Network Intrusion Detection Based on the Improved Artificial Fish Swarm Algorithm

Guo Wang, Dong Dai
Henan Mechanical and Electrical Engineering College, Xinxiang, Henan, 453003, China
E-mail: guo_wang0102@163.com

***Abstract*—In order to predict network anomalies and get rid of the drawbacks of current detection, early prediction of abnormal for detecting early characteristics of the abnormal is introduced in the invasion anomaly detection process. First, the objective functions are constructed according to the feature subset dimensions and the detection accurate rates of the detection model. Then the artificial fish swarm algorithm is used to search the optimal feature subset and the chaotic, feedback mechanisms are introduced to improve the artificial fish swarm algorithm, the excessive intrusion feature rough sets produced in the classification process are simplified to guarantee the simplicity of characteristics and the estimation model for residuals gray level to predicate the early simplified invasion. Finally KDD1999 database is applied to testify the validity of the algorithm. The simulation results illustrate the improved artificial fish swarm algorithm can obtain the optimal intrusion feature subsets and reduce the dimensions of the feature subsets, which not only increase the network intrusion detection rates and reduce the errors, but also speed up the network abnormal intrusion detection.**

***Index Terms*— artificial fish swarm algorithm, feature selection, chaotic mechanisms, intrusion detection, feedback mechanism**

I. Introduction

With the rapid development of the computer network technology, network has penetrated into all fields of society. However, due to the network features of open, without charge and undefended, the network attacks are diversified. The attack amounts and the extent of harm are increasingly serious. The traditional fire wall and anti-virus technologies are unable to meet the modern needs of network security [1]. The network intrusion detection as a kind of security defense technology becomes more and more important, which is the current research focus in network security [2].

There are two types of network intrusion detection technology - misuse detection and anomaly detection [3]. The misuse detection can only detect the known attack types. It can’t identify unknown and variant attacks, so there are no practical meanings for the method[4]. The anomaly detection can detect the "new" attack modes which becomes the main direction of current research. The network intrusion detection is a pattern recognition problem. The original network intrusion features contain redundant features which can not only cause bad effect on the performance of the classifier and increase the probabilities of dimension disaster appearance influencing the efficiency, but also make the detection accuracy more worse when the amount of the features excess some threshold because there are no linear proportional relationship between the amount of intrusion features and detection results [5]. Thus, it is important to select the key features strongly related to the intrusion features and eliminate the redundancy. The current network intrusion feature selection algorithms mainly are sequential selection, rough set method, such as principal component analysis, genetic algorithm, particle swarm optimization algorithm and ant colony algorithm [6, 7, 8, and 9]. The ant colony algorithm is a heuristic algorithm with global optimization ability which becomes the mainstream feature selection algorithm a lot of practice researches show the heuristic algorithms have the defects of searching local optimization and slow convergence speed [10]. It’s difficult to obtain stable feature search results. The detection results are related to the selection of the detection algorithms. The main detection algorithms are neural network and Support Vector Machine (SVM) [11]. Due to the strong generalization ability of the support vector machine, it has become the major network detection algorithm.

First, the objective functions are constructed according to the feature subset dimensions and the detection accurate rates of the detection model. Then the artificial fish swarm algorithm is used to search the optimal feature subset and the chaotic, feedback mechanisms are introduced to improve the artificial fish swarm algorithm, the excessive intrusion feature rough sets produced in the classification process are simplified to guarantee the simplicity of characteristics and the estimation model for residuals gray level to predicate the early simplified invasion. Finally KDD1999 database is applied to testify the validity of the algorithm. The simulation results illustrate the improved artificial fish swarm algorithm can obtain the optimal intrusion feature subsets and reduce the dimensions of the feature subsets, which not only increase the network intrusion detection rates and reduce the errors, but also speed up the network abnormal intrusion detection. However, all the methods above are used for the detection of abnormal situation. Once the accident occurs, we are unable to avoid the loss. Therefore, this paper presents a network abnormal data detection method can be used for early predication.

