A New Fuzzy Neural Network Model and Its Application on Network Operating Situation Awareness

Sihong Liu[1]

University of Electronic Science and Technology of China
Chengdu, 611731, China
E-mail: 15680758984@126.com

Xingming Li

University of Electronic Science nd Technology of China
Chengdu, 611731, China
E-mail: xingmingl@uestc.edu.cn

Xiaobo Fan

University of Electronic Science And Technology of China
Chengdu, 611731, China
E-mail: xiaobof@outlook.com

In order to monitor the network in real time, scientifically analyze the network situation and predict its changing trend, we establish the Network Operating Situation Awareness (NOSA) method based on new Fuzzy Neural Network (FNN) model, which has improved the well-known back propagation (BP) algorithm to accelerate the learning rate; besides, we use Wang-Mendel (WM) method to efficiently extract fuzzy rules. The simulation results show that, in comparison with the traditional models, the new model features superiorities in convergence speed and learning accuracy, and gives a relatively accurate assessment of network operating situation. The new FNN model is a powerful method of NOSA, especially in complex environment, where there're massive uncertain and ambiguous data.

1.Introduction

The Network Operating Situation Awareness (NOSA) is to acquire, understand and analyze the factors that can cause changes in network operating situation, and predict future trend of the network. Tim Bass proposed the concept of Cyberspace Situational Awareness in 1999 for the first time [1]. It has thus become a research hotspot since then. The existing studies of Network Situation Awareness (NSA) are mainly focusing on security problems of network [2] with seldom concerning transmission [3], information superiority and survivability, etc. And there are three main assessment methods of NSA: the mathematical method, the knowledge reasoning and the pattern recognition. Y. Zhuo established NSA system based on pattern recognition method: Rough Set Theory (RST) [4]. X.Z. Chen constructed a mathematical model to evaluate the security situation of network [5].

This paper uses Fuzzy Neural Network (FNN) to assess the network operating situation. As the combination of fuzzy inference system and neural network, it can not only reason ambiguous information in network but also features strong learning ability, which can help extract key influencing factors of situation and give useful conclusions of real-time system through the analysis of complex data. The traditional FNN is mostly based on the five-layer feed-forward neural network structure [6, 7], and the widely-used algorithm of FNN is the BP algorithm [7]. However, the well-known BP algorithm is easy to trap into the local minimum and its global convergence speed is too slow. Another important problem of FNN is the generation of fuzzy rules. The existing research mainly uses genetic algorithm [8, 9] and the fusion self-organizing algorithm [6], etc., but these algorithms either require continual repeated learning or are hard to implement owing to complex mechanisms.

We propose a new four-layer FNN model, which presents the enhanced version of well-known BP algorithm that has been improved based on new activation function—Rectified Linear Units (ReLU), adjusting learning rate and memorial modifying process of weights. The improved BP algorithm can effectively avoid falling into the local minimum and speed up the convergence rate. In addition, we apply Wang-Mendel (WM) method in the hidden layers of FNN model to extract fuzzy inference rules. In comparison with other methods, WM method is simpler, yet useful and more efficient, which implies that it does not continually need repeated learning.

The remainder of this paper is organized as follows: Section 2 introduces the structure of the new FNN and the improved BP algorithm as applied. Section 3 gives a detailed description of how to establish a fuzzy rule base on the WM method. Section 4 gives the simulation of our proposed methods on the NOSA system. We end with a conclusion in Section 5.

2. A New Fuzzy Neural Network Model and Its Learning Algorithm

2.1 Structure of the New Fuzzy Neural Network

This paper presents a new four-layer FNN model, namely input layer, membership generation layer, reasoning layer and reverse fuzzy output layer with its structure shown in Fig. 1.
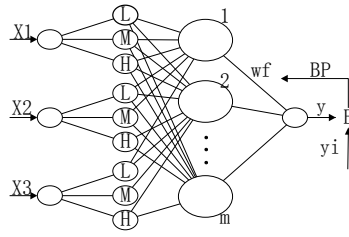
Figure 1 : New Fuzzy Neural Network (FNN) Model

There are many parameters which can reflect the network's situation from different aspects, such as bandwidth, connectivity, packet loss and round trip delay, etc. Taking the independence and comprehensiveness into account, we choose three basic and important parameters: bandwidth, packet loss and round trip delay as inputs in our model.

