Modeling and Analysis of Network Security Situation Prediction Based on Covariance Likelihood Neural

Chenghua Tang[1], Xin Wang[1], Reixia Zhang[1], and Yi Xie[2]

[1] School of Computer Science and Engineering, Guilin University of Electronic Technology, Guilin 541004, China
{tch,wxin,rxzhang}@guet.edu.cn
[2] Department of Information Science and Technology, Sun Yat-Sen University, Guangzhou 510275, China
Xieyi5@mail.sysu.edu.cn

Abstract. Security situation is the premise of network security warning. For lack of self-learning on situation data processing in existing complex network, a modeling and analysis of network security situation prediction based on covariance likelihood neural is presented. With the introduction of the error covariance likelihood function, and considering the impact of sample noise, the network security situation prediction model using the situation sequences as input sequences, and in the back-propagation to achieve the parameters adjustment. Results show that the model can take advantage of the relationship characteristics between the complexity and efficiency in complex neural networks, and the method has good performance of situation prediction.

Keywords: network security, situation prediction, covariance, neural.

1 Introduction

Most of the current situation prediction techniques are a part of situation assessment, which make reference on whether to the early warning after the situation evaluation. Main methods used in Multi-sensor data fusion [1], grey correlation analysis [2], AHP [3], etc., these methods rely on specialists than those given the initial elements of the security situation beginning weights, required during operation of man-made changes to the weights, some algorithm do not have the self-learning ability. Because of the uncertainty, ambiguity and variability characteristics, of the attack information, situation prediction involves computer science, military strategy, political science, and other disciplines, its importance has been related to people's lives and national security. After "9.11"incident, the European Union accelerated the implementation pace of establishment of electronic information security program, requiring strict inspection early warning and emergency response capacity of information network infrastructure and network system. British Institute of King's College London researched the information warfare threat assessment and early warning of attacks, and proposed the decision-making intelligent early warning system [4]. Kijewski studied for early warning and attack of a prototype system identification framework [5]. NetSA Working Group developed the SILK to carry out large-scale real-time

D.-S. Huang et al. (Eds.): ICIC 2011, LNBI 6840, pp. 71–78, 2012.

monitoring of network security situation, the potentially malicious network behavior before they become unable to control the identification, response and early warning [6]. Honeywell Laboratories proposed using the theory of plan recognition for intrusion prediction that the behavior from the observed sequence of reasoning, called the plan recognition or task tracking in the field of artificial intelligence [7]. Honeynet Project proposed to predict the invasion of hacker's intent and possible future theoretical methods using moving average models and other statistical theory [8]. U.S. Department of Defense Advanced Research also heavily subsidized project development such as the purpose of hacking prediction, but due to various reasons, the progress of their projects have not received specific information. In China, Hu Huaping proposed for large-scale network intrusion detection and early warning system architecture [9]. Hu Wei proposed an improved model for prediction of Grey Verhulst [10]. For periodic attacks, Zhang Feng proposed method of network security warning based on intrusion events [11]. An Xifeng analyzed the characteristics of security incidents, security incidents to establish a distributed network of early warning model [12].

These results are mainly stay in the theoretical framework, the traditional situation prediction methods rely on experts to give too much weight, and lack of self-learning. This paper presents a security situation prediction method based on covariance likelihood neural, introduced the concept of state sequences, and the back propagation process to achieve the right value for the specified parameter self-learning adjustment. In order to avoid the situation of all elements of algorithms to interpolate the sample points, abandoning the traditional minimum variance error function, the introduction of maximum likelihood estimation, re-defined error function, considered the impact of sample covariance and noise on the network training, and from the global Point of the network to achieve unsupervised learning.

2 Principle of Situation Prediction Control

Network security situation prediction model based on the past situation input and situation output predicts the future trend of the situation. The past input and output values as the neural network training signal, namely, the scalar adjustable parameters of the transfer function, and output the result, that is, the future of security situation. The basic structure of network security situation prediction mode based on neural network is shown in Fig. 1.

