2014

ISSN : 0974 - 7435 Volume 10 Issue 21

BioTechnology

An Indian Journal

FULL PAPER

BTAIJ, 10(21), 2014 [12767-12773]

Fuzzy analysis method with artificial nerve network applied in information risk assessment

Cheng Yuandong
Anhui University of Science and Technology, Anhui Huainan China, 232001,
(CHINA)
E-mail: andoncheng@foxmail.com

ABSTRACT

The F-BPNN (Fuzz-back-propagation neural network) model of information security risk assessment will also be presented. By using the neural network model effectively the artificial influence in the traditional evaluation process such as the analysis hierarchy process could be avoided. It is hoped that the results of the study will summarize and enrich the implementation model of the information security risk assessment.

KEYWORDS

Fuzzy analysis; Artificial nerve network; Information risk assessment.

INTRODUCTION

Most previous assessment methods were mainly concentrated in the qualitative analysis which could not meet the demand of today's society. Therefore, this paper on the quantitative method of information security risk assessment shows an increasing demand and much more significance.

So this paper would like to empirically test the use of combined methods Fuzzy Back-propagation Neural Network in determining the actual level of system security risk. Specifically, the research would like to answer the following objectives:

How will the Fuzzy Analysis method be combined with Back-propagation Neural Network and what model can be designed in combining the new strength methods?

THE THEORETICAL BACKGROUND

Artificial neural network (ANN), usually called neural network (NN), is a complex network connected by a large number of processing unit (neurons). Information process can learn itself according to the sample and own highly parallel computing ability by using the neural network. Artificial neural network developed rapidly in recent years, applied more and more widely, especially in pattern recognition, image processing, voice recognition, artificial intelligence and combinatorial optimization, etc.

The learning capabilities of neural network allows controller to learn a certain function, which are highly nonlinear, and represent the dynamic of the processes. This is performed during long training period of controllers in supervised and unsupervised manner[1].

A lot of different neural network structures have been studied. The most commonly used structure is formed in three layers, called the input layer, the hidden layer, and the output layer. The small circles represent nodes which consisted as every layer. The lines between the small circles indicate the flow of information from one node to the next. In this type of neural network, the information flows only from the input to the output (that is, from left-to-right), while other types of neural networks may have more intricate connections[2].

There is a branch of neural network structure named Back-propagation neural network which are nonlinear regression structure that represents input-output mapping. The advantage of Back-propagation network is its higher learning speed and it’s more adaptive ability to new data. And also, it can be easily designed and trained than the other neural networks.

Fuzzy system owns the ability of expression to be easily understood, and the neural network has strong adaptive ability. Fuzzy Back-propagation Neural Network combines the fuzzy theory and Back-propagation Neural Network which can improve the whole system learning ability and express ability[3]. There are two kinds of specific combination method: one is to plant the fuzzy classification method in the neural network, the other is to structure network based on fuzzy theory structure. This paper will use the first method.

COMBINING FUZZY ANALYSIS WITH NEURAL NETWORK

In the information security risk assessment, there are the nonlinear relationship and dynamic change regulation between each evaluation index risk factors and risk level of risk factors. Neural network can realize any complex nonlinear mapping relationship from the inputs to the outputs[4]. Accordingly, this paper will apply BP neural network to set up information security risk assessment model.

Model structure

Fuzzy system has the advantage of expression which is easy to be understood, and the neural network has a strong adaptive ability. Fuzzy neural network is combining the fuzzy theory with the neural network. It can improve the whole system learning ability and expressing ability. Specific combination method has two kinds: one is to put the fuzzy classification method into the existing neural network; the other is to establish the neural network structure based on fuzzy theory. This paper will use the first kind, taking fuzzy membership as the neural network's input.