The two hidden layers of the model is the so-called fuzzy reasoning part, they carry out the abstract process of transferring input data to information and again to knowledge. The number of linguistic variables corresponding to the input variables decides the amount of nodes in Layer 2, in this paper, we choose three linguistic variables: low, medium and high. The quantity of nodes in Layer 3 depends on the number of fuzzy inference rules, as generated by WM method.

The output in Layer 4 represents the network's performance after the assessment of FNN.

The most widely used activation function in neural network is the Sigmoid function [3]; however, the Sigmoid function features some disadvantages: it causes the gradient to disappear and its output is not zero-center, which would lead to slow convergence of the whole network and poor learning performance.

We choose ReLU function as the activation function of the nodes, its formula is $f(x)=max(0,x)$. Compared with the Sigmoid function, ReLU function has a great effect on the convergence of random gradient descent [7] and precision of learning. Its advantages are the following: firstly, the function can be suppressed unilaterally; secondly, it has a relatively broad excitement border; finally, it is sparsely activated, that is, the neurons only select a small part of the input signals for response while a large number of signals are deliberately shielded, which can improve the accuracy of learning, better and faster extract sparse .

2.2 Improved BP Algorithm

Most traditional FNN uses BP algorithm as the main learning algorithm, but the well-known BP algorithm has some shortcomings, such as the learning rate is fixed [7], the weights' modifying process has no memory [7], making it easy to fall into the local minimum. We make some improvements to the BP algorithm. One is to adjust the learning rate adaptively, the other is to modify the weights with previous modification amount added at each time. These two improvements can effectively avoid trapping into local minimum points and speed up the convergence rate.

Assume that the error objective function be $E=(y-Y)^2/2$, where y is the actual output and Y is the desired output of the FNN. According to the theory of error back propagation of BP algorithm, we can get the modified formula of three parameters to be adjusted in the network training process. The concrete derivation process is shown as follows:

$$\frac{\partial E}{\partial \omega_i} = (y - Y) \cdot ReLU \left(\frac{\omega_i * O_i^3}{\sum\limits_{i=1}^m O_i^3} \right)' \cdot \frac{O_i^3}{\sum\limits_{i=1}^m O_i^3} \quad , \text{where} \quad ReLU(x)' = \begin{cases} 0, & x \le 0 \\ 1, & x > 0 \end{cases} \quad . \tag{2.1}$$

$$\frac{\partial E}{\partial m_{ij}} = (y - Y) \cdot ReLU \left(\frac{\omega_i * O_i^3}{\sum\limits_{i=1}^m O_i^3} \right)' \cdot \frac{\omega_i}{\sum\limits_{i=1}^m O_i^3} \cdot ReLU \left(\prod_{i=1}^n \mu_{ij} \right)' \cdot \prod_{l=1, l \ne i}^n \mu_{ij} \cdot 2 \cdot e^{\frac{-(x_i - m_{ij})^2}{\delta_{ij}^2}} \cdot \frac{(x_i - m_{ij})}{\delta_{ij}^2}$$

$$\frac{\partial E}{\partial \delta_{ij}} = (y - Y) \cdot ReLU \left(\frac{\omega_i * O_i^3}{\sum\limits_{i=1}^m O_i^3} \right)' \cdot \frac{\omega_i}{\sum\limits_{i=1}^m O_i^3} \cdot ReLU \left(\prod_{i=1}^n \mu_{ij} \right)' \cdot \prod_{l=1, l \ne i}^n \mu_{ij} \cdot 2 \cdot e^{\frac{-(x_i - m_{ij})^2}{\delta_{ij}^2}} \cdot \frac{(x_i - m_{ij})^2}{\delta_{ij}^3}$$

