Network Security Situation Estimation based on Big Data

Xuan Huang
[1]School of Software and Internet of things Engineering, Jiangxi University of Finance and Economics, Nanchang, 330013, China
[2]Management science and engineering, Faculty of Management, Nanchang University, Nanchang, 330031, China

Abstract: The traditional method uses the terminal network monitoring method to carry out network security estimation. due to the strong power attenuation of the terminal of the network communication channel, the accuracy of the security situation estimation is not high and the performance detection effect is poor. Therefore, a network security estimation and situation prediction algorithm based on adaptive data classification and virus infection membership feature extraction is proposed in this paper. The network security estimation model in large data environment is constructed. The adaptive data classification algorithm is used to cluster and evaluate network attack information data, extract the infection membership feature of network attack virus data, and realize network security situation prediction and virus attack. Simulation results show that the algorithm has higher prediction accuracy for virus data flow, can effectively realize network virus flow prediction and data detection in different scenarios, and improves the network's ability to resist virus attacks in large data environments.

Keywords: Big data; Network security; Antiviral ability

1. Introduction

As an important part of network security situation technology, security situation assessment plays a key role in the final decision - making. At present, the research on security situation assessment is relatively hot, but it is not yet mature, and the security situation prediction is still in its infancy. At present, the few methods of situation prediction are mainly neural networks, because neural networks, especially radial basis function neural networks, have good approximation performance and the advantage of processing nonlinear data. Most of them are based on the idea of off-line learning, and the learning process is long and the accuracy is limited, so the prediction effect of large-scale network security situation is not ideal[1].

In view of the above problems, this paper studies the prediction of network security situation. Firstly, starting with the learning algorithm of neural network, through the analysis of related theories and algorithms, the concept of the importance of hidden neurons is introduced, and a lifelong learning algorithm is proposed[2]. The basic idea of life-long learning algorithm is to run the learning process of neural network through the whole use phase of the network and make the network continue to learn while carrying out actual prediction work. The RBF neural network for lifelong learning can be quickly constructed by using the lifelong learning algorithm. the network can learn and adjust itself during operation, thus maintaining the timeliness of the network. In the experiment, lifelong learning RBF neural network is used to predict the alarm times. The results show that the prediction accuracy is continuously improved with the continuous adjustment of the network.

2. Network Security Evaluation Algorithm based on Big Data

Cloud computing uses centralized data management, which is extremely open and complex. in cloud computing network applications, the security risks it faces mainly include identity authentication risk, network layer risk, host layer risk and application risk[3]. The association degree algorithm of cloud computing network security evaluation is based on grey system theory. firstly, the network behavior characteristic quantity is set up as S_a From which factor behavior characteristic quantities are obtained $S_i - S_0$ The observation data on sequence t is defined as $S_0(t)$, $(k = 1,2,...,n)$ The network characteristic behavior sequence is obtained $S_0(t) = S_0(1), S_0(2), ...S_0(n)$. Thus, the network factor behavior sequence can be calculated. $S_i(t) = [S_i(1), S_i(2), \ldots, S_i(n)]$. Which t can be serial numbers of time, index, object, etc[4]. From this, it can be concluded that the correlation coefficient of sequence S_i and S_0 at point k is defined as:

$$\delta_{0i}(t) = \frac{\min_i \min_t |S_0(t) - S_i(t)| + \mu \max_i \max_t |S_0(t) + S_i(t)|}{|S_0(t) + S_i(t)| + \mu \max_i \max_t |S_0(t) - S_i(t)|} \quad (1)$$

According to the above formula is:

$$\delta(S_0, S_i) = \frac{1}{n}\sum_{i=1}^{t} 10\mu_{0i}(t), \mu \in (0,1) \quad (2)$$

Combined with the above formula, it can be obtained that during the t period, the network security evaluation index is :

$$Rs_i(t) = \delta(S_0, S_i)\sum_{i=1}^{t} 10^{\mu_{0i}} v \quad (3)$$

Where R is the attack evaluation value generated by attack δ on service μ, and v is service within time t. The number of types of supplies suffered is several attacks a on service s; The degree of harm caused depends on the type of attack R belongs to S. According to the characteristics of cloud computing network, build a network security evaluation model based on cloud computing as shown in the figure below:

