International Journal of Security and Its Applications
Vol. 10, No. 11 (2016), pp.1-12
http://dx.doi.org/10.14257/ijsia.2016.10.11.01

Network Intrusion Detection Model based on Combination of Fisher Score and ELM Approach

Hong Mei[1], Wang Ju[1], Qi Yao Wu[2] and Zhai Ning[3]

[1]*Changchun University, Chang Chun, China*
[2]*CRRC ChangChun Railway Vehicles Co., LTD, Chang Chun, China*
[3]*State Crid Jilin Electric Power Company Limited, Chang Chun, China*

Abstract

The diversity and concealment of network attack lead to the difficulty of network intrusion detection, in order to further improve the detection accuracy and efficiency of network intrusion detection, this article proposes a novel model FS-ELM, which is based on the combination of Fisher Score (FS) for feature selection and ELM classifiers for network intrusion detection. In the proposed model, FS is used to conduct feature selection to select the most distinguished feature subsets, and then to get diverse training subsets, in terms of these subsets, ELM classifiers are trained. Finally the results are achieved. Experiment on KDD CUP 99 data set, by means of the experimental analysis and comparison with SVM, LS-SVM and KNN, the proposed model not only improves the detection accuracy, but also enhances detection efficiency, it proves that it is an effective model for network intrusion detection.

Keywords*: Feature selection Extreme learning machine Intrusion detection*

1. Introduction

With the constant development and openness of Internet scale, the network attack presents an obvious increasing trend in number and harmful degree. The defense methods based on traditional network security haven't met today's network security needs. The network intrusion detection system, as an active network security defense technology, can collect and analyze system's security audit data from network, extracting various behavioral models and behavioral features, then detect the existing intrusion behaviors in the system, and give the intrusion alert in time. Nowadays the intrusion detection has become the hot issue in modern research in the field of network security [1].

Network intrusion detection systems often use misuse detection and anomaly detection to analyze system's intrusion behaviors. The misuse detection can detect intrusion behaviors of the known patterns, but it is difficult to accurately identify new attack patterns. The anomaly detection can detect unknown attacks through system's abnormal behavior detection, so at present the anomaly detection method is the main research direction of intrusion detection system [2]. In fact, researchers hope that more accurate results are obtained by using artificial intelligence technology, in recent years, researchers have proposed a variety of network intrusion detection researches based on artificial intelligence technology, such as neural networks , RS , artificial immune , support vector machine , and so on.

The intrusion detection model of network security mainly consists of two areas: feature selection and design of classifier [3]. There are a lot of redundant features and irrelevant features in the original network intrusion detection data. These redundant and irrelevant features not only increase the complexity of classifier training and make the training speed slow, but also cause over-fit phenomena easily, which have a negative effect on the test results. However, the selection of classifier is mainly based on machine learning algorithm, in which the network intrusion detection methods based on support vector

ISSN: 1738-9976 IJSIA

machine (SVM) and least square support vector machine (LS-SVM) have got lots of concerns and researches. Because SVM method has good theoretical basis and strong generalization capacity, it can deal with nonlinear data and other practical problems.SVM method shows the superior performance in intrusion detection and makes good effects, but it still exists the following two problems: (1) Most existing methods are based on SVM method and LS-SVM method, but the typical problem is the selection of model's kernel function and parameter, which has an important effect on the results of intrusion detection. There is no unified standard and theory guide for the selection of parameter. Currently most researchers use genetic algorithm to find the parameter iterative optimization. This not only consumes a significant amount of training time, but it also makes local extrema difficult to find the optimal solution. (2) SVM and LS-SVM model can deal with two-category issues, but they cannot be directly used for multi-classification problems of intrusion detection, "one to one" or "one-to-many" complex modes should be used to build classifiers.

Based on the above analysis, in order to overcome these shortcomings and get the intrusion detection system with higher detection rate and efficiency, the intrusion detection system FS-ELM which combines Fisher Score (FS) and ELM is used for network intrusion detection. FS which the simple and efficient way is used in feature selection mechanism, Extreme Learning Machine (ELM) is used in the selection of classifier[4]. ELM only has one hidden nodes of parameter which needs to be set, compared to SVM and LS-SVM, the selection of kernel parameters and the penalty factors has great effect on the results, a lot of training time is needed to select the iteration.