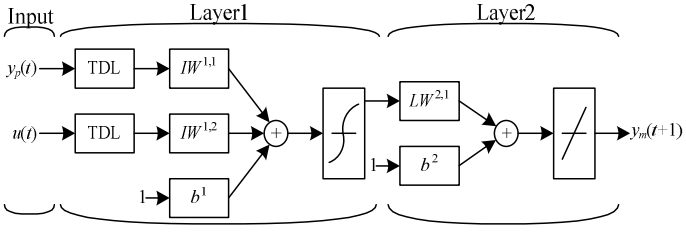

Fig. 1. Network security situation prediction model

The curve (Layer1) and linear (Layer2) transfer function in the basic structure model, respectively, for the network back propagation and linear regression. The prediction model can be trained offline, and the training objects can be a large number of off-line historical security situation data.

Considering a *M* layers feed forward network, the number of neurons of each layer is N_m, (m=1,2,…,M), then the basic network equation can be expressed as:

$$y_i^m = f((W_i^m)^T Y^{m-1}) \tag{1}$$

Where i=1,2,…,N_m, represents the serial number of neurons m layer, y_i^m indicates that the network's first *i-m* layer of output neurons, $f(.)$ indicates the S-type nonlinear function $f(x) = 1/(1+e^{-x})$, $(W_i^m)^T$ is the vector transpose composed of connecting weights about the first *i-m* layer of neurons with the first *m-1* layer all the neurons, and Y^{m-1} is the first *m-1* layer composed of all the output neurons in the vector.

Let the network output is $\hat{y}_i^m = f(X_i, W)$, real output is $\overline{y}_i^m = f(X_i)$, and sample output is $y_i^m = f(X_i) + \xi_i$, then the output layer neuron fitting error is $e_i = y_i^m - \hat{y}_i^m$, and the real bias is $\overline{e} = \overline{y}_i^m - \hat{y}_i^m$, so according to the traditional neural algorithm to obtain the minimum variance method LS type network error function is the following function:

$$E_{LS} = \sum_{i=1}^{N} \phi(e_i) = \sum_{i=1}^{N} e_i^2 \tag{2}$$

Successful application of system identification using neural algorithm depends on the quality of samples, because the training algorithm does not consider the training samples error, using the least square error to get the best value of function E_{LS} only when the deviation submit to the Gaussian distribution. E_{LS} guides the learning process, and the ultimate goal is to make all the opportunity to sample the fitting error tends to zero balance, that approximate the network output for each sample by the noise pollution output, rather than real output, so in the sample case with noise, it interpolates all the training samples in the identification model, rather than approaching the real object model, which will lead to the practical application of the more iterations, the smaller the training error, and generalization capability is worse, but significantly lower efficiency.

3 Neural with Covariance Likelihood

Learning sample of data set (X_i ,Y_i^m), (i=1,2,…, N_m), given a set of weights W, because of the existence of the error, the conditional probability density of the network output vector $\hat{y}_i^m$ relative to the weight W is:

$$l(W) = P(\hat{y}_i^m | W) \tag{3}$$

If the sample is independent between the statistics, the continuous application of the Bayesian formula to get all the joint probability density of the sample, the sample likelihood function is:

$$L(W) = \mathrm{P}(\hat{y}_1^m, \hat{y}_2^m, \ldots, \hat{y}_{Nm}^m | \mathrm{W}) = \prod_{i=1}^{Nm} p(\hat{y}_i^m | W) \tag{4}$$

If the network model is correct, then all the differences between the samples output and the actual output mainly derive from the samples noise, therefore the fitting error must to be considered, then assume the sample output error obeys the Gaussian distribution, the density function of the network output vector with respect to weight W Conditions and Gaussian distribution parameters:

$$P(\hat{y}_i^m | W, \mu, \sigma) = (2\pi | Cy_i |)^{-1/2} \exp\{-\frac{(y_i^m - \hat{y}_i^m)^T (y_i^m - \hat{y}_i^m)}{2Cy_i}\} \tag{5}$$

Where $Cy_i = \mathrm{cov}[\hat{y}_i^m | y_{i-1}^m]$, is the output condition of the sample covariance matrix reflects the error structure of the sample output, and fitting deviation (y_i^m - $\hat{y}_i^m$) is the mean of sample error.

