

Research and Application of Key Technique of Artificial Intelligence to Process Information

Rong Xia

*Science and Technology College, Nanchang Hangkong University, NanChang,
JiangXi, China, 330034
2764284670@qq.com*

Abstract

The continuous development of science technology promotes the great development of the information processing technology based on artificial intelligence. Such technology refers to the combination of various technologies such as the computer technology, communication technique, control technique and others, which is applied to the information collection and processing of information. This thesis, research the key artificial intelligence technology to process information, including the artificial neural network, fuzzy theory and others. Besides, the information processing technology based on artificial intelligence will be applied to the design of the intelligent transportation system in this research. The result shows that the application of the intelligent transportation system to which the information processing technology based on artificial intelligence could relieve the stress during the rush hours efficiently and improve the efficiency of information processing.

Keywords: *Artificial intelligence; information processing; communication technology; artificial neural network; fuzzy theory*

1. Introduction

Artificial information processing technology was originated from 1930's and the research field of information processing is no longer limited to the researches about the general method and theory. At present, the research field gradually expands and more attention is paid to the exploring of new methods and theory. The major research object, the original linear, causal minimum phase system is replaced by the present non-linear, non-causal and non-phase system. In addition, the analysis and processing of time-varying signals became the new research field of modern information processing. The essence of the information processing technology based on artificial intelligence is the process and methods to change the incomplete, unreliable, inconsistent and uncertain information into complete, reliable, consistent and certain information ^[1]. Hence, the research about the information processing technology based on artificial intelligence is of great significance.

2. Key Theory and Methods of Information Processing Information Based on Artificial Intelligence

The information processing based on artificial intelligence is actually the comprehensive application of various science technologies, including expert system, machine learning, artificial neural network, fuzzy system theory, evolutionary calculation, multi-information fusion and other theories and methods ^[2].

2.1. Artificial Neural Network

The artificial neural network is actually a calculation mode that is established in accordance with the structure and operating principles of human neural networks in biology. Besides, such network is a parallel distributed system in which the components are inter-connected. Via imitating the functions of human brain to learn, memorize and dealing with problems, the neural network could acquire relevant knowledge, store knowledge in the connections, make adjustment to the knowledge stored in the system and cope with problems based on the acquired knowledge.

2.1.1. Artificial Neural Network Model

Artificial neuron is the basic processing unit of neural network and the novel is similar to the biological neuron. It is composed of three basic parts ^[3]: connection, summation and activation function, as shown in the Figure 1.

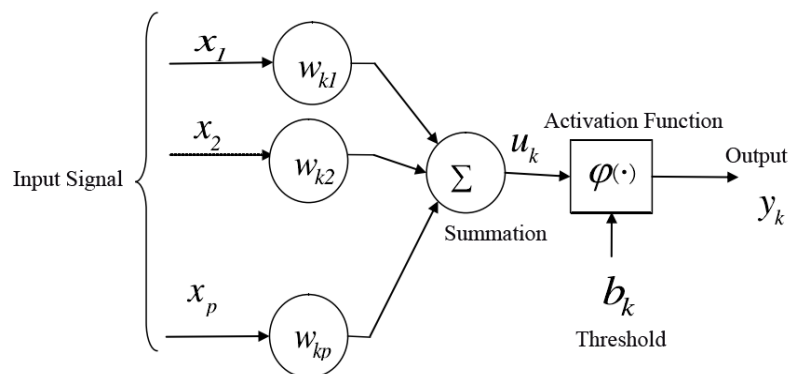


Figure 1. Basic Neuron Model

In the above graph, $X = (x_1, x_2, \dots, x_n)^T$ is the input signal, $W = (w_1, w_2, \dots, w_n)^T$ is the connection value of the neuron; u_k is the sum of the input signals; b_k is the threshold; $\phi(\bullet)$ is the activation function; y_k is the input signal.

$$u_k = \sum_{i=1}^p w_{ki} x_i \quad (1)$$

$$y_k = \phi(u_k - b_k) \quad (2)$$

2.1.2. Learning Method of Artificial Neural Network

One of the important features of artificial neural network is the access of the information from the external and the adjustment to its parameters (including the weight value and threshold) in accordance with the range that was predetermined. Hence, the functions of the neural networks could be continuously improved via the automatic adjustment to the system parameters in accordance with the features and property of the information to be processed. The learning methods of the artificial neural networks could be classified into the following two types:

(1) Supervised Learning (Learning from Teachers) refers to the input data and expected output data that was provided by the external environment for the neural network. Afterwards, the network will calculate the actual output data in accordance with the input data, compare the actual output data and the expected output data, as well as adjust the system parameters in accordance with the difference. At last, the actual output data could satisfy the relevant needs. As shown in the Figure 2.

