Journal of Network Intelligence
Taiwan Ubiquitous Information
 ISSN 2414-8105
Volume **8**, Number **3**, August **2023**

SDN Anomaly Traffic Identification Based on Fruit Fly Optimized TWSVM

Li Yang*
Sichuan Vocational and Technical College of Communications
Chengdu 611130, China
13106554@qq.com

*Corresponding author: Li Yang
Received February 23, 2023, revised April 6, 2023, accepted June 10, 2023.

Abstract. *A novel kind of communication design is called software defined networking (SDN). SDN makes up the network interface and splits the control layer from the data layer, which enables central control of the network and improves its scalability and programmability. However, SDN also faces a number of network security threats. Anomalous traffic identification technology can defend against malicious traffic attacks and thus protect SDN network security. Therefore, this work proposes an SDN anomalous traffic identification algorithm based on machine learning techniques. Firstly, the SDN security architecture system is studied, and the security architecture is targeted to have centralised control to monitor traffic in real time at the source of the attack. Secondly, the Fruit Fly Optimization (FOA) is improved to address the problems of solidification step and single population type. Then, to address the problem that the existing support vector machine (SVM)-based traffic identification algorithm cannot classify imbalanced data, a multi-classification Twin SVM based on the improved FOA, referred to as IFOA-TWSVM, is proposed in combination with the strong merit-seeking capability of the fruit fly algorithm. The penalty parameters and kernel parameters are dynamically optimised according to the number of positive and negative class samples, which results in a highly accurate multi-classification recognition model. The test results show that the proposed algorithm shows a better advantage in accurately describing the operation of the network compared with various other algorithms.*
Keywords: machine learning; cybersecurity; Fruit Fly Optimization; software defined networking; Twin SVM.

1. **Introduction.** With the rapid development of the Internet, global information sharing has gradually become a reality. In their daily work and life, the general public can work and play through various connected devices, which has greatly improved the efficiency of work and the convenience of life. However, with the rapid increase in the number of Internet users, the volume of online information has also grown dramatically, meaning that the world has entered the era of big data informatization [1,2]. However, the increase in the scale of data has also led to a gradual increase in the difficulty of operation and maintenance. Improper maintenance can lead to failures or anomalies that can cause huge financial losses, such as worm attacks [3], misconfiguration of the network or abnormal server loads [4]. How to efficiently maintain the many network facilities in the big data environment has become a vital issue that must be self-addressed within the field of network maintenance [5].

In today's information age, the emergence of cloud computing and big data has made computer networks dynamic, complex, concurrent and real-time, which requires network

administrators to be able to implement advanced policies to automate the configuration of the network. However, traditional network configuration changes are complex and can easily lead to configuration errors. In addition, the Internet's curing mechanism hinders innovation and evolution of the network infrastructure [6,7]. In order to solve the many challenges of traditional TCP/IP network architecture, Clean Slate of Stanford University proposed the concept of software defined networking (SDN) in 2009 [8,9]. SDN can effectively reduce network load, improve network efficiency and achieve network performance scaling.

However, SDNs are exposed to the same threats of network attacks as traditional networks. Network attacks often occur in the form of traffic anomalies, by which is meant that network traffic does not behave in accordance with the expected normal pattern of behaviour. The presence of anomalous traffic means that there may be some unauthorised access to information and data manipulation in the network, such as denial of service (DoS), worms and viruses [10,11]. Privileged access and attacks on hosts are carried out through the network using known vulnerabilities. Therefore, the frequent occurrence of anomalous traffic in the network will jeopardise the effectiveness and reliability of the network.

Abnormal traffic identification technology refers to the identification of abnormal traffic in the network through effective technical means, and is a basic method to ensure network security [12,13]. Timely and accurate abnormal traffic identification can effectively reduce the impact of malicious attacks on the network and network operation services. real-time abnormal traffic identification in SDN can guarantee the confidentiality, integrity and security of SDN network information, and further promote the development and application of SDN [14,15]. Therefore, the study of abnormal traffic identification techniques in SDN has important theoretical and application values.

Many aspects of the current configuration and operation of network security devices require pattern matching systems based on expert knowledge, which are barely supportable in existing systems, but are out of reach in large-scale cloud computing data centres. As a result, a plethora of machine learning-based intrusion identification techniques have recently emerged. Support vector machines [16,17] stand out for their good classification of small samples and high-dimensional data, but they also have obvious drawbacks. Support vector machines do not train satisfactorily with large amounts of data and cannot handle multiple classification tasks with unbalanced categories.

Therefore, the aim of this work is to construct SDN anomalous traffic identification using advanced machine learning techniques, so as to achieve accurate multi-classification training for large-scale unbalanced traffic data. The proposed algorithm is compared with traditional SVM algorithms and other multi-classification algorithms, and shows better performance in accurately describing the operating conditions of the network. The proposed algorithm shows better advantages in accurately describing the operation of the network when compared with traditional SVM algorithms and other multi-classification algorithms.

1.1. **Related Work.** When faced with the task of operating and maintaining network equipment with such a large volume of data, the traditional approach requires manual work based on experience. Engineers need to monitor the hardware and performance status of the network equipment at all times to keep track of the current operating conditions of the network system and to assess the quality of network services. However, the accuracy of manual operations is very dependent on experience and is less stable. Therefore, accurate fault warning of network service equipment is an essential task in operations and maintenance management. For the identification of network traffic anomalies, there have

been many research results at home and abroad, such as time series analysis [18], wavelet analysis [19], SVM, neural network [20], etc.