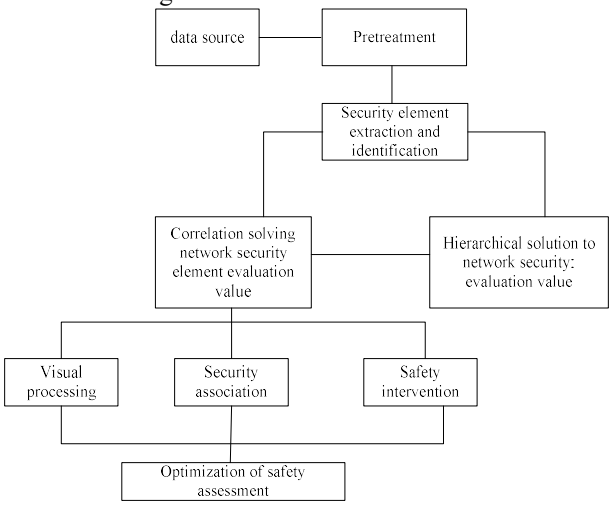


Figure 1. Network security evaluation model

As shown in the figure, cloud computing is used to collect the security audit data, intrusion detection data and network security situation in the network, preprocess the data, unify the data format, extract valuable network security information, and prepare for network security assessment[5-6]. In the evaluation of network security, the association analysis method is used to calculate the threat value of known network attack types and determine the degree of attack damage. Thirdly, the security evaluation value of the network system is calculated hierarchically, and the network security chart is drawn by using the visualization module to facilitate analysis of the network security situation[7].

3. Analysis Method of Network Security Situation Based on Big Data

The function that network security situation technology can tell what kind of danger may occur is reflected through network security situation assessment[8]. Network security situation assessment is to pre-process the original network events, extract the information that has certain correlation and reflects the characteristics of some network security events, and use certain mathematical models and prior knowledge to give a reliable evaluation probability value for reference whether some security events really occur. The network security situation assessment is not a simple process, it contains many concrete steps and is hierarchical[9]. The following figure is a three-level model of network security perception.

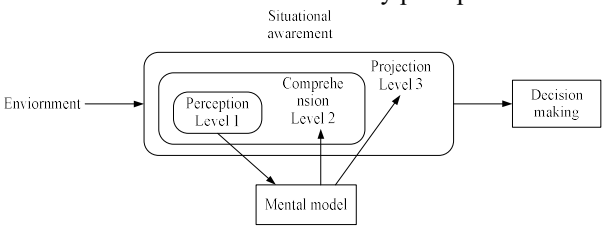


Figure 2. Network security awareness system

There have been different viewpoints on the functional model of data fusion in history, but it was proposed by the data fusion group. as shown in the figure of the model for data fusion results after several modifications, it is being adopted by more and more actual systems[10]. The purpose of constructing data fusion model is to promote better communication and understanding among system managers, theoretical researchers, designers, and evaluators, so that the whole system design, development, and implementation process can be carried out efficiently and smoothly.

The calculation of network security situation value is real - time, while threat assessment is a comprehensive summary and evaluation of network security situation over a period of time. Threat assessment combines the situation of recent security incidents with the mining of historical data to judge the current threat situation of the network. Through the network security threat assessment, the false alarm that may exist in the network security situation value can be eliminated to a certain extent, which can help the continuous correction and improvement of the security situation value algorithm. From the above analysis, we can see the composition of the network security situation awareness system as shown in the Figure.

4. Analysis of Experimental Results

The calculation of situation value in simulation directly uses the statistical value of the number of alarms generated by the network in the data set as its current situation value, because the network security situation value is calculated layer by layer based on the statistical value multiplied by its corresponding weight value. Therefore, as long as the weights are set uniformly, the alarm statistics value can represent the real situation of the network security situation value. According to the statistical analysis of network attacks on the simulation data set, it cor-

responds to the time taken for a complete network attack, that is, the attack period is approximately three days. Therefore, the parameter detection results of the neural network security situation prediction model are as follows.