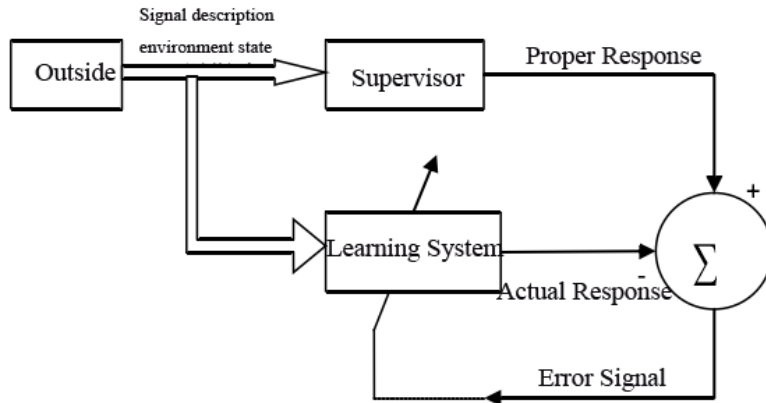


Figure 2. Schematic Flow of Supervised Learning

(3) Unsupervised Learning (Self Learning): In the process of self-learning, there is only input data and no equivalent expected data is output. Hence, it is impossible to make adjustment to the system parameters via calculating difference. Actually, the self-learning makes it possible to adjust the system parameters automatically in accordance with the statistical rules of the output data provided by the external environment. As shown in Figure 3:



Figure 3. Schematic Flow of Unsupervised Learning

2.1.3. Learning Calculation of Artificial Neural Network

The learning methods of artificial neural network are based on mathematics model and network model. In the mathematics model, the structure of the artificial neuron is established according to the neurons in human brain, which was the basic unit of information processing. In a mathematics model, there are plenty of artificial neurons which are combined together to form a complicated neural network to complete various functions. There are two kinds of learning methods of common artificial neural networks.

(1) Fault Rectification Learning: Given that at the time n , the input of k neuron is $x_k(n)$, the expected output is $d_k(n)$ and the actual output is $y_k(n)$, the error signal of output is $e_k(n) = d_k(n) - y_k(n)$. In order to minimize the difference between the actual output value and the expected output value, a difference signal $e_k(n)$ should be preset to be used as the argument function, which could be named as object function. Hence, the object function could be minimized and the fault rectification learning could become a process of solving the optimal value. The common object function includes:

$$J = E(1/2 * \sum_k e_k^2(n)) \quad (3)$$

In the above formula, J represents the expectation operator of the object function. In most cases, the instantaneous value of J at the time of n , $\varepsilon(n)$, could replace J .

$$\varepsilon(n) = 1/2 * \sum_k e_k^2(n) \quad (4)$$

In that case, in order to minimize the difference between the actual output value and the expected output value, the object function $\varepsilon(n)$ should be minimized, the question is changed into the solution extreme of weight value w , under the condition of $\varepsilon(n)$.

(2) **Competitive Learning:** The term competitive learning refers to the learning process of neural network in which the output of every neuron completes with others and the one with the largest or smallest output win the competition. After that, the winner neuron will be activated while the output of other neurons will be restrained. The rule in which the winner is activated and the losers are restrained was named as competitive learning, which could be presented in the following formula.

$$\Delta w_{kj}(n) = \begin{cases} \eta(x_j - w_{ji}) & \text{in case of neuron } j \text{ wins the competition} \\ 0 & \text{in case of neuron } j \text{ fails the competition} \end{cases} \quad (5)$$

2.1.4. Feedforward Neural Network

(1) Linear Threshold Unit

The neural network model could be decomposed into many units and each of them is composed of a neuron and an adjustable weight connected with the neuron. Such a unit is called a linear threshold unit, which could be presented by the M-N Model. Given the linear threshold unit, the input is an n dimensional variable and the output is a 1 dimensional variable. The activation function is a threshold function, as shown in the Figure 4.

