

Evaluating the Network and Information System Security Based on SVM Model

Shaomei Yang

Economics and Management Department, North China Electric Power University, Baoding City, China
Email: yangshaomei77@126.com

Donglei Liu

College of Business, Agricultural University of Hebei, Baoding City, China
E-mail: dongleiliu@163.com

Zhibin Liu

Economics and Management Department, North China Electric Power University, Baoding City, China
Email: liuzhibin771112@126.com

Abstract—With more and more e-business operating on line, the network and information security problem is not only confined to the concept of network technical systems, but at the same time has special significance beyond network itself, which is economic ties, cultural transmission media, the social operation and management way and so on. Based on the analysis of the security importance and security evaluation present, this paper establishes a security evaluation system and describes the evaluation mechanism based on SVM algorithm and model. The security evaluation of 12 network enterprises in Beijing shows that the results given by this model are reliable, and this method to evaluate the network and information system security is feasible.

Index Terms—network and information system, security, security evaluation, SVM

I. INTRODUCTION

With the through computer applications, especially the increasing popularity of home computers, on the one hand, it hopes that many users can share information resources, on the other hand, it also hopes that can transmit information and communicate with each other among the computers. The hardware and software configuration of the personal computers is relatively low in general, and the function is limited, therefore, it requires that the hardware and software resources of the large and super computer and their management information resources should be shared for the large number of micro-computers, in order to take full advantage of these resources. For these reasons, promote the computer's network development; connect the dispersed computer into the network; and make up the computer network. Network is the combination product of the modern communications technology and computer technology. The so-called computer network, is a network system of the large scale and strong functions,

which lines with the computer distributed in different geographical regions and the specialized external equipment communication circuit, so that many computers can transmit information to each other expediently, and share the resources of the hardware, software, Data information, and so on. Popularly, the network is the computer collection which lines with the cable, telephone lines, or wireless communications, and other. Information system is defined technically as an interrelated components collection through the information gathering, processing, storage and distribution supported the decision-making and control in organization. Security evaluation is also said risk evaluation. The definition of security evaluation is: measure and predict the system security through using synthetically the security systems engineering approach, confirm the possibility and order of severity of the system danger through the qualitative and quantitative analysis, and put forward the necessary measures in order to find the lowest accident rate, the smallest incident loss and the best security investment returns.

With the emergence of all kinds of e-mail spread viruses, people attach importance to the network security. More and more customers want to have a clear understanding about the network and information systems security which being used or will be used of their own, but most of them lack the relevant knowledge, expertise and resources, who can't decide whether the confidence level of the networks and information systems security is appropriate, and don't hope to rely entirely on the system developer in this regard, so they want a third party to help them analyze the network system security, that is, conduct security evaluation. Through evaluation, the users can judge whether the networks and information systems is adequate security for their own applications, and whether the concealed security risks are acceptable. In addition, in order to ensure national information infrastructure security, the management department also

call the shots that security evaluation is very necessary for a variety of network and information systems, and to determine the level of safety and reliability. Clearly, security evaluation has the important significance for the network development.

The network and information systems security evaluation is that identifies, analyzes and evaluates the security, identifies all of the security factors in the system and the linkages between them. A comprehensive security analysis including: describe various levels of the security probability and influence, give uniform measurement criterion to evaluate the security indicators of the factors, and then give the security level throughout the system based on some evaluation methods. The calculation methods of the security evaluation have qualitative method and quantitative method. The main evaluation methods including: AHP, decision tree analysis, Markov modeling and so on, however, these methods are subject to stochastic factors in the evaluation, and the evaluation results are influenced by subjective experience and knowledge limitations easily, which affect the accuracy and objectivity of the evaluation results. In this paper, we evaluate security level using the method of combining the qualitative with the quantitative, and overcome the subjectivity in the evaluation through the application of self-adapting regression Support Vector Machine (SVM).

II. THE SECURITY EVALUATION INDEX SYSTEM

According to the national basic requirements of the network and information security system, with the practical experience engaged in information security management work in many years, we believe that should follow the following principles when formulate the system security evaluation index system:

- (1) Integrate the laws and regulations of the state information and information system security.
- (2) Satisfy the security requirements of the users and application environment to the information systems.
- (3) Good maneuverability and implementation expediently.
- (4) Simple, practical, feasible and economic.