© 2013 ACADEMY PUBLISHER
doi:10.4304/jcp.8.11.2990-2996

II. The Establishment of Abnormal Intrusion Detection Model

The structure of the abnormal intrusion detection model is shown as Figure 1. First, wrapper feature is used to select model. The global optimal searching capability based on the improved artificial fish swarm algorithm is used to select the features. Then, SVM classification results will be applied to evaluate the classification performance of these feature combinations and keep updating the selected feature sets until find the feature combination with most optimal classification results.

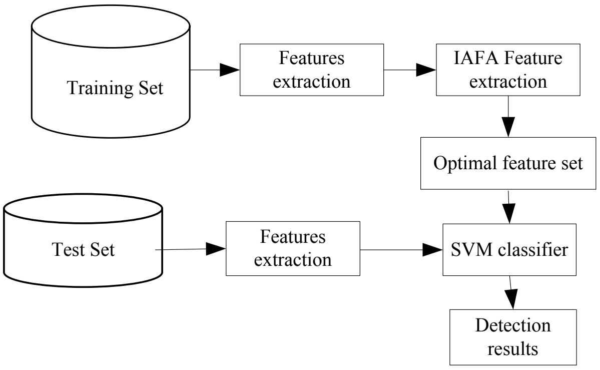

Figure 1 abnormal intrusion detection model

III. The Abnormal Intrusion Detection Model based on Improved Artificial Fish Swarm Algorithm

The artificial fish swarm algorithm is a kind of group intelligent algorithm and the convergence speed is fast. The global optimal searching capability is strong and it doesn't require too many for the initial and objective functions. It simulates the foraging, huddling, following behaviors of the fish swarm in the nature to fulfill the global optimization searching. In order to increase the network intrusion detection accuracy, the Improved Artificial Fish Swarm Algorithm (IAFA) is proposed.

A.Basic Artificial Fish Swarm Algorithm

The artificial fish swarm algorithm is a kind of group intelligent algorithm and the convergence speed is fast. The global optimal searching capability is strong and it doesn't require too many for the initial and objective functions. It simulates the foraging, huddling, following behaviors of the fish swarm in the nature.

(1) Foraging behavior. Assume the current sate of the artificial fish is Xi, the state Xj is randomly selected in the field of view. If Yi<Yj, it will move forward toward this direction $X_i^{t+1} = X_i^t + \frac{X_j - X_i^t}{\left\| X_j - X_i^t \right\|} \cdot Step \cdot Rand()$. If the condition can't be met, it moves randomly by one step $X_i^{t+1} = X_i^t + Step \cdot Rand()$ in which Yi represents the food density in state X i, Step is the moving step length, Rand() represents the random number in the range of (0,1).

(2) Huddling behavior. The current artificial fish Xi will search the amount nf of the partner in the field of view and the state Xc in the central. If Yc/nf>δYi, it illustrates there are more food in the partner center and it's not too crowd, then it will move toward the partner center by one step $X_i^{t+1} = X_i^t + \frac{X_c - X_i^t}{\left\| X_c - X_i^t \right\|} \cdot Step \cdot Rand()$, otherwise, it will execute foraging behavior.

(3) Following behavior. The current artificial fish Xi will search the partner Xj who have the densest food in the field of view. If Yj/nf>δYi, it will move forward toward the artificial fish Xj with densest food. Otherwise, it will execute foraging behavior.

(4) Random behavior. The artificial fish will randomly select a state in the field of view then it move toward the direction, which is default behavior of the foraging behavior.

(5) Bulletin Board. It records the states of the most optimal artificial fish. Each artificial fish will compare the artificial current state with the recorded state after each execution. If the current state is better, it will update the bulletin board. After the algorithm terminated, the elements in the bulletin board are the most optimal and the corresponding states are the optimal solutions.

B.Improved Artificial Fish Swarm Algorithm

Chaotic phenomenon is a specific phenomenon of the nonlinear dynamic system. There are following properties:

(1) Randomness. The rules of the chaotic phenomenon are affected by the initials.

(2) Ergodicity. It will transverse all of the states in some region according its own rules.

