Fuzzy neural network based on a Sigmoid chaotic neuron*

Zhang Yi(张 诣) and Wang Xing-Yuan(王兴元)†

School of Electronic & Information Engineering, Dalian University of Technology, Dalian 116024, *China*

(Received 5 May 2011; revised manuscript received 31 August 2011)

The theories of intelligent information processing are urgently needed for the rapid development of modem science. In this paper, a novel fuzzy chaotic neural network, which is the combination of fuzzy logic system, artificial neural network system, and chaotic system, is proposed. We design its model structure which is based on the Sigmoid map, derive its mathematical model, and analyse its chaotic characteristics. Finally the relationship between the accuracy of map and the membership function is illustrated by simulation.

Keywords: chaos, fuzzy, neural network

PACS: 05.45.Gg, 84.35.+i **DOI:** 10.1088/1674-1056/21/2/020507

1. Introduction

In the course of the rapid development of modern science and technology, theories and tools of intelligent information processing have become urgently required and should lead to breakthroughs in knowledge representation, model calculation, computational efficiency, and adaptive learning. The borderline subjects of fuzzy logic systems, artificial neural network systems, and chaotic systems are devoted to the research on the human brain. There are many fuzzy concepts in human thinking. For example, the human brain can be modeled as a continuous-time nonlinear dynamic system by an artificial neural network, or chaos theory could be used to understand some irregular activities in the human brain.

A fuzzy neural network[1−3] is the combination of fuzzy theory and a neural network. It has the characteristics of a neural network, such as self-learning, self-adaptive, and parallel processing of information. It also has the ability of fuzzy inference to process language information. A chaotic neural network[4−8] is the combination of chaos theory and a neural network. It not only has the characteristics of a neural network but also a searching ability by using chaos. Therefore, it can better simulate the information processing in the human brain.

A fuzzy neural network is a static network, and a chaotic neural network cannot deal with fuzzy language information. The fuzzy chaotic neural network is the combination of fuzzy, chaos, and a neural network,[9,10] and has great potential and promising prospects. The bifurcation parameter in a chaotic neural network is controlled by fuzzy logic rules, which is the earliest combination of fuzzy logic, artificial neural networks, and chaos theory.[11]

Fuzzy chaotic neural networks are grouped into three types according to their method of realization: (i) the chaotic map that is introduced in the fuzzy neural network and has been proposed in Ref. [12]; (ii) the chaotic neural network that is fuzzed and can be used to process fuzzy information;[13] (iii) the model of a fuzzy chaotic neuron that is based on a chaotic neuron and the concept of fuzzy number.[14]

In this paper, we construct a novel fuzzy chaotic neural network which belongs to the first type. It has a neural network structure, chaotic characteristics, and a fuzzy function. The network can combine fuzzy, chaos, and neural networks. This network is a dynamic fuzzy neural network model, so it can approach the dynamic system. The chaotic map is introduced into it, so the ability of chaos recollection is enhanced. The constant transition between chaos and non-chaos in this network is controlled by the fuzzy rules, so the network can stay in a certain model.

*Project supported by the National Natural Science Foundation of China (Grant Nos. 61173183, 60973152, and 60573172), the Superior University Doctor Subject Special Scientific Research Foundation of China (Grant No. 20070141014), and the Natural Science Foundation of Liaoning Province, China (Grant No. 20082165).

†Corresponding author. E-mail: wangxy@dlut.edu.cn

© 2012 Chinese Physical Society and IOP Publishing Ltd

2. Fuzzy chaotic neural network

2.1. Sigmoid chaotic neural network

The kinetic equation of a Sigmoid chaotic neuron is[15]

$$x(t+1) = kx(t) - \alpha f x(t) + \alpha(t),$$
$$y(t+1) = f(x(t+1)), \tag{1}$$

where x is the internal state of neuron, y is the output signal of neuron, $f(x) = 1/(1 + \mathrm{e}^{-\beta x})$, $a(t) = (A(t) - \theta)(1 - k)$, A is the strength of external input, θ is the threshold of neuron, $k \in [0, 1]$ is refractoriness decay rate, and α is proportional coefficient.

