Efficient Intrusion Detection Scheme based on SVM

ZHOU Guangping
Zhejiang University of Science & Technology, Zhejiang, 310023, China
E-mail: zhouzhou@zust.edu.cn

Anup Shrestha
University of Southern Queensland, Toowooba, QLD 4350, Australia
E-mail: anup.shrestha@usq.edu.au

***Abstract*—The network intrusion detection problem is the focus of current academic research. In this paper, we propose to use Support Vector Machine (SVM) model to identify and detect the network intrusion problem, and simultaneously introduce a new optimization search method, referred to as Improved Harmony Search (IHS) algorithm, to determine the parameters of the SVM model for better classification accuracy. Taking the general mechanism network system of a growing city in China between 2006 and 2012 as the sample, this study divides the mechanism into normal network system and crisis network system according to the harm extent of network intrusion. We consider a crisis network system coupled with two to three normal network systems as paired samples. Experimental results show that SVMs based on IHS have a high prediction accuracy which can perform prediction and classification of network intrusion detection and assist in guarding against network intrusion.**

***Index Terms*—Support Vector Machine (SVM), Improved Harmony Search (HIS), Intrusion Detection, Comprehensive Evaluation**

I. INTRODUCTION

In recent years, the Internet is gaining rapid development and the wave of information has swept across the globe [1]. Applications of information technology provide the necessary conditions for improvement of work efficiency in all walks of life; transform people's working style, living environment and thinking concept; and promote development of a connected society. However, wide application and increasing popularity of network technologies used in the Internet have also brought various risks and threats to the network system, and thereby network system security is a serious challenge in today's world. Research shows that information systems in the digital society is a tempting target for intrusion by hacker groups, organized crime and so on that threatens all types of information systems [2]. Since the network system developed from dedicated and closed networks into an open and distributed architecture, computer infrastructure and networks are facing all kinds of security threats such as hacking which exploits weaknesses of computer networks, data tampering and deletion, unauthorized system access and modification, computer viruses and malwares, impersonated communication, network eavesdropping, deliberate human sabotage and unexpected natural disasters [3]. Complex and volatile security threats and potential safety hazards are very hard for people to safeguard. Kisi (2004) reported that public security bodies in China uncovered cases of more than 16,700 people using networks to commit crimes in the period between the years 2000 to 2009 and arrested 19,300 criminal suspects. The total expenses incurred in this investigation was nearly 9.5 billion Yuan [4]. It is therefore clear that people are facing serious information security problems while ironically enjoying incredible benefits brought by the digital society.

Due to the reasons discussed previously, network invasion problems have been drawing more and more attention in recent years. Many leading countries in Europe and America are engaged in promoting digitization and broadening network connections with a focus on taking required information security measures to strengthen information security management of networks [5-7]. Due to continuing expansion of networks and increasing risks of networked information systems, research on problems faced by existing network intrusion detection systems and building a better information systems security system have become significant in order to protect information security across networks.

There are several types of network attacks. Five prominent types of network attacks are briefly discussed next. (1) Passive attacks. In this type of attack, unauthorized information is passively accessed by the attacker but such information are not modified. Passive attacks include unauthorized network traffic analysis, eavesdropping unprotected network communication, decrypting data that is encrypted using a weak key and gaining unauthorized information (including identification theft by using passwords). Passive attacks lead to leaking of information or data files to the attacker without consent and understanding of the communicating parties thereby destroying confidentiality of information exchange. For example, leakage of credit card numbers, medical documents and/ or other personal information in

doi:10.4304/jnw.8.9.2128-2134

networks [8]. (2) Active attacks. Active attacks include intentional act of the attacker tampering information in order to destroy existing protection mechanism and data often by introducing malicious application codes and stealing or modifying information. Examples of active attack approaches include attacking network infrastructure, misusing unauthorized information, electronically penetrating network perimeters or attacking a legitimate network user trying to connect to an enclave. The consequences caused by active attacks include leakage of transmitted data files, blockage of network services and unintended data modification [9-10]. (3) Physical close-in attacks. It refers to unauthorized individuals physically accessing networks, systems or equipments for the purpose of changing, collecting or blocking access to data in the network. The major approaches of real physical close-in attacks are secretly entering in physical networks or enabling unauthorised access to physical network devices; or concurrently using these two approaches [11]. (4) Insider attacks. Insider attacks could be malicious or non-malicious, wherein malicious attacks typically mean that internal staff working in network systems intercepts or destroys information in a planned way to use such information for fraudulent purposes and interdicts access of other authorized users. Conventional non-malicious attacks are often caused by actions such as carelessness, lack of technical knowledge or inadvertently bypassing security policies in order to "finish work" [12]. (5) Distribution attacks. This means malicious modification of hardware or software within the network or in the processes of a distributed software or hardware. These attacks may introduce a backdoor component or other malicious codes embedded to a hardware or software product originally distributed for a different purpose in order to access information or perform unauthorized system function at a later time [13].