According to maximum likelihood theory, the parameters μ , σ determining conditions, so that the W * for the largest value of likelihood function $L(W)$ is the most optimal estimation of W.

$$L(W^*, \mu, \sigma) = \max L(W, \mu, \sigma) \tag{6}$$

Apparently equivalent to the following error function is minimized:

$$E(W, \mu, \sigma) = -2\ln L(W, \mu, \sigma) \tag{7}$$

Therefore, the new network error function is:

$$E(W, \mu, \sigma) = \sum_{i=1}^{Nm} \frac{(y_i^m - \hat{y}_i^m)^T (y_i^m - \hat{y}_i^m)}{Cy_i} + 2\sum_{i=1}^{Nm} \ln | Cy_i | + 2N_m \ln \sqrt{2\pi} \tag{8}$$

Constant term omitted, the final error function is:

$$E(W, \mu, \sigma) = \sum_{i=1}^{Nm} \frac{(y_i^m - \hat{y}_i^m)^T (y_i^m - \hat{y}_i^m)}{Cy_i} \tag{9}$$

According to this building method the network is called covariance likelihood neural in this paper.

Compared with the conventional error function, where taking the sample covariance into account. In addition, when $L(W^*)$ to take the maximum, that is, the deviation of training samples tends to the expected error in the center of the surface, the learning intensity maximum deviation tends to zero faster, its resistance to errors and noise samples interference.

4 Security Situation Prediction Model

According to the service, host and network security system in the threat situation provided by the target network, the situation assessment model can be established. Because of the space limitations, we only assess the security situation on the service as example.

Definition 1. The function F_S said the security situation in the target network service status, denoted as:

$$F_S(S,C,N,D,t) = N(t) \cdot 10^{D(t)} \tag{10}$$

Where S represents a service the target network provided; C indicates the type of service attacks on the; N is said that services by the number of attacks; D is said that the severity of the attack; $N(t)$ is said that the severity of attacks in t time; $D(t)$ is said that the number of attacks occurred in t time.

Definition 2. The function of F_H said the security situation in the host status of the network, denoted as:

$$F_H(H,V,F_S,t) = V \cdot F_S(t) \tag{11}$$

Where H represents the target hosts on the network; V indicates that the service's weight of all opened services.

Definition 3. Assuming in t (as small as possible) time period, select a state sequence from the state database, as the future network security situation prediction model of the input sequence, denoted by $X^{(0)} = (x^{(0)}_1, x^{(0)}_2, ..., x^{(0)}_n)$, where $x^{(0)}_t \geq 0$, t = 1,2, ..., n; $X^{(1)}$ is $X^{(0)}$ of 1-AGO sequence, denoted as $X^{(1)} = (x^{(1)}_1, x^{(1)}_2, ..., x^{(1)}_n)$, where:

$$x^{(1)}t = \sum_{i=1}^{t} x^{(0)}i \text{ , } t=1,2,\ldots,n. \tag{12}$$

By definition 1 and definition 3, the services available for the target network security situation prediction function model:

$$F_S(t+(n+1)) = f(\sum_{i=1}^{n} F_S(t+i)) \tag{13}$$

Thus, parameters $F_S(t+(n+1))$ and $\sum_{i=1}^{n} F_S(t+i)$ can be recognized the existence of a nonlinear relationship.