A Entities and environment security index

The entities and environment security index includes ten aspects: Without dangerous building within 100m around the motor room (U_1), dangerous building is that the places of existing flammable, explosive, toxic gases, etc., such as the such as gas stations, gas pipelines and so on; Monitoring system (U_2), that is, the monitoring implementation facilities of the external environment and operating environment during the system running; Fireproof and waterproof measures (U_3), fireproof is that equipped with automatic fire alarm system in the computer room, or the fire-fighting equipment, emergency plans and related systems applied to the computer room; waterproof is that no water seepage and leakage in the computer room, for example, it needs waterproof layer if having water facilities on the upper room; Environment monitoring and control facilities in

the motor room (temperature, humidity and cleanliness)(U_4), temperature control is that the computer room has air-conditioning, and the temperature maintained at 18°C to 24°C; humidity control is that the relative humidity remained at 40% ~ 60%, cleanliness control is that the computer room and equipment should be kept clean and health, take off shoes into the computer room, the doors and windows with closed performance; Thunder proof measures (having thunder proof devices, the good grounding)(U_5); Standby power supply and owned generators (U_6); Using UPS (U_7); Anti-static measures (adopt anti-static flooring, equipment grounding is good)(U_8); Special circuit supply electricity (separate it from air-conditioning and lighting electricity)(U_9); Guard against theft measures(U_{10}), someone is on duty, install anti-theft security doors in the import and export, install metal defended equipment to the window, equip with the radio-controlled anti-theft networking facilities in the computer room.

B Organization management and security system index

The organization management and security system index includes eight aspects: specialized information security organizations and full-time information security personnel (U_{11}), the information security organizations establishment and the information security personnel appointment, which must have an official document of the relevant units; Perfect information security management rules and regulations (U_{12}); Strict management system about information security personnel providing or transferring (U_{13}); The management system of equipment and data is comprehensive, and which is on the wall (U_{14}); The detailed the manual and the integrity work records(U_{15}); The emergency treatment plan(U_{16}); The integrity plans and system of the information security training (U_{17}); The security responsibilities is clear for all types of employee and managers, and security management system is strict(U_{18}).

C Security technology index

The security technology index includes seven aspects: disaster recovery technology countermeasures (U_{19}); Separation measures of the development work and operational work (U_{20}); Having application business and system security audit function (U_{21}); Having the system operation log (U_{22}), that is the written records of opening and shutting down the computer, and the equipment operation condition, etc.; Server backup measures (U_{23}); Anti-hacking facilities (U_{24}), including set up a firewall, have intrusion detection and other facilities; Computer virus prevention measures (U_{25}), including the software and hardware products of preventing and eliminating the virus, and then regular upgrades.

D Network and communication security index

The network and communication security index includes five aspects: having eye-catching signs where placed communications facilities (U_{26}); Backup of important communication lines and control devices (U_{27}); Encryption measures (U_{28}); Security audit tracking measures of systems running (U_{29}); Access control

measures of the networks and information systems (U₃₀), is that, divide the system user's access based on the work nature and the rank.

E Software and information security index

The software and information security index includes five aspects: access control measures of operating system and database (U₃₁); Damage prevention measures of application software and information system (U₃₂); Monitoring facility of the database and system state (U₃₃); User identification measures (U₃₄); Remote backup of system user information (U₃₅).

III. THE PRINCIPLE AND MODEL CONSTRUCTION OF SELF-ADAPTING REGRESSION SVM

SVM is a new machine learning method proposed by Vapnik based on the statistical theory learning, and is also a regression method with the good generalization ability, which likes the neural network, has the capacity of approximating any continuous limited nonlinear function, and SVM method has many advantages that the neural network has not. SVM is from the Optimal Hyper-plane under the circumstances of the linear separable. The so-called Optimal Hyper-plane, is such a separating hyper-plane, which will not only be able to classify correctly for all training sample, but make the distance (defined as the interval)largest where from the proximate point to the Hyper-plane in the training samples. The vector of the distance Optimal Separating Hyper-plane is called support vector.