With continuous improvement and progress in the fields such as statistical learning and artificial intelligence, many statistical approaches and intelligent computation methods are being applied into the research of network intrusion. In current premises of rapidly developing network information particularly due to massive adoption of the Internet, increasing challenges due to mounting network system intrusions and continuously improving science and technology in computing, this study will construct a mixed classification model for network intrusion detection using the Support Vector Machine (SVM) combined with Improved Harmony Search (IHS) algorithm. Moreover, this study puts forward the corresponding network intrusion prevention countermeasures, in order to promote the ability to guard against network intrusion.

II. Proposed Approach for Intrusion Detection

In this section, we present the proposed intrusion detection approach based on SVM model with IHS algorithm [14]. The framework is composed of the following procedures: data collection, preprocessing, feature extraction, detection algorithm and evaluation. The graphical illustration of the proposed framework can be found in Figure 1.

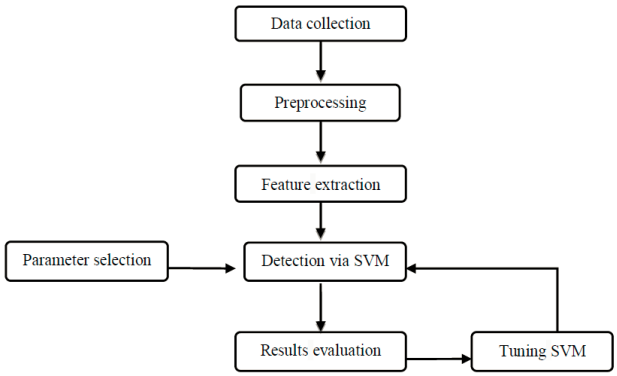

Figure 1. Framework of the proposed intrusion detection approach

A. Data Collection and Feature Extraction

The data used in our experiments were collected from the general mechanism network system of a certain city in China between 2006 and 2012. We divide the mechanism into normal network system and crisis network system according to the harm extent of network intrusion. A crisis network system is defined as the network which has more network intrusion frequency, more invasion cases, larger threat to its network system and is close to collapse. Seven categories of network system are classified and illustrated in Table 3.

TABLE I. Samples Distribution List

	Normal Network System	Crisis Network System	Sub-total	In Total
Training Samples	370	161	531	884
Verification Samples	92	40	132	
Testing Sample	154	67	221	

TABLE II. Intrusion Category Distribution of Network System

Invasion Categories			
Virus attack	Management system is not sound	Key safety is missing	Software upgrade is slow
System vulnerability	Management personnel quality is low	Hardware configuration is relatively low	Network planning is not reasonable

During data collection, a crisis network system coupled with two to three normal network systems are used as paired samples. Data collection for paired samples is conducted in the following way: after obtaining the category of invasion suffered by this system from the basic properties of the crisis network system, two or three normal network systems which are similar in nature and size to the crisis network system are determined to represent the paired sample. In this study, the samples involved 884 network systems, of which 60% are training samples, 15% are verification samples and 25% are test samples. The sample distribution is demonstrated in Table 1. We use eight invasion

categories for classification as shown in Table 2. We use the normalized version of the collected data as features.

B. Intrusion Detection Using Support Vector Machine

The Support Vector Machine (SVM) was put forward by Vapnik et al. in 1995 and its main theory is the theory of Structural Risk Minimization of statistical learning theory. The SVM is primarily to use the Separating Hyper-plane to separate two-classes data and process the classification problems in data. Depending on the SVM classification boundaries, it can be divided into two types: linear support vector machine and nonlinear support vector machine.