(3) Deterministic. The chaotic trajectory is determined by iteration equations. Because the chaotic searching can be easily executed and can avoid the local extremes and it's better than random searching, the chaotic variables have advantages for local searching.

The chaotic variables choose Tent projection.

$$x_{i+1} = \begin{cases} 2x_i & x_i \in (0, 0.5] \\ 2(1 - x_i) & x_i \in (0.5, 1] \end{cases} \tag{1}$$

According to Tent projection, the artificial fish i generate chaotic sequences of the points according to following steps.

Every dimension X_{ik}, k=1,…,n of the artificial fish state X_i is projected to the region [0,1].

cXik=(Xik-ak)/(bk-ak) (2)

In the equation, a_k, b_k represent the minimum and maximum of the k-dimension variable Xik.

(2) The equation (1) is iterated M times to generate the chaotic sequence $cX_{ik}^1, cX_{ik}^2, \dots, cX_{ik}^M$.

© 2013 ACADEMY PUBLISHER

(3) The states in the chaotic sequence are projected to the original space according to the equation (3).

$$X_{ik}^{s} = a_k + cX_{ik}^{s}(b_k - a_k) \quad (3)$$

(4) The chaotic sequence of Xi projected through Tent can be obtained from the previous sequence.

$$X_i^s = (X_{i1}^s, X_{i2}^s, ..., X_{in}^s) \quad (4)$$

(5) Evaluate the new artificial fish states X_i^s.

(6) If the new artificial fish state X_i^s is better than Xi, X_i^s will be used as the chaotic local searching result. Otherwise, let s=s+1 and return to step (2).

The feedback strategy is a feedback behavior defined for the artificial fishes. The artificial fishes will move toward the most optimal states in the bulletin board in some probability. At the late stage of the basic artificial fish swarm algorithm optimization, because the random behavior of the artificial fishes will reduce the optimization precisions and efficiency, the artificial fishes are designed to execute the random behavior in the probability of Pfb, and the feedback behavior in the probability of 1-Pfb to ensure better optimization precisions and efficiency at the late stage of the optimization. Pfb will reduce according to the equation (5).

$$P_{fb} = \theta P_{fb} \quad (5)$$

In the equation, θ is the attenuation factor of the feedback probability.

Because the chaotic search has the feature of ergodicity, it can be used as an effective optimization algorithm to prevent falling into local optimal. When it is introduced to the artificial fish swarm algorithm to search globally, it can avoid the artificial fishes keeping turn around the local optimal value. This can improve the global convergence of the artificial fish and add the feedback strategy into the artificial fish swarm algorithm to increase the optimal searching efficiency.

C. The Intrusion Feature Selection of the Improved Artificial Fish Swarm Algorithm

The detailed steps of the intrusion feature selection combining chaotic searching and feedback strategy are as follows.

(1) Collect network states information to form learning samples and the samples are pre-processed.

(2) Extract the features of the network states.

(3) The designed artificial fish parameters are: positions, maximum move step length Step, vision radius Visaul, population size n, congestion factor δ, feedback probability Pfb, attenuation factor of the feedback probability θ, maximum iteration times max_iterate and bulletin board;

(4) In the feasible region, n artificial fishes are randomly generated. Each artificial fish represents one feature subset, the initial iteration times is passed_iterate =0；

(5) The adapt functions are computed according to equation (6). All of the artificial fish states when f is maximum are recorded in the bulletin board. The definition of the adapt function is:

$$f(s) = \lambda P_{error} + (1-\lambda)\frac{d}{D} \quad (6)$$

In the equation, d is the dimension of the selective feature subset s; D is the dimension of the network intrusion detection candidate feature sets; Perror is the classification errors; λ is the weight coefficients of the classification errors.

The computation equation of the weight coefficients λ is:

$$\lambda = \frac{100}{100 + D \times x} \quad (7)$$

In the equation, x represents reduction percentage (x%) of the network intrusion detection error when the feature increase by one dimension.