The model is the BP neural network model with single hidden layer which is constituted by the input layer, hidden and output layers[5]. Network's input Eigen vector is all the evaluation index of risk factors, including asset confidentiality, integrity, availability, complexity of vulnerable transferred into remuneration, severity of vulnerability, availability of vulnerability and the technical content of threat. These indexes will be quantized and normalized to be a neural network’s inputs. After the BP neural network learning and training process, the output Eigen vector of the network will be the level of risk[3].

The number of neurons in hidden layer will be more complicated if Network applies a single hidden structure. Experiments are repeated many times to get a proper number of neurons. The transfer function Sigmold usually is adopted to be neurons function of hidden layer, and the transfer function Purelin usually is adopted to be neurons function of output layer.

Input processing

The BP neural network is suitable for quantitative data but there is a lack of corresponding processing ability for qualitative index analysis. It is not easy to calculate the value of risk factors index; so we use the fuzzy evaluation method mentioned to quantify the indexes of information security risk factors. Take the output of the fuzzy system as the inputs of a neural network[6]. The specific implementation method is:

(1) According to the analysis of the Chaper1, we can analyze the correlation of the assets, threat, vulnerability and threats to find out the information security risk factors.

(2) According to the description about fuzzy evaluation method, we can establish risk factors set $U=\{u_1,u_2,\ldots,u_n\}$.

(3) Construct the evaluation sets. Evaluate the risk factors from the following aspects: unauthorized access, unauthorized access system resources, data divulged, refused to service illegally modify data and software, system collapse, etc. Index comments of risk degree of risk factors are given by experts. Each index comments will be divided into m grades, and judgment sets $V=\{v_1,v_2,\ldots v_m\}$.

(4) Experts give the comments of each risk factor, according to the method of Chpter2.1.2, the membership vector of the risk factor u_i to judgment set V is $R_i=\{r_{i1},r_{i2},\ldots,r_{in}\}$,$i=1,2,\ldots,n$. Structure the membership matrix R.

(5) Set weight distribution sets $A=(a_1,a_2,\ldots,a_m)$. The operations of fuzzy transform $B=A\cdot R$, B is the weight of risk factors in a certain judgment, reflect the value of the risk factors, and the value is in (0, 1). B can be used as the inputs of the BP neural network.

APPLICATION IN RISK ASSESSMENT OF INFORMATION SECURITY

Database server is the most important and most valuable assets of an IT company. There are plenty of the previous and future sensitive financial data in the database service, including trade record, commercial contract and accounting data, etc. Some confidential information must be regarded as the company's secret, like the ownership of the technology, engineering data, and market planning and decision-making, to prevent unauthorized access[6]. Database server also includes detailed customer information, such as financial accounts, credit card number, and credit information of business partners, etc.

There are six security risk factors about database server: unauthorized accessing (M_1), unauthorized accessing system resources (M_2), divulging data (M_3), denying service (M_4), unauthorized modifying data and software (M_5), system function collapse (M_6). Because there are six risk factors about the database server security, we can set six neurons in the input layer and one neuron in the output layer. Take this matrix as BP neural network's input, take the security level assessment as the output of the BP neural network, according to the empirical formula, calculate the number of hidden neurons $log_2 n+8=log_2 6+8\approx 3+8=11$ (n is the number of hidden) input level to the hidden, take "Sigmoid" as the learning function for the hidden layer and for the output layer, set a= 0. 05 as learning rate of the sample training set $e=10^{-4}$ as the training goal. Finally, we can get the assessment model through BPNN's adaptive learning.

The historical assessment data of the database server of each quarter from January 2007 to March 2012 are shown in Appendix A in the appendices.

There are 21 group history assessment data of the company's database server in the above mentioned tables until April 2012. So, in this fuzzy neural network security risk assessment, we will take the first 16 sets of data to train the neural network, take the left 5 sets of data to simulate the neural network and to test the accuracy of the neural network. Training and simulation are all designed in the "MATLAB software", the concrete code and the explanation of key part is shown in Appendix B.