The SVM method evolved from the optimal separating hyper-plane in the linearly separable case. The so-called optimal separating hyper-plane is required to be capable of not only separating two kinds of samples correctly but also maximizing the classification interval. Assuming that the training data is as follows:

$(x_1, y_1), \cdots, (x_l, y_l)$ where $x_i \in R^n$ is the input sample; $y_i \in \{+1, -1\}$ is the category table;

$i = 1, 2, ...l$ If x_i belongs to the first category, y_i will be $y_i = 1$, or else, it will be $y_i = -1$. The learning goal is to construct a discriminant function, to separate these two categories as correctly as possible.

$(w \cdot x_i) + b \geq 1 \rightarrow y_i = +1$

$(w \cdot x_i) + b \leq 1 \rightarrow y_i = -1$

where w is the normal vector of the super-plane, and x_i is a deviation value. To solve the separating hyper-plane, you need to resolve the following quadratic programming (QP) problem which has the following constraint conditions:

$$y_i (w \cdot x_i) + b \geq 1, i = 1, 2, ..., l \quad (1)$$

We need to solve the minimum value of the following functions:

$$\min \phi(X) = \frac{1}{2} \|W\|^2 \quad (2)$$

Because Equation (5) is a quadratic function and the Constraint (4) is a linear expression, the objective function belongs to the typical quadratic programming problem. Using Lagrange multiply to solve this quadratic programming problem with linear constraints, we will obtain:

$$L[w, b, a] = \frac{1}{2} W^2 - \sum_{j=1}^{j} a_j \left[y_i \left(W \cdot x_j + b \right) - 1 \right] \quad (3)$$

$$s.t. \alpha_i \geq 0 \quad (4)$$

But this form of the support vector machine is still difficult to solve. The solution is to find out its dual problem. The new problem is as follows:

$$L_D = \sum \alpha_i - \frac{1}{2} \sum_{i,j} \alpha_i \alpha_j y_i y_j x_i x_j \quad (5)$$

While it was originally a problem to solve the minimum value of L, its dual problem has now turned into a problem to solve the maximum value, in which a set of constraints $\alpha_i \geq 0$ applies.

When you address the optimal solution of dual problem, every Lagrange coefficient q will correspond to a training sample. When the coefficient is greater than 0 or the threshold, it means the sample is a support vector, and will fall into the boundaries of the separating plane. To substitute it into Equation (8), and furthermore using the Karush-Kuhn-Tucker conditions given by Fletcher, we can obtain the value of b. Finally we get the final classification function as follows:

$$f(x) = \text{sign} \left[\sum_{i=1}^{l} \alpha_i y_i x_i^T x + b \right] \quad (6)$$

In this Equation, α_i is a Lagrange multiplier, and b is the classification threshold of the equation.

C. Efficient Detection Using the Improved Harmony Search Algorithm

As the collocation and combination of musical instrument tones determine the achievements of music performance, such phenomenon controlled by IHS algorithm creates a new calculation structure. The executive routine of IHS algorithm that contains five steps is illustrated in Figure 2.

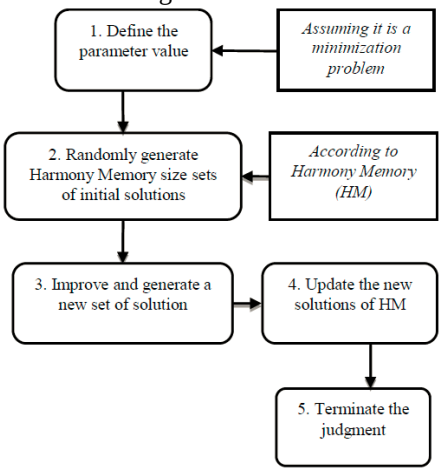

Figure 2. Executive routine of IHS algorithm

Step 1: Define the problem and the parameter value, assuming it was a minimization problem, and its model is as follows:

$$\min \ mizef(x)$$

$$\text{subject to } x_i \in X_i \ for \ \ i = 1, 2, ...N \quad (7)$$

where $f(x)$ is the objective function value, and x_i is the i -th decision variable value, and X_i is the discrete variable value, which is $X_i = \left(x_i(1), x_i(2), \cdots, x_i(k) \right)$. If X_i is a continuous variable, then $x_i^l \leq X_i \leq x_i^U$, where N is the number of decision variables, K is the number of possible values in the discrete variables. The parameters of IHS algorithm include: (1) Harmony

Memory Size (HMS), i.e., in each generation, reserving the feasible solutions of HMS set to the Memory for the next generation to evolution, (2) HMCR (Harmony Memory Considering Rate), among them, $0 \le HMCR \le 1$; (3) PAR (Pitch Adjustment Rate), among them, $0 \le PAR \le 1$; (4) bw is the parameter used to control the variable moving step size.