The mathematical model of a Sigmoid neural network is

$$\begin{aligned} x_i(t+1) &= kx_i(t) + \sum_{j=1}^{n} T_{ij} h_j(f(x_j(t))) \\ &\quad + \sum_{j=1}^{q} \omega_{ij} A_j(t) - \alpha(f_i(x_i(t))) \\ &\quad - \theta_i(1-k), \\ y_i(t+1) &= f_i(x_i(t+1)), \end{aligned} \tag{2}$$

where $i = 1, 2, \ldots, n$, with n being the number of chaotic neurons in the network, q is the number of external inputs, T_{ij} is the connection weight of i-th neuron and j-th neuron, ω_{ij} is the connection weight of the j-th external input and the i-th neuron, f_i is the continuous output function of the i-th neuron, h_i is the internal feedback of the j-th neuron to the i-th neuron.

2.2. Structure of network

This fuzzy chaotic neural network is the combination of a Sigmoid chaotic neuron and fuzzy theory. It can realize fuzzy functions, such as the generation of membership functions, normalization, rule-based reasoning and de-fuzzy. The model of this system is the mathematical expression of a fuzzy logic system, and its state evolution has chaotic characteristics.

The model of a fuzzy chaotic neural network is shown in Fig. 1. There is premise part and a conclusion part in this model. The premise part is made up of a layer of generating membership functions, a product layer, and a normalization layer. The conclusion part is composed of an input layer, a rule conclusion layer, and an output layer. The weights of the rule conclusion layer and the output layer are produced by the normalization layer of the premise part. The rule conclusion layer is realized by the chaotic neural network, which can be constituted by a one layer or a multi-layer network.

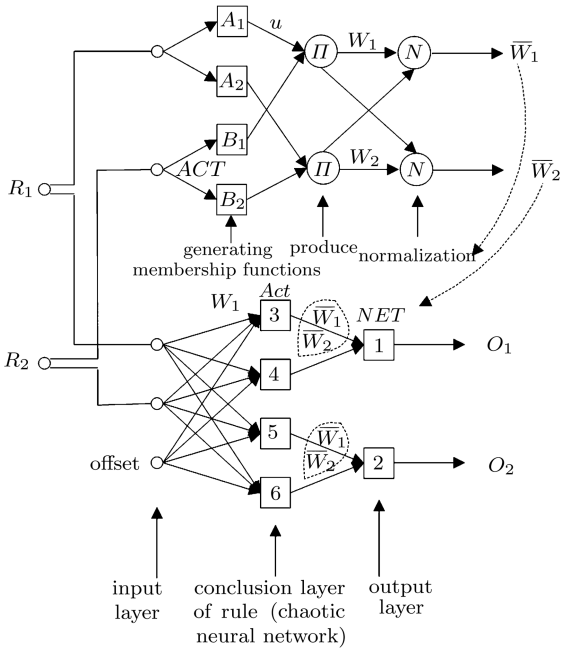


Fig. 1. Model of a fuzzy chaotic neural network.

2.3. Deduction of the mathematical model

2.3.1. Mathematical model of the premise part

The layer of generating membership functions is described by a Gaussian function,[16] so the membership function is

$$u_{qm}(R_q(\tau+1)) = \exp\left(-\left(\frac{R_q(\tau+1) - st_{qm}}{\sigma_{qm}}\right)^2\right). \tag{3}$$

The product layer output is

$$W_m(\tau+1) = \prod_q u_{qm}(\tau+1). \tag{4}$$

The output of the normalized layer is

$$\bar{W}_m(\tau+1) = \frac{W_m(\tau+1)}{\sum\limits_m W_m(\tau+1)}. \tag{5}$$

2.3.2. Mathematical model of the conclusion part

The input of the conclusion layer is

$$net_{mp}(\tau+1) = \sum_l w_{mpl}(\tau+1) act_{ml}(\tau+1). \tag{6}$$

The output of the conclusion layer is

$$ACT_{mp}(\tau+1) = f_{mp}[x_{mp}(\tau+1)],$$
$$f_{mp}(x) = \frac{1}{(1 + \mathrm{e}^{-\beta_{mp} x})}. \tag{7}$$