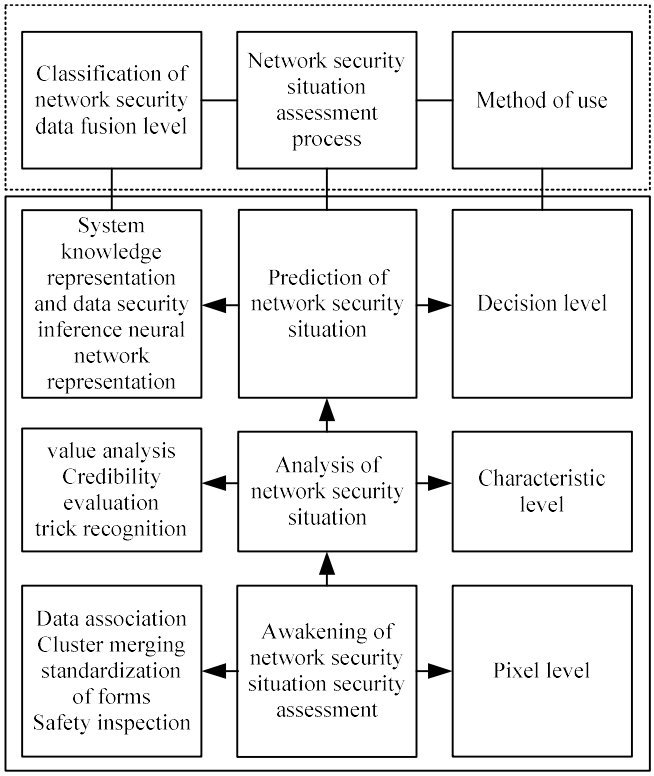


Figure 3. Network security situation assessment system module based on big data

In order to test the superiority of the model in this paper, the current classic network security situation estimation model l is selected for comparative experiments. the root mean square absolute error (RMSE) and average relative error of their fitting results and prediction results are shown in the table.

Table 1. Comparison of Absolute Error Data in Network Security Assessment and Detection

Actual Value	Predicted Value	Error Value
0.24	0.22	0.02
0.45	0.45	0.00
0.41	0.40	0.01
1.02	1.00	0.02
1.45	1.45	0.00
0.84	0.86	0.02
0.54	0.55	0.01

As can be seen from table 1, of all the network security situation estimation models, the fitting and prediction errors of this model have decreased, which improves the estimation accuracy of the network security situation and verifies the feasibility and superiority of this model.

For a network security situation estimation model, the fitting performance can only describe the fitting effect of the model. the most important thing is its generalization ability. the model predicts the test sample set and the results are shown in the figure.

The value measured before the repair began after the sa failure occurred. Ratio of su failures that are not available before repair begins; Si Ratio of x after t1 repair;T0 when the fault occurred; Ti the time to repair the fault; Tr time after all repairs are completed As can be clearly seen from the figure $Rt = (t_1 - t_0)Sa + Su = 1$,and $Sa < Sa + Sr \le 1$, as if $Sa + Sr = 1$, Then the system has 100 % service recovery capability. As can be seen from the above figure, compared with the fitting results of the network security situation, the deviation of the network security situation prediction in this paper increases greatly, but the deviation between the network security situation prediction value and the real network security situation value is small, and the change trend of the two is consistent. The experimental results show that this model is a network security situation estimation model with strong generalization ability and high prediction accuracy.

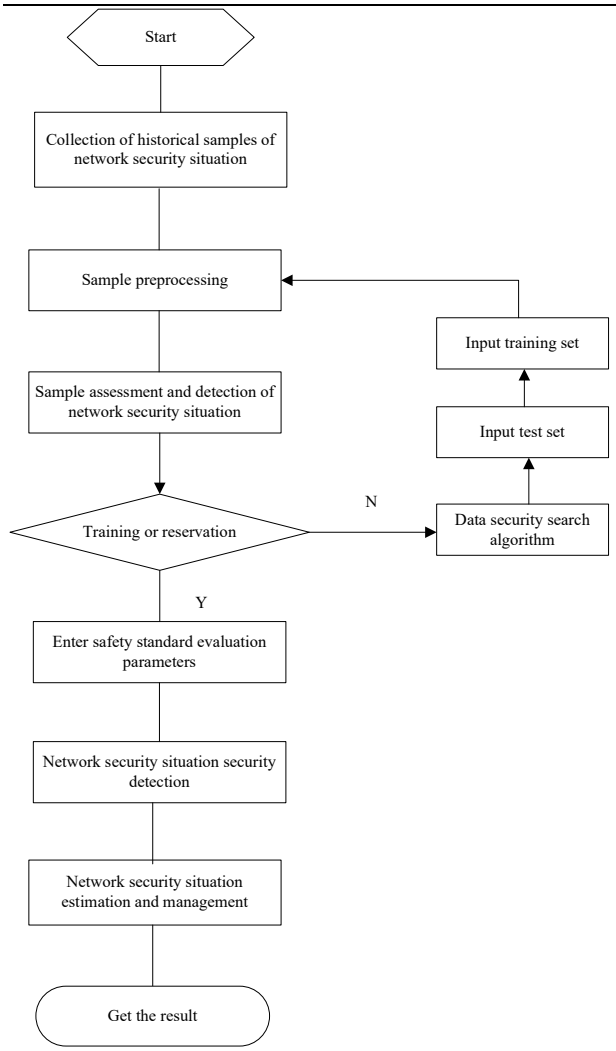
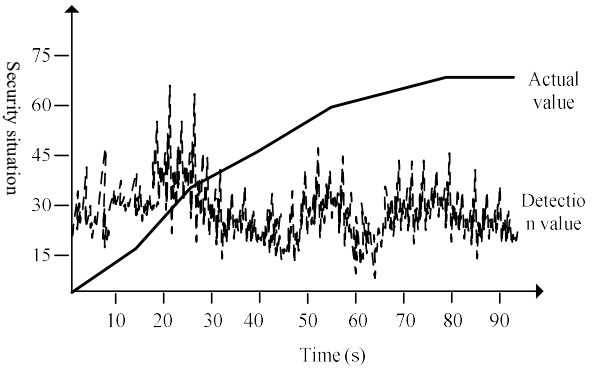
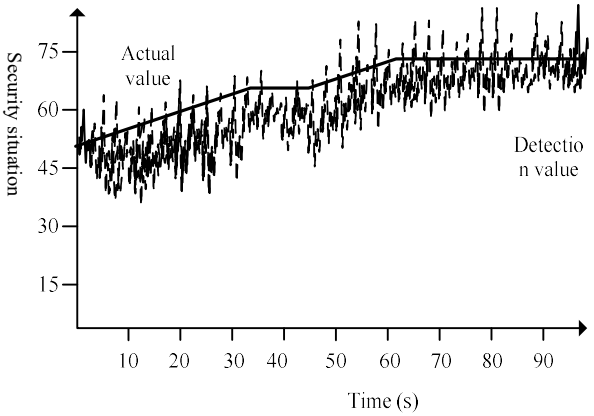