The proposed FS-ELM model first uses FS method for feature selection, scoring the importance of features, then remove the redundant or irrelevant features in the data, get the feature subset with separating capacity and use different training data set to get ELM model, finally the last classification results are obtained. The simulation experiment shows that the proposed method gets feature sets with separating capacity and higher detection accuracy rate, which is better than SVM model, LS-SVM and KNN model based on grid computing. Further more, the efficiency of training and test is also greatly improved and the effectiveness and feasibility of this method is validated.

2. Introduction to the Theory

2.1. Fisher Score (FS)

FS method[14] is a kind of supervised feature selection method, which is usually used to determine the most relevant feature and then to make classification. It selects good features based on the fractional value of distinguishing ability measure which is defined by the Fisher criterion. Assume that a given data sample set is (xi, ti), ti∈{1, 2, …, m}, m is the number of classification. Suppose ni represents sample's number in type i, ui represents sample's mean in type i, μ represents the mean of global sample, σi2 represents sample's variance in type i, the calculating formula which is mainly used for every feature value is as follows:

$$Score = \frac{\sum_{i=1}^{m} n_i (\mu_i - \mu)^2}{\sum_{i=1}^{m} n_i \sigma_i^2} \tag{1}$$

FS method can directly calculate the score of every feature. If a certain numerical value has higher between-class scatter and lower within-class scatter, then this feature score is higher, conversely, it is lower.

2.2. Extreme Learning Machine (ELM)

ELM algorithm is proposed by Huang and others [6], which is a new single implicit layer feedforward neural network. Studies have proved that the ELM has the same global approach ability as neural network. The network structure consists of three layers: the input layer, the single implicit layer and output layer. The input weights and hidden biases of network generate randomly and it can't be adjusted once it generates. The output weights is calculated directly by least-squares estimation method. ELM has good generalization abilities. Because it doesn't need to iteratively adjust the weight value of network, it avoids the local extremum which is generated by gradient descent algorithm, long time study, the impact on learning rate and some other questions, improving the speed of training and testing, so the research and application of ELM has received wide attention.

The network structure of ELM is shown in Figure 1.

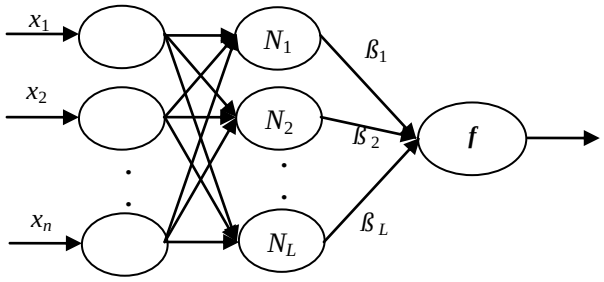

Figure 1. The Network Structure of ELM

For N different training sample sets(xi, ti), i =1, …, N,其中 xi = [xi1, xi2, …, xid]T ∈ Rn, ti = [ti1, ti2, …, tim]T ∈ Rm, the single implicit layer feedforward neural network of L hidden nodes and activation function g(x) and its expression is as follows:

$$\sum_{j=1}^{L} \beta_j g(w_j \cdot x_i + b_j) = o_i \tag{2}$$

wj = [wj1, wj2, …, wjn]T (j = 1, 2, …, L) shows the input weight vector value which connects No. j hidden node and input layer node, bj shows the bias of No. j hidden node, wj · xi shows the inner product of wj and xi, βj = [βj1, βj2, …, βjL]T shows the output weight vector value which connects No. j hidden node and output layer node. Oi is the real result of relevant xi in the network.

Huang and others [7] have proved that if the activation function g(x) is infinitely differentiable, then input weight and hidden layer Bias values in the single hidden layer network can be randomly generated, and once the fixed adjustment, that makes:

$$\sum_{j=1}^{L} \beta_j g(w_j \cdot x_i + b_j) = t_j \ , \ j = 1, 2, \ldots, L \tag{3}$$

In fact, the above formula is to find least-square solutions β' of a linear system Hβ= T. Among them, T = [t1, t2, …, tm]T is the target matrix (expectation output value).