(6) Evaluate the results of the foraging behavior, following behavior and huddling behavior of some artificial fish. After the fish executes some behavior, if the state of the artificial fish is better than previous state, the artificial fish will move forward by one step and turn to step (8).

(7) On random number r is generated. If r<Pfb, the artificial fish will execute random behavior, otherwise it will execute feedback behavior to move one step toward the optimal direction.

(8) All of the optimal artificial fishes execute chaotic searching according to equation (1) ~ (3) to obtain the best artificial fish state in current solution region.

(9) The bulletin board is updated to record the best artificial fish sate in step (8).

(10) The feedback probability is updated according to equation (4).

(11) Judge the algorithm termination conditions. If it reaches the maximum iteration times, the algorithm will be terminated and the artificial fish state in the bulletin board will be outputted, namely the abnormal intrusion detection optimal features subset. Otherwise, passed_iterate +1 and turn to step (5).

D.Constraints of Rough Set

According to the features of the network collected, SVM optimal hyper plane applied can be represented as:

$$y = w^T \varphi(x) + b \quad (8)$$

In the equation, w the normal vector of the hyper plane, b is the offset vector of the hyper plane.

If it's nonlinear classification problem, the nonlinear classification problem can be transferred to quadratic optimization problem.

$$\min J(w,\xi) = \frac{1}{2}\|w\|^2 + c\sum_{i=1}^{n}\xi_i \quad (9)$$

The relative constrains are:

$$y_i(w \bullet \Phi(x_i) + b) \geq 1 - \xi_i$$
$$\xi \geq 0, i = 1, 2 \cdots, n \quad (10)$$

© 2013 ACADEMY PUBLISHER

In the equation, $\xi = (\xi_i, \ldots, \xi_l)^T$, c is penalty coefficient.

For the classification problem of large amounts of samples, the learning speed of SVM is slow. Lagrange multiplier can transfer SVM classification problem to dual problem. The solution of the dual problem can solve the hyper plane optimization problem to speed up the classification. The SVM decision function:

$$f(x) = sign(\sum_{i=1}^{n} \alpha_i y_i (\varphi(x) \bullet \varphi(x_i)) + b) \qquad (11)$$

In equation, $sign$ is signed function, α_i is Lagrange multiplier.

Because RBF only needs to determine one parameter, namely the width parameter σ of the kernel functions to benefit the parameter optimization. Thus, RBF kernel functions are used to establish the support vector machine. The definition of the RBF kernel functions is:

$$k(x_i, x_j) = \exp\left(\frac{-\left\| x_i - x_j \right\|}{2\sigma^2} \right) \qquad (12)$$

The purpose of the attribute reduction is to remove the invalid and redundant features and reduce the sample dimension under the premise of no loss of effective information network intrusion feature. The reduction results to training samples are directly used to the feature selection of testing samples. The theory formula of rough sets is

Definition 1: An information system S can be expressed as $S = < U, A, V, f >$ in which U represents the collection of non-empty finite objects also called domain $U = \{x_1, x_2, \ldots x_n\}$; A is non-empty finite set called the set of attributes $A = \{a_1, a_2, \ldots a_m\}$; V is a collection of the attributes' range $V = \{V_1, V_2, \ldots V_m\}$ in which V_i indicates the range of attribute a_i , $a_i \in A$; f is a function of the attributes and objects, $f(x_i, a_j) \in V_j$. If A is composed by conditional attribute set C and decision attribute set D and satisfies the condition $C \cap D = \Phi, C \cup D = A$, $S = (U, C \cup D)$ is called a decision system.

Definition 2: For the decision system $S = (U, CUD), R \subseteq C$,

ind（R）$= \{(x_i, x_j) \in U \times U \mid f(x_i, r) = f(x_j, r), \forall r \in R\}$

is called the R indiscernible relation on domain U. Indiscernible relation family divided from domain U according to attribute set R is described as U / R .

Definition 3: In decision system $S = (U, CUD)$, attributes' set P reduced from conditional attributes' set C should meet the following condition:

$(i) ind(P, \{d\} S) = ind(C, \{d\});$

$(ii) Non \quad P' \subset C \quad makes \quad ind(P', \{d\}) = ind(C, \{d\})$

All the reduction sets of C are recorded as $RED_d(C)$. $CORE_d(C) = \bigcap RED_D(C)$ the intersection of all the reduction sets of C is called the core.