The internal state of the conclusion layer is

$$x_{mp}(\tau+1) = kx_{mp}(\tau) + \sum_{rl} T_{mprl} h_{rl} f_{mp}(x_{rl}(\tau)) + net_{mp}(\tau) - \alpha f_{mp}(x_{mp}(\tau)) - \theta_{mp}(1-k). \tag{8}$$

The input of the output layer is

$$\begin{aligned} NET_j(\tau+1) &= \frac{\sum\limits_m W_m(\tau+1) Act_{mp}(\tau+1)}{\sum\limits_m W_m(\tau+1)} \\ &= \frac{W_1(\tau+1)Act_{1j}(\tau+1) + W_2(\tau+1)Act_{2j}(\tau+1) + \cdots + W_m(\tau+1)Act_{mj}(\tau+1) + \cdots}{W_1(\tau+1) + W_2(\tau+1) + \cdots + W_m(\tau+1) + \cdots}. \end{aligned} \tag{9}$$

The network output is

$$\begin{aligned} O_j(\tau+1) &= F_j(NET_j(\tau+1)), \\ F_j &= \frac{1}{(1+\mathrm{e}^{-\beta_j NET_j})}, \end{aligned} \tag{10}$$

where W_m is the m-th membership value of the m-th rule, Act_{mp} is the p-th output of the m-th rule, act_{ml} is the l-th input of m-th rule, T_{mprl} is the connection weight of the rt-th chaotic neuron to the mp-th chaotic neuron, ω_{mpl} is the connection weight of the l-th external input to the mp-th chaotic neuron, h_{rt} is a characteristic function of axon transform of the rt-th chaotic neuron, that is, internal feedback, and expressed as

$$h(x) = \begin{cases} 1, & x \geq 0.5, \\ 0, & x < 0.5. \end{cases} \tag{11}$$

In formula (10), we can see that the final output of this network is a mathematical expression of the fuzzy logic system, so this system can realize the function of the fuzzy system.

2.4. Learning algorithm of a fuzzy chaotic neural network

The learning process is to minimize the squared error and E by steepest descent algorithm. Premise part and conclusion part are trained at the same time. The training process of the conclusion part is a dynamic recursion process, so this learning algorithm is dynamic.

The definition of the error function is

$$E = \frac{1}{2}\sum_{j=1}^{L} (\bar{O}_j - O_j)^2, \tag{12}$$

weight coefficient should make the adjustment in the opposite direction of the E function gradient, and then the network output can be close to the expected output.

The renewal process of parameters (s t or σ) in the premise part is

$$\begin{aligned} \Delta s(\tau) &= -\eta_s \frac{\partial E(\tau)}{\partial s(\tau)} \\ &= -\eta_s \frac{\partial E(\tau)}{\partial O_j(\tau)} \cdot \frac{\partial O_j(\tau)}{\partial s(\tau)} \\ &= -\eta_s \frac{\partial E(\tau)}{\partial O_j(\tau)} \cdot \frac{\partial O_j(\tau)}{\partial NET_j(\tau)} \\ &\quad \times \frac{\partial NET_j(\tau)}{\partial W_m(\tau)} \cdot \frac{\partial W_m(\tau)}{\partial s(\tau)} \\ &= -\eta_s \frac{\partial E(\tau)}{\partial O_j(\tau)} \cdot F_j'(NET_j(\tau)) \\ &\quad \cdot \left[\frac{\partial}{\partial W_m(\tau)} \frac{\sum\limits_m W_m Act_{mp}(\tau)}{\sum\limits_m W_m} \right] \cdot \frac{\partial W_m(\tau)}{\partial s(\tau)} \\ &= -\eta_s \left\{ -(\bar{O}_j(\tau) - O_j(\tau)) F_j'(NET_j(\tau)) \right\} \\ &\quad \cdot \frac{Act_{mp}(\tau) - NET_j(\tau)}{\sum\limits_m W_m} \cdot \frac{\partial W_m(\tau)}{\partial s(\tau)}, \end{aligned} \tag{13}$$

where η_s is the learning rate, and $\bar{O}_j$ is the expected output.