Step 2: Establish an initial solution to randomly generate HMS sets of initial solutions, and calculate the respective objective function values. After sorting according to size of $f(x)$, store the solutions to the Harmony Memory, i.e. $HM = \left(x^1, ..., x^{HMS}\right)^T$, where x^i is the i-th set of solution after IHS algorithm of function assessment values are sorted $x^i = \left(x_1^i, x_2^i, \cdots, x_N^i\right)$.

$$HM = \left(x^1, ..., x^{HMS}\right)^T$$

Step 3: According to the generation of Harmony Memory, gradually improve and generate a new set of solution $x' = \left(x'_1, x'_2, \cdots, x'_N\right)$. Each variable in x' is generated mainly by relying on these three mechanisms: (1) memory consideration: the variable values are randomly drawn from HM; (2) pitch adjustments: the variable values will be fine adjusted after being drawn from HM; (3) randomization: generated in random. For example, the first newly-solved variable x'_1 can have HMCR probabilities to be selected from $\left(x'_1 \sim x_1^{HMS}\right)$ of HM; one of the values has l – HMCR probabilities to be randomly generated in the variable scope. Similarly, other variables have the following general equation:

$$x'_i \leftarrow \begin{cases} x'_i \in \left\{x'_i, x'_2, \cdots, x_i^{HMS}\right\} & w.p. \quad HMCR \\ x'_i \in X_i & w.p. \quad (1 - HMCP) \end{cases} \tag{8}$$

Each variable value has the probabilities of HMCR to be selected out of the historical values of number stored in HM, and has the probabilities of 1 – HMCR to be randomly generated in the variable scope. So, when HMCR = 1, the selected variable values mainly stem from the historical values of number stored in the HM, i.e., giving up the opportunities besides values of number in the HM. And the probabilities of l - HMCR will be randomly generated in the variable scope; this approach is similar to the significance of genetic algorithm mutation. Yet the variable values in new solutions from HM have the probability of PAR for Pitch Adjustment to find the nearest solutions.

$$x'_i \leftarrow \begin{cases} Yes & w.p. \quad PAR \\ No & w.p. \quad (1 - PAR) \end{cases} \tag{9}$$

Therefore, HMCR and PAR in IHS algorithm respectively play the roles of Global Search and Local Search weights.

Step 4: Update the new solutions of HM generated in Step 3, which will be better than the worst set of solutions among function evaluation values of HM via the function evaluation, then update the new solutions into HM.

Step 5: Terminate the judgment. If it fails to meet the termination conditions, return to Step 3 ; if it has met the termination conditions, stop the algorithm or go back to Step 3 and Step 4. When IHS algorithm generates the new solutions $x^i = \left(x_1^i, x_2^i, \cdots, x_N^i\right)$ in Step 3, each variable x_f is generated by depending largely on these three mechanisms: (1) using the variable values-memory consideration from the Memory: the variable values are randomly drawn from HM; (2) pitch adjustments: the variable values will be fine adjusted after being drawn from HM; and (3) randomization: generated in random.

The two parameters in IHS Algorithm, PAR and bw are the key factors in the process of adjustment and optimization. The traditional Harmony Search algorithms have these two parameters fixed, and hence cannot vary with the update to music. Therefore if the two parameters PAR and bw are improperly set in the beginning, such as a smaller PAR coupled with a larger bw, it will degrade the optimizing performance and increase the number of iterations required for optimizing performance, and therefore the convergence time will be prolonged. Obviously, in the process of using IHS algorithm for optimization, choosing a bigger bw can promote the diversity of fine adjustment search at the beginning stage of optimization, but as the optimization process evolves, bw shall be smaller and smaller. This is the only way it can slowly approach to the global optimal solution. The fine-adjustment probability PAR, on the other hand, shall also be gradually increasing along with the evolution of optimization in order to slowly approach to the global optimal solution.

