pre*(%)	0	3.48	26.70	71.51	38.82	99.70	95.81	37.68	30.88
	Rec(%)	0	7.08	13.61	88.95	53.55	98.31	83.51	33.19	24.90
	*F*1(%)	0	4.17	17.10	79.16	44.27	99.00	89.07	34.33	26.06
	Acc(%)					89.44				
	n					5				

To observe the convergence of the BWO-BOA in multi-classification, Figure 7 shows the changes of the fitness of the algorithms during 50 iterations in the four datasets.



(**a**)



(**b**)

Figure 7. *Cont.*

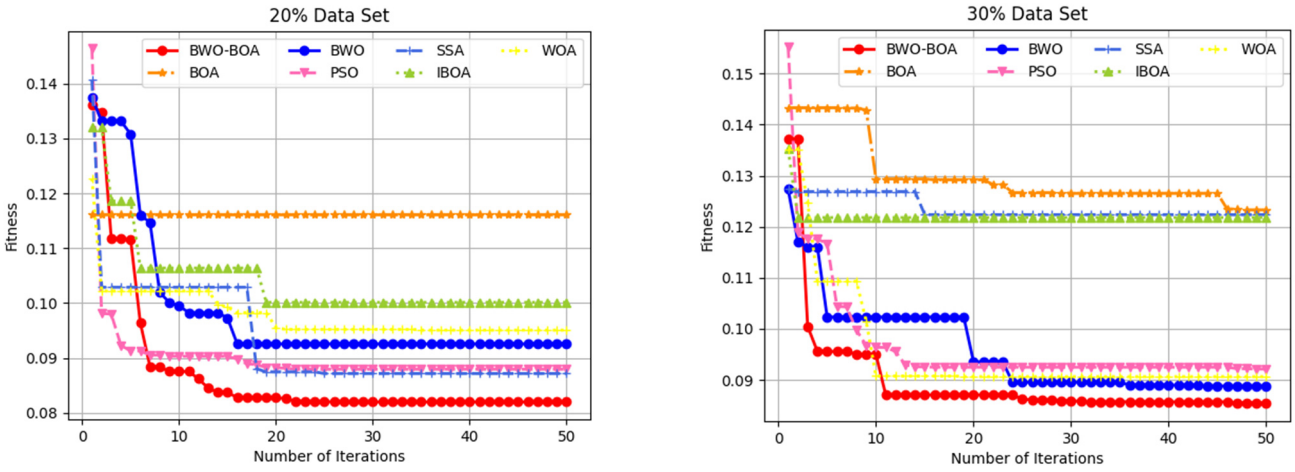


Figure 7. (**a**) Fitness curves of multi-classification for each algorithm on 5% datasets; (**b**) Fitness curves of multi-classification for each algorithm on 10% datasets; (**c**) Fitness curves of multi-classification for each algorithm on 20% datasets; (**d**) Fitness curves of multi-classification for each algorithm on 30% datasets.

It can be seen in Figure 7 that the BWO-BOA algorithm has the greatest number of fitness changes in the iterative process, which proves that the BWO-BOA algorithm has a strong global search ability and does not easily fall into the local optimum. The BWO-BOA algorithm has the lowest fitness among the seven algorithms in the four datasets, which shows that, compared with the other six algorithms, the BWO-BOA algorithm not only has higher classification accuracy in multi-classification, but it can also reduce the redundancy of features based on guaranteeing the important features.

6. Conclusions

In order to improve the classification accuracy and reduce the redundancy of feature selection in intrusion detection models, this paper proposes an improved BOA combined with the BWO algorithm, namely, the BWO-BOA algorithm, which utilizes the dynamic adaptive search strategy and movement search strategy of BWO and the small differential mutation search strategy to improve the original BOA to solve the problems of low precision, slow convergence and easily falling into the local optimum. This paper uses the BWO-BOA algorithm to find the optimal feature subset and proposes a feature selection model based on the BWO-BOA algorithm for network intrusion detection. Experiments on the UNSW-NB15 dataset show that, compared to the BOA, BWO, PSO, SSA, WOA and IBOA algorithms, the proposed model can not only obtain a better feature subset, but also obtain a higher classification accuracy, which proves that the proposed model based on the BWO-BOA algorithm can effectively alleviate the problems of low classification accuracy and redundancy of feature selection in intrusion detection for network security.

