Advances in Intelligent Systems Research, volume 147
International Conference on Network, Communication, Computer Engineering (NCCE 2018)

Application of Improved Fireworks Algorithm Optimized SVM in Intrusion Detection

He Li[1, a)], Xiang Ji[2, b)] and Jingmei Li[1, c)]

[1]*Harbin Engineering University, Harbin 150001, China.*
[2]*No.713 Research Institute of CSIC, Zhengzhou 450015, China.*

[a)] 1280096748@qq.com
[b)] jixiangheu@163.com
[c)]lijingmei@hrbeu.edu.cn

Abstract. In order to further improve the effectiveness and accuracy of intrusion detection and avoid the security risks brought by network attacks, this paper proposes an intrusion detection strategy based on improved fireworks algorithms—TFWA-SVM. This algorithm first optimizes the fireworks algorithm, constrains the initial fireworks position in the fireworks algorithm, avoids the waste of computing resources caused by the initial concentration of the fireworks. At the same time, it uses the fitness function stretching technology to make the algorithm have more superior global exploration capabilities. The improved fireworks algorithm is applied in the selection of SVM penalty factor and kernel function parameters, and then the powerful classification ability of SVM is used to classify the data packets in the network. Using the BPNN, SVM, FWA-SVM and TFWA-SVM to simulate the KDD99 data set. The experimental results show that TFWA-SVM has obvious advantages in convergence speed and classification accuracy. It can improve the quality of intrusion detection to a certain extent and has extensive research value.

Keywords: intrusion detection, fireworks algorithm, support vector machine, kdd-99 data set.

INTRODUCTION

In recent years, with the rapid development of information technology, information technology has affected all aspects of individuals, society, and countries. In order to effectively control the illegal invasion of the network, scholars have conducted a lot of researches. The neural network based on error back propagation algorithm BP is widely used in intrusion detection. However, the BP neural network is based on the principle of minimizing empirical risk. Insufficient learning sample is likely to cause overfitting. Excessive learning sample is more likely to fall into the dimension disaster and the generalization ability is not high.

Various types of swarm intelligence optimization algorithms are also used in intrusion detection. In order to improve the accuracy of network intrusion detection, Feng Y. and others combined the K-nearest neighbor algorithm with the improved particle swarm optimization algorithm and proposed a network intrusion detection model that combines the above algorithms to improve the detection effect of network intrusion [1]. Yang h. and others used the probabilistic jump and strong local search ability of simulated annealing (SA) algorithm to improve the genetic algorithm, and double optimized the network weight and structure of the neural network. However, due to the combination of multiple optimization algorithms, the computational complexity increases drastically and the detection speed decreases [2]. Support vector machine (SVM) adopts the principle of minimizing structural risk, which can effectively solve the regression problems of small samples, nonlinear. It also has strong generalization ability, can find the global optimal solution, overcoming the problem of local extremum of neural network [3, 4].

FWA (Fireworks Algorithmis) [5] is a new swarm intelligence optimization algorithm published by Professor Tan Ying at the first international community intelligence conference in 2010. Fireworks algorithms not only have many advantages as existing swarm intelligence algorithms has, but also has many of its own characteristics. It has

the characteristics of strong global optimization, rapid convergence, and simple implementation. This design uses the improved global search ability and excellent convergence ability of the improved fireworks algorithm to optimize the SVM parameter model and uses the strong classification ability of the SVM in conjunction with regression to determine whether the data packet in the network is an attacking data packet. The experimental results show that this improved SVM-based intrusion detection model with improved fireworks algorithm has high detection accuracy.

METHOD PRINCIPLE

Support Vector Machine.

Support vector machine classify the data by mapping the data samples from the original space to the higher dimensional feature space through nonlinear mapping functions, and then constructing the optimal hyperplane in the higher dimensional feature space. Maximize the distance of the two kinds of sample in the hyperplane. Assuming training data:

$$\begin{aligned} &(x_1, y_1),(x_2, y_2),\cdots,(x_n, y_n) \quad x \in \mathrm{R}^n \quad y \in (-1,+1) \\ &\textit{Hyperplane is marked as} \quad (\omega \bullet x+b)=0 \end{aligned} \tag{1}$$

In order to construct the optimal hyperplane, let two kinds of sample intervals are maximized. The maximum separation interval can be converted to the following optimization problem.

$$\varphi(w) = \frac{1}{2} \| w \|^2 + C\sum_{i=1}^{n} \xi_i \tag{2}$$

$$s.t \quad y_i \ [(\omega \bullet x_i+b)]-1+\xi_i \geq 0 \quad i=1,\cdots, n \tag{3}$$