As mentioned above, the first improvement of BP algorithm in this paper is to adjust the learning rate based on the two successive iterative gradient directions, that is,

$$\eta_a^{k+1} = \eta_a^k \cdot 2^\lambda \quad , \quad \lambda = sign \left(\frac{\partial E}{\partial a^{k+1}} \cdot \frac{\partial E}{\partial a^k} \right) \tag{2.2}$$

The second improvement is to modify the weights with momentum term, that is,

$$\Delta a_i^{k+1} = \beta \cdot \Delta a_i^k + (\beta - 1) \cdot \eta^{k+1} \cdot \frac{\partial E}{\partial a_i^{k+1}} \tag{2.3}$$

where $a = \omega$, m or δ , and β is the momentum coefficient.

In the practical application, we should train the FNN until the value of error objective function is less than the specified value or the training times exceeds certain number.

3. Application of Wang-Mendel Method in Fuzzy Rule Extraction

NOSA needs to analyze parameters of network, then mine the potential association of them, finally give an assessment result of the network situation. However, the main problem in the analysis of NOSA is the difficulty of judging the normality of the network. The traditional analysis and assessment methods divide the normal and abnormal behavior of network too sharply, which thus result in a large number of false alarms around the boundaries. Fuzzy inference system can help us process ambiguous information in network, as shown in the model of Fig. 1. Layer 2 and Layer 3 refer to the "IF-THEN" part of fuzzy inference system. Layer 2 is responsible for blurring the input numeric variables, and Layer 3 is the rule part of the "IF-THEN" inference rule.

There are many methods for extracting fuzzy rules, such as genetic algorithm [8], fusion self-organizing algorithm [6], etc. We use WM method to extract fuzzy inference rules. It is proposed by Wang and Mendel to predict a chaotic time-series [10]. The principle of WM algorithm is to extract fuzzy rule by calculating the support degree of each data pair in the data set. Compared with other methods, WM method is simpler, yet useful and more efficient, which signify it does not need continually repeated learning. The detailed steps of WM algorithm are as follows:

1. Define the domain intervals of each variable, then divide each domain interval into 2N + 1 regions, and assign each region a fuzzy membership function.

2. Determine the degrees of given variables in different regions, and assign a given variable to the region with maximum degree.

3. Calculate the support degree of each rule. The support degree is calculated as follows:

Suppose there is a data pair in the N input one output data set

$$\left(x_1^l, x_2^l, \ldots, x_N^l ; y^l\right) \quad , \quad l=1,2,\ldots$$

It corresponds to a fuzzy rule R^l , which can be expressed as: IF x_1^l is A_1^l and x_2^l is A_2^l … and x_N^l is A_N^l THEN y is B^l . A_i^l , B^l are the fuzzy regions of the input and output variables in rule R^l .

We use Gaussian function as the membership function. The membership degree of the fuzzy region corresponding to each variable can be calculated as follows:

$$\mu\left(x_i\right)=e^{\frac{-\left(x_i-m_{c_i}\right)^2}{\delta_{c_i}^2}} \tag{3.1}$$

The support degree for each rule is:

$$Supp\left(R^l\right)=\mu_{A_1^l}\left(x_1^l\right)\cdot\mu_{A_2^l}\left(x_2^l\right)\cdot\ldots\cdot\mu_{A_N^l}\left(x_N^l\right)\cdot\mu_{B^l}\left(y^l\right) \tag{3.2}$$

4. The redundancy merging and conflict elimination of fuzzy rules. If there are some conflicting rules, which have the same IF part but a different THEN part, we accept only the rule that has the maximum support degree. For the repeated rules, we combine them into one rule.

5. Create a fuzzy rule base. In practice, we often have some priori information about the data pairs; therefore, we can assign a support degree to each data pair which represents our belief of its usefulness. Thus, the expert's linguistic rules and the rules extracted from sampled input-output data pairs constitute the fuzzy rule base.