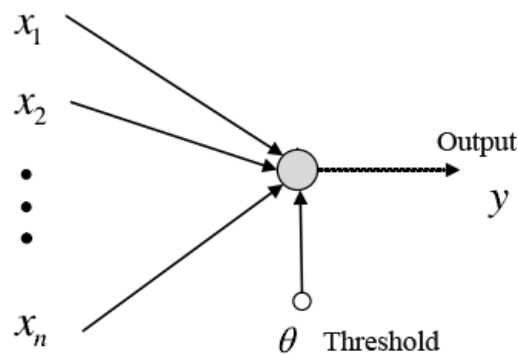


Figure 4. Linear Threshold Unit

The mathematical formula of the above graph is as followed:

$$x_j \in R, X = (x_1, x_2, \dots, x_n)^T \quad (6)$$

$$x_j \in R, W = (w_1, w_2, \dots, w_n)^T \quad (7)$$

In case of $v = \sum_{i=1}^n w_i x_i - \theta, \theta \in R$ is the threshold, $y = \text{sgn}(v)$, and $\text{sgn}(v) = \begin{cases} +1, v \geq 0 \\ -1, v < 0 \end{cases}$

The relevant research reveals that the linear threshold unit could calculate any Boolean product or sum. The feed-forward network composed of linear threshold units could be used to calculate any Boolean functions.

(2) Multiple-Layer Feed-forward Network

Multiple-layer feed-forward Network is composed of input layer, hidden layer and output layer. The hidden layer could be one or more layers. The single layer feed-forward network could only solve separable problems but the multiple-layer feed-forward Network with hidden layer could cope with non-linear separable issues. The Figure 5 shows a simple three-layer feed-forward neural network.

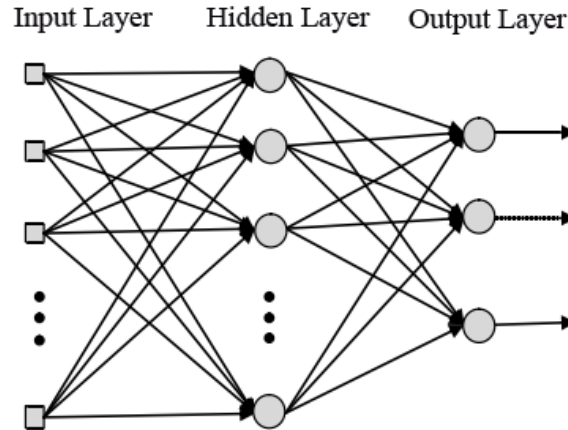


Figure 5. Three-Layer Feed-Forward Neural Network

In the feed-forward neural network, if the activation functions in the calculation layer is a continuous function (such as sigmoid function), the final result output of the output layer could be extremely similar to a continuous function. To be specific, given that a feed-forward neural network has a P input and a Q output, the whole neural network could be regarded as a non-linear mapping from P dimensional space to Q dimensional space.

(3) BP Calculation

In specific situation, the feed-forward neural network could be similar to any continuous function so that it is possible to use supervise learning to train feed-forward neural network in accordance with the Difference-Minimizing Principle until the network is capable of automatic adaption and solving. In addition, it is necessary to consider how to adjust weight and to determine the own parameters of the network and others. The signal flow of BP network is show in the Figure 6.

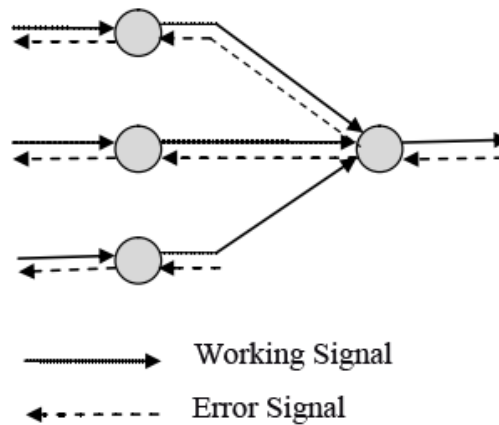


Figure 6. Signal Flow of BP Network

Error signal refers to the difference between the actual output signal and the expected output signal of network. The signal of such difference is transferred from the output port and calculated. At last, the network parameters are adjusted accordingly.