Figure 4. Network security situation estimation process



Results of traditional methods



The results of this method are as follows

Figure 5. Comparison and detection results of network security situation assessment

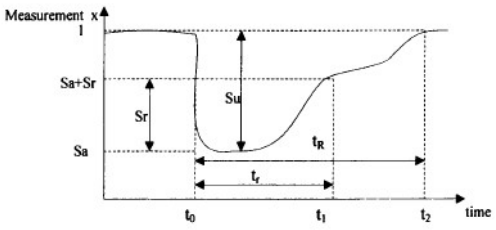


Figure 6. N etwork security attribute detection

5. Concluding Remarks

The network security assessment uses the security element extraction model of cloud computing architecture to carry out gray correlation analysis on the known network attack types, and calculates the network security assessment values by layers, and graphically presents the network security situation by using the calculated results, thus the network security can be intuitively analyzed and predicted. Through the research, we can carry out a preliminary security assessment and prediction analysis of network attacks during cloud computing operations, so as to grasp the security status of the entire network system in real time, and to take timely countermeasures to prevent further expansion of the harm in the event of a security threat to the network system.

6. Acknowledgments

The authors would like to thank all the referees for their constructive and insightful comments that allowed us to improve this paper. The contributions of this research are the following: The contributions of this research are the following: Primary Research & Development Plan of Jiangxi Province (No. 20181ACE50033); the Jiangxi Province Social Science Planning Project (No.18YJ28);

Jiangxi Province Science and Technology Innovation Platform Project (No. 20181BCD40005).

References

[1] Wang X., Zhou X. M. Research and simulation on reasonable distribution of big data in cloud computing environment [J]. Computer Simulation. 2016, 33 (3), 292 - 295.

[2] Guang H., Wen W., Xing H., et al. Complex Cyber-Physical networks: from cybersecurity to Security Control[J]. Journal of Systems Science & Complexity. 2017, 30(1), 46-67.

[3] Vladimir, Kirill, Kravtsov. Principles of network security protocols based on dynamic address space randomization[J]. Communications and Computers: Chinese and English Versions. 2016(2), 77-89.

[4] Guan Y. H., Qiao P. L. High dimensional differential evolutionary algorithm based on cloud population for network security prediction[J]. Journal of Jilin University. 2016, 46(2), 568-577.

[5] Zhang J. Biodiversity science and macroecology in the era of big data[J]. Biodiversity Science. 2017, 25(4), 355-363.

[6] Cao Z. New trends of information security—how to change people's life style?[J]. Science China. 2016, 59(5), 1-3.

[7] Benke K. K. Uncertainties in big data when using internet surveillance tools and social media for determining patterns in disease incidence[J]. Jama Ophthalmol. 2017, 135(4), 402.

[8] Mouza C. D., AS-Index: A structure for string search using n-grams and algebraic signatures abstract[J]. Journal of Computer Science & Technology. 2016, 31(1), 147-166.

[9] Wu B. F., Feng G., He G. J., et al. Big data on global changes: Data sharing platform and recognition[J]. Journal of Remote Sensing. 2016, 20(6), 79-98.

[10] Zhang X. G., Jiang R., Wang X. W., et al. From big biological data to big discovery: The past decade and the future[J]. Chinese Science Bulletin, 2016, 61(36), 3869-3877.