In recent years, experts and scholars at home and abroad have studied various techniques for identifying anomalous traffic in networks. Abdulhammed et al. [21] proposed a machine learning-based scheme for detecting anomalous traffic in networks, arguing that the behaviour of anomalous traffic is fundamentally different from other applications and providing guidelines for subsequent research on anomalous traffic detection. Murthy et al. [22] designed a hybrid intelligent intrusion detection system based on Bayesian and genetic algorithms (BAGA). Chiba et al. [23] proposed an intrusion detection algorithm scheme based on Back Propagation (BP) neural network to solve the problem of large training sample size not easy to converge and improve the efficiency and correctness. Paulheim and Meusel [24] proposed a statistical-based anomalous traffic value detection scheme with parameters, in which traffic that does not conform to a set probability distribution range is judged as anomalous traffic. In contrast, Lee et al. [25] propose a statistical-based scheme for detecting anomalous traffic values without parameters. The common drawback of these two schemes is that they cannot efficiently solve the problem of high-dimensional data.

Clustering algorithms, a classical and important unsupervised learning algorithm, are also clearly suitable as a solution. Velea et al. [26] proposed an unsupervised clustering scheme to cluster web users based on the characteristics of their log data using the k-means algorithm, but the scheme has a low detection rate of anomalous traffic and a more complex algorithm. Most of the aforementioned algorithms rely heavily on the correctness of the labels, due to the fact that the labels can determine the final results derived from the model built. However, the current complex network attacks are constantly being updated, when the labels often fail and lead to misclassification of clusters, hence the need to change the labels of the data in real time.

Sakr et al. [27] introduced the traditional SVM to the field of network intrusion detection. the SVM is less dependent on data dimensionality because it solves the sample dimensionality problem. the basic idea of TWSVM is to transpose the traditional quadratic programming problem into two smaller-scale problems. the classification performance of TWSVM is excellent and does not have the disadvantages of sparsity and low generalization, which has become a hot research area in machine learning. TWSVM reduces the training time and improves the classification accuracy. Compared to SVM, TWSVM can cut down the training time by 3/4. Therefore, TWSVM has a clear advantage when dealing with high traffic network monitoring tasks. This work applies the advantages of TWSVM to SDN anomalous traffic identification, in combination with FOA, to effectively handle high-dimensional unbalanced datasets, providing a reference and effective solution for the cyber security field.

1.2. **Motivation and contribution.** Currently, population intelligence algorithms are widely used to solve many complex practical problems, and breakthroughs have been made. The common population intelligence algorithms include Fruit fly optimization algorithm (FOA) [28], particle swarm algorithm, artificial fish swarm algorithm, ant colony algorithm and ant lion algorithm, etc. FOA is a global optimization algorithm based on the foraging behaviour of fruit fly population, and compared with other common population intelligence algorithms, this optimization algorithm has the advantages of simple algorithm principle, less adjustment parameters, less computation, and stronger global optimization capability. Compared with other common population intelligence algorithms, this algorithm has the advantages of simple principle, fewer adjustment parameters, less

computational effort and better global search capability.
The main innovations and contributions of this study are shown below.

(1) FOA has a simple structure, high computational efficiency and is easy to implement, but it also has a few shortcomings, for example solidification step and single population type. Consequently, this work analyses the drawbacks of FOA and proposes corresponding improvement methods.

(2) To address the practical needs of the diversity of attack types, this work proposes the use of a multi-classification structure to implement TWSVM by analyzing the shortcomings of traditional classification recognition algorithms, and proposes a multi-classification TWSVM algorithm based on the improved FOA, referred to as IFOA-TWSVM, by combining the characteristics of the Drosophila algorithm with its strong merit-seeking ability, using IFOA to perform dynamic merit-seeking on the penalty parameters and kernel parameters of TWSVM according to the different numbers of positive and negative class samples The result is a highly accurate multi-classification recognition model.

2. **SDN security architecture system.** When users need to deploy security, they can connect network security devices to the corresponding protected virtual machines at the management level and deploy relevant security policies.
The SDN network security architecture works in a way that the user starts the appropriate virtual machine through the protection process and then, through the SDN, pulls traffic to the entry point of the next layer of security devices and sends down the policy for security protection. In this deployment method, the user only needs to know a small amount of information. The SDN network security architecture is shown in Figure 1.

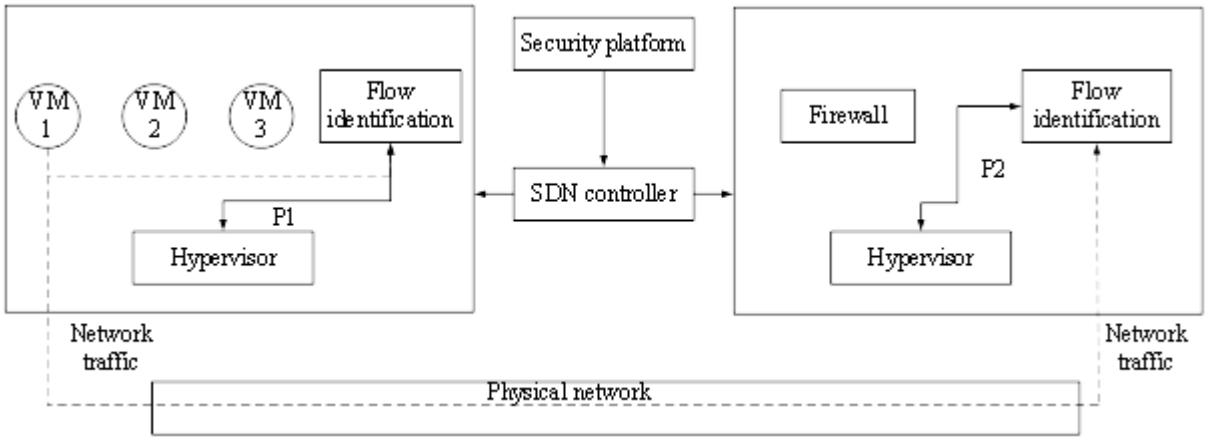

Figure 1. SDN Network Security Architecture

When it is determined that traffic from VM1 to VM2 is to be monitored, the security centre identifies an IPS device from the compute node or security node. An IPS is an intrusion prevention system. The Security Centre then issues a redirected flow command to all switches on path P between the switch where VM1 is located and the switch where the IPS is located. If the IPS on the compute node is selected, the path is P1; if the IPS on the security node is selected, the path is P2. This way all traffic with source VM1 and destination VM2 reaches the IPS device along path P. After inspection the packets are lost from the IPS input port. At this point, the SDN controller calculates the path from the IPS to VM2 based on the topology and issues a flow command. Along the path the switch transmits the traffic to VM2.
Historical data such as target IP addresses and service ports are used as diagnostic indicators in the network layer, application layer and data layer. In exploring source address