The output error and recycle number of the BP neural network training is shown in Figure 1. In the figure, the x-coordinate denotes times of the iterative cycle; y-coordinate denotes the change of error. In Figure 1, we find the error precision of the neural network is below the $e=10^{-4}$ after the weights 11493 times iteration are adjusted.

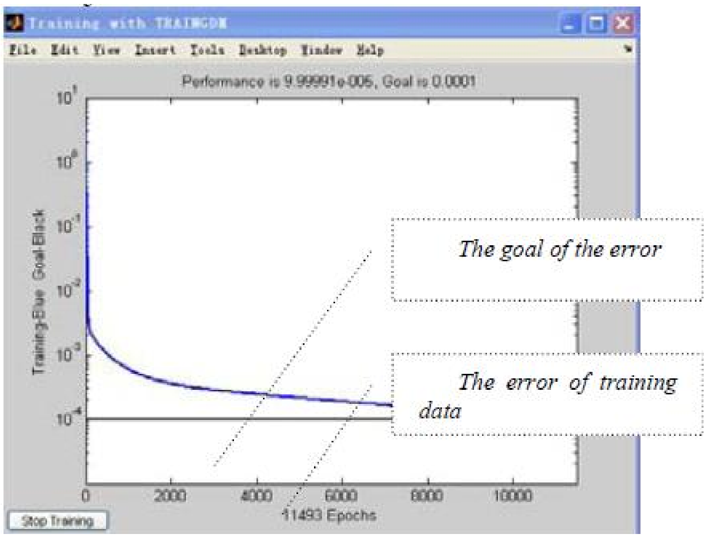

Figure 1 : Output error and cycle times of BPNN

The comparison between the results of experts' assessment and the simulation results of BPNN is shown in TABLE 1.

TABLE 1 : Result comparison

No.	201101	201102	201103	201104	201201
Result of Experts	0.5600	0. 4100	0.5400	0.4700	0.6100
Results of Simulation	0.5188	0.3773	0.5844	0.4236	0.6214
Errors	0.0412	0.0327	-0.0444	0.0464	-0.0114

Figure 2 is the comparison between the results of experts' assessment and the simulation results of BPNN. "+" stands for BP neural network simulation results, "O" stands for experts' assessment results. We can tell that the fitting is very well from the graph which means the BPNN has a strong adaptive ability. And the simulation I designed, the result can be used to assess the security risk level of the database server of the company.

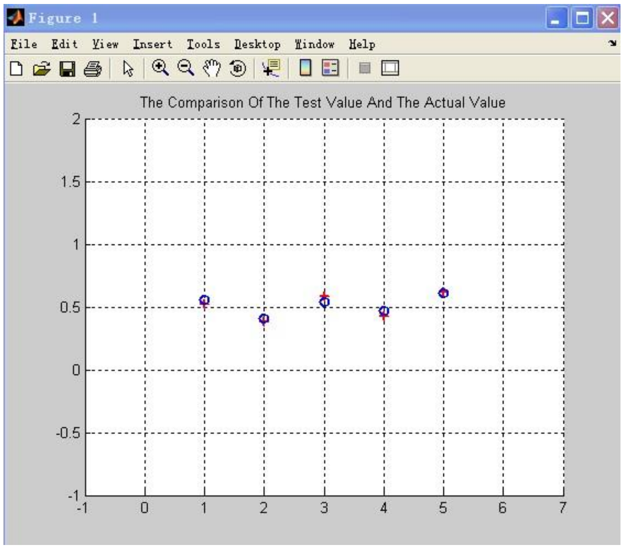

Figure 2 : Comparison of the test value and the actual value

Calculate the BP neural network's inputs with fuzzy comprehensive evaluation method. First, construct the fuzzy set $U=\{C_1,C_2,...,C_6\}$, of which $C_1,C_2,...,C_6$ are risk factors: unauthorized accessing (M_1), unauthorized accessing system resources (M_2), divulging data (M_3), denying service (M_4), unauthorized modifying data and software (M_5), system function collapse (M_6). And then build the evaluation set. Build different evaluation set to the different criterion.