Definition 4: In the decision system S, $a_i(x_j)$ is the value of sample x_j in the conditional attribute a_i, and $d(x_j)$ represents the value of sample x_j under the decision attribute d . Identified matrix $M = \{m_{ij}\}$ is defined as:

$$m_{ij} = \begin{cases} a_t \mid a_t \in C \wedge a_t(x_i) \neq a_t(x_j), d(x_i) \neq d(x_j) \\ 0, d(x_i) = d(x_j) \end{cases} \qquad (13)$$

In intrusion detection systems, U is the training sample set, C represent 41 features of the samples that also called set of attributes, D represents sample classification whose range is {-1, 1}, where 1 means normal network connection, -1 abnormal network connection. V represents a range set of 41 attributes. Attribute reduction of invasive sample obtains the minimum reduction set P in the set C, thereby reducing the sample dimension of the input support vector machine.

Main steps for the attribute constraints algorithm based on discernibility matrix adopted are:① Calculate the core of C with respect to D; ② Calculate the identified matrix MR determined by R; ③ Remove the items contain core in the matrix and add the element appears most in the remaining items of the matrix; ④ Repeat the steps above until the matrix is empty.

Definition 5: Assume $S = (U, A), R \subseteq A, X \subseteq U$, $\overline{R}(X)$, $\overline{R}(X)$ and R boundary of sub-set X are defined as:

$$\begin{aligned} &\overline{R}(X) = \bigcup \{Y \in U / R \mid Y \cap X \neq \Phi\} \\ &\underline{R}(X) = \bigcup \{Y \in U / R \mid Y \subseteq X\} \\ &BN_R(X) = \overline{R}(X) - \underline{R}(X) \end{aligned} \qquad (14)$$

The elements in $\underline{R}(X)$ is definitely belongs to set X. Elements in $\overline{R}(X)$ may belongs to X which reflects the fuzziness of the elements. $BN_R(X)$ is called the boundary set of X relative to R composed by elements neither totally included in X nor U which reflects the uncertainty of the set. The boundary set of the attribute is the intersection of each attribute boundary set.

$$BN_A(X) = BN_{a1}(X) \cap BN_{a2}(X) \ldots \cap BN_{a|A|}(X) \qquad (15)$$

$BN_R(X)$ is calculated through formula (15). Because the support vectors determining the SVM optimal classification surface are mainly distributed in the border region, selecting the border set as SVM training subset can significantly reduce the number of training samples.

The boundary portion obtained from the rough set is usually the place noise samples exist. A feature of SVM

© 2013 ACADEMY PUBLISHER

is unusually sensitive to noise which may cause over-learning. The paper adopts method in literature [12] to assess the border sample.

Definition 6: Neighborhood matching operator $N\_M(x',k)$ is defined as

$$N\_M(x',k)=\frac{|\{x\,|\,label(x)=label(x'),x\in kNN(x')\}|}{x}$$

$kNN(x')$ is the k-order nearest neighbor set of x' which reflects the distribution consistency of x' with the neighboring points. The smaller the $N\_M(x',k)$ is, the more possible it is noise sample. Set proper domain value, the noise points of match operators less than domain value can be removed.

IV. Predictive Judgment of Invasion

Categorize intrusion data can be described with the following data sequence:

$$Y^{(0)}(u)=\{y^{(0)}(1),y^{(0)}(2),\cdots,y^{(0)}(p)\} \qquad (16)$$

Set $Y^{(1)}(u)=\sum_{j}^{u}y^{(0)}(j),u=1,2,\cdots,p$, Then the new ordered sequence of data can be got:

$$Y^{(1)}(u)=\{y^{(1)}(1),y^{(1)}(2),\cdots,y^{(1)}(p)\}$$

History intrusion data equation is as follow:

$$\frac{dy^{(1)}(u)}{du}-by^{(1)}(u)=v$$

Wherein b and v are required parameters for Residual Error Gray estimated model.