Given that in the n iteration, the NO.J neuron in the output layer is $y_j(n)$ so the difference signal of that unit is

$$e_j(n) = d_j(n) - y_j(n) \tag{8}$$

Given that the average error is $\frac{1}{2}e_j^2(n)$, the instantaneous value of squared error is

$$\varepsilon(n) = \frac{1}{2} \sum_{j \in m} e_j^2(n) \quad (9)$$

In that formula, m represents the number of the neuron at the output layer. Given that there are n groups training samples, the average errors is

$$\varepsilon_{AV} = \frac{1}{N} \sum_{n=1}^N \varepsilon(n) = \frac{1}{2N} \sum_{n=1}^N \sum_{j \in m} e_j^2(n) \quad (10)$$

ε_{AV} is the object function of network learning and aim of learning is to minimize the value of ε_{AV} . Besides, the BP calculation is the continuous calculation of the differences between the actual output signal and the expected output. During such a process, the wright is adjusted in accordance with the difference until the object ε_{AV} is minimized. In the whole learning process, the gradient information adjustment method could be used to adjust the wright in accordance with the following formula. Given that is the connection between Neuron i to Neuron j, then

$$\Delta = w_{ji} = -\eta \delta_j(n) y_i(n) \quad (11)$$

In that formula, η is the step size of learning; y_i is the input signal from Neuron i to Neuron j; δ_j is the local gradient and the minus sign represents that the weight correction is moving downward along the gradient.

2.2. Fuzzy Neural Network

The fuzzy theory is applicable for the research of the uncertain phenomena. The uncertainty of the phenomena is determined by the essence of such phenomena so it is not the research objects of dualism principle in the traditional mathematics. Hence, it is difficult to classify the research objects clearly and the objects are determined to be uncertain. The central part of the fuzzy theory is the fuzzy system which is a dynamical system with the capability of coping with fuzzy information. The fuzzy system is composed of four parts: the fuzzy rules database, fuzziness generating device, anti-fuzziness device and fuzziness predicting device.

2.2.1. General Model of Fuzzy Neural Network

The fuzzy logical system refers to the system which is closely related with fuzzy concepts and fuzzy logic [4]. Such a system is composed of four parts, including fuzziness generating device, fuzzy rules database, fuzziness predicting device and anti-fuzziness device, as shown in Figure 7.

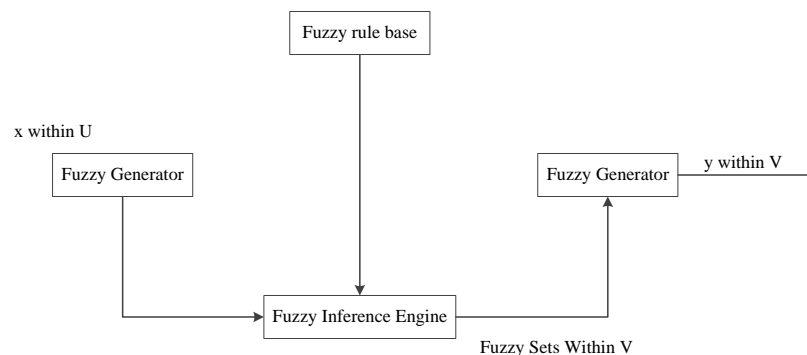


Figure 7. Fuzzy Logical System

Each point in the Domain U is mapped to the fuzzy collection in the Domain U via the fuzziness generating device. The anti-fuzziness device maps the fuzzy collection in the

Domain V map to certain points in the Domain V. The fuzziness predicting device is capable of predicting the fuzzy conclusion in accordance with the fuzziness predicting knowledge in the fuzzy rule database and the fuzzy collection generated by the fuzziness generating device. That is to say, the fuzziness predicting knowledge in the fuzzy rule database and the fuzzy collection generated by the fuzziness generating devices could be used to predict the fuzzy theories by means of the fuzziness generating device, that is, the fuzzy collection in the Domain V, which is further input into the anti-fuzziness device.

2.2.2. Fuzzy Rule Database

The fuzzy rule database is composed of a series of fuzziness predicting rules and the forms of such rules could be presented by the following formula:

$$R^{(l)} : \text{If } x_1 \text{ is } F_1^l, \dots, x_n \text{ is } F_n^l \text{ then } y \text{ is } G^l \quad (12)$$

In that formula, F_1^l is the fuzzy collection of $U_i \subset R$ and G^l is the fuzzy collection of $V \subset R$. Besides, $x=(x_1, \dots, x_n) \in U_1, \dots, U_n$ and $y \in V$ are language variables. In addition, $l=1, 2, \dots, M$, which means that M represents the sun of all the rules. It is obvious that x and y represent the input and output of the logic system. The multiple input and output system could be decomposed into multiple-input single-output systems so only the multiple-input single-output systems are taken into the consideration of this thesis [5].