Based on the above consideration, Mahdavi et al. put forward an improved harmony search (IHS) method, namely using the changing bw and PAR in Step 3, which will dynamically change with the evolution algebra of new music, as shown below.

$$PAR(gn) = PAR_{\min} + \frac{\left(PAR_{\max} - PAR_{\min}\right)}{NI} \times gn \tag{10}$$

$$bw(gn) = bw_{\max} \cdot \exp(c \cdot gn) \tag{11}$$

$$c = \frac{ln\left(\frac{bw_{\min}}{bw_{\max}}\right)}{NI} \tag{12}$$

where $x_{\max}$ and $x_{\min}$ respectively represent the maximum and minimum values of x, and gn and NI respectively represent evolution algebra and total evolution algebra.

III. Experimental Results

In this section, we will evaluate the proposed intrusion detection approach. The experimental procedures of the proposed intrusion detection method can be found in Figure 3. It contains seven main procedures: data collection, data preprocessing, feature extraction,

parameter selection, intrusion detection, results evaluation and retraining.

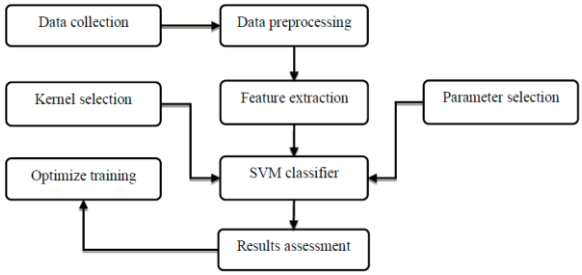

Figure 3. The experiment flow

The data used in our experiments were collected from the general mechanism network system of a certain city during 2006 and 2012[15-16]. The networks are divided into normal network system and crisis network system, according to the extent of network intrusion. A crisis network system is coupled with two to three normal network systems as paired samples, of which there are 268 crisis network system and 616 normal network systems, with a total of 884. As for the collection of paired samples, it was achieved in the following way: after obtaining the category of invasion suffered by this system from the basic properties of the crisis network system, two or three normal network systems which are similar in nature and size as the crisis network systems were found out to act as its the paired sample. The categories of network system include seven categories, as shown in Table 3.

TABLE III. CATEGORY DISTRIBUTION OF NETWORK SYSTEM

Invasion Categories	Crisis Network System	Normal Network System	Total Number of Samples
Middle School LAN	31	60	91
Primary School LAN	25	45	64
University LAN	25	67	92
Community LAN	51	122	173
Institution LAN	59	134	193
Small Business LAN	22	49	71
Big Business LAN	34	83	117
Total number of samples in total	268	616	884

A. Model Accuracy Assessments

Far from being determined by the size of mean square error (MSE), he advantages and disadvantages of classification prediction model should depend on the Identification Correct Rate (CorrRate) to test samples in the model test phase, and its definition is as follows:

$$CorrRate = \frac{TN + TP}{TN + TP + FN + FP}$$

Among them, TN = True normal network system (True Negative), TP = True crisis network system (True Positive), FN = False normal network system (False Negative) and FP= False crisis network system (False Positive) are four performance parameters in the composition of the Confusion Matrix, and this Matrix style is as shown in Table 4. However it must be noted that the four performance parameters is based on a partition threshold value TH and the TH value used in this study will take the intermediate value of 1.5 which is between 1(for crisis network system) and 2 (for normal network system).

TABLE IV. CONFUSION MATRIX

		The actual situation	
		Crisis Network System	Normal Network System
prediction	Crisis network system	TP	FP
	Normal network system	FN	TN

In addition, we can also use the Receiver - Operating Characteristics (ROC) curve for judgment. The ROC curve drawing is based on Sensitivity and Specificity for mappings, which are respectively defined as:

$$Sensitivity = \frac{TP}{TP + FN}$$

$$Specificity = \frac{TN}{TN + FP}$$

Taking (1 - Specificity) as the horizontal axis and Sensitivity as the longitudinal axis, we begin to gradually adjust the threshold value from 1 to 2. Each threshold value can obtain the results of Sensitivity and Specificity respectively, which is a point in the map, so having these many points connected to form a curve, namely the ROC curve. If it is an effective prediction model, then the area under ROC curve (Area Under the Curve, AUC) will be greater than 0.5. We know that better the prediction model is, higher the AUC will is. To a first approximation, the prediction accuracies between 0.5 ~ 0.7 are not good , those between 0.7 and 0.9 belong to medium prediction accuracies, and those between 0.9 and 1.0 mean that the prediction accuracy is very high, so ROC (or AUC) is a threshold - independent index for prediction model accuracy assessment .

