Research on Risk Assessment of Information System Based on Fuzzy Neural Network

Guangliang ZHU,Yuanbao WANG
(Xi'an Peihua University, P. R. China)
guangliangzhu@163.com

Abstract：In this paper, a risk assessment model based on fuzzy comprehensive evaluation and neural network is constructed, which is applied to information system risk assessment. Firstly, the risk of information system is established. Secondly, the fuzzy comprehensive evaluation and neural network algorithm are analyzed, and the risk assessment method based on fuzzy neural network is established. Finally, an example is given to verify the feasibility and scientific nature of the method.
Keywords：Fuzzy comprehensive evaluation; Neural network; Risk assessment

1 Introduction

With the development of information technology, different countries are now paying more and more attention to information security risk assessment. In view of information security and risk assessment research process, some European countries and the United States comparatively go further. With the increasing understanding of concept and technology of information security, countries pay much attention to information system risk assessment. Risk assessment methods can be divided into qualitative assessment, quantitative assessment, knowledge-based assessment and model-based assessment, manual and tool-assisted assessments, static and dynamic assessments [1-3]. Information network security risk assessment is complex, nonlinear, uncertain, and of strong timing, while the traditional mathematical model for risk assessment has greater subjective randomness and fuzziness. This problem can be well solved by using neural network methods with intelligent characteristics and ability to deal with uncertain problems. When no reference case can be obtain from the rule base, the evaluation of the unknown problems can be well solved by the neural network. However, the neural network does not have the ability to deal with the qualitative index, so it is necessary to input the quantitative data needed by the system, and the fuzzy evaluation method can be used to quantify all kinds of uncertain risk factors. In this way, the output of the fuzzy system can be treated as a neural network in order to solve the input problem of neural network. By combining fuzzy evaluation method with neural network, the problem of quantitative assessment in security risk can be solved better.

2 Quantitative evaluation model

Being the most mature and widely used artificial neural network, Back Propagation (BP) [4-5] can simulate any nonlinear input-output relation, approximate any continuous function and realize nonlinear mapping. With the advantages of simple structure and strong maneuverability, BP is suitable for risk analysis and evaluation of complex information network system. BP neural network is composed of several neurons, in which one node represents a neuron, and each layer consists of several nodes. The input layer, a number of hidden layers, and an output layer, which all are consisted of nodes, make the BP neural network, in which only the adjacent layer nodes have connections, while the same layer or cross-layer nodes are not connected to each other.

Because of the complexity of information system risk, there are many uncertain factors in risk analysis of information system, and these factors may be related to each other. These influencing factors may lead to nonlinear and dynamic changes to the system risk. BP neural network has the characteristics of clear thinking, rigorous structure, strong maneuverability, etc. The introduction of hidden nodes enables a three-layer nonlinear network with Sigmoid neurons to approach any continuous function with arbitrary accuracy. Using neural network to establish a risk analysis and evaluation model of information system can solve some problems in risk assessment. Based on the layer structure of information system risk, the BP neural network is designed. The risk influencing factors of the information system are processed by fuzzy quantification, and the result is regarded as the input quantity of the neural network input layer, and the output system risk assessment value is calculated by the output layer by the learning algorithm. Its model structure is shown in

Copyright © 2019, the Authors. Published by Atlantis Press.
This is an open access article under the CC BY-NC license (http://creativecommons.org/licenses/by-nc/4.0/).

figure 1.