2.2.3. Fuzziness Predicting Device

The fuzziness predicting device carries out the function of converting the “If-then” rule in the fuzzy rule database into specific map in accordance with fuzzy logical rules. That is to say, such device functions to map the fuzzy collection in the formula $U=U_1 * U_2 * \dots * U_n \subset R^n$ into the fuzzy collection in Domain V. The fuzzy rules are shown in Graph 6.

Such a process could be shown through the following formula in which the fuzzy container functions $F_1 * F_2 * \dots * F_n \rightarrow G^l$ exist in the Domain $U * V$. Given that the fuzzy collection A is the input of fuzziness predicting device, if the synthetic calculation sup-* is used, the fuzzy collection B l in Domain V deducted from each fuzziness predicting rule is shown in the following formula:

$$u_{B^l}(y) = \sup_{x \in U} (u_{F_1^l * \dots * F_n^l}(x, y) * u_A(x)) \quad (13)$$

2.2.4. Fuzziness Generating Device and Anti-fuzzy Device

The function of the fuzziness generating device is to map a certain point in the Domain $U \subset R^n$, $x=(x_1, \dots, x_n)T$, to a fuzzy collection A' . There are two ways of mapping:

(1) Single-value fuzziness generating device: if the fuzzy collection A is a fuzzy single value for the supporting collection x, the statement of $x'=x$ and $u_{A'}(x) = 1$ is true to specific point. For the rest of $x' \neq x$, the statement of $u_{A'}(x) = 0$ is true.

(2) Multiple-value fuzziness generating device: in case of $x'=x$, $u_{A'}(x) = 1$. However, when the differences between x and x increases, the value of $u_{A'}(x)$ decreased from 1.

In most applications, the first method, the single-value fuzziness generating device is adopted [6].

2.3. Evolution Calculation

The evolution calculation s established in accordance with the natural selection principle in the nature and provides important instructions about the refinement of machine and learning. Such calculation could be named as genetic algorithm by which the research objects are scanned in a optimal way in accordance with the genetic features of various creatures. Such calculation is comparatively simple and applicable in a wide scope. By means of such calculation, relevant information could be parallel processed.

Due to the fact that the major research object of the genetic algorithm is individual phenomenon, which are processed by means of selection, intersection and transformation, such calculation could be different from the traditional calculations. During the recent years, many scholars have made great efforts to apply the evolution calculation to various fields such as recognize graphs, machine learning and automatic control. Hence, such a method could be used as the most common methods in the information processing via computer intelligence.

2.4. Information Fusion Technology

Due to the lack of the application of actual multiple resource information system, a complete theoretical system is not established despite of the fact that the information fusion technology has attracted much attention and been applied to various fields. For the sake of faster development of such technology, it is necessary to make further research about theories. In accordance with the difficult parts of the present information fusion field, the following research topics are worth investigating^[7-8].

(1) Collection, organization and management of the data in the information fusion system.

(2) General theory of the distributive multiple-targets sensor tracking based on the concept of information scheme.

(3) Application of artificial intelligence to information fusion.

(4) The establishment of an index system to evaluate the performance of information fusion system.

Specifying the fusion of distributive system C3I so as to lay foundation for the comprehensive information system of research fields and the combat platform arrangement C³I.

3. Application of Artificial Intelligence to Information Processing

The transportation system is a complex typical distributive system which is characterized by non-linearity, fuzziness and uncertainty. The ITS (Intelligent Transportation System) is a hot research field, the primary concern of which is to investigate the application of advanced intelligence technology, automatic control technique, computer technique, information and communication technique, as well as the electronic sensor techniques^[9].

3.1. Application of BP Neural Network to the Predication of Traffic Flow

During the process of prediction, the east-to-west traffic flow statistics of an intersection in Zhenjiang City from 7:00 to 21:00 in May, 2015 is investigated. The traffic flow in every single quarter is regarded as one unit and every 8 units was regarded as an input sample and listed in the table. That is to say, after inputting 8 neurons, the statistics of the traffic flow is used as an output sample. In other words, a neuron in the output layer forms a sample pair. After the first statistic of traffic flow is removed, the one following that statistic will take its place. Eight statistics will form an input sample. Then the next statistic of traffic flow will be used as an output sample and further forms a sample pair. The normalized value of element could be shown by the following formula^[10]:

$$a = (x_i - \min(X)) / (\max(X) - \min(X)), i = 1, 2, \dots, n \quad (14)$$

As for the BP Neural Network Calculation, it is necessary to set up the parameters from the perspectives of number of network layers, original weight, learning rate and the standards to judge differences *etc.*

Even though the application of the traditional BP neural network could lead to the reaching of expected value, the convergence rate is slow and the training time is long. Hence, the improved BP neural network calculation is applied.