IP verification, source address verification is achieved through information interaction between the SDN controller and the OpenFlow switch. SDN technology can control the traffic flow of the network, including the release, blocking, redirection and mirroring of traffic, in addition to statistical traffic information. The combination of abnormal traffic identification algorithms and SDN technology applied to the network access layer allows for real-time forwarding restrictions on abnormal ports. In short, in OpenFlow SDN, the SDN control system can issue FLOW-MOD commands to network devices [29]. The matching traffic is controlled by action actions, as in Figure 2. anomalous traffic identification is an important method in network traffic control. This paper proposes an algorithm for port scanning traffic anomaly identification, which can play a good role in supporting future research on network security posture assessment models.

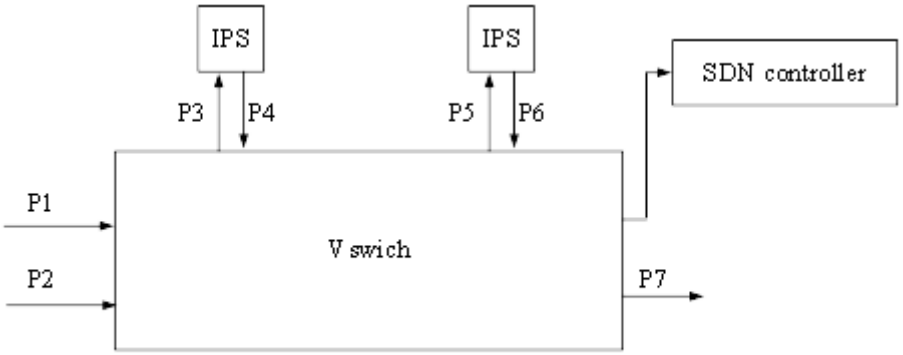

Figure 2. Traffic control on virtual Swich in SDN

3. **Problems and shortcomings of network traffic identification.**

3.1. **Principle of network traffic identification.** Let there exist a directed connectivity graph $N =< V, E >, n = |V|, m = |E|$ in an SDN.
Each edge $< i, j >$ has a non-negative statistical traffic $C(i, j)$. n has two special vertices s and t. s is called the downlink point, t is called the receive point, and the remaining vertices are called intermediate points. We call N the statistical network traffic of an SDN and denote it as $N =< V, E, c, s, t >$.
Let $f : E \to R*$, where R is the set of non-negative numbers. The capacity limit and the equilibrium condition are each satisfied by the following 2 conditions.

$$\forall < i, j >\in E, f(i, j) \leq c(i, j) \tag{1}$$

$$\forall i \in V - \{s, t\}, \sum_{<j,i>\in E} f(j, i) = \sum_{<i,j>\in E} f(i, j) \tag{2}$$

We call f a feasible flow on N. The net flow at the downstream point s is the flow of the feasible flow, denoted as $v(f)$.

$$v(f) = \sum_{<s,j>} f(s, j) - \sum_{<j,s>} f(j, s) \tag{3}$$

This work uses OpenFlow technology based on the directed graph model to implement SDN anomaly traffic collection. As a result, all paths of attack traffic can be traced without the need to query IP addresses. In addition, the number of flow entries generated by the relevant switch during the tracing process is relatively small. The switch traceability rules are highly scalable.

3.2. **Shortcomings of traditional network traffic anomaly identification algorithms.** In practice, anomaly identification based on expert knowledge definition has the problem of high cost and reliance on human experience. Nowadays, machine learning based network traffic anomaly identification is the future trend.

At this stage, researchers are needed to use SVM for classification models and show better applicability than neural networks.SVM is an excellent general machine learning method based on statistical learning theory. The main idea of SVM is to find the hyper plane with the maximum interval distance so as to separate the positive and negative class samples.SVM solves the SVM solves the problem of high data dimensionality.

Let the sample set be $(x_1, y_1), \ldots, (x_n, y_n)$, the sample database be $X \in R^{l \times n}$, and the category label database be $Y \in R^{l \times l}$. where l represents the number of samples and n represents the feature dimension. The support vector machine finds the best hyper plane (maximum interval) over the two categories of samples.

$$w^T \varphi(x) + b = 0 \tag{4}$$

Where w is the optimal hyperplane normal vector and b is the bias value.
The optimal hyperplane is obtained by solving the constrained optimisation problem below.

$$\min_{w,b,\xi} \frac{1}{2}||w||^2 + c\sum_{i=1}^{l} \xi_i \;\; s.t\; y_i \left(w^T \phi\left(x_i\right) + b\right) + \xi_i \geq 1, \xi_i \geq 0, i = 1, 2, \cdots, l \tag{5}$$

Where ξ is the amount of slack and c is the penalty parameter.
The problem of constructing an optimal hyperplane becomes a quadratic programming problem. The quadratic programming problem is then transformed into a dyadic problem by introducing Lagrange multipliers. The role of the kernel function is to achieve the low-dimensional to high-dimensional transformation without any increase in computational effort. A commonly used SVM kernel function is the RBF kernel function.

$$K\left(x_i, x_j\right) = \exp\left(\frac{\left\|x_i - x_j\right\|^2}{-2\sigma^2}\right) \tag{6}$$

Through the introduction of classical SVM, it can be understood that SVM presents an effective solution for solving the high-dimensional and non-linear traffic anomaly identification problem. However, in practical applications, SDN traffic anomaly identification is going to need to address data imbalance, training time and multiple classification problems.
The training time of SVM increases significantly, so it does not meet the real-time requirement well. In addition, SVM is not ideal for classification of unbalanced datasets. SVM is a classical two classifiers. However, in network data, it is mainly divided into various types such as normal data, Oos, Probe, U2R and R2L. In order to be more relevant to practical needs, multi-classification detection algorithms are one of the key elements of this work.