The renewal process of the conclusion part is

$$\begin{aligned} \Delta\omega_{mpl} &= -\eta_{w_{mpl}} \frac{\partial E(\tau)}{\partial \omega_{mpl}} \\ &= -\eta_{w_{mpl}} \frac{\partial E(\tau)}{\partial NET_j(\tau)} \cdot \frac{\partial NET_j(\tau)}{\partial \omega_{mpl}} \\ &= -\eta_{w_{mpl}} \frac{\partial E(\tau)}{\partial O_j(\tau)} \cdot \frac{\partial O_j(\tau)}{\partial NET_j(\tau)} \cdot \frac{\partial NET_j(\tau)}{\partial x_{mp}(\tau)} \\ &\quad \times \frac{\partial x_{mp}(\tau)}{\partial net_{mp}(\tau)} \cdot \frac{\partial net_{mp}(\tau)}{\partial \omega_{mpl}} \\ &= -\eta_{w_{mpl}} \frac{\partial E(\tau)}{\partial O_j(\tau)} \cdot F_j'(NET_j(\tau)) \frac{W_m}{\sum\limits_m W_m} \\ &\quad \times f_{mp}'(x_{mp}(\tau)) \cdot \frac{\partial x_{mp}(\tau)}{\partial net_{mp}(\tau)} \\ &\quad \times \frac{\partial}{\partial \omega_{mpl}} \sum \omega_{mpl} act_{mp} \\ &= -\eta_{w_{mpl}} \left\{ -(\bar{O}_j(\tau) - O_j(\tau)) \right\} \end{aligned}$$

$$\times F'_j(NET_j(\tau))\frac{W_m}{\sum\limits_m W_m}\cdot f'_{mp}(x_{mp}(\tau))$$
$$\times \frac{\partial x_{mp}(\tau)}{\partial net_{mp}(\tau)}\cdot act_{mp}, \tag{14}$$

where

$$\frac{\partial x_{mp}(\tau)}{\partial net_{mp}(\tau)} = k\frac{\partial x_{mp}(\tau-1)}{\partial net_{mp}(\tau-1)} + 1 - \alpha f'_{mp}(x_{mp}(\tau-1)) \times \frac{\partial x_{mp}(\tau-1)}{\partial net_{mp}(\tau-1)},$$

so the training process is dynamic.

$$\begin{aligned}\Delta T_{mprt} &= -\eta_{T_{mprt}}\frac{\partial E(\tau)}{\partial T_{mprt}}\\ &= -\eta_{T_{mprt}}\frac{\partial E(\tau)}{\partial NET_j(\tau)}\cdot\frac{\partial NET_j(\tau)}{\partial T_{mprt}}\\ &= -\eta_{T_{mprt}}\frac{\partial E(\tau)}{\partial O_j(\tau)}\cdot\frac{\partial O_j(\tau)}{\partial NET_j(\tau)}\\ &\quad\times\frac{\partial NET_j(\tau)}{\partial Act_{mp}(\tau)}\cdot\frac{\partial Act_{mp}(\tau)}{\partial T_{mprt}}\\ &= -\eta_{T_{mprt}}\frac{\partial E(\tau)}{\partial O_j(\tau)}\cdot F'_j(NET_j(\tau))\cdot\frac{W_m}{\sum\limits_m W_m}\\ &\quad\times\frac{\partial Act_{mp}(\tau)}{\partial x_{mp}(\tau)}\cdot\frac{\partial x_{mp}(\tau)}{\partial T_{mprt}}\\ &= -\eta_{T_{mprt}}\frac{\partial E(\tau)}{\partial O_j(\tau)}\cdot F'_j(NET_j(\tau))\cdot\frac{W_m}{\sum\limits_m W_m}\\ &\quad\times f'_{mp}(x_{mp}(\tau))\cdot\frac{\partial x_{mp}(\tau)}{\partial T_{mprt}},\end{aligned} \tag{15}$$