$$A^{(1)}(u)=\frac{1}{3}y^{(1)}(u)+\frac{1}{3}y^{(1)}(u+1)$$,

$u=0,1,2,\cdots,p$. Then get the designated invasion data matrix:

$$C=\begin{bmatrix}1 & -A^{(1)}(0) & 1 & 0\\ 0 & 1 & 1 & -A^{(1)}(1)\\ \vdots & \vdots & \vdots & \vdots\\ -A^{(1)}(p) & 1 & 0 & 0\end{bmatrix} \qquad (17)$$

The components of historical incursions data are described as follows

$$z=[y^{(0)}(0),y^{(0)}(1),\cdots,y^{(0)}(p)]^{T}$$

Using data reverse transform method to obtain the following formula to calculate the predicted parameters：

$$(b,v)^{T}=(C^{T}C)^{-1}C^{T}z$$

Using the following formula to calculate time spent in the intrusion prediction:

$$y^{(1)}(u-1)=[y^{0}(1)-\frac{v}{b}]f^{-1}-\frac{v}{b} \qquad (18)$$

Iterative processing the predicted time, the estimated initial model for Residual Error Gray can be obtained

$$\hat{y}^{0}(u+1)=(1+f^{b})[y^{0}(1)+\frac{v}{b}]e^{-1}$$,

$u=0,1,\cdots,p+1$.

Calculate the difference between the initial historical intrusion data sequence and the forecast data sequence of actual invasion:

$$f^{(0)}(u)=y^{(0)}(u)-\hat{y}^{(0)}(u)$$

Residual data sequence is got：

$$f^{(0)}(u)=\{f^{(0)}(1),f^{(0)}(2),\cdots,f^{(0)}(p)\}$$

Assuming that the value of the residual data sequence is less than zero, then the iterative summing transform to the difference is carried out by the following formula

$$f^{(0)}(u)=f^{(0)}(u)+3\left|\max f^{(0)}(u)\right|$$

Get the data sequence difference greater than zero:

$$f^{(1)}(u)=\{f^{(1)}(1),f^{(1)}(2),\cdots,f^{(1)}(p)\}$$

The initial residential gray estimated mathematical model calculated is

$$f^{(1)}(u+1)=\left[f^{(0)}(1)-\frac{v_f}{b_f}\right]f^{-b_f u}-\frac{v_f}{b_f} \qquad (19)$$

Carry out iterative differential transform to the data mentioned above, restored historical intrusion data in an orderly sequence can be obtained.

$$\hat{f}^{(0)}(u-1)=[f^{(0)}(1)+\frac{v_f}{b_f}](f^{-b_f(u-1)}-f^{-b_f u})$$

$$-2\left|\max f^{(0)}u\right|$$

Wherein $u=0,1,\cdots,p+1$. Residual Error Gray estimated prediction model can be obtained as follows through the above method

$$\hat{y}_1^{(0)}(u+1)=\hat{y}^{(0)}(u+1)+\lambda(u-1)\hat{y}^{(0)}(u+1)$$

Wherein $u=1,2,\cdots,p$.

A.Prediction Error Compensation

According to estimation model of Residual Error Gray, the initial sequence of historical intrusion data can be obtained.

$$Y^{(0)}(u)=\{y^{0}(1),y^{0}(2),\cdots,y^{0}(p)\}$$

Using the following formula, error coefficient of historical intrusion data can be calculated.

$$\gamma=(y^{(0)}(1)-\frac{v}{b})f^{b}$$

The designated historical invasion data is divided into p different states. Status transfer uses the following formula

$$\eta_j=[\eta_{1j},\eta_{2j}],\eta_{ij}\in\eta_j,j=0,1,\cdots,p \qquad (19)$$

Wherein

$$\eta_{1j}=Z(u)-B_j$$

© 2013 ACADEMY PUBLISHER

$$\eta_{2j} = Z(u) - B_j$$

The values of η_{1j} and η_{2j} change with time.