4. SDN anomaly traffic identification based on IFOA-TWSVM..

4.1. **Construction of a multi-classification TWSVM.** Based on the traditional SVM principle, the core idea of TWSVM is to construct two hyper planes for positive and negative class samples respectively. TWSVM is a great improvement in the solution of the heterogeneous problem, so the training time is greatly reduced [30]. TWSVM is more excellent in processing classification tasks in real time, so it is more suitable for traffic anomaly identification than SVM.

The goal of a linear bipartite support vector machine is to seek two hyperplanes in two dimensions that satisfy the following conditionshyper plane.

$$\begin{cases} \omega_+ \cdot x + b_+ = 0 \\ \omega_- \cdot x + b_- = 0 \end{cases} \tag{7}$$

Since most solution problems are linearly inseparable, TWSVM also introduces kernel functions. Then the TWSVM seeks the hyperplane as shown below.

$$K(x^T, C^T)u_+ + b_+ = 0, K(x^T, C^T)u_- + b_- = 0 \tag{8}$$

Where $C = [A; B] \in R^{l \times n}$ and K are the kernel functions. Let $A = (x_1, \cdots x_p)^T$ be all the positive samples in the training sample set T and $B = (x_{p+1}, \cdots x_{p+q})^T$ be all the negative samples in the training sample set T.

Similarly, the plane dividing the positive and negative classes respectively satisfies the following conditions.

$$\min_{u_+ b_+ \xi_-} \frac{1}{2}\left\|K(A, C^T)u_+ + e_+ b_+\right\|^2 + c_1 {e_-}^T \xi_- \quad s.t. -(K(B, C^T)u_+ + e_- b_+) + \xi_- \geq e_-, \xi_- \geq 0 \tag{9}$$

$$\min_{u_- b_- \xi_+} \frac{1}{2}\left\|K(B, C^T)u_- + e_- b_-\right\|^2 + c_2 {e_+}^T \xi_+ \quad s.t. (K(A, C^T)u_- + e_+ b_-) + \xi_+ \geq e_+, \xi_+ \geq 0 \tag{10}$$

To further simplify equations (9) and (10), they are pairwise transformed.

$$\max_{\alpha} {e_-}^T \alpha - \frac{1}{2}\alpha^T R(S^T S)^{-1} R^T \alpha \quad s.t. 0 \leq \alpha \leq c_1 e_- \tag{11}$$

$$\max_{\gamma} {e_+}^T \gamma - \frac{1}{2}\gamma^T S(R^T R)^{-1} S^T \gamma \quad s.t. 0 \leq \gamma \leq c_2 e_+ \tag{12}$$

Where $R = [K(B, C^T)e_-]$ and $S = [K(A, C^T)e_+]$. Solving equation (11) and equation (12) yields

$$(u_+^T, b_+)^T = -(S^T S)^{-1} R^T \alpha \tag{13}$$

$$(u_-^T, b_-)^T = -(R^T R)^{-1} S^T \gamma \tag{14}$$

The resulting classified hyperplane is shown below.

$$classlabel = \arg \min_{k=+,-} |K(x^T, C^T)u_k + b_k| \tag{15}$$

TWSVM itself is a binary classifier, whereas the traffic anomaly identification studied in this work has a greater need for multi-classification capability, so the one-versus-one approach is used to construct a suitable multi-classification TWSVM algorithm.

Initially, a TWSVM is created among any 2 categories of samples, resulting in k(k-1)/2 TWSVMs for k clusters of samples. When categorizing an unknown sample, the class that receives the most votes is deemed the category of that unknown sample.

Suppose there are four classes: A,B,C,D. During training, the data corresponding to A,B; A,C; A,D; B,C; B,D; C,D are selected as the training set and six training results are obtained. In testing, the corresponding data are put to test against each of the six results. A voting format is then taken and a final set of results is obtained.

The voting process [31] is shown below.

a=b=c=d=0.

(A, B)-classifier: if A wins, then A=A+1; otherwise, B=B+1;

(A, C)-classifier: if A wins, then A=A+1; otherwise, C=C+1;

...

(C, D)-classifier: if A wins, then C=C+1; otherwise, D=D+1;

The decision is the Max(A, B, C, D).

It can be seen that multi-classification TWSVM solves the problems of long practice time and binary classification that exist in traditional SVM, but the problem of unbalanced data classification still exists. For example, the number of positive class samples is much larger than the number of negative class samples. Therefore, FOA will be used subsequently to adjust the penalty parameters of the multi-classification TWSVM so as to solve this problem

4.2. **Improved FOA.** FOA is a global foraging algorithm based on the foraging behaviour of a fruit fly population.

Compared with other common population intelligence algorithms, FOA has the advantages of a simple algorithm principle, fewer adjustment parameters, less computational effort and better global search capability. However, FOA also has a few shortcomings, for example solidification step and single population type. Therefore, in this work, an improved FOA (IFOA) is proposed by analysing the shortcomings of FOA.

In a FOA, the population is made up of a number of individual fruit flies, each with a positional coordinate. All four individuals are in the best possible position if the weight vector is the highest in the population. In other words, the entire population will gravitate toward one of the four positions, creating a deceptive position.

IFOA employs a series of angle vectors to describe the location coordinates X of individual fruit flies in order to overcome these problems.

$$X = [\theta_1, \theta_2, \cdots, \theta_i] \tag{16}$$

where i denotes the dimension of the variable and denotes the phase angle. Since the tan function is monotonically increasing in the range $[-\pi/4, \pi/4]$ and its value domain is [-1,1]. The flavour concentration of an individual can thus be represented by the judged value Smell.

$$\text{Smell} = \tan(X) \tag{17}$$

The principle of encoding the tangent function of individual positions in IFOA is shown in Figure 3.