3.2. Improving BP Neural Network to be applied to the Prediction of Traffic Flow

In such a research, the improved method of Levenberg-Marquardt in the BP neural network is employed and the research direction could be shown through the following formula.

$$S(X^{(k)}) = -(H^{(k)} + \lambda^{(k)}I)^{-1} \nabla f(X^{(k)}) \quad (15)$$

After the function `initff` is used to initialize the weight and threshold of the network, the function `trainlm` is used to train the network. The result of such calculation is the relationship between the square errors and the number of training steps. As shown in Figure 8.

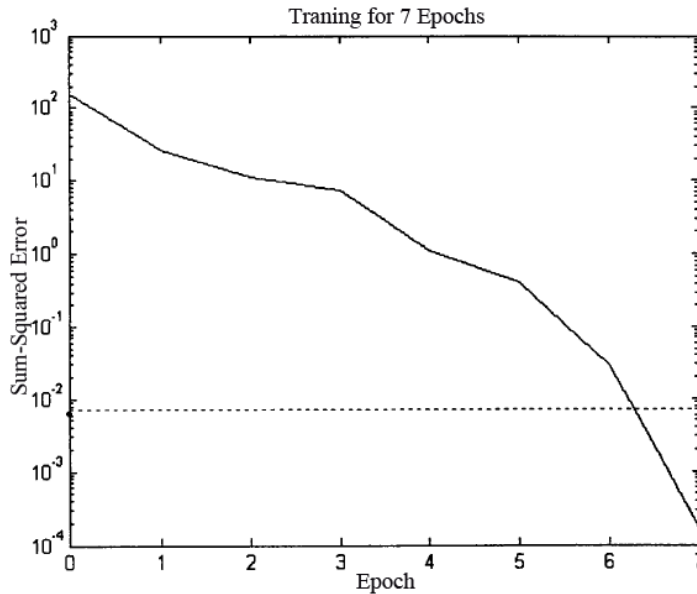


Figure 8. Relationship between the Square Errors and the Number of Training Steps

The weight and threshold of the network after training are as followed:

```
w1=[
-1.9511 5.1513 5.4712 6.0927 -1.4154 -4.5452 -2.6301 -2.2957
4.1814 2.6375 1.0407 3.0318 4.1252 -1.4267 -1.1518 4.6138
.
.
3.2831 3.3638 -3.9925 0.3732 -0.9538 -4.9724 3.6667 -2.1737
-3.5586 1.6954 4.7412 -1.1795 -0.3883 3.7689 -2.3702 3.5086 ]
b1=[
-5.0538 -7.3815 3.0091 6.1634 5.1800 -1.7538 1.2221 1.1930
-3.9142 -2.9588 1.1508 6.2020 8.4618 -6.6427 -1.8875 -8.3998
],
w2=[
1.2164 -0.2369 1.8636 0.5041 -1.3616 1.0431 -0.9159 0.0885
0.6383 0.0821 0.0636 1.0275 -0.9151 -0.2444 -0.4101 1.2574
-1.1256]
b2=[0.6582]
```