For the criterion: B_ "database server", the meaning of the judgment set $V=\{V_1,V_2,...,V_8\}$ of the risk factor set U is shown in TABLE 2.

TABLE 2 : Definition of risk level

Risk Level	Description
V_1 Can be ignored	It is unlikely to occur.
V_2 Very Low	Might occur 1 to 2 times every quarter.
V_3 Low	Might occur 3 to 5 times every quarter.
V_4 Medium	Might occur 6 to 9 times every quarter.
V_5 High	Might occur 3 to 4 times every month.
V_6 Higher	Might occur 5 to 7 times every month.
V_7 Very High	Might occur 8 to 9 times every month.
V_8 Extremely High	Might occur more than 10 times every month.

The experts will evaluate the probability of the risk factor U according to TABLE 1. Each risk probability given by the experts should be one kind of the judgment set $V_1, V_2, \ldots, V_8$. Experts' assessment form is shown in TABLE 3:

TABLE 3 : Experts' assessment form of database server

	Risk Factors	Fuzzy Assessment Level							
		V$_1$	**V$_2$**	**V$_3$**	**V$_4$**	**V$_5$**	**V$_6$**	**V$_7$**	**V$_8$**
Risk Probability B	unauthorized accessing (M$_1$)	5%	4%	6%	5%	10%	30%	23%	27%
	unauthorized accessing system resources (M$_2$)	4%	3%	10%	7%	6%	10%	30%	40%
	divulging data (M$_3$)	13%	15%	10%	20%	5%	7%	16%	14%
	denying service (M$_4$)	20%	20%	10%	12%	8%	8%	12%	10%
	unauthorized modifying data and software (M$_{5)}$	13%	10%	5%	8%	17%	20%	13%	14%
	system function collapse (M$_6$)	18%	22%	10%	17%	13%	8%	5%	7%

Remarks: The meaning of V$_1$ to V$_8$ is shown in TABLE 1 which should be attached below in the actual operation.

According to the evaluation opinion of each expert, calculate each risk probability of the relative index, and we will get the membership matrix R of B_C.

$$R = \begin{bmatrix} 0.05 & 0.04 & 0.06 & 0.06 & 0.10 & 0.30 & 0.23 & 0.27 \\ 0.04 & 0.03 & 0.10 & 0.07 & 0.06 & 0.10 & 0.30 & 0.40 \\ 0.13 & 0.15 & 0.10 & 0.20 & 0.05 & 0.07 & 0.16 & 0.14 \\ 0.20 & 0.20 & 0.10 & 0.12 & 0.08 & 0.08 & 0.12 & 0.10 \\ 0.13 & 0.10 & 0.05 & 0.08 & 0.17 & 0.20 & 0.13 & 0.14 \\ 0.18 & 0.22 & 0.10 & 0.17 & 0.13 & 0.08 & 0.05 & 0.07 \end{bmatrix}$$

Calculate the sorted weight vector of the membership matrix. Determine the weights standard $V_1, V_2, \ldots, V_8$ of the B_C membership matrix is: 1/36,2/36,3/36,4/36,5/36,6/36,7/36,8/36. According to the formula $B = A \bullet R^T$, calculate the relative weight under the criterion B$_{1\_}$ "risk probability": (0.183,0.191,0.123,0.108,0.136,0.101), and then normalize the relative weight to get the sorted weight vector: (0.217,0.227,0.146,0.129,0.161,0.120).