$$H = H(w_1, w_2, \ldots, w_L, b_1, b_2, \ldots, b_L, x_1, x_2, \ldots, x_n)$$

$$= \begin{bmatrix} g(w_1 \cdot x_1 + b_1) & \cdots & g(w_L \cdot x_1 + b_L) \\ \vdots & \cdots & \vdots \\ g(w_1 \cdot x_N + b_1) & \cdots & g(w_L \cdot x_N + b_L) \end{bmatrix}_{N \times L} \tag{4}$$

$$\beta = \begin{bmatrix} \beta_1^T \\ \vdots \\ \beta_L^T \end{bmatrix}_{L\times m}, \quad T = \begin{bmatrix} t_1^T \\ \vdots \\ t_N^T \end{bmatrix}_{N\times m}$$

Among them, H = {hij} (i =1, …, N; j =1, …, L) was known as the hidden layer output matrix, column j of the hidden layer node j is corresponding to the input of x1, x2, …, xn, the line i of H is corresponding to the output vector of input xi.

If the number of hidden layer nodes is equal to the number of training samples N, then H is an invertible square matrix, the single hidden layer feedforward neural network approached the training sample at zero error. But in most cases, the number of hidden layer nodes is far less than the number of training samples, that is $L \square N$, then H is a N×L matrix, Usually we use the least squares method to determine a linear system weights of the output:

$$ß' = H^{+}T \tag{5}$$

Among them H+ is Moore-Penrose generalized inverse matrix of hidden layer out matrix H.

The learning process of ELM algorithm is mainly divided into three steps:

1. The random input produces weight value and the hidden layer(wi, bi), i = 1, 2, …, L;

2. Calculate the hidden layer and output matrix H according to the formula (4).

3. Calculate the output weights ß : ß'= H+T, complete the ELM.

Generally it is not more than two lines.

3. FS-ELM Model

In this section, we will introduce the FS-ELM intrusion detection model in detail. FS-ELM model mainly includes two parts: the feature selection and classifier learning. The whole operation process includes four stages: in the first stage, grade the importance of intrusion detection data characteristics by using FS feature selection methods, thus obtain feature subset which has the most differentiable ability. In the second stage, train on ELM classifier in different training data set by using the optimal feature set which was provided by the first stage as the input. In the third stage, test the performance of the model on the test data set, so as to realize function of intrusion detection. The overall framework of the model is shown in Figure 2.

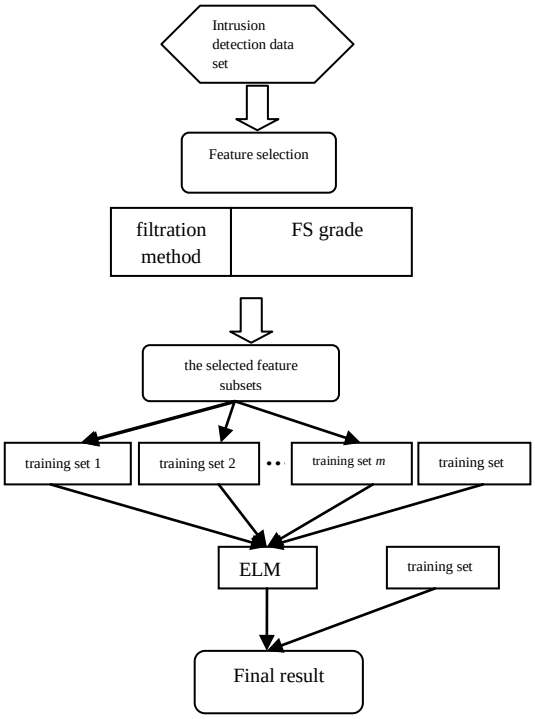

Figure 2. The Overall Flow Chart of the FS-ELM Model

3.1. The Generation of Training Subsets

For training method, we need to consider the diversity of samples besides the classification accuracy. Samples were abstracted from the original data samples by random sampling method, then were processed and converted to produce new samples. Assuming that the original data set of samples for X, class label for Y, new training sample for m, the algorithm of specific processes of the training sample as shown below:

```
Input: initial intrusion detection training data X
Output: generate new training subsets (New_sub1, New_sub2, …, New_subm)
Begin
  Load X;  /*Loading training set X*/
   For i = 1 to m
    [AsubX, AsubY] = randomsub (X);
    trainX_subnew = bootstrapal(AsubX, AsubY);
                                        /* Making a sampling */
    Coeff = pcasky(trainX_subnew); /*Processing to produce new samples */
    R_coeff = sort (Coeff); /* To reorde */
    New_sub (i) = trainX_subnew* R_coeff;
   End For
End
Output: the final training subsets (New_sub1, New_sub2, …, New_subm)
```

3.2. The Creation of Classifier

The definition of the effective classifier is pointed out in the literature: a sufficient and necessary condition of getting higher classification accuracy must be accurate and exist differences. That is to say, the classifier model with large differences has higher generalization ability. So it is an important problem to build classifier with big differences. It can be illustrated from the following three aspects for the ELM model:

1. The initial conditions: Due to the nature of the ELM (that is random input and hidden layer weights bias), the initialization conditions can get a lot of classifiers;

2. Network structure: Get different network structures by setting the number of hidden layer neurons for ELM in the experiment.