According to Kolmogorov mapping existence theorem of multi-layer neural network, the nonlinear mapping relationship can be three layers feed-forward artificial neural network approximation achieved. Then the known time situation series act as network input, and the network output is to identify the situation prediction. This paper requires $F_S(t+(n+1))$ and time $(t+(n+1))$ within the range, by definition given in 1 and 2, network services, host situation assessment methods to obtain network training samples required for the application covariance likelihood neural algorithm to train the network, follow these steps:

1. According to the history and current situation of network security information data, the service and the host of multi-input single-output prediction of artificial neural network model and corresponding network error function $E(W)$ are established:

$$P(\hat{y}_i^m \mid W) = (2\pi \mid Cy_i \mid)^{-1/2} \exp\{-\frac{(y_i^m - \hat{y}_i^m)^T (y_i^m - \hat{y}_i^m)}{2Cy_i}\} \tag{14}$$

$$E(W) = \sum_{i=1}^{Nm} \frac{(y_i^m - \hat{y}_i^m)^T (y_i^m - \hat{y}_i^m)}{Cy_i} \tag{15}$$

Where $\hat{y}_i^m$ and y_i^m, respectively, for the actual output and expected output of the m layer i neurons, corresponding to the situation prediction value; for every single point of communication during the process input parameters, where W can represent the situation assessment model Severity of attacks D, the importance of services V parameters.

2. Training the neural network for the fit error (y_i^m - $\hat{y}_i^m$) tending to zero, the weights of self-learning adjustment, to find the optimal parameters, and output the prediction model finally. The network output is what we want recognition the next moment on the service or host situation prediction values. In the real world, real-time training the network, the fitting curve of situation prediction can be available.

5 Experiments and Results

Purpose of the experiment is to verify the effective and reasonable of the security situation prediction model.

Experimental environment is configured with Ubuntu 10.04 LTS / Inter Core 2 Duo E7200/2G/250G host, SUN, IBM and other large servers, and multi-layer routers, Gigabit switches, IDS, firewall, and fiber optic cable to construct more complex Network. Using Domain 3.5, Namp3.5 and Trinity V3 to attack a protected server, collect IDS, firewall and system log information, and assess the service or host security situation every 10 hours. The value of each assessment with the previous values are to be established the time series, input into the covariance likelihood neural, finally, output the situation prediction value each moment in turn, and calculate the actual situation value in the next time point for comparison.

Firstly, Experiment gets the number of attacks on the ftp, telnet, rpc, dns, socks, www, etc., and then assesses each service's security situation by definition 1. Trial of 60 consecutive made the security situation in the value of rpc services, as with the previous 45 training samples, and after 15 testing samples, after pretreatment, input into covariance likelihood neural to train. Pre-set training speed factor = 0.6, target error goal = 0.0001, after running about 32.116s, achieve the target error for the 6107 iterations. At the same time using traditional neural to test security situation prediction for comparative experiments, found that after iteration 10989 times the error has not reached the goal, and time has been occupied 149.973 seconds. Fig. 2 shows the situation prediction results about rpc service security situation based on covariance likelihood neural.

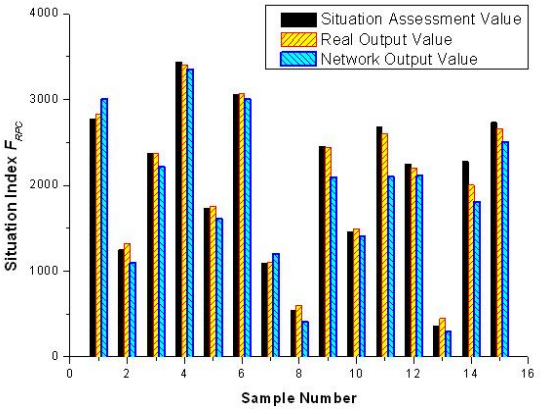

Fig. 2. Prediction of rpc service security situation

Finally, considering the various services on the server, by definition 2, evaluate the value of the server host's security situation, the same method draw neural prediction on the host in Fig. 3.The figure shows that the value of the neural with covariance likelihood prediction can better approximate the true assessed value, with a good prediction effect.