where

$$\frac{\partial x_{mp}(\tau)}{\partial T_{mprt}} = \frac{\partial x_{mp}(\tau-1)}{\partial T_{mprt}} + \sum_{rt} h_{rt}(f_{mp}(x_{rt}(\tau-1))) - \alpha f'_{mp}(x_{mp}(\tau-1))\frac{\partial x_{mp}(\tau-1)}{\partial T_{mprt}},$$

so the adjustment of weight is a dynamic process.

3. Chaotic characteristics of a fuzzy chaotic neural network

The system of two chaotic neurons in the conclusion layer with a single input–single output is discussed; the equations of fixed points are

$$\begin{aligned}x_1 &= kx_1 + w_1A - \alpha_1\cdot\frac{1}{1+\mathrm{e}^{-\beta_1x_1}} - \theta_1(1-k_1) + T_{21}h_2\cdot\frac{1}{1+\mathrm{e}^{-\beta_2x_2}} + T_{11}h_1\cdot\frac{1}{1+\mathrm{e}^{-\beta_1x_1}},\\ x_2 &= kx_2 + w_2A - \alpha_2\cdot\frac{1}{1+\mathrm{e}^{-\beta_2x_2}} - \theta_2(1-k_2) + T_{12}h_1\cdot\frac{1}{1+\mathrm{e}^{-\beta_1x_1}} + T_{22}h_2\cdot\frac{1}{1+\mathrm{e}^{-\beta_2x_2}}.\end{aligned} \tag{16}$$

When

$$\mathrm{e}^{-\beta x}\approx\frac{1}{1+\beta x},$$

then

$$\frac{1}{1+\mathrm{e}^{-\beta x}} = \frac{1+\beta x}{2+\beta x}.$$

The fixed points of the system are

$$\begin{aligned}x_1^{1,2} &= \frac{-(2-2k_1+\alpha_1\beta_1-\beta_1d_1)}{2\beta_1(1-k_1)} \pm \frac{\sqrt{(2-2k_1+\alpha_1\beta_1-\beta_1d_1)^2-4\beta_1(1-k_1)(\alpha_1-2d_1)}}{2\beta_1(1-k_1)},\\ x_2^{1,2} &= \frac{-(2-2k_2+\alpha_2\beta_2-\beta_2d_2)}{2\beta_2(1-k_2)} \pm \frac{\sqrt{(2-2k_2+\alpha_2\beta_2-\beta_2d_2)^2-4\beta_2(1-k_2)(\alpha_2-2d_2)}}{2\beta_2(1-k_2)},\end{aligned} \tag{17}$$

where

$$\begin{aligned}d_1 &= w_1A - \theta_1(1-k_1) + T_{21}h_2\cdot\frac{1}{1+\mathrm{e}^{-\beta_2x_2}} + T_{11}h_1\cdot\frac{1}{1+\mathrm{e}^{-\beta_1x_1}},\\ d_2 &= w_2A - \theta_2(1-k_2) + T_{12}h_1\cdot\frac{1}{1+\mathrm{e}^{-\beta_1x_1}} + T_{22}h_2\cdot\frac{1}{1+\mathrm{e}^{-\beta_2x_2}}.\end{aligned} \tag{18}$$

The equation of the Jacobian matrix multiplier λ of this system is

$$|\lambda I - A| = \begin{vmatrix}\lambda - k_1 + \dfrac{\alpha_1\beta_1}{(2+\beta_1x_1)^2} & 0\\ 0 & \lambda - k_2 + \dfrac{\alpha_2\beta_2}{(2+\beta_2x_2)^2}\end{vmatrix} = 0. \tag{19}$$

When there exists at least one $|\lambda| > 1$, the fixed points are unstable. Next we will discuss an unstable condition with fixed points.