(1)Two-stock strategy.

In FOA, the whole population is considered as a whole, and each individual in the population will follow the same rules for visual search, which will lead to a reduction in the retrieve ability of the community as a whole. The entire community may settle into a local optimal location if the location of the global optimal solution to the objective function is far or the location of the optimal individual is not perfect, resulting in the final search result of the algorithm is also a local optimal solution.

To address this problem, IFOA records the optimal and worst positions after each iteration based on the dimensions of the individual fitness value and figures out the distance among each individual. An individual will be categorized as belonging to a subpopulation with more search capacity if it is situated nearer to the ideal place, otherwise it will be classified as a subpopulation with lower search power. Different search strategies are also assigned to each population according to its characteristics.

(2) Adaptive search radius.

According to the two-population strategy, IFOA has two small subpopulations. For the subpopulation with stronger search capability, the Iteration step size should be progressively reduced to get a local, precise search around the global optimum position. For the subpopulation with poorer search ability, the search radius should be expanded, thus enhancing its global search ability. In summary, the search radius has been adapted in

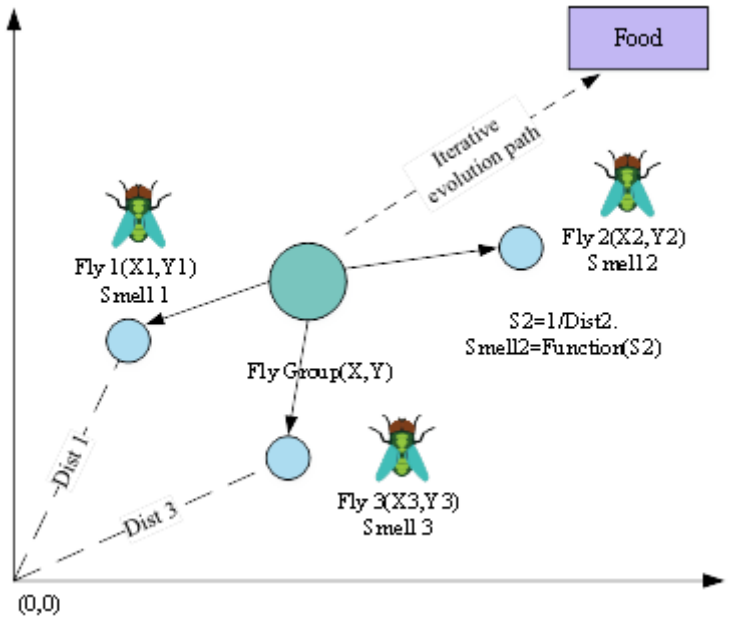

Figure 3. Principle for encoding the tangent function of individual positions in IFOA

order to enable it to be adjusted adaptively.

$$\begin{cases} R_{Best} = R \cdot \left(1 - \frac{g_i}{m \cdot g_{\max}} - \frac{\text{fitness}_{i-1}}{n \cdot \text{fitness}_i}\right) \\ R_{\text{worst}} = R \cdot \left(1 + \frac{g_i}{m \cdot g_{\max}} + \frac{\text{fitness}_{i-1}}{n \cdot \text{fitness}_i}\right) \end{cases} \tag{18}$$

Where g_i is the present generation, g_{max} is the maximum generation, R_{Best} is the iteration step of the subgroup with better ability, R_{worst} is the iteration step of the subgroup with worse ability, R is the initial iteration step, m and n are both parameters of the adaptive radius, and i is the number of iterations.

4.3. **Classification and identification process based on IFOA-TWSVM.** The SDN anomaly traffic identification in this work is divided into four main phases: feature acquisition, parameter seeking phase, training phase and prediction phase.

(1) We input the dataset with all the features into the TWSVM model for training to obtain the classification accuracy R. Then, after removing one feature f in turn, the feature set is input into the TWSVM model for training again to obtain the classification accuracy R'. If $R > R'$, f is a useful feature and is retained. If $R < R'$, f is an irrelevant or negative feature and is not retained. Finally, the resulting set of features is the set of features we want.

(2) TWSVM will find the best parameter by IFOA for 3 parameters (kernel function parameter σ, penalty variables c_1 and c_2). The position of an individual can be represented by the horizontal and vertical coordinates in two-dimensional coordinates. The steps taken to optimise the TWSVM parameters by adopting IFOA are shown in Table 1.

(3) In the training and prediction phases, the RBF kernel function is mainly used in this paper. Solving the quadratic programming problem will be implemented using the functions in the Medium Optimisation Toolkit.

(4) As the one-versus-one method is used to construct a suitable multi-classification TWSVM algorithm, the training *k(k-1)/2* TWSVM classification hyperplanes are stored

in V as column vectors. In the prediction phase, the predicted samples are substituted into the two hyper plane equations separately to derive the distance calculation results. The hyper plane with the smaller distance is the predicted classification result for that sample. Voting is carried out according to the one-versus-one method to determine the final classification result.

Table 1. Pseudocode for IFOA-TWSVM.

IFOA selection penalty parameters and kernel function parameters

```
Inputs: training dataset T and test dataset P;
Outputs: c1 , c2 , σ and maximum recognition rate;
X_axis=10*rands(1,3);
Y_axis=10*rands(1,3);
Set the number of iterations maxgen=100 and the population size sizepop=20
for i=1:sizepop

   // Initial Drosophila individual flight distance.
    X(i, :)=X_axis+2*rand()-1;Y(i, :)=Y_axis+2*rand()-1;
    D(i, 1) = (X(i, 1))^2 + (Y(i, 1))^2)^0.5.
    // Find the distance from the origin.
    D(i, 2) = (X(i, 2))^2 + (Y(i, 2))^2)^0.5.
    D(i, 3) = (X(i, 3))^2 + (Y(i, 3))^2)^0.5.
    //Determine the flavour concentration determination value
    Smell(i, 1) = 1/D(i, 1);
    Smell(i, 2) = 1/D(i, 2);
    Smell(i, 3) = 1/D(i, 3);
    // Adjustment of three parameters using IFOA
    c1=Smell(i, 1);
    c2=Smell(i, 2);
    σ = Smell(i, 3);
    Inputting the training and test sets into the TWSVM model to obtain the accuracy (Small(i)).
    Record bestSmall = max(Small) and the corresponding best positions X_axis and Y_axis.
end
for gen = 1:maxgen
    for i = 1:sizepop
    Increasing the flight distance of individual Drosophila at the best position of the previous generation.
    end
    Record bestSmall = max(Small) and the corresponding best positions X_axis and Y_axis.
end
Output best: c1 , c2 , σ and maximum recognition rate.
```