4. Simulation of NOSA System

Among the network assessment methods based on mathematical model, the weight analysis is the most common evaluation approach, and its assessment function is usually an exponential expression, determined by the situation factors and their importance weights, e.g., we can define an assessment function of network operating situation as follow:

$$y=\left(1+x_1^{0.5}+x_2^{-1}+x_3^{-1.5}\right)^2 \tag{4.1}$$

where x_1 , x_2 and x_3 represent the network parameters of round trip delay, the packet loss rate and bandwidth, the exponential of each variable depends on its importance in our evaluation system. We select these three parameters for the following reasons: (1) they can fully reflect the state of the whole network; (2) they are easy to measure in practice; (3) their relevance is very small. As the three parameters are different from many aspects, in practical application, we must carry out dimensionless process for them after data collection.

The simulation in this paper is executed on Eclipse IDE of Java language, our FNN model is designed on this IDE to evaluate the performance of network. Through simulation, we compare the new FNN with the traditional five-layer FNN in learning speed and error rate.

We generated 40 pairs of data according to Formula 4.1 to imitate the changing process of network's parameters over a period of time, the data is shown in Table 1. We divide the sample into two groups, 1 to 20 for the group A, 20 to 40 for the Group B.

No.	Y	x_1	x_2	x_3	No.	Y	x_1	x_2	x_3
1	11.11	1	3	1	21	9.545	1	1	5
2	6.521	1	5	2	22	6.043	1	3	4
3	10.19	1	1	3	23	5.724	1	5	3
4	6.043	1	3	4	24	11.25	1	1	2
5	5.242	1	5	5	25	11.11	1	3	1
6	19.02	5	1	4	26	14.36	5	5	2
7	14.15	5	3	3	27	19.61	5	1	3
8	14.36	5	5	2	28	13.65	5	3	4
9	27.42	5	1	1	29	12.43	5	5	5
10	15.39	5	3	2	30	19.02	5	1	4
11	5.724	1	5	3	31	6.38	1	3	3
12	9.766	1	1	4	32	6.521	1	5	2
13	5.87	1	3	5	33	16	1	1	1
14	5.406	1	5	4	34	7.219	1	3	2
15	10.19	1	1	3	35	5.724	1	5	3
16	15.39	5	3	2	36	19.02	5	1	4
17	19.68	5	5	1	37	13.39	5	3	5
18	21.06	5	1	2	38	12.68	5	5	4
19	14.15	5	3	3	39	19.61	5	1	3
20	12.38	5	5	4	40	15.39	5	3	2

Table 1 : Data Samples

x_1	x_2	x_3	Y	y_1^{tra}	y_1^{new}	y_2^{tra}	y_2^{new}
1	3	1	11.11	10.484	11.111	10.647	11.127
1	5	2	6.521	6.503	6.515	6.407	6.545
1	1	3	10.19	9.969	10.187	10.127	10.206
1	3	4	6.043	6.081	6.169	6.673	6.124
1	5	5	5.242	5.466	5.267	5.495	5.27
5	1	4	19.02	18.417	18.983	18.592	19.048
5	3	3	14.15	14.197	14.244	13.92	14.129
5	5	2	14.36	14.267	13.697	14.293	14.34
5	1	1	27.42	25.624	27.523	24.404	27.456
5	3	2	15.39	15.343	16.075	15.651	15.437
1	5	3	5.724	5.523	5.729	5.606	5.712
1	1	4	9.766	9.447	9.69	10.204	9.914
1	3	5	5.87	6.531	5.883	6.094	5.865
1	5	4	5.406	5.411	5.318	5.586	5.513
1	1	3	10.19	9.969	10.187	10.127	10.206
5	3	2	15.39	15.343	16.075	15.651	15.437
5	5	1	19.68	20.028	19.759	20.166	19.719
5	1	2	21.06	22.373	20.035	22.73	21.121
5	3	3	14.15	14.197	14.244	13.92	14.129
5	5	4	12.38	13.023	12.661	12.168	12.77

Table 2 : Simulation Results of Experiment I and II

We conduct two experiments on both traditonal FNN and new FNN. Experiment I: taking Group A as training data; Experiment II: taking both Group A and Group B as the training data. The two experiments both select Group A as the testing data. The simulation results of experiment I and II are as Table 2 shows.

y_i^{tra} , y_i^{new} are the outputs of testing data in experiment I, II (when $i=1,2$), which represent the performance of network in traditional FNN and new FNN respectively. Y is the desired output of network's performance. The output reflects the dynamic change of network operating situation over a period of time.