Finally, take the sorted weight vector as the BPNN's inputs to get the assessment result. The code to calculate the result is shown in the following:

```
pause
clc
Assesstment_inputdata= [0.217; 0.227; 0.146; 0.129; 0.161; 0.120];
Assesstment_result=sim (net, Assesstment_inputdata);
Assesstment_result % output assessment result
```

The result of output is：Assesstment_result = 0.1766

If according to TABLE A-10 in the appendices risk classification criteria of security levels, the security risk level of company's database server is class 2, which means the database server has a higher security and the risk coefficient is smaller.

SUMMARY

Using F-BPNN, the security risk value of company's database server is 0.1766, the risk level is 2. The result is reliable based on the following two steps: (i) Taking the result of Fuzzy Comprehensive Evaluation as the input of BPNN. (ii) Realizing the error goal after the simulation experiment to BPNN.

This paper put forward Fuzzy Back Process Neural Network method of risk assessment. The artificial neural network is applied in the information system risk assessment, the neural network's input is pretreated, and the output of the fuzzy system as a neural network's input is taken. Artificial neural network can estimate real-time the level of risk factors after been trained. This model can be used for information security risk factors evaluation after been trained.

Appendix A

TABLE 4 : Historical Assessment Data from January 2007 to March 2012

Serial Number	No.	M_1	M_2	M_3	M_4	M_5	M_6	Assessment
1	200701	0.5	0.5	0.4	0.5	0.6	0.4	0.48
2	200702	0.73	0.6	0.7	0.8	0.5	0.7	0.72
3	200703	0.3	0.2	0.4	0.5	0.3	0.2	0.35
4	200704	0.5	0.5	0.6	0.8	0.4	0.5	0.56
5	200801	0.4	0.3	0.4	0.3	0.3	0.4	0.38
6	200802	0.4	0.3	0.4	0.6	0.5	0.3	0.41
7	200803	0.5	0.6	0.2	0.8	0.7	0.5	0.66
8	200804	0.3	0.2	0.3	0.2	0.2	0.4	0.28
9	200901	0.7	0.6	0.8	0.7	0.8	0.6	0.75
10	200902	0.4	0.3	0.4	0.6	0.5	0.4	0.47
11	200903	0.6	0.6	0.7	0.6	0.8	0.6	0.64
12	200904	0.7	0.5	0.7	0.6	0.8	0.5	0.64
13	201001	0.5	0.5	0.6	0.6	0.4	0.7	0.61
14	201002	0.3	0.4	0.3	0.4	0.3	0.2	0.34
15	201003	0.4	0.5	0.4	0.6	0.6	0.4	0.54
16	201004	0.5	0.6	0.4	0.8	0.7	0.5	0.68
17	201101	0.4	0.3	0.4	0.6	0.6	0.5	0.56
18	201102	0.5	0.2	0.3	0.4	0.2	0.4	0.41
19	201103	0.5	0.4	0.5	0.7	0.6	0.5	0.54
20	201104	0.3	0.5	0.4	0.5	0.5	0.3	0.47
21	201201	0.5	0.4	0.6	0.7	0.6	0.5	0.61

Appendix B

TABLE 5 : % *this program of dissertation in Chapter 4_BP network by He Jidong and ChengYuandong (Andon) can be pasted directly into MATLAB to have a test.*