A The basic principle and learning process of self-adapting regression SVM

Self-adapting support vector regression algorithm based on RBF kernel function, with regard to no sensitivity coefficient ϵ , value

$$\epsilon = 0.5D \tag{1}$$

Among them, D is variance of the noise distribution function (random process), during the iteration, the no sensitive coefficient ϵ remains unchanged. In the k-th iteration, the width coefficients of punishment factor and kernel function are recorded as $C^{(k)}$ and $\sigma^{2(k)}$, using these parameters, calling SVM learning algorithm, we can carry on regression estimation for the training sample, and then calculate the relative fitting error of the training samples:

$$E_i^{(k)} = \frac{|y_i - f(x_i)|}{y_i} \tag{2}$$

The average relative fitting error is:

$$ME^{(k)} = \frac{1}{n} \sum_{i=1}^n E_i^{(k)} \tag{3}$$

The regression relative accuracy is regard as δ , and the adjustment step of the punishment factor and width coefficient are regard as ΔC and $\Delta\sigma^2$.

Before the training, first of all, we should arrange all samples in an orderly manner, that is regarding a sample optionally as m dimension sample $x_1(x_{11}, x_{12}, \dots, x_{1m})$, choose the sample which has the smallest distance to x_1 as x_2 from the remaining samples except for x_1 , and then

choose the sample which has the smallest distance to x_2 as x_3 from the remaining samples except for x_1 and x_2 , the same token, with regard to x_i , choose the sample which has the smallest distance to x_i as x_{i+1} from the remaining samples except for x_1, x_2, \dots, x_i , at last, we can arrange all samples in an orderly manner as x_1, x_2, \dots, x_n , the sample x_{i+1} is the adjacent point of the sample x_i , and the median of the adjacent points:

$$x_i^{mid} = (x_i + 0.5(x_{i+1} - x_i)) \tag{4}$$

The distance between the sample adjacent points:

$$d_i = \|x_{i+1} - x_i\| \tag{5}$$

The distance mean is:

$$\bar{d} = \frac{1}{n-1} \sum_{i=1}^{n-1} d_i \tag{6}$$

For orderly sample set $\{x_i, y_i\}$ ($i=1, 2, \dots, n; x_i \in R^m, y_i \in R$), the regression function is $f(x)$, suppose regression accuracy is $\delta > 0$. As to the sample point x_i ($i=1, 2, \dots, n-1$), if the distance between the adjacent points is not too big, that is $d_i < \bar{d}$, and the dependent variable between the two adjacent points is not too close, that is $|y_{i+1} - y_i| > \delta$, we check the adjacent median value $f(x_i^{mid})$ of the regression function, if it isn't in between y_i and y_{i+1} , then the regression function is too complicated, which can lead to excessively fitting.

B The particular steps of self-adapting regression SVM algorithm

The particular steps of support vector regression method based on parameter self-adapting adjustment are as follows:

(1) Select the parameters initial value, the iteration initial value $k=0$, the width coefficient initial value is:

$$\sigma^{2(0)} = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^T (x_i - \bar{x}) \tag{7}$$

In the formula, n is the number of samples; \bar{x} is the mean of input vector, and then

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \tag{8}$$

The punishment factor initial value is:

$$C^{(0)} = 2(\max(y_i) - \min(y_i)) \tag{9}$$

At the same time value $\Delta C > 0$ and $\Delta\sigma^2 < 0$ by experience, proportion factor $\beta = (0, 1)$. We can carry on support vector regression evaluation using the parameters initial value, and then obtain the average fitting relative error $ME^{(0)}$.

(2) $k=k+1$, adjust parameters $C^{(k)}$ and $\sigma^{2(k)}$, carry on support vector regression evaluation based on the adjusted parameters, and then obtain the average fitting relative error $ME^{(k)}$. Parameters adjustment formulas are:

$$\begin{aligned} C^{(k+1)} &= C^{(k)} + \Delta C \\ \sigma^{2(k+1)} &= \sigma^{2(k)} + \Delta\sigma^2 \end{aligned} \tag{10}$$

(3) Check if the average fitting relative error is to be reduced, that is judging whether $ME^{(k+1)} < ME^{(k)}$ is

tenable, if tenable, then go to step (5); if not tenable, order $C^{(k+1)}=C^{(k)}$, $\sigma^{2(k+1)}=\sigma^{2(k)}$, and then go to step (4).