5. **Testing and analysis.**

5.1. **Experimental configuration.** In order to verify the performance of the SDN anomaly traffic identification method proposed in this paper in a big data environment, performance test analysis was conducted on a Hadoop platform.
The test cluster machine environment parameters are: 10 web server nodes with Intel i7 processors, 3.2GHz CPU and 8GB RAM. All service nodes communicate with each other via 1000M fibre. One of the nodes is set up as a Jobtracker and the other 9 compute nodes are set up as tasktrackers, each with one reduce job slot and two map job slots. The hardware and software parameters of each service node are shown in Table 2.

5.2. **Evaluation metrics.** In order to verify the energy saving effect of the proposed IMV-MLF algorithm, this paper compares In order to quantify the performance of anomalous traffic identification, the two most commonly used comprehensive evaluation metrics

Table 2. Software and hardware parameters of the experimental environment.

Hardware	Software environment
Intel CPU 16 cores	Ubuntu 32-bit
16 GB of memory	Java 1.7
Hard Drive 500G	
Network bandwidth 100M	

were selected: the relative error and the model accuracy , both of which were calculated as follows.

$$\mathrm{E} = \frac{|y(t) - \hat{y}(t)|}{y(t)} \tag{19}$$

$$\mathrm{P} = \sqrt{\sum_{t=1}^{n} \frac{|y(t) - \hat{y}(t)|^2}{n}} \tag{20}$$

Where $y(t)$ is the true value, $\hat{y}(t)$ is the predicted value and n is the predicted length of time.

5.3. **Data pre-processing.** An example of network traffic data cached on one web server node over a 24-hour period.
The sample is the average traffic rate every 5 minutes, with a total of 288 sample points, whose actual network traffic profile is shown in Figure 4. It can be seen that the traffic is

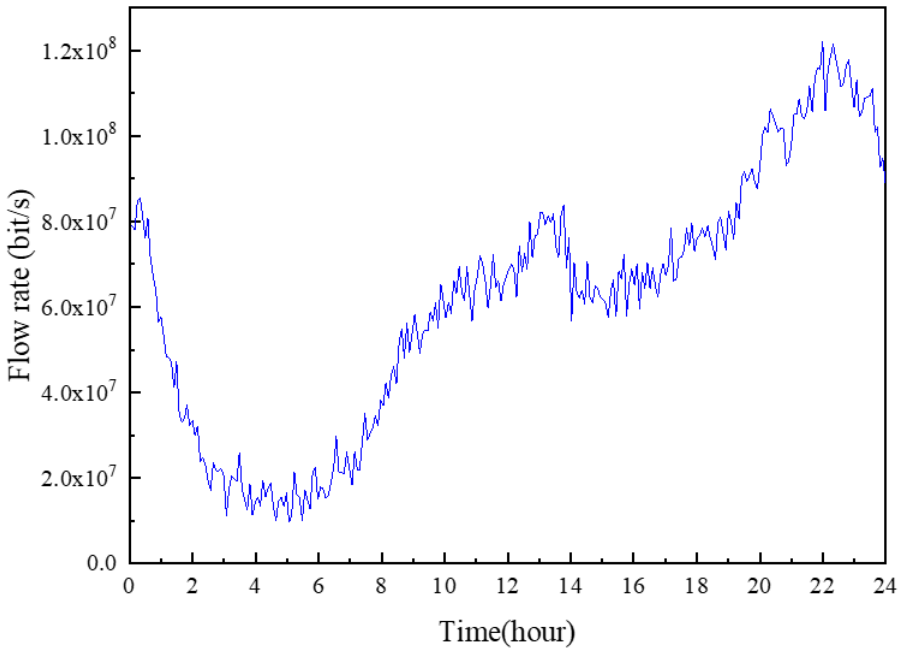

Figure 4. Example of a day's flow curve.

lowest between 1am and 6am and highest between 8pm and 12pm, which is in line with the actual normal operation of the web server. However, due to the large values of the network traffic, the collected data series will be logarithmically processed to reduce the standard deviation as described earlier, and the smoothed series is shown in Figure 5.

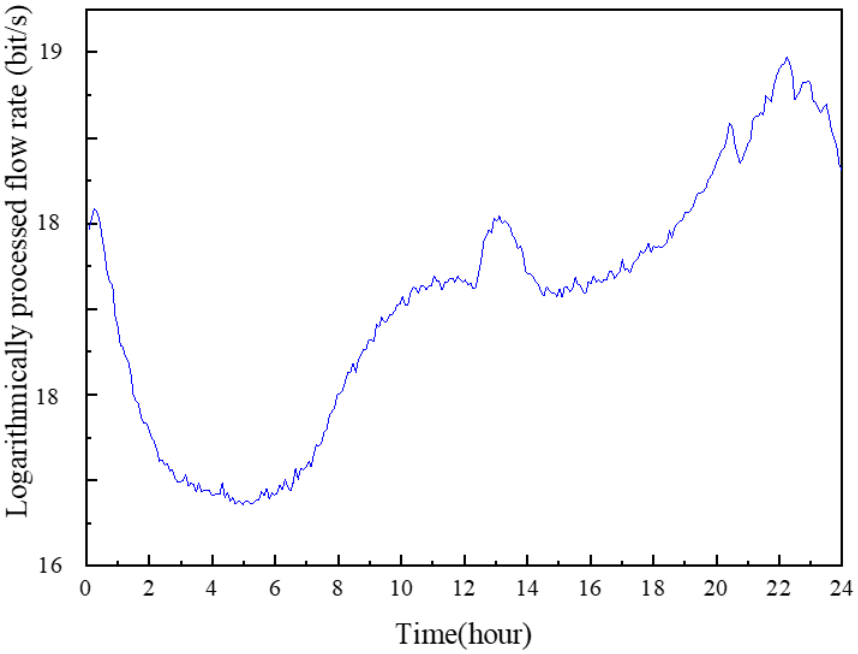

Figure 5. Flow curve after treatment

5.4. **Anomaly test results.** Ten days of traffic data were collected as a dataset with a total of 2880 data points using the average traffic rate every 5 minutes as a sampling point.