The relaxation variable $\xi_i \geq 0$, C is a constant greater than zero, determines the penalty for misclassifying samples [6]. Introducing the Lagrange multiplier to convert it to a dual problem:

$$Q(\alpha) = \sum_{i=1}^{n} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{n} \alpha_i \alpha_j y_i y_j K(x_i, x_j) \tag{4}$$

$$s.t \quad \sum_{i=1}^{n} y_i \alpha_i = 0 \quad 0 \leq \alpha_i \leq C \quad i=1,\cdots, n \tag{5}$$

$$f(x) = \{\mathrm{sgn}\}\{ \sum_{i=1}^{n} \alpha_i^* \ y_i k(x_i \bullet x) + b^* \} \tag{6}$$

The final decision function can be obtained by solving the above quadratic programming problem:

The RBF radial basis kernel function has the advantages of wide convergence domain and good nonlinear mapping ability. Therefore, this paper uses RBF radial basis kernel function as the kernel function of the support vector machine classifier, combined with the support vector machine penalty factor C and the kernel parameter g affects the premise of the classification precision and generalization performance of the support vector machine, and an optimization method based on these two parameters is established [7]. The penalty factor C is used to balance the confidence range of the SVM with the empirical risk ratio. The larger the value of C, the greater the punishment for the empirical error, and the classifier is easily affected by the noise and error in the sample. Although the

classification accuracy of the sample is very high, it is easy to cause overfitting, leading to a decrease in accuracy when applied to the test set. Similarly, when the penalty factor C value is too low, the under-fitting is likely to occur.

The choice of kernel parameter g will directly affect the characteristics of the kernel function. When the value of g changes, it will change the complexity of sample data distribution in the high-dimensional feature space, that is, change the dimension of the high-dimensional feature space, the dimension is too large. Too small will make the structural risk cannot be minimized.

Therefore, for each characteristic subspace, there is a set of the most suitable penalty factor C and kernel parameter g, which makes the classification of SVM the best.

The Basic Fireworks Algorithm.

Fireworks algorithm is a swarm intelligence algorithm proposed by Prof. Tan Ying in 2010. Compared with other swarm intelligence algorithms, it has the advantages of high parallelism and strong global optimization ability. In each iteration of the algorithm, fireworks and sparks generated by explosions and Gaussian mutation sparks together constitute the current feasible solution. The fitness function is used to judge the quality of the fireworks. In the next round of iterations, good-quality firework explosions produce more sparks and nearer explosion distances; fireworks with poor quality produce fewer sparks and have longer explosion distances.

The standard fireworks algorithm is described as follow steps: Find a point x in the D-dimensional feasible domain Ω so that it corresponds to the fitness function f(x), which has a global minimum fitness value. The specific steps are:

(1) Random initialize N individual positions xi in the feasible domain Ω serve as the initial fireworks.

(2) Assess the fitness value of each fireworks and detonate them. The radius of explosion Radi of each firework and the number of fireworks SSi resulting from the explosion are calculated according to formula (7) (8).

$$Rad_i = RC \times \frac{f(x_i) - y_{\min} + \varepsilon}{\sum_{i=1}^{N}(f(x_i) - y_{\min}) + \varepsilon}; \tag{7}$$

$$SS_i = M \times \frac{y_{\max} - f(x_i) + \varepsilon}{\sum_{i=1}^{N}(y_{\max} - f(x_i)) + \varepsilon}; \tag{8}$$

Among them, ymin=min (f(x)) is the minimum value of the current fireworks population under the fitness function, that is, the optimal fitness value, and ymax=max (f(x)) is the maximum fitness function in the current fireworks population, that is, the worst fitness, RC is a constant, used to adjust the explosion radius, M is a constant, used to determine the basic value of the number of sparks. ε is a minimum value that prevents the division by zero in the formula.

(3) From the sparks generated by the explosion, z dimensions are randomly chosen as the set DS, z=round (D*(0, 1)), where D is the dimension of the solution space and the round function is a rounding function. In the dimension k, the explosive operation is performed using the formula (9) (10) and the boundary treatment is performed, and the processed result is stored in the explosion spark population.

$$h = Rad_i \times rand(-1, 1) \tag{9}$$

$$ex_{ik} = x_{ik} + h \tag{10}$$

In the formula: h is the positional offset, xik is the kth dimension of the i-th fireworks individual, and exik is the explosion spark after the explosion operation of xik.