B. Main Results

In the first experiment, we use the SVM methods for intrusion detection, and employ the general mechanism network system of a certain city between 2006 and 2012 as the sample to conduct the experiment. This study divides the mechanism into normal network system and crisis network system, and then it used Type I Error (Normal theory accuracy) and Type II Error (Abnormal theory accuracy) as the evaluation criteria to identify the effectiveness of SVM in network intrusion detection. During the experiment we also used the Improved Harmony Search (IHS) algorithm to determine the parameters of Support Vector Machine (SVM) model. Then we used the identified parameters as the input index to perform the training. The test was performed for ten iterations on this model, and the predictive performance of the test is as shown in Table 5. In the experiment the following IHS and SVM parameters setting were applied: HMS = 30, HMCR = 0.7, NI = 1000, PARmin = 0.35, PARmax = 0.9, bwmax = 10-4, bwmax = 2, C = 1540, r = 0.0826. We used Type I Error (Normal theory accuracy) and Type II Error (Abnormal theory accuracy) as the evaluation criteria to identify the effectiveness of SVM

and IHS in network intrusion detection. From the experiment result, we can observe that the Type I Error and Type II Error are lower than 5% in both the normal network system and the crisis network system with the training as well as the sample data. The average rates are 3.26% and 3.58%. respectively. The reasons counting for the above results are mainly from the following three aspects: first, the SVM has the ability to map the non-linear data in the low dimensional to the line high dimensional by a Kernel function which can make the classifier adapt to data distribution well. The adaptability essentially comes from the flexibility of the Kernel parameters. Second, the selection method for kernel parameters can adapt to dataset much better in comparison with empirical parameter selection. Third, the SVM method can be applied to the conditions where sample data is large-scale, complex dimension, and containing a large number of heterogeneous information.

TABLE V. PREDICTION RESULTS OF NETWORK SYSTEM INTRUSION CATEGORY (SVM-IHS MODEL)

	Type I Error	Type II Error
Run 1	2.33%	4.52%
Run 2	4.95%	2.26%
Run 3	3.24%	4.52%
Run 4	1.50%	4.98%
Run 5	2.60%	4.07%
Run 6	4.79%	3.62%
Run 7	2.60%	2.33%
Run 8	3.14%	4.98%
Run 9	1.76%	2.71%
Run 10	4.67%	1.81%
Avg	3.26%	3.58%

In the second experiment, we divide the mechanism into normal network system and crisis network system according to the harm extent of network intrusion, and use a crisis network system coupled with two to three normal network systems as paired samples. We used the SVM method to training the data and the IHS method to determine the Kernel parameters. Through this experiment we can prove the effectiveness of the IHS method in determining the Kernel parameters. The test was performed with ten iterations on this model, and the predictive performance of the test is shown in Table 6. In the experiment, IHS and SVM parameters are set with the following parameters: HMS = 30, HMCR = 0.7, NI = 1000, PARmin = 0.35, PARmax = 0.9, bwmax = 10-4, bwmax = 2, C = 1540, r = 0.0826, TH=0.5. The evaluation criteria are standard deviation and the error variability of Type I and Type II. CorrRate is the model accuracy of the threshold value TH at 0.5, and the average value (Avg) is 87%. Demonstrated by the standard deviation (Std), the error variability of Type I and Type II is greater than that the CorrRate suggesting that measuring the model performance with CorrRate is more reliable. Furthermore it was found that the average value is 92% (> 90%) for the AUC pointer suggesting that SVM-IHS model is an excellent prediction model. Table 6 demonstrates each of the ten execution results where AUC is always greater than the CorrRate values and there are relations of monotonous change between the two indexes. The three possible reasons for this observation are discussed next: firstly, AUC has larger variability than CorrRate and it should be more reliable than the average value of CorrRate. Secondly, CorrRate is just a performance pointer of TH = 0.5, however AUC is the average performance pointer of all possible THs. Thirdly, the selection method IHS for kernel parameters can adapt to datasets much better in comparison with empirical parameter selection.