3. Training data set: Get different training data set by sampling and transformating, and gain ELM training model in different training data set.

3.3. The Result of Classification

According to the above work, we can get a combination of multiple ELM classifier. Next step is to choose a proper strategy to combine different classifiers into one classifier. The combination of the classifier determines the final output. For a given sample data x, (x = x1, x2,…, xn), If there is the M number of classifier model Tm(x), m =1, 2, …, M, then the computation formula of samples' final classification results is as follows:

$$T(x) = \arg\max \sum_{m=1}^{n} \mu_{x,y} \tag{7}$$

In $\mu_{i,j} = \begin{cases} 0, & i \neq j \\ 1, & i = j \end{cases}$, $y \in \{-1, 1\}$ is the output class label of classifier.

4. Experiment Design

4.1. The Description and Pretreatment of Data

The data set of network intrusion detection standard comes from KDD CUP99public data set. This data set includes 41 feature properties, 7 character type fields and 34 numeric type fields. It includes 4 kinds of intrusion types: Dos, Probe, U2R, U2L as well as the normal condition (Because the space is limited, the detail information of 41 features is not given, see the KDD CUP 99 data set.). The data of network intrusion detection data set is too large, this experiment chose 5000 normal samples and intrusion samples respectively at random. Among the rest, 8000 samples are training sets and 2000 samples are test sets.

At first, the data should be pretreated. In order to reduce the size differences among feature values, the normalization method is used to pretreat the data. The normalization method is a kind of method to classify data mining algorithm or forecast former data. This method will converse all the feature values to [0, 1]. Its purpose is to remove the order differences among data, avoid the interference caused by larger order data to smaller order data, and guarantee the effectiveness of results. The calculation formula of normalization method for entered data is as follows:

$$x_i = (x_i - x_{\min}) / (x_{\max} - x_{\min}) \tag{8}$$

$x_{\min}$ is the minimum of data concentration, $x_{\max}$ is the maximum of data concentration. Moreover, in the experiment process, in order to avoid the occurrence of over-fitting and low-fitting phenomenon, and make results more persuasive, 50 percent authorization method is used. 4 data subsets are taken as training set, the remaining 1 data subset is taken as checking set. Because the cross authorization which is only conducted once can't guarantee the fairness of results, the data is sampled at random, the training set and checking set produced every time are not completely same. In this experiment, the model ran 5 times, then seek its average value as final results.

4.2. Experiment Setting

The FS-ELM model which was put forward is designed and achieved under the circumstances. ELM uses ELM toolkit, the adopted machine configuration is intel processor, the dominant frequency 1.83G, the memory is 1G.

The following FS-ELM model's experiment results are compared with the methods of SVM, LS-SVM and KNN. The detailed parameters of the model are set as follows: in order to make a fair comparison, the parameter settings of SVM and SVM model are the same, using grid computing methods. The hunting zone of C and γ in the model is form $C \in \{2^{-10},...,2^{15}\}$ to $\gamma \in \{2^{-15},...,2^{5}\}$. In the KNN method, K value is obtained from the highest accuracy rate through experiments. K value is 1. Because SVM and LS-SVM can only handle two-category problems, while the intrusion detection is a multi-classification problem, so SVM and LS-SVM use one-to-one mode to construct the classifier and achieve the comparative experiment of network intrusion detection.