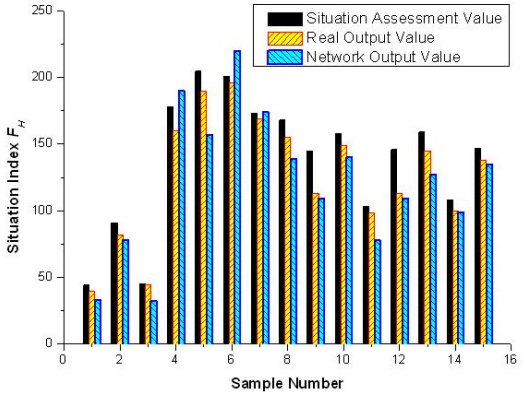

Fig. 3. Prediction of host security situation

6 Conclusions

This paper introduced neural and studied its improvement to establish a security situation prediction model based on covariance likelihood neural. The traditional error function is replaced by the maximum likelihood error function. The impact of sample covariance and noise on the network training is considered. The situation sequences established through the situation assessment model are used as the training input

sequences, and the self-learning adjustment of the appointed parameters' values is implemented in the process of back propagation training. The new method can make full use of the characteristics of the network more complex, finer grain size, the higher the efficiency, and results show it can effectively predict security situation and provides an effective way of network security strategic early warning.

Acknowledgments. This work was supported by the National Natural Science Foundation of China under Grant No.60970146, the China Postdoctoral Science Foundation under Grant No.20070420793, and the Department of Education research project in Guangxi, P.R. China under Grant No.201012MS088. The helpful comments from anonymous reviewers are also gratefully acknowledged.

References

1. Onwubiko, C.: Functional requirements of Situational Awareness in Computer Network Security. In: Proc of the IEEE International Conference on Intelligence and Security Informatics (ISI 2009), Dallas, Texas, USA (2009)
2. Zhao, G.S., Wang, H.Q., Wang, J.: Study on Situation Evaluation for Network Survivability Based on Grey Relation Analysis. Journal of Chinese Computer Systems 27(10), 1861–1864 (2006)
3. Chen, X.Z., Zheng, Q.H., Guan, X.H.: Quantitative Hierarchical Threat Evaluation Model for Network Security. Journal of Software 17(4), 885–897 (2006)
4. US Infrastructure Assurance Strategic Roadmaps. Strategies for Preserving Our National Security. Sandia National Laborato-ries, Sand Report, 98-1496 (1998)
5. Kijewski, P.: ARAKIS-An early warning and attack identification system. In: Proc. of the 16th Annual First Conference, Dudapest, Hungary (2004)
6. Carrie, G., Michael, C., Michael, D.: More Netflow Tools: for Performance and Security. In: Proc. of the 18th Large Installation Systems Administration Conference (LISA 2004), Atlanta, GA, USA (2004)
7. Christopher, W.G., Goldman, R.P.: Honeywell Labs. Plan Recognition in Intrusion Detection Systems (2001)
8. Das, S., Lawless, D.: Trustworthy Situation Assessment via Belief Networks. In: Proc. of the 5th International Conference on Information Fusion, USA (2002)
9. Hu, H., Zhang, Y., Chen, H.T., Xuan, L., Sun, P.: The Study of Large Scale Networks Intrusion Detection and Warning System. Journal of National University of Defense Technology 25(1), 21–25 (2003)
10. Hu, W., Li, J.H., Chen, X.Z.: Network Security Situation Prediction Based on Improved Adaptive Grey Verhulst Model. Journal of Shanghai Jiaotong University (Science) 15(4), 408–413 (2010)
11. Zhang, F., Qin, Z.g., Liu, J.d.: Intrusion Event Based Early Warning Method for Network Securiyt. Computer Science 31(11), 79–81, 131 (2004)
12. An, X.f., Li, W.H., Liu, Z.: Research on early-alert, orientation and rapid isolation control system for large-scale networks. Computer Engineering and Design 29(8), 78–81 (2008)