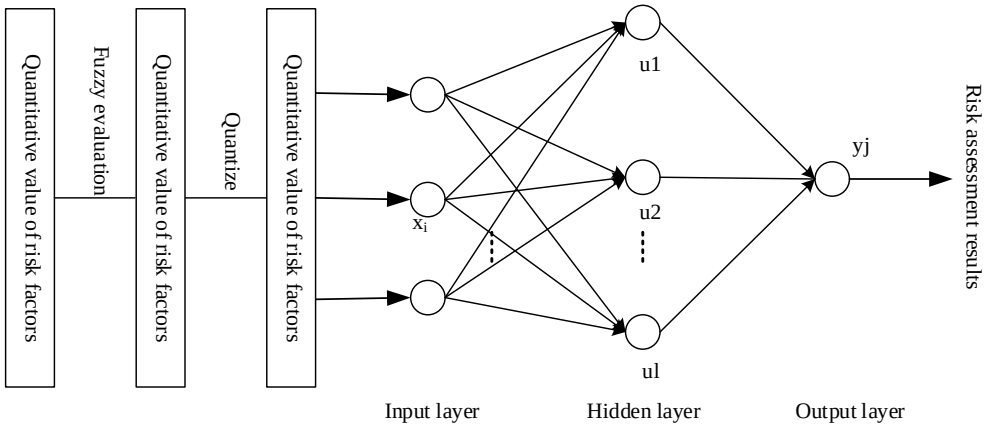

Fig. 1 Risk quantitative evaluation model of information system based on fuzzy neural network

In information systems, in addition to quantitative factors that can be quantitatively evaluated by numerical values there are some other influencing factors which cannot be directly expressed by concrete number, or its size or severity after occurrence cannot be determined by accumulative data. These factors are qualitative factors, which are usually expressed by natural language with fuzzy characteristics. The fuzzy mathematics is usually used to study the problems with fuzzy uncertainty, while the qualitative influencing factors in information system are usually analyzed quantitatively by using fuzzy evaluation theory.

3 Information system risk Assessment based on Fuzzy Neural Network

3.1 Establishment of information systems risks

Information system risks can be expressed by probability function, that is to say, the probability of uncertain threat event and its possible security hazard can be expressed by probability function. When the probability of an uncertain event and its loss function are known, the following expressions can be used to represent the risk measurement: $R(x) = R(C, P) = \sum_{i=1}^{n} p_i u(c_i)$ Where x denotes system risk; $c = (c_1, c_2, \Lambda\ , c_n)^T$, $p = (p_1, p_2, \Lambda\ , p_n)^T$, $c_1, c_2, \Lambda\ , c_n$ denotes all n possible consequences of an uncertain event; $p_1, p_2, \Lambda\ , p_n$ refers to the probability of $c_1, c_2, \Lambda\ , c_n$ and $\sum_{i=1}^{n} p_i = 1$; $u(c_i)$ is a quantitative function for the value of consequences. Whether to determine the probability of the event and the degree of damage caused by the security attribute or not is the key to analysis of information system security risk.

In information systems, the main causes of uncertain events are the vulnerability (or leaks) of the system and its resources, as well as the threat to both information system and its resources. And the occurrence probability of uncertain events is closely related to the vulnerability and threat. In hardware, software, network and communication protocols, there are different vulnerabilities or defects, which are the main source of vulnerability of information systems. Threats to systems usually include active threats and passive threats, such as destruction of information systems and information resources, malicious misuse or tampering, theft of information resources, deletion or loss of information, disclosure of information, prohibition and interruption of services, and so on. Consequences of risk events mainly include the influence of assets, the influence of ability, the cost of system recovery, the leakage of data, the deterioration of environment, the interference of communication and the loss of information. The impact on capacity is mainly being delayed, interrupted and weakened. System recovery costs include information recovery costs and service recovery costs. Figure 2 shows the probability and consequence hierarchy of uncertain events.