5. The Results and Analysis of Experiment

In order to verify the validity of this approach, according to the intrusion detection data sets, the experiment firstly puts forward comparative classification results among FS-ELM model and models based on SVM(Lin-SVM) model, SVM(RBF-SVM), LS-SVM and KNN respectively, as shown in Table 1. ACC represents the accurate rate of classification, EACC represents the accurate rate of classification which is obtained based on classifier method, Std represents standard deviation. From the experimental results, in these five models, FS-ELM method gets the highest average classification accuracy rate 95.51%, while models based on SVM(Lin-SVM) model, SVM(RBF-SVM), LS-SVM and KNN get 91.59%、93.95%, 92.80% and 92.95% respectively. Through further analysis, under the function of feature selection, the classification accuracy rates of these five models improved 2.56%, 0.56%, 1.32%, 1.06% and 2.57% respectively compared with the classification accuracy rates without the function of feature selection. At the same time, under the condition of feature selection and non-feature selection, the classification accuracy rates of classifier method have improved respectively, representing the good performance of the learning method, so the classification accuracy rates are improved.

In order to analyze FS-ELM model roundly, the experiment further analyzes the influence of feature selection and the number of hidden layer node on models. Table 2 shows the change situation of different hidden layer node numbers with different feature selection, the number of feature subset takes 5, 10, 15, 20, 25, 30, 35 respectively and all features, the number of hidden layer node takes 5, 10, 20, 30 and 40. From the Table, we can see that when the number of hidden layer node is 5, the classification results of ELM increases as fluctuations state with the increase of characteristic set. When the number of hidden layer node is 10 and the number of feature is 10, the classification accuracy is 94.87%. However, the classification accuracy doesn't improve further with the increase of feature number. Because feature selection can wipe off the irrelevant features to improve classification precision, feature selection mechanism worked. Meanwhile, with the increase of the hidden node number, the classification accuracy doesn't improve further, which indicates that the excessive increase of ELM node number generates fitting easily, causing the decrease of classification accuracy.

FS-ELM approach not only achieves classifier model, but it also achieves feature selection mechanism synchronously. The data set of network intrusion detection consists of 41 features. As can be seen from Table 1, not all features are helpful for classification accuracy, the feature selection improves classification precision, which has the same results as shown in the Fig. In order to study the feature selection process of FS-ELM approach roundly, which features in invasion detection system are more important? The importance of each feature is gained by FS approach. As is shown in Table 3, 10 most

important features are gained, the scores of these 10 feature are relatively high. Then the arranged distribution from high to low according to the importance of features is F7, F21, F2、F10, F22, F23, F6, F3, F26 and F20, illustrating that the relative level between these important features and network security invasion detection is higher than other features. Studying these features further can provide more powerful proof for the analysis and defense of intrusion detection, so the network security is safeguarded effectively.

The only determined parameter for ELM classifier is the number of hidden layer node, the setting of number of hidden layer node ranges [1,41], increasing gradually from 1, the interval is 1, the changes of the number of hidden layer node have influence on the classification results, as shown in Figure 3. Through observation, we can find that when the number of hidden layer node decreases, the classification accuracy rate is very low, that is because the low number of node cause low fitting phenomenon. The classification accuracy rate increases gradually with the increase of node number. When number of node is 13, the highest classification result is got. After that, the results have some fluctuations with the gradual increase of node number, finally the results tend to be stable gradually. So when the number of hidden layer node is 13, the highest classification accuracy rate is obtained (when we get the same results, take the minimum of number of node, so the more compact network structure can be obtained.).

Because the learning method is put forward to study the network security invasion detection, the selection of classifier number maybe has significant influence on the results of network invasion detection. But facing to this problem, there is no unified selection criteria and theory. Experimental study conducted on the number of classifier, the setting of value ranges [1, 10] through many experiments by manual setting. The parameter value which is corresponding to the best classification result is selected as the default of this experiment. The results are shown in Figure 4. From this Figure, we can see that the number of classifier begins from 1, then with the increase of number, the classification accuracy rate improves obviously. When the number of classifier is 5, the highest classification accuracy is obtained. Then with the increase of number, the classification results decrease gradually with fluctuation state. It shows that the excessive number of classifier isn't helpful for the improvement of classification accuracy rate.