The number of observed anomalies was 47. Network traffic monitoring was performed on this dataset using SVM, BP neural network, convolutional neural network (CNN), and IFOA-TWSVM. The relative errors and model accuracies of the four models were calculated as shown in Table 3.

It can be seen that the proposed IFOA-TWSVM can obtain better recognition results

Table 3. Error comparison of different models.

Date	**Relative error**			
	SVM	**BP Neural Networks**	**CNN**	**IFOA-TWSVM**
2018-02-01	0.0209	0.0060	0.0160	0.0060
2018-02-02	0.0196	0.0173	0.0182	0.0160
2018-02-03	0.0114	0.0072	0.0098	0.0054
2018-02-04	0.0691	0.0533	0.0603	0.0413
2018-02-05	0.0192	0.0092	0.0183	0.0072
2018-02-06	0.0508	0.0104	0.0421	0.0043
2018-02-07	0.0341	0.0118	0.0228	0.0057
2018-02-08	0.0172	0.0091	0.0143	0.0076
2018-02-09	0.0992	0.0868	0.0975	0.0723
2018-02-10	0.0872	0.0076	0.0813	0.0066

compared to the other three classification models. the relative error of IFOA-TWSVM is smaller and the model accuracy is the best. In addition, the number of anomalies correctly detected by SVM, BP neural network and CNN were 33, 40 and 37 respectively, while IFOA-TWSVM correctly detected more anomalies, reaching 44, and the accuracy

of anomaly recognition reached 93.6%. Taken together, the above results conclude that IFOA-TWSVM is significantly effective in SDN traffic anomaly identification.
The training time of SVM was compared with that of IFOA-TWSVM for various types of attacks. the training time of IFOA-TWSVM was significantly lower than the sample training time of SVM, with a ratio of basically 1:2, and sometimes even up to 1:4.
From the analysis of the experimental results, the multi-classification IFOA-TWSVM has an advantage in terms of recognition accuracy. At the same time, the training time of IFOA-TWSVM is almost 3/4 less than that of traditional SVM, so IFOA-TWSVM has better multi-classification results in terms of both detection accuracy and training time. From the experiments for large amount of data processing, the training time advantage provides great support for SDN anomaly detection in a large data environment.

6. **Conclusion.** This work proposed a multi-classification IFOA-TWSVM based SDN anomaly traffic identification algorithm. By analyzing the shortcomings of traditional classification recognition algorithms, this work uses a multiclassification structure to implement TWSVM.FOA has the features of high computational efficiency and easy implementation, but it also has some shortcomings, for example solidification step and single population type. Therefore, IFOA is proposed in this work, and the penalty and kernel parameters of TWSVM are dynamically optimised according to the number of positive and negative class samples, which results in a highly accurate multi-classification recognition model. There is still room for refinement in the mechanism of multi-classification, as the one-to-one approach may result in the same number of votes for several classifications, which ultimately affects the detection rate. Therefore, there is still much room for improvement in the recognition rate in multi-classification anomaly detection.

Funding Statement. This work supported by the project "Sichuan Research Center of Higher Vocational Education Project, NO. GZY22C21"and "National Industrial and Information Technology Vocational Education and Teaching Steering Committee, NO. GXHZWC15750".

REFERENCES

[1] T.-Y. Wu, L. Wang, X. Guo, Y.-C. Chen, and S.-C. Chu, "SAKAP: SGX-Based Authentication Key Agreement Protocol in IoT-Enabled Cloud Computing," *Sustainability*, vol. 14, no. 17, 11054, 2022.
[2] T.-Y. Wu, Q. Meng, S. Kumari, and P. Zhang, "Rotating behind Security: A Lightweight Authentication Protocol Based on IoT-Enabled Cloud Computing Environments," *Sensors*, vol. 22, no. 10, 3858, 2022.
[3] Y. Li, G. Huang, C. Wang, and Y. Li, "Analysis framework of network security situational awareness and comparison of implementation methods," *EURASIP Journal on Wireless Communications and Networking*, vol. 2019, no. 1, 2019.
[4] T.-Y. Wu, L. Yang, J.-N. Luo, and J. Ming-Tai Wu, "A Provably Secure Authentication and Key Agreement Protocol in Cloud-Based Smart Healthcare Environments," *Security and Communication Networks*, vol. 2021, pp. 1–15, 2021.
[5] Y. Tang and M. Elhoseny, "Computer network security evaluation simulation model based on neural network," *Journal of Intelligent & Fuzzy Systems*, vol. 37, no. 3, pp. 3197-3204, 2019.
[6] M.-E. Wu, J.-H. Syu, and C.-M. Chen, "Kelly-Based Options Trading Strategies on Settlement Date via Supervised Learning Algorithms," *Computational Economics*, vol. 59, no. 4, pp. 1627–1644, 2022.
[7] C.-M Chen, Q. Miao, S. Kumar, and T.-Y. Wu, "Privacy-preserving authentication scheme for digital twin-enabled autonomous vehicle environments," *Emerging Telecommunications Technologies*, 2023, [Online]. Available: https://doi.org/10.1002/ett.4751
[8] M. Angelini, G. Blasilli, T. Catarci, and S. Lenti, "Vulnus: Visual Vulnerability Analysis for Network Security," *IEEE Transactions on Visualization and Computer Graphics*, vol. 25, no. 1, pp. 183-192, 2019.