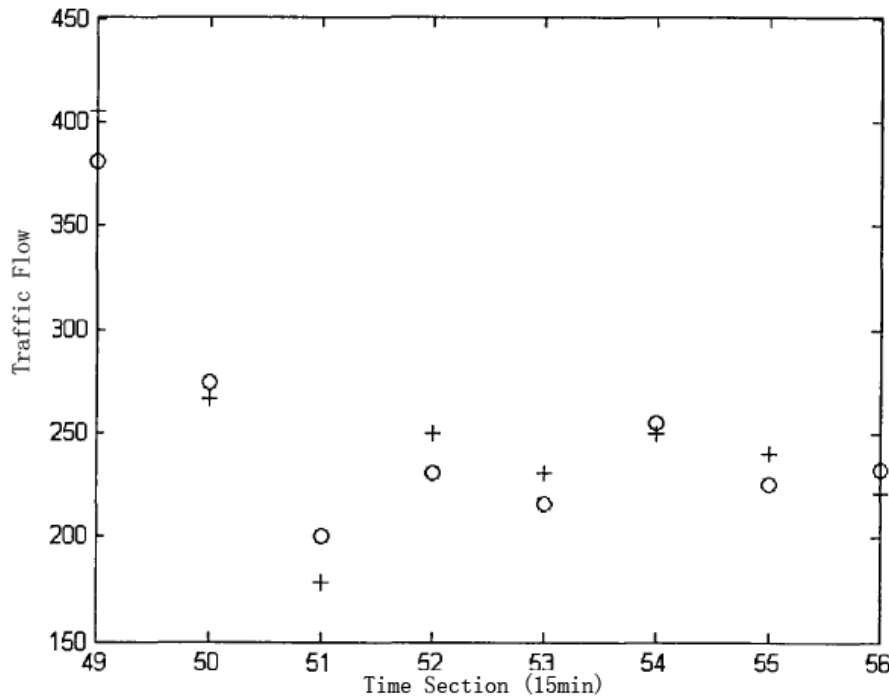


Figure 9. Result of the Prediction of the Employment of the Levenberg-Marquardt Method

As shown in Graph 7, the convergence speed of the Levenberg-Marquardt method is the fastest. According to the relevant experiences, the employment of the Levenberg-Marquardt method could lead to satisfying results when it is used to solve most problems.

3.3. Simulation Result of the Prediction of Traffic Flow

The network made up of intelligent neurons is called high-order neural network, which is capable of completing the following two tasks: the first capability of such network is called EIC (External Intelligent Capability), which refers to the ability to making adjustment to the strength of the connection among neurons. Such intelligence is shown through the network that is composed of plenty of neurons, so it is named as EIC. The second kind of ability is IIC (Internal Intelligent Capability), which is shown though the reversible function $F(\cdot)$ of a neuron and irrelevant to the connection among neurons. The information of high-ordered neural network is stored in the joint weight and transfer function. However, the traditional neurons are incapable of storing information [11].

The input vectors of the four nodes at the input layer are as followed: the traffic vector of the section of research t-80s, the traffic flow of the section t-60s, the traffic flow of the section t-40s and the traffic flow of the section t-20s. The node at the output layer is the traffic flow of the section of research ts.[9-10]

The number of the nodes at the hidden layer is 6, which could be adjusted in accordance with the actual situation of simulation.

The simulation statistics is collected from an intersection in Zhenjiang city in the rush hour from 8:00 to 12:00 in May, 2015. The loop detector buried under the road is used to detect the number of the vehicles passing this section every 20 seconds and the number is marked by k ($k = 0, 1, 2, \dots, 26$). The detection last for 720 units of detection.