Table 1. The Result Comparison of 5 Models

Model	10 features		All features	
	ACC	Std	ACC	Std
FS-ELM	93.59	0.42	91.38	0.53
Lin-SVM	91.03	0.55	89.74	0.63
RBF-SVM	92.31	0.67	89.10	0.72
LS-SVM	92.80	0.43	90.92	0.70
KNN	92.95	0.21	91.31	0.37

Table 2. The Change of FS-ELM Model Results in Numbers of Different Hidden Layer Nodes

Feature number / The number of node	5	10	15	20	25	30	35	41
5	83.33	89.31	87.18	87.82	83.97	89.10	83.33	86.54
10	90.03	94.87	89.46	90.38	79.49	89.59	80.77	85.26
20	91.67	93.59	90.38	87.18	81.41	91.67	80.77	88.46
30	87.82	91.03	91.49	88.46	83.33	83.69	82.05	91.67
40	87.18	91.67	87.18	87.82	85.26	79.49	86.54	91.67

Table 3. The Selected Feature Subset of FS-ELM Model from Intrusion Detection Data Set

Model	Selected feature	The number of feature
FS-ELM	F7, F21, F2, F10, F22, F23, F6, F3, F26, F20	10

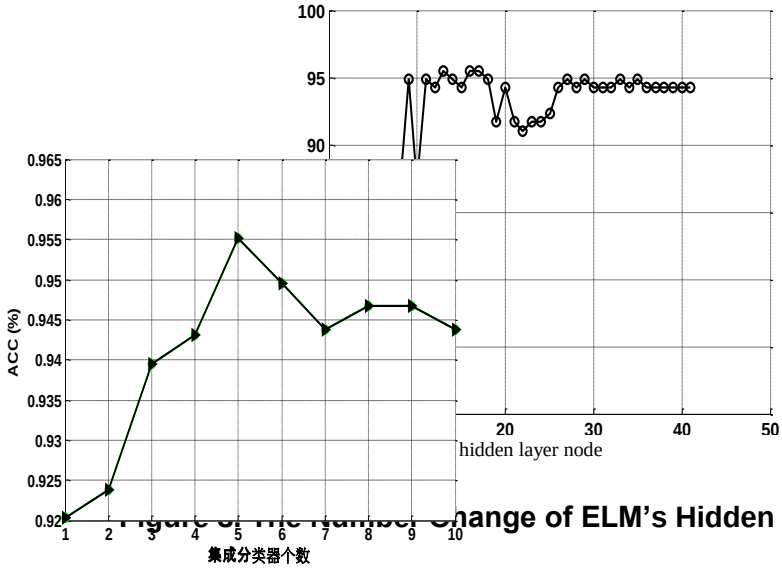

Figure 3. The Number Change of ELM's Hidden Layer Node

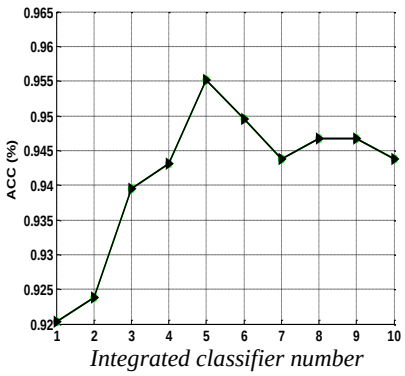

Figure 4. The Influence of Classifier Number on Detection Results

Figure 5 shows the comparative situation among five different models and single classifier with feature selection, classifier with feature selection, single classifier without feature selection, integrated classifier without feature selection. From the Figure, we can see that the former two results are obtained based on feature selection. The classification accuracy of five models under the condition of single classifier improved. The latter two results are obtained without feature selection. By contrast, the average classification accuracy rates based on feature selection improved than those without feature selection. Feature selection method plays an important effect, and shows the effectiveness of FS-ELM model.

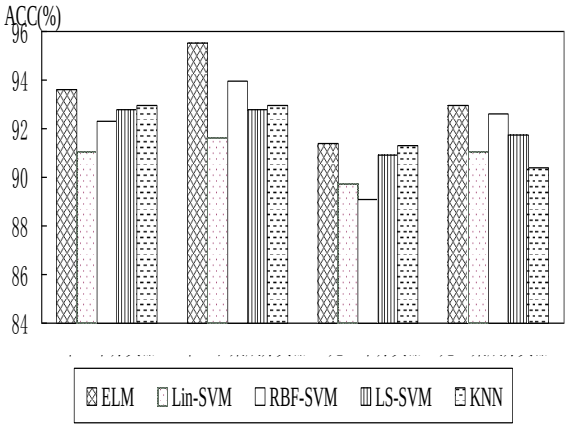

Figure 5. The Integrative Comparative Situation of Models

In addition, the training time, test time, detection accuracy rate and misstatement rate of these five kinds of algorithmic methods were compared too, as is shown in Table 4. From the Table, the training time and test time of FS-ELM algorithmic method is less than other four methods. The detection effectiveness is the highest, the accuracy rate is the highest and the misstatement rate is the lowest. That's because feature selection mechanism saved the time of learning time of classifier, eliminated irrelevant features, improved classification accuracy rate. Meantime, ELM algorithmic method was used, the effective learning ability itself improved training and test speed greatly, showing that FS-ELM can meet the real-time requirement of network intrusion detection better, further verifying that FS-ELM model is the effective model of improving intrusion detection system.