Data= [0.5 0.5 0.4 0.5 0.6 0.4 0.48						
0.73	*0.6*	*0.7*	*0.8*	*0.5*	*0.7*	*0.72*
0.3	*0.2*	*0.4*	*0.5*	*0.3*	*0.2*	*0.35*
0.5	*0.5*	*0.6*	*0.8*	*0.4*	*0.5*	*0.56*
0.4	*0.3*	*0.4*	*0.3*	*0.3*	*0.4*	*0.38*
0.4	*0.3*	*0.4*	*0.6*	*0.5*	*0.3*	*0.41*
0.5	*0.6*	*0.2*	*0.8*	*0.7*	*0.5*	*0.66*
0.3	*0.2*	*0.3*	*0.2*	*0.2*	*0.4*	*0.28*
0.7	*0.6*	*0.8*	*0.7*	*0.8*	*0.6*	*0.75*
0.4	*0.3*	*0.4*	*0.6*	*0.5*	*0.4*	*0.47*
0.6	*0.6*	*0.7*	*0.6*	*0.8*	*0.6*	*0.64*
0.7	*0.5*	*0.7*	*0.6*	*0.8*	*0.5*	*0.64*
0.5	*0.5*	*0.6*	*0.6*	*0.4*	*0.7*	*0.61*
0.3	*0.4*	*0.3*	*0.4*	*0.3*	*0.2*	*0.34*
0.4	*0.5*	*0.4*	*0.6*	*0.6*	*0.4*	*0.54*
0.5	*0.6*	*0.4*	*0.8*	*0.7*	*0.5*	*0.68*
0.4	*0.3*	*0.4*	*0.6*	*0.6*	*0.5*	*0.56*
0.5	*0.2*	*0.3*	*0.4*	*0.2*	*0.4*	*0.41*
0.5	*0.4*	*0.5*	*0.7*	*0.6*	*0.5*	*0.54*
0.3	*0.5*	*0.4*	*0.5*	*0.5*	*0.3*	*0.47*
0.5	*0.4*	*0.6*	*0.7*	*0.6*	*0.5*	*0.61];*

```
[M, N] = size (Data);
% select data
InputData = Data (:, 1:N-1)';
TargetData = Data (:, N)';
TrainNumber = 16;
Input_train = InputData (:, 1: TrainNumber); % Training data input
Output_train = TargetData (:, 1: TrainNumber); % Training data output
Input_test = InputData (:, TrainNumber+1: M); % Forecasting data input
Output_test = TargetData (:, TrainNumber+1: M); % Testing data output
net = newff(minmax(Input_train),[11,1],{'tansig','purelin'},'traingdm')
inputWeights=net.IW{1,1};
inputbias=net.b{1};
layerWeights=net.IW{1,1};
layerbias=net.b{2};
% set the BP network parameter
net.trainParam.epochs = 30000; % Iteration times
net.trainParam.goal = 1e-4; % Learning goal
net.trainParam.lr = 0.05; % Learning rate
net.trainParam.mc = 0.9; % Momentum factors
net.trainParam.show = 50; % Show times
[net,tr] = train (net, Input_train,Output_train);
Pause
Clc
% simulation on BP network
A = sim (net,Input_test);
% comput simulation error
E = Output_test - A;
Output_test % output the actual data in order to test the error
A % output the simulation data
E % output the error
Pause
Clc
% Draw the comparison of test value and actual value
set (gca,'XLim',[-1 7]);
set(gca,'XGrid','on');
set(gca,'YLim',[-1 2]);
set(gca,'YGrid','on');
hold on;
plot(A,'r+','LineWidth',2);
plot(Output_test,'bo','LineWidth',2);
title('The Comparison Of The Test Value And The Actual Value');
```

REFERENCES

[1] Philip D.Wasserman; “Neural computing: Theory and practice” [M], Van Nostrand Reinhold, **(1989)**.

[2] R.Adler.; “Ergodic and mixing properties of innate memory channels”, [J], Proc.Amer.Math.Soc, 924-930 **(1961)**.

[3] R.Csutoraa, J.J.Buckleyb.; “Fuzzy hierarchical analysis: The lambda-max method”, [J], Fuzzy Sets and Systems, 181-195 **(2001)**.

[4] Rahib Hidayat Abiyev; “Fuzzy back-propagation neural network for control of dynamic plants”, [J], Mathematical Sciences, **(2005)**.

[5] Wang Wei; “Artificial neural network theory introductory and application” [M], Beijing: Beijing University of Aeronautics and Press, **(1995)**.

[6] Zhao ZhenYu, Xu Yong; “The fuzzy theory of the neural network and the basic and applied” [M], Beijing Tsinghua university press, **(1996)**.