3.1. d_1 is the bifurcation parameter

(I) when $x_1 = x_1^1$,

$$d_1 > \max\left(\frac{[(1-k_1)(4\alpha_1\beta_1 - 2 - 2b_1) - c_1^2]}{\beta_1(9-8k_1)},\ \ -\alpha_1 + \frac{2(1-k_1)(1-b_1)}{\beta_1}\right),$$
$$d_1 < \min\left(\frac{[(1-k_1)(4\alpha_1\beta_1 - 2 - 2b_1) - c_1^2]}{\beta_1(9-8k_1)},\ \ \alpha_1 - \frac{2(1-k_1)(1-b_1)}{\beta_1}\right); \tag{20}$$

(II) when $x_1 = x_1^2$,

$$d_1 > \max\left(\frac{[(1-k_1)(4\alpha_1\beta_1 - 2 - 2b_1) - c_1^2]}{\beta_1(9-8k_1)},\ \ \alpha_1 - \frac{2(1-k_1)(1-b_1)}{\beta_1}\right),$$
$$d_1 < \min\left(\frac{[(1-k_1)(4\alpha_1\beta_1 - 2 - 2b_1) - c_1^2]}{\beta_1(9-8k_1)},\ \ -\alpha_1 + \frac{2(1-k_1)(1-b_1)}{\beta_1}\right), \tag{21}$$

where $b_1 = \sqrt{\alpha_1\beta_1/(1+k_1)}$, and $c_1 = (2 - 2k_1 + \alpha_1\beta_1 - \beta_1 b_1)^2$.

3.2. d_2 is the bifurcation parameter

I) when $x_2 = x_2^1$,

$$d_2 > \max\left(\frac{[(1-k_2)(4\alpha_2\beta_2 - 2 - 2b_2) - c_2^2]}{\beta_2(9-8k_2)},\ \ -\alpha_2 + \frac{2(1-k_2)(1-b_2)}{\beta_2}\right),$$
$$d_2 < \min\left(\frac{[(1-k_2)(4\alpha_2\beta_2 - 2 - 2b_2) - c_2^2]}{\beta_2(9-8k_2)},\ \ \alpha_2 - \frac{2(1-k_2)(1-b_2)}{\beta_2}\right); \tag{22}$$

II) when $x_2 = x_2^2$,

$$d_2 > \max\left(\frac{[(1-k_2)(4\alpha_2\beta_2 - 2 - 2b_2) - c_2^2]}{\beta_2(9-8k_2)},\ \ \alpha_2 - \frac{2(1-k_2)(1-b_2)}{\beta_2}\right),$$
$$d_2 < \min\left(\frac{[(1-k_2)(4\alpha_2\beta_2 - 2 - 2b_2) - c_2^2]}{\beta_2(9-8k_2)},\ \ -\alpha_2 + \frac{2(1-k_2)(1-b_2)}{\beta_2}\right), \tag{23}$$

where $b_2 = \sqrt{\alpha_2\beta_2/(1+k_2)}$, and $c_2 = (2 - 2k_2 + \alpha_2\beta_2 - \beta_2 b_2)^2$.

Therefore, the attractor is the stable focus at the beginning, and then it becomes an invariant stream form, so the system enters chaos in a short time.

4. Relationship between map accuracy and membership function

In this model, there is a smooth transformation of a fuzzy inference system which is completely known and a network with black box neurons. It could keep interpretability and improve accuracy of the map, so its learning process should reach a peak. However this is not the case in reality. There is contradiction between interpretability and accuracy.

In the simulation, we fit a curve with two inflection points: the upper left corner and the lower right corner, which is based on the relationship between map accuracy and membership function. The results are shown in Fig. 2. Figure 2(a) shows the curve, Figs. 2(b) and 2(c) display the fitting curves with two membership functions, figures 2(d) and 2(e) indicate the fitting curve with three membership functions, figure 2(f) exhibits the training process of three membership functions, with number 1 denoting the expected curve, and number 2 the practical curve.