Author Contributions: Conceptualization, H.X.; methodology, H.X. and Y.L.; software, Y.L. and Q.G.; validation, H.X., Y.L.; formal analysis, H.X.; investigation, Y.L. and Q.G.; resources, Y.L. and Q.G.; data curation, Y.L. and Q.G.; writing—original draft preparation, Y.L.; writing—review and editing, Y.L. and H.X.; supervision, H.X.; project administration, H.X.; funding acquisition, H.X. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Natural Science Foundation of China (No. 61602162).

Data Availability Statement: The dataset is available at: https://research.unsw.edu.au/projects/unsw-nb15-dataset, accessed on 16 September 2022.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Zhou, Y.; Cheng, G.; Jiang, S.; Dai, M. Building an efficient intrusion detection system based on feature selection and ensemble classifier. *Comput. Netw.* **2020**, *174*, 107247. [CrossRef]
2. Deng, M.; Wang, K.; Zhang, C.; Zhang, Y. A comparative study of network intrusion detection evaluation data sets. *Mod. Comput.* **2020**, *20*, 20–26.
3. Alazzam, H.; Sharieh, A.; Sabri, K.E. A feature selection algorithm for intrusion detection system based on pigeon inspired optimizer. *Expert Syst. Appl.* **2020**, *148*, 0957–4174. [CrossRef]
4. Alshamy, R.; Ghurab, M. A review of big data in network intrusion detection system: Challenges, approaches, datasets, and tools. *Int. J. Comput. Sci. Eng.* **2020**, *8*, 62–74.
5. Zarei, M.; Rezai, A.; Hamidpour, S.S.F. Breast cancer segmentation based on modified gaussian mean shift algorithm for infrared thermal images. *Comput. Methods Biomech. Biomed. Eng. Imaging Vis.* **2021**, *9*, 574–580. [CrossRef]
6. Darabi, N.; Rezai, A.; Hamidpour, S.S.F. Breast cancer detection using RSFS-based feature selection algorithms in thermal images. *Biomed. Eng. Appl. Basis Commun.* **2021**, *33*, 2150020. [CrossRef]
7. Salimian, M.; Rezai, A.; Hamidpour, S.; Khajeh-Khalili, F. Effective features in thermal images for breast cancer detection. In Proceedings of the 2nd National Conference on New Technologies in Electrical and Computer Engineering, Isfahan, Iran, 6 February 2019; pp. 1–7.
8. Li, Y.; Qin, J. Application of a feature-selection method based on improved genetic algorithm in coronary heart disease detection. *J. Huaihua Univ.* **2021**, *40*, 57–62.
9. Shi, Q.; Pan, F.; Long, F.; Li, N.; Gou, H.; Su, H.; Xie, Y. Summary of research on feature selection methods. *Microelectron. Comput.* **2022**, *39*, 1–8.
10. Liang, B. A Review of the Application of swarm intelligence in network intrusion's feature selection. *Comput. Knowl. Technol.* **2020**, *16*, 30–32.
11. Tawil, A.A.; Sabri, K.E. A feature selection algorithm for intrusion detection system based on moth flame optimization. In Proceedings of the 2021 International Conference on Information Technology (ICIT), Amman, Jordanm, 14–15 July 2021; pp. 377–381.
12. Chen, T.; Dong, H. Industrial anomaly intrusion detection using ant colony algorithm and deep reinforcement learning. *J. Chin. Comput. Syst.* **2022**, *43*, 779–784.