(4) Adjust parameter steps ΔC and $\Delta\sigma^2$, the parameter step adjustment formula are:

$$\begin{aligned} \Delta C &= \beta\Delta C \\ \Delta\sigma^2 &= \beta\Delta\sigma^2 \end{aligned} \tag{11}$$

Determine whether the step is less than the minimum level, if not, go to step (2); if so, then go to steps (7).

(5) Judge whether $ME^{(k+1)} < \delta$ is tenable, if not, go to step (2); if so, then go to step (6).

(6) Determine whether the regression function is too complicated. Check the adjacent median value $f(x_i^{mid})$ of the regression function, if the value is in between y_i and y_{i+1} , if not, then the regression function is too complicated, while the number of iterations isn't the maximum, order $C^{(k+1)}=C^{(k)}$, $\sigma^{2(k+1)}=\sigma^{2(k)}$, and then go to step (4); if the regression function isn't complicated or the number of iterations is the maximum, then go to step (7).

(7) The iteration termination, at this time the regression estimation function is the final result.

IV. APPLICATION EXAMPLES

We take the security evaluation based on network enterprises survey data in Beijing as an example, and select 12 enterprises' data as the sample, including the excellent level I, the good level II, the general level III, the poor level IV, take 1-8 samples as the training set, the four residual samples as testing samples. In order to enhance the speed and accuracy, we treat all samples data differentially. Construct a four levels SVM classification; use the RBF kernel function, the function width $\sigma=0.4$, value $C=1000$, use Matlab6.5 programming, treat the four levels SVM classification with the training samples shown in table 1, the training time is less than 0.09S. We test the trained SVM classifier through using test samples, the test results are in table 2. In order to check the method function, firstly, we evaluate the test sample based on the fuzzy neural network (FNN) method, and the result is the exact same as the self-adapting regression SVM; then design a BP artificial neural network (ANN) category, the input layer neurons is 35, which corresponding separately 35 security evaluation parameters, the output layer neurons is 4, which corresponding separately four security evaluation levels, the hidden layer node number is 30. We train and test base on the same samples and the same configuration computer, the results show that ANN classifier has a misjudgment which takes a III level security sample as II, and other classifications are correct, the ANN classification necessary training time is 1.48s, which is much higher than the SVM classifier training time.

TABLE I.
NETWORK AND INFORMATION SYSTEM SECURITY EVALUATION DATA

No	Level	U_1	U_2	U_3	U_4	U_5
----	-------	-------	-------	-------	-------	-------

1	I	0.96	0.93	0.96	0.91	0.84
2	I	0.86	0.95	0.97	0.90	0.92
3	II	0.83	0.86	0.88	0.90	0.77
4	II	0.87	0.85	0.91	0.86	0.83
5	II	0.88	0.92	0.75	0.83	0.87
6	III	0.78	0.80	0.82	0.72	0.74
7	III	0.76	0.90	0.81	0.78	0.72
8	IV	0.65	0.79	0.63	0.67	0.83

CONTINUED TABLE

No	Level	U_6	U_7	U_8	U_9	U_{10}
1	I	0.89	0.98	0.90	0.94	0.93
2	I	0.97	0.88	0.95	0.99	0.93
3	II	0.89	0.81	0.89	0.87	0.93
4	II	0.88	0.85	0.83	0.92	0.81
5	II	0.79	0.84	0.91	0.79	0.83
6	III	0.70	0.76	0.82	0.83	0.72
7	III	0.76	0.77	0.91	0.86	0.70
8	IV	0.76	0.66	0.76	0.63	0.69

CONTINUED TABLE

No	Level	U_{11}	U_{12}	U_{13}	U_{14}	U_{15}
1	I	0.86	0.92	0.98	0.98	0.90
2	I	0.90	0.93	0.94	0.90	0.93
3	II	0.79	0.92	0.86	0.86	0.84
4	II	0.87	0.91	0.84	0.86	0.86
5	II	0.88	0.79	0.91	0.84	0.84
6	III	0.74	0.71	0.90	0.76	0.78
7	III	0.75	0.77	0.73	0.74	0.73
8	IV	0.82	0.76	0.68	0.64	0.67