The error probability of historical invasion data can be calculated by the following formula.

$$P_{jk}(l) = \frac{N_{jk}(l)}{N_j}$$

Therefore, the error matrix of predication for invasion abnormal is expressed as follow:

$$P(l) = \begin{bmatrix} P_{11}(l) & P_{12}(l) & \cdots & P_{1p}(l) \\ P_{21}(l) & P_{22}(l) & \cdots & P_{2n}(l) \\ \vdots & \vdots & \vdots & \vdots \\ P_{n1}(l) & P_{n2}(l) & \cdots & P_{nn}(l) \end{bmatrix} \quad (20)$$

$N_{jk}(l)$ is the changing parameter of the network intrusion incident data, η_j is the data changing parameter after l times data iterative processing to the original data and η_k is the number of the initial data collection.

Based on the above matrix, network intrusion prediction error can be calculated, so as to complete the compensation.

Through the methods above, Residual Error Gray estimated mathematical model can be established to compensate for the prediction error, thus prediction for early abnormalities of the network can be completed.

V. Simulation Experiments

A.Data Source

The experiments are carried on the environment of P4 dual-core 3.0 GMHZ CPU, 2G RAM, Unix operation system and VC language development kits. The data is from KDD CUP 99 abnormal detection database. The database contains 4 intrusion types-Probe, DoS, U2R and R2L. In the database, each connection totally has 41 features in which 9 are discrete attributes and 32 are continuous attributes. In order to compare the detection results of the improved artificial fish swarm algorithm, the generic algorithm and particle swarm optimization algorithm are applied as the comparable experiments. The model performance evaluation indexes are average detection accuracy and average detection speed.

B.The Effects of the Features on the Network Intrusion Detection Performance

In order to test the effects of different feature subsets on the network intrusion detection performance, the network intrusion detection observation model established by Support Vector Machine is used and the test results are shown in figure 2.

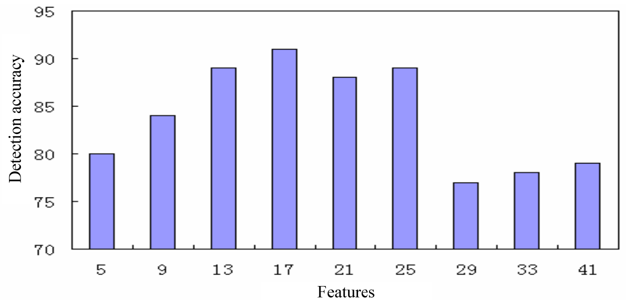

Figure 2 Average intrusion detection accurate rates of different feature subsets

Form figure 2, more network features are not necessary to obtain more optimal detection results because there are some redundancy and useless features in the network features. All the features are inputted to the classifier to train. The computation complexity is quite high which will reduce the network intrusion detection efficiency and accuracy. Thus before detecting the intrusion, the features should be selected.

C. Detection Accurate Rates Comparison

The generic algorithm, particle swarm optimization algorithm and the proposed improved artificial fish swarm algorithm will run in each database for 10 times. Their average network intrusion detection accuracies (%) are computed separately. The results are shown in table 1. From table 1, the improved artificial fish swarm algorithm can increase the abnormal intrusion detection accuracies compared with generic algorithm and particle swarm optimization algorithm. The results illustrate the improved artificial fish swarm algorithm can obtain more optimal feature subsets, accurately describe the network state change information compared with generic algorithm and particle swarm optimization algorithm which can effectively eliminate redundancy and useless features.