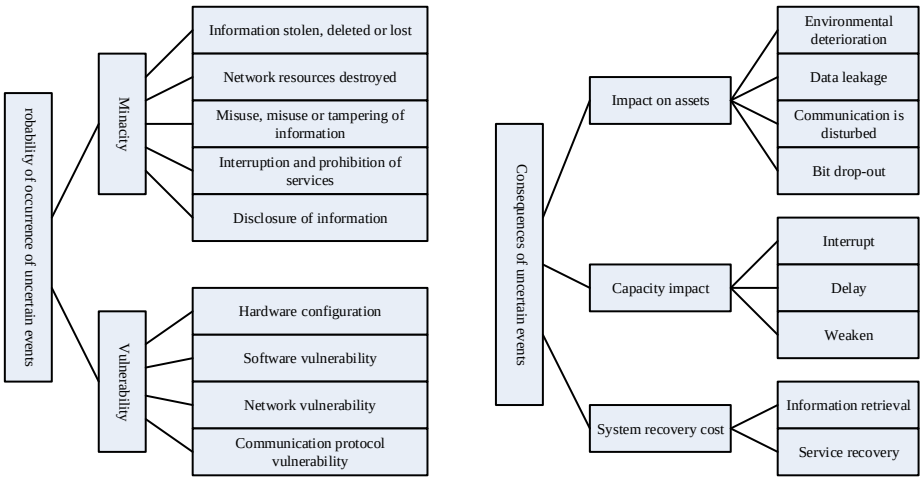

Figure 2 probability and consequence hierarchy of uncertain events

3.2 Fuzzy Comprehensive Evaluation of Information system risk

In the process of information system risk quantification, it is difficult to determine the probability of uncertain events in the risk because of the complexity of system risk and many uncertain factors that affect the quantification. It is also difficult to judge its consequences directly and accurately. In order to solve this problem, the fuzzy judgment theory can be used to analyze risk factors which are not easy to determine in the system risk. To a certain extent, the influence intensity and scope of each risk factor of system risks can be accurately described. The methods are as follows:

1. To establish a common set of different factors. The set is set to $U=(u_1,u_2,\Lambda\ ,u_n)$, including various factors that have an impact on the probability and consequences of uncertain events.

2. To establish an evaluation set. The set is set to $V=(v_1,v_2,\Lambda\ ,v_n)$, referring to the assessment assurance level (EAL) in the Common criteria, definable evaluation set V =(very large, larger, average, smaller, very small).

3. To evaluate influencing factors. According to the Delphi Method, an evaluation experts group is formed, with the task to establish a fuzzy relationship for the evaluation factors. By using this relationship, the risk factor set U is linked to the risk assessment set V, so as to obtain the membership degree of each influence factor in the evaluation set. Composition membership matrix $R=(r_{ij})_{n\times m}$, which $r_{ij}=\dfrac{v_{ij}}{\sum_{j=1}^{m}v_{ij}}$,

v_{ij} indicates that there are v_{ij} v_j comments on the evaluation factor u_i.

4. Quantization of evaluation vector. Definition C=(0.9,0.8,0.7,0.6,0.5,0.4,0), then the numerical risk evaluation value of the i factor can be expressed as $x_i=R_i\times C^T$, in which, R_i is the membership vector of the i factor.

3.3 Neural Network training for risk Analysis of Information Systems

Based on the learning ability of neural network, the evaluation ability of the system is improved, and the weight coefficient of risk assessment is obtained by neural network. Risk assessment value is calculated by converting the weight coefficient of risk assessment into the weight value of fuzzy comprehensive evaluation. Taking the three-layer BP neural network as an example, the input layer is composed of n nodes, the hidden layer is composed of one node and the output layer is a single node. The input of the system is the quantitative value of each risk influence factor, and the output risk evaluation value. The neural network is trained with m samples. The learning steps are as follows:

Step1: The initial learning error is e and the learning rate is η, η'. The initial value and threshold of neuron connection weight are interval random numbers.

Step2: To calculate the input and output of each layer.

Input layer: The numerical input $x_i^{(q)}$ of node i through fuzzy evaluation is used as the quantitative value of the i risk factor of the q sample. $x_i^{(q)}$ through the input layer node is $a_{ij}^{(q)}$, still in the input layer. The input and output are equal, that is to say $x_i^{(q)} = a_{ij}^{(q)}$, $i = 1,2,\Lambda\ ,n$, $j = 1,2,\Lambda\ ,l$, $q = 1,2,\Lambda\ ,l$.

Hidden layer: The weighted sum of the output variables of the input layer is input to the hidden layer as input to the hidden layer, $u_k^{(q)} = \sum_{i=1}^{n} w_{ik} a_{ik}^{(q)}$, w_{ik} is the connection weight of node i to k, $w_{ik} \geq 0$, $k = 1,2,\Lambda\ ,l$; The output of k is $b_k = f(u_k)$, f is sigmoid function $f(x) = 1/(1+e^{-x})$.

Output layer: The output layer is a single node, and the system risk assessment value $y^{(q)}$ trained by the neural network is output by the single node of the output layer. $y^{(q)} = f(\sum_{k=1}^{l} v_k b_k^{(q)})$, v_k is node k to output node connection weight.