(4) In order to increase randomness, G Gaussian mutation sparks are generated. Gaussian mutations are randomly selected from the solution space by z dimensions z= round (D*(0, 1)). Where D is the dimension of the

solution space, for each dimension k in the new dimension space, Gaussian variation is calculated using formula (11), and the transboundary part is projected into the solution space and mxik is saved to the spark population.

$$mx_{ik} = x_{ik} * e \tag{11}$$

In the formula, e~N (1,1) is a random number with a mean of 1 and a variance of 1; mxik is a Gaussian variation spark after a xik mutation.

(5) N individuals from fireworks, explosion sparks, and Gaussian mutation spark populations are selected as the next-generation iteratively calculated fireworks population. Fireworks with the best current fitness value are placed, and other fireworks form a candidate set. Other candidate fireworks are placed selectively using the roulette algorithm. The probability of each fireworks being selected is

$$p(xi) = \frac{R(x_i)}{\sum_{x_j \in K} R(x_i)} \tag{12}$$

$$R(x_i) = \sum_{x_j \in K} d(x_i - x_j) = \sum_{x_j \in K} \| x_i - x_j \| \tag{13}$$

In the formula, R (xi) is the distance sum of the current spark and all other sparks in the candidate set k. In the candidate set, if there are multiple sparks around the individual, the current density is higher, the spark is selected as the next generation. The probability of participating in iterative fireworks is greatly reduced.

(6) Determine if the current state satisfies the termination condition of the iteration. If it satisfies, it will end. Otherwise, skip to step (2).

INTRUSION DETECTION MODEL OF TFWA-SVM

Optimization of SVM Parameters Based on TFWA.

The new algorithm proposed in this paper mainly improves the algorithm in two aspects.

Introduction of the Concept of Fireworks Explosion

The biggest advantage of fireworks algorithms over other swarm intelligence algorithms is the fast convergence speed, and fireworks algorithms are easily trapped in local optimal solutions [8]. The initial firework generation position of fireworks algorithms has an important influence on the solution of the algorithm. The appropriate initialization method can be used. Avoid the algorithm falling into the local optimal solution [9]. In order to achieve a better initialization position, this paper proposes a method to calculate the initial fireworks explosion.

The area occupied by N two-dimensional particles in space is calculated using formula (14):

$$(max(x) - min(x)) * (max(y) - min(y)) \tag{14}$$

Where x is the abscissa of the N particles, y is the ordinate of the particle, and so on. In the D-dimensional solution space, the formula for the area occupied by the N particles is shown in (15).

$$S_D = \prod_{i=1}^{D} (max(x_i) - min(x_i)) \tag{15}$$

Where xi is the ith dimension of the particle. The area of the solution space is calculated in the same way and is denoted as RD. Defining the initial fireworks explosion $\phi = \frac{S_D}{R_D}$, If ϕ is not greater than a threshold, the population will be reinitialized so that the initial fireworks of the explosion fireworks algorithm can be distributed throughout the search space as much as possible.

Stretching technology of fitness function

During the operation of the fireworks algorithm, although the global optimization ability is strong, it may still cause the algorithm to fall into a local optimal solution near the local optimal value or cause the algorithm to converge slowly in the late iteration. This paper proposes a stretching technique of the fireworks algorithm and improves the calculation formula of the number of fireworks explosions to avoid the problem that the fireworks algorithm falls into the local optimal solution and the lack of power in the late iteration.

Assume to find the minimum value in the interval [-2.5, 2.5] of a univariate quartic equation y=-x4+5x2+x, the function itself is used as the fitness evaluation function of the fireworks algorithm. The fitness function is shown in the Fig. 1.

Stretching the fitness function using the formula (16)

$$G(u) = f(u) + \gamma(sign(f(u) - f(\bar{u})) + 1) * (f(u) - f(\bar{u})) \tag{16}$$

In the formula (16), f (u) is the original fitness evaluation function. $\bar{u}$ is the position of the optimal solution found by the current fireworks algorithm in the solution space. γ is a positive integer greater than 1, and the larger the value is, the larger the fitness function stretch is, and G(u) is a new fitness evaluation function of the algorithm. The graph of the stretched function is shown in the Fig. 2.

The stretching technology can ensure that when there is a better position than the current global optimum, the advantage of the current position can be expanded, thereby increasing the number of explosive sparks at this position, guarantees the convergence speed and reduces the possibility that the algorithm converges to the local optimum.