CONTINUED TABLE

No	Level	U_{16}	U_{17}	U_{18}	U_{19}	U_{20}
1	I	0.97	0.84	0.96	0.90	0.86
2	I	0.97	0.90	0.96	0.97	0.95
3	II	0.88	0.74	0.91	0.85	0.84
4	II	0.82	0.83	0.92	0.75	0.86
5	II	0.89	0.84	0.79	0.84	0.82
6	III	0.76	0.77	0.71	0.76	0.90

7	III	0.76	0.72	0.76	0.90	0.81
8	IV	0.69	0.82	0.76	0.62	0.65

CONTINUED TABLE

No	Level	U ₂₁	U ₂₂	U ₂₃	U ₂₄	U ₂₅
1	I	0.96	0.96	0.95	0.89	0.94
2	I	0.90	0.95	0.86	0.97	0.87
3	II	0.80	0.90	0.87	0.89	0.86
4	II	0.85	0.91	0.80	0.87	0.84
5	II	0.84	0.86	0.85	0.85	0.86
6	III	0.80	0.76	0.76	0.86	0.74
7	III	0.76	0.71	0.78	0.73	0.79
8	IV	0.73	0.78	0.69	0.78	0.64

CONTINUED TABLE

No	Level	U ₂₆	U ₂₇	U ₂₈	U ₂₉	U ₃₀
1	I	0.91	0.97	0.90	0.87	0.95
2	I	0.96	0.98	0.94	0.91	0.95
3	II	0.87	0.85	0.91	0.78	0.90
4	II	0.86	0.90	0.84	0.86	0.90
5	II	0.90	0.78	0.85	0.89	0.78
6	III	0.80	0.81	0.75	0.76	0.70
7	III	0.90	0.84	0.75	0.70	0.79
8	IV	0.78	0.65	0.68	0.80	0.78

CONTINUED TABLE

No	Level	U ₃₁	U ₃₂	U ₃₃	U ₃₄
1	I	0.93	0.95	0.96	0.91
2	I	0.88	0.94	0.99	0.92
3	II	0.87	0.86	0.84	0.92
4	II	0.85	0.85	0.91	0.83
5	II	0.87	0.91	0.73	0.88
6	III	0.75	0.81	0.80	0.73
7	III	0.78	0.92	0.83	0.72
8	IV	0.66	0.79	0.63	0.65

CONTINUED TABLE

No	Level	U ₃₅	Decision
----	-------	-----------------	----------

1	I	0.89	1
2	I	0.90	1
3	II	0.74	-1,1
4	II	0.87	-1,1
5	II	0.88	-1,1
6	III	0.75	-1,-1,1
7	III	0.72	-1,-1,1
8	IV	0.78	-1,-1,-1,1

TABLE II.
THE ACTUAL EVALUATION RESULTS COMPARED WITH THE NETWORK TRAINING RESULTS AND CLASSIFICATION

No	SVM1	SVM2	SVM3	SVM4
1	1			
2	-1	1		
3	-1	-1	1	
4	-1	-1	-1	1

CONTINUED TABLE

No	Evaluation results	FNN	ANN
1	I	I	I
2	II	II	II
3	III	III	II
4	IV	IV	IV

V. CONCLUSIONS

SVM is a common learning method based on VC dimension theory of the statistical learning theory and the structure risk minimize (SRM) principle, through the limited sample information, which can explore the best compromise between the complexity and learning ability of the model, and then, in order to receive the best outreach capacity. According to the analysis, inquire into the impact which different parameters to SVM regression function, in this paper put forward regression SVM method based on parameter self-adapting adjustment, in order to avoid select the best parameters through the complicated cross-certification steps. By the application of self-adapting regression SVM in the network and information system security evaluation, make up for the shortcomings which the sample data is few, and then overcome the defects which the traditional neural networks may converge to the local minimum points and the network structure determination can only depend on experience, enhance the generalization ability, so improve the system convergence speed and evaluation accuracy. Examples show that self-adapting SVM is an effective method to security evaluation, and then it will be good development prospects in other areas.

ACKNOWLEDGMENT

This research was supported by the Scientific Research Foundation for Young Teachers of North China Electric Power University. The item No. is 200611034.