Table 4. The Comparison of Training Time, Test Time, Detection Accuracy Rate and Misstatement Rate among Five Models

Model	training time	Test time	detection accuracy rate(%)	misstatement rate (%)
FS-ELM	9.291	3.795	95.51	4.49
Lin-SVM	12.685	7.314	91.59	8.41
RBF-SVM	14.979	7.143	93.95	6.05
LS-SVM	13.212	6.962	92.80	7.20
KNN	12.867	12.867	92.95	7.05

In order to collect statistics whether the classification feature of five models has outstanding differences, the test of signed rank Wilcoxon is conducted based on the intrusion detection data set [8], as is shown in Table 5. In order to ensure fairness, the experiment got results by conducting the model 5 times independently, results can be seen from the table. FS-ELM compared with other four algorithmic methods. There were obvious differences in the statistics. It shows that the method has improved in the detection accuracy rate on the problem of intrusion detection.

Table 5. The Feature Comparison of Five Models

t detection	Five models				
	FS-ELM	Lin-SVM	RBF-SVM	LS-SVM	KNN
FS-ELM	-	0.042	0.048	0.043	0.042
Lin-SVM		-	0.465	0.043	0.042
RBF-SVM			-	0.138	0.138
LS-SVM				-	0.223
KNN					-

6. Conclusion

The intrusion detection is a hot issue in the field of network security research, a more accurate, stable and effective intrusion detection model can meet real-time network intrusion detection of complex changes and frequent attacks better. So network intrusion detection model which combines Fisher score and ELM is put forward. In this model, FS approach can eliminate superfluous features and irrelevant features from network data, select the feature subset which has higher correlation degree with detection results, reducing feature dimension and improving the accuracy rate of classifier's training time and classification. Meanwhile, using ELM classification device improves the accuracy rate of detection. The experiment selects the data of KDD CUP 99 database as experiment data, experiment results showed that the proposed model is better than models based on SVM, LS-SVM and KNN. It overcomes the shortcoming of SVM because the nuclear function parameter and punishment factor SVM need iterative regulation many times. It not only gets the higher detection accuracy rate, but it also quickens detection efficiency and lowers misstatement rate. So it is a kind of effective network invasion detection model.

Of course, there are still many places which are worthy of further study. First of all, ELM is used as a classifier, other models can be used in this data set to know whether they have better performances. This is the issue that needs to study.

References

[1] Tang Z. J. and Li J. H., "Intrusion Detection Technology[M]", Beijing: Tsinghua University Press, **(2004)**.
[2] Beghdad R., "Modeling and solving the intrusion detection problem in networks[J]", Computers & Security, vol. 23, no. 8, **(2004)**, pp. 687-696.
[3] Yu L. and Liu H., "Efficient feature selection via analysis of relevance and redundancy[J]", Journal of Machine Learning Research, **(2004)** May, pp. 1205-1224.
[4] Huang G. B., Zhou H. and Ding X., "Extreme learning machine for regression and multiclass classification[J]", IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics, vol. 42, no. 2, **(2012)**, pp. 513-529.
[5] Chen H. L., Yang B. and Wang G., "A three-stage expert system based on support vector machines for thyroid disease diagnosis[J]", Journal of Medical Systems, vol. 36, no. 3, **(2012)**, pp. 1953-1963.
[6] Huang G. B., Zhu Q. Y. and Siew C. K., "Extreme learning machine: a new learning scheme of feedforward neural networks[C]", IEEE International Joint Conference on Neural Networks, **(2004)** February, pp. 985-990.
[7] Huang G. B., Zhu Q. Y. and Siew C. K., "Extreme learning machine: theory and applications[J]", Neurocomputing, vol. 70, no. 1, **(2006)**, pp. 489-501.
[8] Rey D. and N. M. Wilcoxon, "Signed Rank Test [M]", International Encyclopedia of Statistical Science, Springer Berlin Heidelberg, **(2011)**, pp. 1658-1659.