13. Syarif, I.; Afandi, R.F.; Saputra, F.A. Feature selection algorithm for intrusion detection using cuckoo search algorithm. In Proceedings of the 2020 International Electronics Symposium (IES), Surabaya, Indonesia, 29–30 September 2020; pp. 430–435.
14. Mahboob, A.S.; Moghaddam, M.R.O. An anomaly-based intrusion detection system using butterfly optimization algorithm. In Proceedings of the 2020 6th Iranian Conference on Signal Processing and Intelligent Systems (ICSPIS), Mashhad, Iran, 23–24 December 2020; pp. 1–6.
15. Tubishat, M.; Alswaitti, M.; Mirjalili, S.; Al-Garadi, M.A.; Rana, T.A. Dynamic butterfly optimization algorithm for feature selection. *IEEE Access* **2020**, *8*, 194303–194314. [CrossRef]
16. Selvakumar, B.; Muneeswaran, K. Firefly algorithm based feature selection for network intrusion detection. *Comput. Secur.* **2019**, *81*, 148–155.
17. Saheed, Y.K. *A Binary Firefly Algorithm Based Feature Selection Method on High Dimensional Intrusion Detection Data*; Springer: Cham, Switzerland, 2022; Volume 109, pp. 273–288.
18. Almomani, O. A feature selection model for network intrusion detection system based on PSO, GWO, FFA and GA algorithms. *Symmetry* **2020**, *12*, 1046. [CrossRef]
19. Al-Yaseen, W.L. Improving intrusion detection system by developing feature selection model based on firefly algorithm and support vector machine. *Int. J. Comput. Sci.* **2019**, *46*, 534–540.
20. Li, X.; Yi, P.; Wei, W.; Tian, L. LNNLS-KH: A feature selection method for network intrusion detection. *Secur. Commun. Netw.* **2021**, *3*, 1–22. [CrossRef]
21. Chen, H.; Ma, X.; Huang, S. A feature selection method for intrusion detection based on parallel sparrow search algorithm. In Proceedings of the 2021 16th International Conference on Computer Science & Education (ICCSE), Lancaster, UK, 17–21 August 2021; pp. 685–690.
22. Yang, J.; Ye, Z.; Yan, L.; Wang, R. Modified naive bayes algorithm for network intrusion detection based on artificial bee colony algorithm. In Proceedings of the 2018 IEEE 4th International Symposium on Wireless Systems within the International Conferences on Intelligent Data Acquisition and Advanced Computing Systems (IDAACS-SWS) IEEE, Lviv, Ukraine, 20–21 September 2018; pp. 35–40.
23. Alsaleh, A.; Binsaeedan, W. The influence of salp swarm algorithm-based feature selection on network anomaly intrusion detection. *IEEE Access* **2021**, *9*, 112466–112477. [CrossRef]
24. Alzaqebah, A.; Aljarah, I.; Al-Kadi, O.; Damaševičius, R. A modified grey wolf optimization algorithm for an intrusion detection system. *Mathematics* **2022**, *10*, 999. [CrossRef]
25. Xu, H.; Przystupa, K.; Fang, C.; Marciniak, A.; Kochan, O.; Beshley, M. A combination strategy of feature selection based on an integrated optimization algorithm and weighted k-nearest neighbor to improve the performance of network intrusion detection. *Electronics* **2020**, *9*, 1206.
26. Li, J.; An, Q.; Lei, H.; Deng, Q.; Wang, G.-G. Survey of lévy flight-based metaheuristics for optimization. *Mathematics* **2022**, *10*, 2785. [CrossRef]