TABLE VI. PREDICTION RESULTS OF NETWORK SYSTEM INTRUSION CATEGORY (SVM-IHS MODEL)

	CorrRate	AUC
Run 1	89.14%	0.9394
Run 2	87.78%	0.9380
Run 3	88.24%	0.9360
Run 4	85.52%	0.9277
Run 5	87.33%	0.9328
Run 6	89.59%	0.9493
Run 7	85.07%	0.9147
Run 8	86.88%	0.8928
Run 9	85.52%	0. 9033
Run 10	85.52%	0.8655
Avg	87.06%	0.9200
Std	1.63%	0.0259

In the third experiment, we investigated the prediction results of ROC Curve in the SVM-IHS model. The ROC curve after performing the third iteration is as shown in Figure 4. Taking (1 - Specificity) as the horizontal axis and Sensitivity as the longitudinal axis, we begin to gradually adjust the threshold value from 1 to 2. Each threshold value can obtain the results of Sensitivity and Specificity respectively, which is a point in the map. Having several points connected to form a curve, we plotted the ROC curve. The test were performed for ten iterations on this model, and the ROC curve after performing the third iteration is as shown in Figure 4. In the process the parameters are the punish factors. Moreover we also have the following parameters set: HMS = 30, HMCR = 0.7, NI = 1000, PARmin = 0.35, PARmax = 0.9, bwmax = 10-4, bwmax = 2, C = 1540, r = 0.0826, TH=0.5. We used the Receiver - Operating Characteristics (ROC) curve as the evaluation standard. The ROC curve drawing is based on Sensitivity and Specificity for mappings. From the result we can see the area under ROC curve (Area Under the Curve, AUC) is be greater than 0.5, which is much bigger than the Random classifier. Every point of the ROC Curve is higher than the Random Classifier Curve. Likewise, those between 0.7 and 0.9 belong to medium prediction accuracies, and those between 0.9 and 1.0 mean that the prediction accuracy is very high, so ROC (or AUC) is a threshold - independent index for prediction model accuracy assessment. The reasons to explain the above results are discussed next: in comparison with the traditional machine learning methods, the SVM can be applied to the condition when the sample data is large-scale, complex dimension and contains a large number of heterogeneous information. Furthermore, the application of the IHS method according to the distribution information of the input data to select the Kernel parameters of the SVM provides better adaptability for the SVM model.

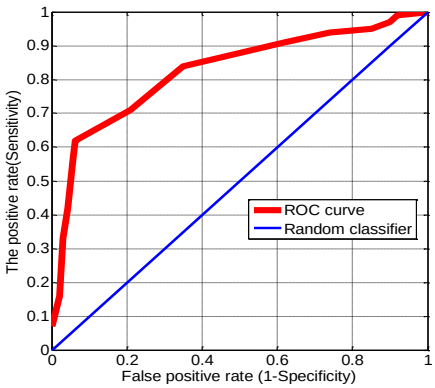

Figure 4. ROC Curve in SVM-IHS Model

IV. CONCLUSION

With high-speed development and widespread use of network information in today's connected world, network invasion is becoming an increasingly serious challenge for network engineers and managers. Therefore, this experiment has vital significance for the security of network information to build a robust network intrusion detection and security system. Network intrusion detection is a big concern for network information system security in general and guarding against network intrusion is a basic, preliminary guarantee for the development and management of secure network information. This study establishes a support vector machine (SVM) model which is used to identify and predict the network intrusion problems; meanwhile, the improved harmony search (IHS) algorithm is introduced to determine the parameters of the model of support vector classification in order to make the SVM have better classification accuracy.

This study provides an in-depth analysis of the SVM model and the Improved Harmony Search (IHS) Algorithm theory in order to evaluate the model accuracy. Taking the general mechanism network system of a certain city between 2006 and 2012 as the sample, this study carries on the simulation analysis. Research findings show that the Support Vector Machine-Improved Harmony Search (SVM-IHS) model has better prediction accuracy and this can be used for prediction and classification of network intrusion detection. This study enriches the method to guard against network intrusion in theoretical aspects. On a practical level, this study establishes the SVM model, introduces the Improved Harmony Search (IHS) algorithm, constructs the SVM-IHS model to identify and detect the network intrusion problems, and provides reference for guarding against network intrusion, which can ultimately play a positive role to promote high-speed development of network information over the Internet.