[9] V. Kumar and O. P. Roy, "Enhanced Network Security for Improved Trustworthiness of VoIP Applications via Cuckoo Search and Machine Learning," *Indian Journal of Science and Technology*, vol. 15, no. 15, pp. 677-688, 2022.
[10] J. Xiao, B. Q. Zhang, and F. Z. Luo, "Distribution Network Security Situation Awareness Method Based on Security Distance," *IEEE Access*, vol. 7, pp. 37855-37864, 2019.
[11] D. Kreutz, F. M. V. Ramos, P. Esteves Verissimo, and C. Esteve Rothenberg, "Software-Defined Networking: a Comprehensive Survey," *Proceedings of the IEEE*, vol. 103, no. 1, pp. 14-76, 2015.
[12] W. F. Xia, Y. G. Wen, C. H. Foh, D. Niyato, and H. Y. Xie, "A Survey on Software-Defined Networking," *IEEE Communications Surveys & Tutorials*, vol. 17, no. 1, pp. 27-51, 2015.
[13] B. A. A. Nunes, M. Mendonca, X.-N. Nguyen, K. Obraczka, and T. Turletti, "A Survey of Software-Defined Networking: Past, Present, and Future of Programmable Networks," *IEEE Communications Surveys & Tutorials*, vol. 16, no. 3, pp. 1617-1634, 2014.
[14] H. Kim and N. Feamster, "Improving network management with software defined networking," *IEEE Communications Magazine*, vol. 51, no. 2, pp. 114-119, Feb. 2013, doi: 10.1109/mcom.2013.6461195.
[15] S. H. Yeganeh, A. Tootoonchian, and Y. Ganjali, "On scalability of software-defined networking," *IEEE Communications Magazine*, vol. 51, no. 2, pp. 136-141, Feb. 2013.
[16] J. Nalepa and M. Kawulok, "Selecting training sets for support vector machines: a review," *Artificial Intelligence Review*, vol. 52, no. 2, pp. 857-900, 2018.
[17] J. Cervantes, F. Garcia-Lamont, L. Rodríguez-Mazahua, and A. Lopez, "A comprehensive survey on support vector machine classification: applications, challenges and trends," *Neurocomputing*, vol. 408, pp. 189-215, 2020.
[18] P. M. Maçaira, A. M. Tavares Thomé, F. L. Cyrino Oliveira, and A. L. Carvalho Ferrer, "Time series analysis with explanatory variables: A systematic literature review," *Environmental Modelling & Software*, vol. 107, pp. 199-209, 2018.
[19] A. Zaremba, Z. Umar, and M. Mikutowski, "Inflation hedging with commodities: a wavelet analysis of seven centuries worth of data," *Economics Letters*, vol. 181, pp. 90-94, 2019.
[20] F. C. Liu, W. J. Huo, and Y. Han, "Study on Network Security Based on PCA and BP Neural Network Under Green Communication," *IEEE Access*, vol. 8, pp. 53733-53749, 2020.
[21] R. Abdulhammed, H. Musafer, A. Alessa, M. Faezipour, and A. Abuzneid, "Features Dimensionality Reduction Approaches for Machine Learning Based Network Intrusion Detection," *Electronics*, vol. 8, no. 3, p. 322, 2019.
[22] Y. Murthy, K. Harish, D. Varma, K. Sriram, and B. Revanth, "Hybrid Intelligent Intrusion Detection System using Bayesian and Genetic Algorithm (BAGA): a Comparitive Study," *International Journal of Computer Applications*, vol. 99, no. 2, pp. 1-8, 2014.
[23] Z. Chiba, N. Abghour, and K. Moussaid, "A novel architecture combined with optimal parameters for back propagation neural networks applied to anomaly network intrusion detection," *Computers & Security*, vol. 75, pp. 36-58, 2018.
[24] H. Paulheim and R. Meusel, "A decomposition of the outlier detection problem into a set of supervised learning problems," *Machine Learning*, vol. 100, no. 2-3, pp. 509-531, 2015.
[25] S. Lee, H. G. Kim, and Y. M. Ro, "BMAN: Bidirectional Multi-Scale Aggregation Networks for Abnormal Event Detection," *IEEE Transactions on Image Processing*, vol. 29, pp. 2395-2408, 2020.
[26] R. Velea, C. Ciobanu, and I. Bcia, "Network Traffic Anomaly Detection Using Shallow Packet Inspection and Parallel K-means Data Clustering," *Studies in Informatics and Control*, vol. 26, no. 4, 2017.
[27] M. Sakr, M. Tawfeeq, and A. El-Sisi, "Network Intrusion Detection System based PSO-SVM for Cloud Computing," *International Journal of Computer Network and Information Security*, vol. 11, no. 3, pp. 22-29, 2019.
[28] A. Darvish and A. Ebrahimzadeh, "Improved Fruit-Fly Optimization Algorithm and Its Applications in Antenna Arrays Synthesis," *IEEE Transactions on Antennas and Propagation*, vol. 66, no. 4, pp. 1756-1766, 2018.
[29] J. Tourrilhes, P. Sharma, S. Banerjee, and J. Pettit, "SDN and OpenFlow Evolution: A Standards Perspective," *Computer*, vol. 47, no. 11, pp. 22-29, 2014.
[30] N. Goyal and K. Gupta, "A hierarchical laplacian TWSVM using similarity clustering for leaf classification," *Cluster Computing*, vol. 25, no. 2, pp. 1541-1560, 2022.
[31] J. Sun, Q. Li, M. Q. Chen, and L. Ren, "Optimization of models for a rapid identification of lithology while drilling - A win-win strategy based on machine learning," *Journal of Petroleum Science and Engineering*, vol. 176, pp. 321–341, 2019.