27. Zhang, X.; Tan, H. Butterfly optimization algorithm with dynamic variance gaussian mutation. *J. Yunnan Norm. Univ. (Nat. Sci. Ed.)* **2022**, *42*, 31–36.
28. Xu, H.; Lu, Y.; Ye, Z. Applying cooperative search strategies to improve butterfly optimization algorithm. In Proceedings of the 2022 IEEE 25th International Conference on Computer Supported Cooperative Work in Design (CSCWD), Hangzhou, China, 4–6 May 2022.
29. Mojtahedi, A.; Sorouri, F.; Souha, A.N.; Molazadeh, A.; Mehr, S.S. Feature selection-based intrusion detection system using genetic whale optimization algorithm and sample-based classification. *arXiv* **2022**, arXiv:2201.00584.
30. Yuan, Q.; Lu, L. Network intrusion detection method based on combination of improved ant colony optimization and genetic algorithm. *J. Chongqing Univ. Posts Telecommun. (Nat. Sci. Ed.)* **2017**, *29*, 84–89.
31. Xu, H.; Fu, Y.; Liu, X.; Fang, C.; Su, J. Applying improved grey wolf optimizer algorithm integrated with cuckoo search to feature selection for network intrusion detection. *Adv. Eng. Sci.* **2018**, *50*, 160–166.
32. Kang, Y.; Wang, H.; Tao, L.; Yang, H.; Yang, X.; Wang, F.; Li, H. Hybrid improved flower pollination algorithm and gray wolf algorithm for feature selection. *Comput. Sci.* **2022**, *49*, 125–132.
33. Arora, S.; Singh, S. Butterfly optimization algorithm: A novel approach for global optimization. *Soft Comput.* **2019**, *23*, 715–734. [CrossRef]
34. Peña-Delgado, A.F.; Peraza-Vázquez, H.; Almazán-Covarrubias, J.H.; Cruz, N.T.; García-Vite, P.M.; Morales-Cepeda, A.B.; Ramirez-Arredondo, J.M. A novel bio-inspired algorithm applied to selective harmonic elimination in a three-phase eleven-level inverter. *Math. Probl. Eng.* **2020**, *2020*. [CrossRef]
35. Chen, J.; He, Q. Improved butterfly optimization algorithm based on cosine similarity. *J. Comput. Appl.* **2021**, *41*, 2668–2677.
36. Gao, W.; Liu, S.; Xiao, Z.; Yu, J. Butterfly optimization algorithm based on convergence factor and gold sinusoidal guidance mechanism. *Comput. Eng. Des.* **2020**, *41*, 3384–3389.
37. Zheng, H.; Feng, W.; Zhou, Y. Butterfly optimization algorithm based on sine cosine algorithm. *Guangxi Sci.* **2021**, *28*, 152–159.
38. Liu, J.; Ma, Y.; Li, Y. Improved butterfly algorithm for multi-dimensional complex function optimization problem. *Acta Electron. Sin.* **2021**, *49*, 1068–1076.
39. Shao, L.; Han, R. Beetle antenna search flower pollination algorithm. *Comput. Eng. Appl.* **2018**, *54*, 188–194.
40. Moustafa, N.; Slay, J. Proceedings of the UNSW-NB15: A Comprehensive Data Set for Network Intrusion Detection Systems (UNSW-NB15 Network DataSet), 2015 Military Communications and Information Systems Conference (MilCIS), Canberra, ACT, Australia, 10–12 November 2015; pp. 1–6.
41. Moustafa, N.; Slay, J. The evaluation of network anomaly detection systems: Statistical analysis of the UNSW-NB15 data set and the comparison with the KDD99 data set. *Inf. Syst. Secur.* **2016**, *25*, 18–31. [CrossRef]
42. Al-Daweri, M.S.; Ariffin, K.; Abdullah, S.; Senan, M. An analysis of the KDD99 and UNSW-NB15 datasets for the intrusion detection system. *Symmetry* **2020**, *12*, 1666. [CrossRef]
43. Kennedy, J.; Eberhart, R. Particle swarm optimization. In Proceedings of the ICNN'95-International Conference on Neural Networks, Perth, WA, Australia, 27 November–1 December 1995; pp. 1942–1948.
44. Mirjalili, S.; Gandomi, A.H.; Mirjalili, S.Z.; Saremi, S.; Faris, H.; Mirjalili, M.S. Salp swarm algorithm: A bio-inspired optimizer for engineering design problems. *Adv. Eng. Softw.* **2017**, *114*, 163–191. [CrossRef]
45. Mirjalili, S.; Lewis, A. The Whale optimization algorithm. *Adv. Eng. Softw.* **2006**, *95*, 51–67.
46. Ren, J.; Zhang, Y.; Zhang, B.; Li, S. Classification method of industrial internet intrusion detection based on feature selection. *J. Comput. Res. Dev.* **2022**, *59*, 12.