This research was supported by the Philosophy and Social Science Research Topics of Baoding City. The item is "Research on Current Situation and Countermeasures of Agricultural Information Degree Based on New Rural Construction in Baoding City", and the item No. is 20080223.

REFERENCES

- [1] Jingwen Tian, Meijuan Gao, The research and application on artificial neural network algorithm[M], Beijing institute of technology press, Beijing, 2006,7, pp.257-259.
- [2] Weiqing Cheng, Jian Gong, Network Security Evaluation [J], Computer Engineering, February 2003, pp: 182-186.
- [3] Zhenmin GUO, Xuelong HU, Huiliang JIANG, Security Evaluation of Network and Information System and Its Index System S Research [J], modern electron technology Sep 2003, pp:9-11.
- [4] Dongmei ZHAO, Haifeng LIU, Chenguang LIU, Risk assessment of information security based on BP neural network[J], Computer Engineering and Applications, 2007, 43(1), pp: 139-141.
- [5] Yaoping Jiang, Research on Space Network Security Evaluation Index System in China [J], management world, 2005, 4, pp: 1-4.
- [6] Zhihui Wang, Fuhua Shu, The improved SVM model and its application on coal mine security evaluation system [J], Mining industry security and environmental protection, 2007,2,pp:82-84,87.
- [7] Graham Francis, Matthew Hinton, Jacky Hollo—way, et al. Best practice benchmarking: a route to competitiveness [J], Journal of Air Transport Management, 1999 (5), pp.105-112.
- [8] Jinli Xu, Application of Support Vector Machine in Water quality Evaluation [J], China countryside water conservancy and hydraulic power generation, 2007,3,pp:7-9.
- [9] KEERTHI S, CHIH J, Asymptotic behavior of support vector machines with gaussian kernel[J], Neural Computation, 2003, 15,pp: 1667-1689.
- [10] CHAPELIE O, VAPINK V N, Choosing multiple parameters for support vector machines [J], Machine Learning, 2002, 46, pp: 131-159.
- [11] Brown, M, H. G. Lewis, S. R. Gunn, Linear spectral mixture models and support vector machines for remote sensing, (submitted to) IEEE Trans, Geoscience and Remote Sensing. 2000.
- [12] Dong Xin and Zhaohui Wi, Speaker recognition using continuous density support vector machines, Electronics letters, 2001,37(17-16), pp: 1099-1101.
- [13] Bahlmann C, B. Haasdonk and H. Burkhardt, On-line handwriting recognition with support vector machinesa kernel approach, Proceedings of Eighth International Workshop on Frontiers in Handwriting Recognition, 2002, pp: 49-54.
- [14] Kim K. I, Jim and K. Jung, Recognition of facial images using support vector machine, Proceedings of the 11th IEEE Signal Processing Workshop on Statistical Signal Processing, Singapore, 2001, pp: 468-471.
- [15] Junfeng Gao, Wengang Shi, Jianxun Tan et al, Support vector machines based approach for fault diagnosis of valves in reciprocating pumps, Proceedings of IEEE Conference on Electrical and Computer Engineering, Canadian, 2002,3, pp: 1662-1627.

Shaomei Yang, was born in Handan City, China, and graduated from the agricultural university of Hebei in 2003, gained the master's degree of management. The author's major field of study is the business management.

Since 2003, she is always working at the North China Electric Power University, Baoding City, China. And she has published more than 12 papers and 1 book. Such as the Electric Power Enterprise Management (Beijing: Chinese Electric Power Publishing Company).

Donglei Liu, was born in Baoding City, China, and graduated from the agricultural university of Hebei in 2002, gained the master's degree of management. The author's major field of study is the management.

Since 2002, she is always working at the agricultural university of Hebei, Baoding City, China. And she has published more than 10 papers and 3 books.

Zhibin Liu, was born in Luannan County, China, and graduated from the agricultural university of Hebei in 2004, gained the master's degree of management. The author's major field of study is the information management.

Since 2004, he is always working at the North China Electric Power University, Baoding City, China. And he has published more than 20 papers and 3 books. Such as the Research on the Advancing Front Topic of Asset Valuation (Beijing: Chinese Finance and Economical Publishing Company).