Step3: To calculate the output error. Let the expected output t_j be a corresponding sample q, the mean square error E_j of a single sample q and the mean error E of m training samples are calculated. $E_j = \frac{1}{2}(t_j - y_j)^2$, $E = \frac{1}{m}\sum_{q=1}^{m} E_q$, In order to minimize E and E_j, the steepest descent method is used to calculate the deviations of each layer.

Output layer deviation $\delta^{(q)}$, $\delta^{(q)} = \frac{\partial E_q}{\partial v_k} = y^{(q)}(1-y^{(q)})(t^{(q)} - y^{(q)})$;

Hidden layer deviation $e_k^{(q)}$, $e_k^{(q)} = \frac{\partial E_q}{\partial v_{ij}} = b_k^{(q)}(1-b_k^{(q)})\sum_{k=1}^{l} v_k \delta^{(q)}$, $k = 1,2,\Lambda\ ,l$。

Step4: To calculate weight correction $\Delta v_k = \eta b_k^{(q)} \delta^{(q)}$, $\Delta v_{ik} = \eta' a_{ik}^{(q)} e_k^{(q)}$ to determine whether it meets $\delta^{(q)} < e$ or not. If it does, the training can stop. Otherwise, return to the second step.

Through the above neural network training, the mean square error of the network output risk evaluation value and the actual risk estimation of the training sample population can converge to the global minimum by using the iterative algorithm. When the mean square error converges to a global minimum, a stable connection weight can be obtained. The adaptive analysis and evaluation of information system risk and the risk analysis and evaluation of untrained information system data are realized. The neural network is used to study and train the samples. Then the method of combining neural network and fuzzy comprehensive evaluation is used to evaluate the risk.

3.5 Application examples

Based on previous analysis, the risk analysis and evaluation of a certain information system are carried out, in which b1~b8 are defined in order as the probability index of uncertain events occurrence, and b9~b18 are defined in order as the impact index of the outcome of the uncertain events in the system. Firstly, the evaluation value of each risk index is determined, and then experts are organized to give the evaluation value of each index under different risk conditions. Secondly, the sample set of system risk evaluation is determined, which is composed of the target valuation of system risk assessment. The objective valuation of system risk assessment is calculated by using the method of fuzzy comprehensive evaluation.

On the basis of sample collection, a three-layer neural network is constructed, in which the input layer

contains 18 neurons and receives the quantification values of 18 risk impact indicators respectively; the hidden layer contains three neurons; the output layer is a single neuron, in charge of outputting system risk evaluation results. Logarithmic sigmoid function is used to deal with logarithmic uniformity of input data.

An early termination method is adopted to improve the generalization ability of the neural network, and the sample set is divided into three sets: a training set, a confirmation set and a test set. The training set, consisted of 15 samples, is used to store the gradient, the updated network weight value and the threshold value of the network performance function. The confirmation set, consisted of 5 samples, is used to train sample's value and determine the final weight value and threshold value. While the test set uses 2 samples to verify the training results. Suppose the average output squared error of sample training (Mean-Squared Errors，MSE=10^{-4}) and learning rate($\eta = \eta' = 0.05$). Table 1 gives the training results obtained from the fuzzy comprehensive evaluation (FCE) and the fuzzy neural network (FNN) obtained from training set samples.

Table 1 Results of fuzzy comprehensive evaluation and fuzzy neural network training of training set samples

Evaluation results of FCE	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
	0.3804	0.4151	0.4707	0.5095	0.5743	0.6154	0.6496	0.6846	0.7185	0.759	0.7826	0.8247	0.8765	0.8937	0.9219
Training results of FNN	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
	0.3837	0.4164	0.4553	0.5312	0.5514	0.5761	0.6578	0.6892	0.7127	0.7674	0.775	0.8314	0.8616	0.8875	0.897

Table 2 shows the results of fuzzy comprehensive evaluation and fuzzy neural network prediction for samples of validation set and test set.