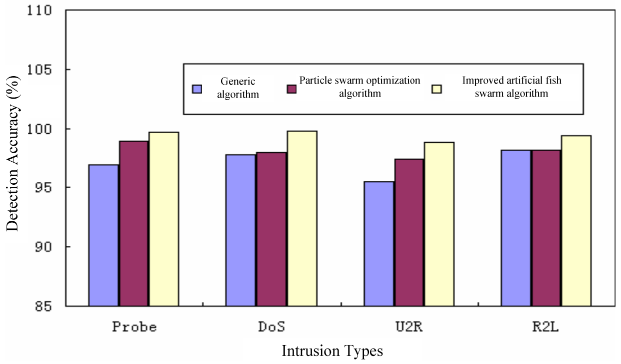

Figure 3 Detection accurate rates comparison by several feature selection algorithm

© 2013 ACADEMY PUBLISHER

D. Intrusion Detection Time Comparison

The average detection time (s) of the generic algorithm, particle swarm optimization algorithm and improved artificial fish swarm algorithm are shown in table 1. From table 1, the detection speed of ACO-SVM is fastest which illustrates ACO can reduce the amount of features and computation time while selecting network features and speed up the algorithm convergence which make the algorithm can better meet the real-time requirements of the network intrusion detection.

TABLE 1

AVERAGE DETECTION TIME COMPARISON BY SEVERAL FEATURE SELECTION ALGORITHM

Intrusion types	Generic algorithm	Particle Swarm algorithm	Improved artificial fish swarm algorithm
Probe	11.24	11.05	9.55
DoS	13.44	12.73	11.21
U2R	4.09	3.20	1.30
R2L	3.45	2.71	2.66

VI. CONCLUSIONS

Aiming at the abnormal intrusion detection feature selection problem and the shortcomings of current selection algorithm, this paper proposes a network abnormal detection method based on predication. The simulation results illustrate the improved artificial fish swarm algorithm can obtain more optimal feature subsets and increase the abnormal detection accurate rates and speeds which can be widely applied in network security.

REFERENCES

[1] D. Gollmann, “Computer security”, *Wiley Interdisciplinary Reviews: Computational Statistics*, vol. 2, NO. 5, 2010, pp. 544–554.

[2] L. T. Heberlein, B. Mukherjee, K. N. Levitt, “Network intrusion detection”. *IEEE Network journal*, 1994, pp. 26-41.

[3] R. A. Kemmerer, G.Vigna, “Intrusion detection: a brief history and overview”, *Computer*, vol. 35, NO. 4, 2002, pp. 27-30.

[4] O. Depren, M. Topallar, E. Anarim, M. K. Ciliz, “An intelligent intrusion detection system (IDS) for anomaly and misuse detection in computer networks”, *Expert systems with Applications*, vol. 29, NO. 4, 2005, pp. 713-722.

[5] A. Patcha, J. M. Park, “An overview of anomaly detection techniques: Existing solutions and latest technological trends”, *Computer Networks*, vol. 51, NO. 12, 2007, pp. 3448-3470.

[6] G. Liu, Z. Yi, S. Yang, “A hierarchical intrusion detection model based on the PCA neural networks”, *Neurocomputing*, vol. 70, NO. 7, 2007, pp. 1561-1568.

[7] D. Dasgupta, F. González, “An immunity-based technique to characterize intrusions in computer networks”, *Evolutionary Computation, IEEE Transactions on*, vol. 6, NO. 3, 2002, pp. 281-291.

[8] R. Poli, “Analysis of the publications on the applications of particle swarm optimization”, *Journal of Artificial Evolution and Applications*, 2008, pp. 3.

[9] S. Banerjee, C. Grosan, A. Abraham, P. K. Mahanti, “Intrusion detection on sensor networks using emotional ants”, *International Journal of Applied Science and Computations*, vol. 12, NO. 3 2005, pp. 152-173.

[10] M. Dorigo, G. Di Caro, “Ant colony optimization: a new meta-heuristic”, *Proceedings of the 1999 Congress on Evolutionary Computation, IEEE*, vol. 2, 1999.

[11] Mukkamala S., Janoski G., Sung A., “Intrusion detection using neural networks and support vector machines”, *Proceedings of the 2002 International Joint Conference on Neural Networks, IEEE*, vol. 2, 2002, pp. 1702-1707.

[12] B. Demir, S. Erturk, “Clustering-based extraction of border training patterns for accurate SVM classification of hyperspectral images”, *Geoscience and Remote Sensing Letters, IEEE*, vol. 6, NO. 4, 2009, pp. 840-844.