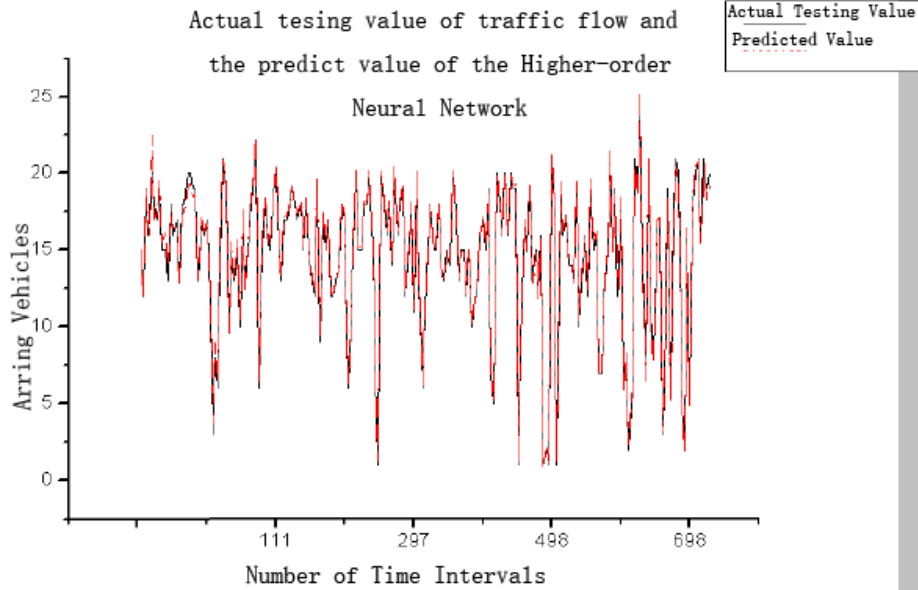


Figure 10. Result of the Traffic Flow Simulation of High-Order Neural Network

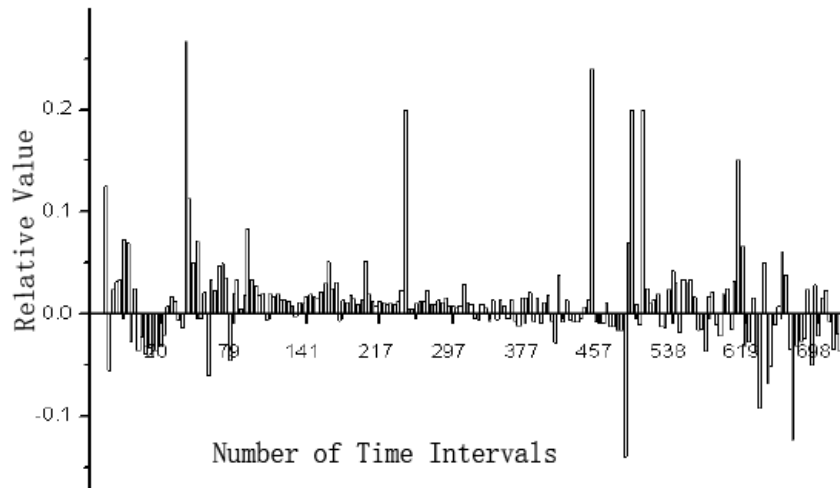


Figure 11. Relative Errors of the Traffic Flow Prediction of High-Order Neural Network

In the above graph, $EC=0.94561$, $tpred=31s$, the average relative error is 3.07%. The number of the intervals among the high errors in relative errors that exceeds 20% is 6.

The simulation result shows that the prediction time of BP neural network is shorter and the precision is improved. The number of the intervals among high errors is within the acceptable range. It is obvious that the application of the high-order neural network is applicable to the traffic flow prediction^[12].

4. Conclusion

Even though the intelligent information processing technology has been developing fast, there are still some problems waiting to be solved and investigated. In order to make sure that such technology could be improved and applied to productive activities, it is necessary to combine such technology with the advanced science and to investigate the

issues in productive practice, which should be summarized as the theories and methods. Due to the fact that the information system to be processed becomes more and more complicated, the dependence on single intelligent information system has both advantages and disadvantages so it is difficult to satisfy all the needs of complex system. As a result, it could be predicted that the necessary trend is to combine various intelligent information processing technology and to investigate comprehensive intelligent information processing technology in the future.

References

- [1] Incomplete information: Rough set analysis[M]. Physica, (2013).
- [2] N. J. Nilsson, "Principles of artificial intelligence", Morgan Kaufmann, (2014).
- [3] E. Wenger, "Artificial intelligence and tutoring systems: computational and cognitive approaches to the communication of knowledge", Morgan Kaufmann, (2014).
- [4] E. Wenger, "Artificial intelligence and tutoring systems: computational and cognitive approaches to the communication of knowledge", Morgan Kaufmann, (2014).
- [5] T. H. Davenport, "Process innovation: reengineering work through information technology", Harvard Business Press, (2013).
- [6] L. C.Lynnworth, "Ultrasonic measurements for process control: theory, techniques, applications", Academic press, (2013).
- [7] L. J. Krajewski, L. P. Ritzman and M. K. Malhotra, "Operations management: processes and supply chains", Pearson, (2013).
- [8] Á. Rebuge and D. R. Ferreira, "Business process analysis in healthcare environments: A methodology based on process mining", Information Systems, vol. 37, no. 2, (2012), pp. 99-116.
- [9] "Soft Computing in Information Retrieval: techniques and applications", Physica, (2013).
- [10] X. S. Yang and X. He, "Bat algorithm: literature review and applications", International Journal of Bio-Inspired Computation, vol. 5, no. 3, (2013), pp. 141-149.
- [11] A. Rehman and T. Saba, "Evaluation of artificial intelligent techniques to secure information in enterprises", Artificial Intelligence Review, (2014), vol. 42, no. 4, pp. 1029-1044.
- [12] "Readings in distributed artificial intelligence", Morgan Kaufmann, (2014).

Author



Rong Xia, she was born in January of 1982 in Luzhou city, Sichuan province. She graduated from University of Electronic Science and Technology of China and got the Bachelor Degree. Now, she is a lecturer of Science and Technology College of Nanchang Hangkong University, and her main research direction is computer software design and database.