Table 2 The results of fuzzy comprehensive evaluation and fuzzy neural network training for the samples of validation set and test set

	Confirmation set sample					Test set sample	
Result	1	2	3	4	5	1	2
Evaluation results of FCE	0.4271	0.5108	0.6255	0.7091	0.7242	0.8126	0.8764
Training results of FNN	0.4373	0.4807	0.6173	0.6945	0.7037	0.8168	0.8997

Figure 3 shows the error output and the number of cycles of the samples of the training set and the confirmation set of the fuzzy neural network.

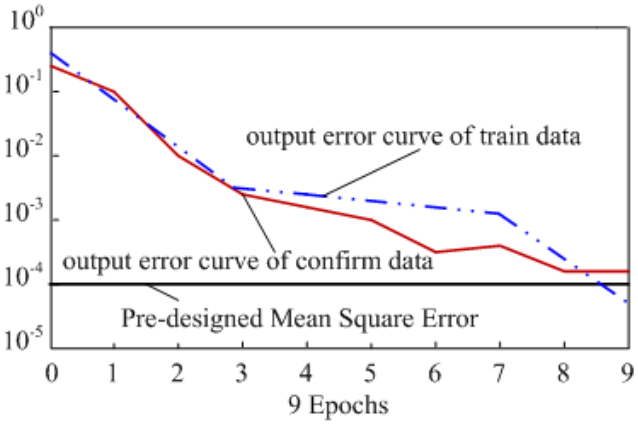

Fig. 3 Output error and cycle number of fuzzy neural network

In figure 3, the horizontal coordinate is the number of iterated cycles, and the vertical coordinate is the accuracy of the error. It can be seen from the diagram that after 9 cycles, the error accuracy of the training set has been set up in advance, and the sample of the confirmation set is also less than 10-4, which conforms to the requirements.

4 Conclusion

It is the beginning of risk assessment to analyze the information system security risk and to establish the

hierarchical structure model of uncertain events occurrence probability and risk events which may cause losses. The fuzzy neural network method, formed by the combination of fuzzy comprehensive evaluation and neural network technology, can be used to accurately quantify the risk index of the system in information system risk assessment. Furthermore, if the fuzzy neural network method is used in assessment feasibility, the choice of risk control strategies can be realized, and risks can be controlled effectively.

References

[1] Noureldien A Noureldiea，Izzeldin M Osman A Stateful Inspection Module Architechmre [EB/OL]. Httl://intl.ieexplore.ieee.org Xplore/conhome jsp， 2000.

[2] Izzeldin M Osma，Noureldien A Noureldien A Method for defeating DoS/DDoS TCP SYN Flooding Attack [EB/OL]. http://www.4law.co il/Le288 htm, 2002-06.

[3] Smith R.N. and Bhattacharya S.， A protocol and simulation for distributed communicating firewalls，Computer Software and Applications Conference，1999，pp. 74-79.

[4] Cheng Li, Li Yong. Evaluation of equipment maintenance Resource support ability based on Neural Network [J]. Military Operations Research and Systems Engineering. 2006 (3), 77-80.

[5] Wang Zhenxing, Liu Chenyu. Fault Diagnosis for Certain Equipment Based on Improved BP Neural Network [J].Computer and Modernization.2010, 17 (2): 200-203.

[6] Wu Shang. Analysis of Network Security Evaluation method [J]. Network Security Technology and its Application，2014（12）：145-146.

[7] Cheng Jian Hui, Zhao Ran. Enlightenment from Australian network security plan to Chinese information security [J]. Advanced Materials Research, 2013, 756: 2542-2546.

[8] Wang Jingguo, Xiao Nan. Drivers of information security search behavior: An investigation of network attacks and vulnerability disclosures [J]. ACM Transactions on Management Information Systems, 2010, 1(1): 546-551.

[9] Xin Dong Dong, Lan Qiu. Cloud-based network information security research and application [J]. Advanced Materials Research, 2013, 791: 1686-1689.

[10] Huang Luji, Luan Jiangfeng, Xiao Jun. Analysis of a defense-in-depth model of information security in computer networks [J]. Journal of Beijing Normal University (Natural Science), 2012, 48(2):138-141.

[11] Zhihui Yan, Xiaobin Li, Gengsheng Deng.[J]. International Journal of Digital Content Technology and its Applications, 2013, 7 (2): 422-429.