For the new model, we extracted 16 rules in fuzzy rule base on Experiment I, thus its original structure is: 3-9-16-1. On experiment II, we extracted 27 rules and the structure is: 3-9-27-1.

The error rate of the model is defined as:

$$err = \frac{1}{n} \cdot \frac{\sum |y_i - Y_i|}{y_i} \cdot 100\%$$
(4.2)

Through the simulation of the traditional five-layer FNN and new FNN, it can be found that the former is about twice the latter of the training time. As for error rate, we can see the advantages of new model more intuitive through Table 3.

	Training set size	
Type of method	20 pairs	40 pairs
Traditional FNN	4.36%	3.66%
New FNN	1.31%	0.45%

Table 3 : comparison of the Error Rates between the Two Mmethods

Upon analysis and simulation, we can conclude that when the training set is sufficient and complete, the new FNN is superior to the traditional five-layer FNN, both in learning efficiency and prediction accuracy.

5. Conclusion

In this paper, we propose the concept of NOSA. Considering the network features of complexity, dynamics as well as ambiguity, we then establish the NOSA method based on the new FNN model, which completes the analysis of situation factors, the extraction of assessment rules as well as the prediction of operating situation very well and basically meets the requirements of NOSA. Therefore, it gives out a beneficial reference to applying the new FNN model in NOSA.

References

[1] H.Q. Wang, J.B. Lai, L. Zhu and Y. Liang. *Survey of Network Situation Awareness System*[J]. Computer Science, vol. 33 pp. 5-10(2006)

[2] D. Adenusi, B.M. Kuboye, B.A. Alese, A.F.B. Thompson. *Development of cyber situation awareness model*[C].Cyber Situational Awareness, Data Analytics and Assessment (CyberSA), 2015 International Conference on. IEEE, 1-11(2015)

[3] Y. Zhuo, Q. Zhang, Z. Gong. *Research and implementation of network transmission situation awareness*[C]. Computer Science and Information Engineering, 2009 WRI World Congress on. IEEE, 4: 210-214(2009)

[4] Y. Zhuo, Q. Zhang, Z. Gong. *Network situation assessment based on RST*[C]. Computational Intelligence and Industrial Application, 2008. PACIIA'08. Pacific-Asia Workshop on. IEEE, 1: 502-506(2008)

[5] X.Z. Chen, Q.H. Zheng, X.H. Guan, C.G. Lin. *Quantitative hierarchical threat evaluation model for network security*[J]. Journal of Software, 17(4):885−897(2006)(In Chinese).

[6] Leng, Gang, T.M. McGinnity, and G. Prasad. *An approach for on-line extraction of fuzzy rules using a self-organising fuzzy neural network*[J]. Fuzzy sets and systems 150.2: 211-243(2005)

[7] D.J. Kim, Z. Bien, and K.H. Park. *Fuzzy neural networks (FNN)-based approach for personalized facial expression recognition with novel feature selection method*[C]. Fuzzy Systems, 2003. FUZZ'03. The 12th IEEE International Conference on (Vol. 2, pp. 908-913). IEEE, 2003.

[8] R.J. Kuo, C.H. Chen, and Y.C. Hwang. *An intelligent stock trading decision support system through integration of genetic algorithm based fuzzy neural network and artificial neural network*[J]. Fuzzy sets and systems118.1: 21-45(2001)

[9] Nair, Vinod, and G.E. Hinton. *Rectified linear units improve restricted boltzmann machines*[C]. Proceedings of the 27th international conference on machine learning (ICML-10). 807-814(2010)

[10]L.X. Wang, J.M. Mendel. *Generating fuzzy rules by learning from examples*[J]. IEEE Transactions on systems, man, and cybernetics, 22(6): 1414-1427(1992)