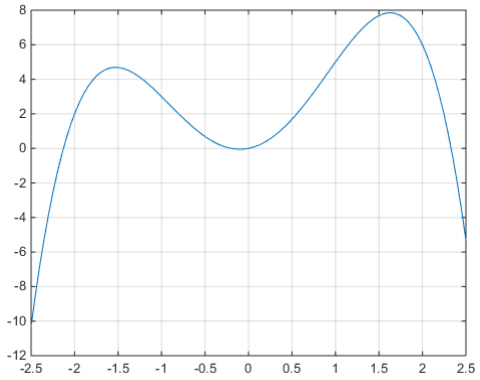

FIG. 1. The function graph of $y = -x^4 + 5x^2 + x$

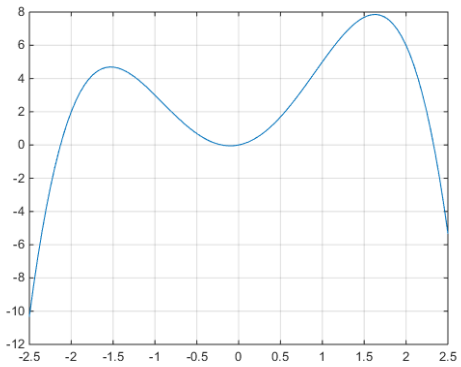

FIG. 2. Stretched function graph of $y = -x^4 + 5x^2 + x$

Intrusion Detection Structure of IFWA-SVM.

As an excellent classifier, support vector machines (SVM) are very useful when solving small sample, nonlinear and high dimensional pattern recognition problems. However, when SVM is used for pattern recognition or regression, the selection of nuclear parameters c and penalty parameters σ is a NP-hard problem. At present, no unified solution model has been formed in the world. That is, the selection of the optimal SVM algorithm parameters can only be performed based on experience, experiment comparison, large-scale search, or use of the interactive test function provided by the software package. At the same time, in training time, the higher training time of this method needs further research and exploration.

The improved fireworks algorithm has the advantages of strong global optimization ability and fast convergence speed. It is combined with the support vector machine. Using IFWA to optimize the kernel parameter c and penalty parameter σ selection of SVM kernel function can improve the search efficiency and introduce it. Intrusion detection can greatly improve the performance of network intrusion detection. The IFWA-SVM-based network intrusion detection model process is shown in Fig. 3. The process can be summarized as follows.

(1) Pre-processing network data packets, extracting their characteristics and laying the foundation for intrusion detection models.

(2) Initialize IFWA related parameters, including the number of initial fireworks, iteration termination conditions, and so on.

(3) Initialize the IFWA population, randomly generate N fireworks in the solution space.

(4) Evaluate the fireworks in the population for fitness.

(5) According to the conditions of IFWA, judge whether the current firework satisfies the iterative planting conditions. If it satisfies, the optimal solution of the fireworks is output as the input of the SVM. If it is not satisfied, sparks and Gaussian mutation sparks are generated based on the fitness value, and the next generation of fireworks is selected and jumped to (4).

(6) Use SVM to classify the optimal network intrusion features to determine whether the packets are threatened or not.

SIMULATION EXPERIMENT

The experimental data in this paper is from the intrusion detection dataset in the KDD99 database of MIT Lincoln Laboratory. The dataset has large data volume and many attributes. There are four types of intrusion data in the data set: Dos, Probe, R2L, U2R. The simulation experiment uses some samples in the data set. Randomly extract 8000 data from training samples and test samples, including normal data packets and 4 types of abnormalities, and make the ratio between normal records and abnormal records 7:3.

The parameters in the algorithm are set as the number of initial fireworks is 100, the number of basic explosion sparks is 80, the basic explosion is 300, and the algorithm is iterated 3000 times. Using IFWA-SVM, common SVM, BPNN (41 input nodes, 40 hidden nodes, 5 output nodes) and IFWA-SVM, the detection accuracy is compared. The

detection results are shown in Table 1. It can be seen that IFWA-SVM classification accuracy is higher than FWA-SVM and ordinary SVM, BPNN classifier.