ACKNOWLEDGMENT

This work is supported by the project: Principles of Information Security foreign course construction (No. F527106C01).

REFERENCES

[1] Bryant, S. M.. “A case-based reasoning approach to bankruptcy prediction mode1ing, Intelligent Systems in Accounting," *Finance and Management*, 6, pp. 195-214, 1997

[2] Coulibaly, P., Anctil, F. and Bobee, B. ‘‘Daily reservoir inflow forecasting using artificial neural networks with stopped training approach’’, *Journal of Hydrology*, Vol. 230, pp. 244–257, 2009

[3] Kisi, O. ‘‘Multi-layer perceptrons with Levenberg–Marquardt training algorithm for suspended sediment concentration prediction and estimation’’, *Hydrological Sciences Journal*, Vol. 49, No. 6, pp. 1025–1040, 2009

[4] Kisi, O. ‘‘River flow forecasting and estimation using different artificial neural network techniques’’, *Hydrology Research*, Vol. 39, No. 1, pp. 27–40, 2008

[5] Cannas, B., Fanni, A., See, L. and Sias, G. ‘‘Data processing for riverflow forecasting using neural networks: wavelet transforms and datapartitioning’’, *Physics and Chemistry of the Earth*, Vol. 31, pp. 1164–1171, 2006

[6] Karunasinghe, D. S. K. and Liong, S. Y. ‘‘Chaotic time series prediction witha global model: artificial neural network’’, *Journal of Hydrology*, Vol. 323, pp. 92–105, 2006

[7] Shuang Xu, Jifeng Ding, Palmprint Image Processing and Linear Discriminant Analysis Method, *Journal of Multimedia*, Vol. 7, No. 3, pp. 269-276, 2012

[8] Wang, W., Van Gelder, P., Vrijling, J. K. and Ma, J. ‘‘Forecasting dailystreamflow using hybrid ANN models’’, *Journal of Hydrology*, 324, pp. 383–399, 2006

[9] Gang Liu, Jing Liu, Quan Wang, Wenjuan He, The Translation Invariant Wavelet-based Contourlet Transform for Image Denoising, *Journal of Multimedia*, Vol. 7, No. 3, pp. 254-261, 2012

[10] Yan Zhao, Hexin Chen, Shigang Wang, Moncef Gabbouj, An Improved Method of Detecting Edge Direction for Spatial Error Concealment, *Journal of Multimedia*, Vol. 7, No. 3, 262-268, 2012

[11] Yan-pei Liu, Yuesheng Gu, Jun Chen, A New Control Structure Model Based on Object-oriented Petri Nets, *Journal of Networks*, Vol. 7, No. 4. pp. 746-753, 2012

[12] Jing Ma, Shuli Sun, "Centralized Fusion Estimators for Multi-sensor Systems with Multiplicative Noises and Missing Measurements", *Journal of Networks*, pp. 1538-1545, Vol. 7, No. 10, 2012

[13] Yiran Wang, Bo Zhang, Shujian Li, "Differential Coherent Algorithm for Interferometric Acquisition of GNSS-R Software Receiver", *Journal of Networks*, Vol. 7, No. 10, pp. 1687-1695, 2012

[14] Li Yi, Huachun Zhou, Fei Ren, Hongke Zhang, "Analysis of Route Optimization Mechanism for Distributed Mobility Management", *Journal of Networks*, Vol. 7, No. 10, pp. 1662-1669, 2012

[15] Guangming Zhang, Zhiming Cui, Pengpeng Zhao, Jian Wu, A Novel De-noising Model Based on Independent Component Analysis and Beamlet Transform , *Journal of Multimedia*, Vol. 7, No. 3, 247-253, 2012

[16] Gang Wang, lin Xiao Gui, DRTEMBB: Dynamic Recommendation Trust Evaluation Model Based on Bidding, *Journal of Multimedia*, Vol. 7, No. 4, pp. 279-288, 201