When the third membership function is introduced, the changes of MF2 and MF3 appear at the beginning of evolution process, and then their tracks merge into one. This shows that the membership func-

tion sometimes could induce some insignificant fuzzy rules.

This simulation results show the contradiction between interpretability and precision: the network with two membership functions may be able to complete the task well, the network with three membership functions can fit curve better and obtain a higher accuracy in a fixed period of time.

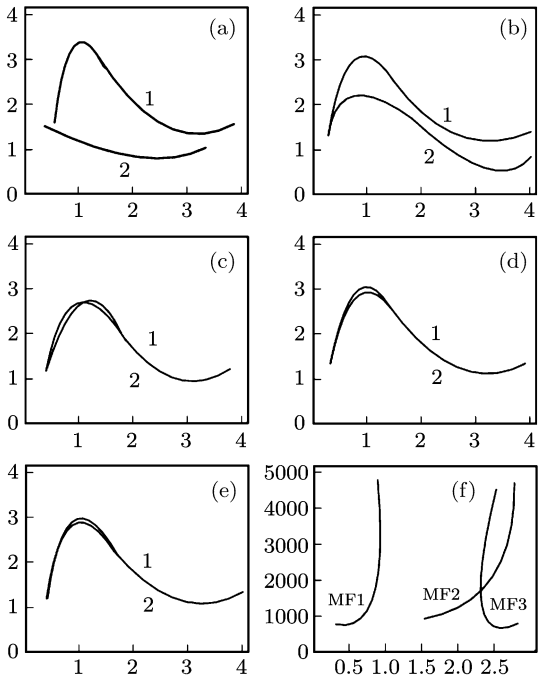


Fig. 2. Simulation results: (a) the curve, (b) the 33-th time, (c) the 25000-th time, (d) the 25-th time, (e) the 3500-th time, and (f) the changes of three membership functions.

5. Conclusion

In this paper, we designed a novel fuzzy chaotic neural network which is the combination of fuzzy, chaos, and neural networks. The structure of this network is a neural network, and it has chaotic characteristics and fuzzy functions. We designed its model structure, derived its mathematical model and weight training formulas by using BP learning algorithm, and analysed its chaotic characteristics by unstable state of its fixed points. Finally, the simulation results showed that more membership functions can improve the map accuracy, but this may lead to morbid fuzzy rules.

References

[1] Yoichi H, James J B and Ernest C. 1993 *Int. J. Intell. Syst.* **8** 527

[2] Juang C F and Tsao Y W 2008 *IEEE T. Fuzzy Syst.* **16** 1411

[3] Gang L, Thomas M M and Girijesh P 2005 *Fuzzy Set. Sys.* **150** 211

[4] Zhang H G, Xie Y H and Wang Z L 2007 *IEEE T. Neural Networks* **18** 1841

[5] Lu H T 2002 *Phys. Lett.* A **298** 109

[6] Adach M and Aihara K 1997 *Neural Networks* **10** 83

[7] Lee S M, Kwon O M and Park J H 2010 *Chin. Phys.* B **19** 050507

[8] Zhao L, Sun M, Cheng J H and Xu Y Q 2009 *IEEE T. Fuzzy Syst.* **20** 735

[9] Ren H P and Liu D 2002 *Acta Phys. Sin.* **51** 982 (in Chinese)

[10] Tan W, Wang Y N, Zhou S W and Liu Z R 2003 *Acta Phys. Sin.* **52** 795 (in Chinese)

[11] Huert Y C 1999 *Int. J. Bifur. Chaos* **9** 1597

[12] Yeh I C 1999 *Fuzzy Set. Sys.* **105** 421

[13] Yeh J P 2007 *Chaos, Solitons and Fractals* **32** 1178

[14] Lin D, Wang X Y, Nian F Z and Zhang Y L 2010 *Neurocomputing* **73** 2873

[15] Ito Y 1991 *Neural Networks* **4** 385

[16] Li J and Zhang Y P 2011 *Acta Phys. Sin.* **60** 070513 (in Chinese)