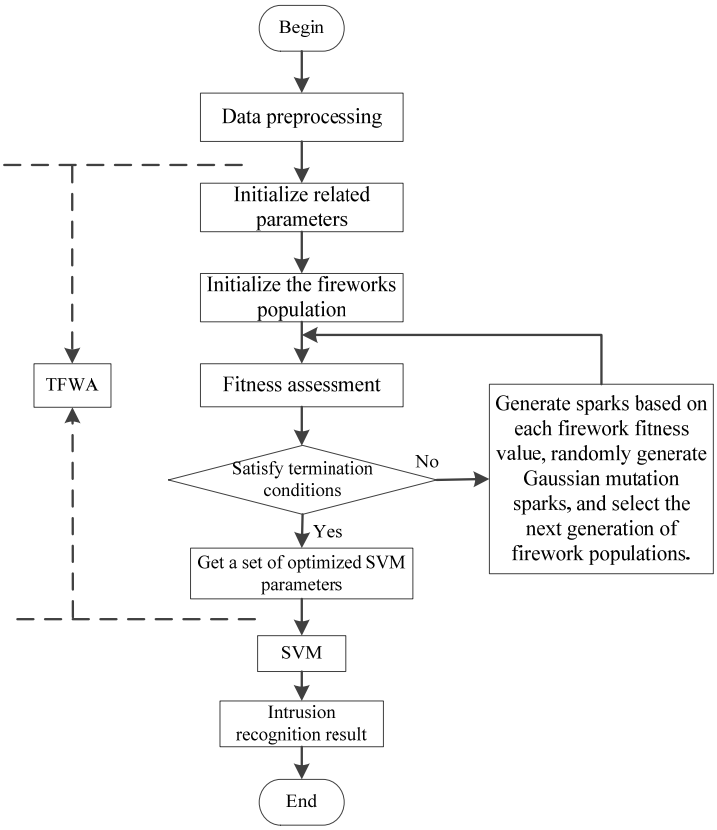

FIG. 3. Network intrusion detection model process of IFWA-SVM

TABLE 1. Comparison of intrusion detection precision

Test set	IFWA-SVM	FWA-SVM	SVM	BPNN
Test set 1	93.18	91.87	86.13	82.28
Test set 2	94.21	92.66	88.02	81.57
Test set 3	95.36	93.95	90.34	85.36
Test set 4	94.29	92.02	85.06	82.10
Test set 5	93.44	90.58	86.71	83.53
Test set 6	92.23	89.19	84.64	79.84

SUMMARY

This paper improves the basic fireworks algorithm in two aspects, and obtains an improved fireworks algorithm, which is used to optimize the kernel parameters and penalty factors of the support vector machine. The improved support vector machine is used in intrusion detection. The new fireworks algorithm optimizes the situation where the initial fireworks population does not cover the solution space and cannot converge to the global optimal solution. At the same time, the fitness function is stretched to prevent the algorithm from converging to the local optimal solution. Finally, the KDD99 dataset is used to train and test the algorithm. The results show that the new detection model can effectively improve the detection rate of intrusion detection and reduce the false alarm rate. At the same time, the algorithm also has the advantage of fast convergence and not easy to converge to the local optimal solution.

REFERENCES

1. Feng Y, Yu S, Liu H. Network intrusion detection based on KNN-IPSO selection feature. Computer Engineering and Applications. Vol. 50 (2014) No. 17, p. 95-99.
2. Yang H, Zhao M, Xie L. Intrusion detection based on adaptive evolutionary neural network algorithm. Computer Engineering and Science. Vol. 36(2014) No. 8, p. 1469-1475.
3. Li X, Han S, Liu X et al. Fireworks algorithm based on reverse learning and dynamic memory feedback. Computer Engineering. Vol. 43(2017) No. 12, p. 203-210.
4. Jyoti Singh Kirar,R .K. Agrawal. Composite kernel support vector machine-based performance enhancement of brain computer interface in conjunction with spatial filter [J]. Biomedical Signal Processing and Control. (2017), p. 33.
5. Tan Y, Zheng S. Research Progress of Fireworks Algorithms [J]. Journal of Intelligent Systems, 2014(5): 515-528.Tan Y, Zhu Y. Fireworks algorithm for optimization[C]// International Conference on Advances in Swarm Intelligence. Springer-Verlag, 2010, p. 355-364.
6. Hui Liu,Xi-wei Mi, Yan-fei Li. Wind speed forecasting method based on deep learning strategy using empirical wavelet transform, long short term memory neural network and Elman neural network[J]. Energy Conversion and Management. (2018), p. 156.
7. Wang C, Chen B, Liu T. Deterministic Learning and Data-Based Modeling and Control [J]. Acta Automatica Sinica. Vol. 35 (2009) N0. 6, p. 693-706.
8. Cao J, Li T, Jia H. Firework Explosion Optimization Algorithm with Genetic Operators [J]. Computer Engineering. Vol. 36 (2010) No. 23, p. 149-151.
9. Fang L. Research and Application of Dynamic Search Fireworks Algorithm [D]. Anhui University, China 2017. p. 17.