Fourier Series Chaotic Neural Networks

Jia-hai Zhang[1], Chen-zhi Sun[1], and Yao-qun Xu[2]

[1] College of Electrical and Automatic Engineering, Sanjiang University, 210012, Nanjing, China
zhangjh6688@sohu.com, sunchenzhi@sju.js.cn
[2] Institute of System Engineering, Harbin University of Commerce, 150028, Harbin, China
xuyq@hrbcu.edu.cn

Abstract. In this paper, Fourier series chaotic neural network model is presented to improve the ability to escape the local minima so that it can effectively solve optimization problems. 10-city traveling salesman problem was given and the effects of the non-monotonous degree in the model on solving 10-city traveling salesman problem were discussed, the figures of the reversed bifurcation and the maximal Lyapunov exponents of single neural unit were given. The new model is applied to solve several function optimizations. Seen from the simulation results, the new model is powerful than the common chaotic neural network.

1 Introduction

The Hopfield network, proposed by Hopfield and Tank [1], has been extensively applied to many fields in the past years. Several chaotic neural networks with non-monotonous activation functions have been proved to be more powerful than Chen's chaotic neural network in solving optimization problems, especially in searching global minima of continuous function and traveling salesman problems [2, 7~8]. The reference [3] has pointed out that the single neural unit can easily behave chaotic motion if its activation function is non-monotonous. And the reference [4] has presented that the effective activation function may adopt kinds of different forms, and should embody non-monotonous nature. The chaotic mechanism of this new model is introduced by the self-feedback connection weight. The activation function of the new chaotic neural network model is composed of Sigmoid and trigonometric function, therefore the activation function is non-monotonous. And because trigonometric function is a kind of basic function, the model can solve optimization problems more effectively. Finally, the new model is applied to solve both function optimizations and combinational optimizations and the effects of the non-monotonous degree in the model on solving 10-city TSP are discussed. Seen from the simulation results, the new model is powerful than the common chaotic neural network.

2 Fourier Series Chaotic Neural Network (FSCNN)

Fourier series chaotic neural network is described as follows:

$$x_i(\mathrm{t}) = f(y_i(\mathrm{t})) \tag{1}$$

L. Zhang, J. Kwok, and B.-L. Lu (Eds.): ISNN 2010, Part I, LNCS 6063, pp. 136–143, 2010.

$$y_i(t+1) = ky_i(t) + \alpha\left[\sum_{\substack{j=1\\ j\neq i}}^{n} w_{ij}x_j(t) + I_i\right] - z_i(t)(x_i(t) - I_0) \tag{2}$$

$$z_i(t+1) = (1-\beta)z_i(t) \tag{3}$$

$$f(u) = S_1(u) + S_2(u) \tag{4}$$

$$S_1(u) = \frac{1}{1 + \exp(-u/\varepsilon_0)} \tag{5}$$

$$S_2(u) = \omega_1 \cos(\varepsilon_1 u) + \omega_2 \sin(\varepsilon_2 u) \tag{6}$$

Where i is the index of neurons and n is the number of neurons, $x_i(t)$ the output of neuron i, $y_i(t)$ the internal state for neuron i, W_{ij} the connection weight from neuron j to neuron i, I_i the input bias of neuron i, α the positive scaling parameter for inputs, k the damping factor of the nerve membrane ($0 \le k \le 1$), $z_i(t)$ the self-feedback connection weight, $\varepsilon_0, \varepsilon_1, \varepsilon_2$ the steepness parameters of the activation function, β the simulated annealing parameter of the self-feedback connection weight $z_i(t)$, I_0 a positive parameter and ω_1, ω_2 the parameters of the trigonometric function.

In this model, the variable $z_i(t)$ corresponds to the temperature in the usual stochastic annealing process and the equation (3) is an exponential cooling schedule for the annealing. The chaotic mechanism is introduced by the self-feedback connection weight as the value of $z_i(t)$ becomes small step by step. The chaotic behavior plays a global search role in the beginning. When the value of $z_i(t)$ decreases to a certain value, the network functions in a fashion similar to the Hopfield network which functions in gradient descent dynamic behavior. Finally, the neurons arrive at a stable equilibrium state. The reference [5] shows that both the parameter β governed the bifurcation speed of the transient chaos and the parameter α could affect the neuron dynamics; in other words, the influence of the energy function was too strong to generate transient chaos when α was too large, and the energy function could not be sufficiently reflected in the neuron dynamics when α was too small. So in order for the network to have rich dynamics initially, the simulated annealing parameter β must be set to a small value, and α must be set to a suitable value, too.

In this model, the parameter ω_1, ω_2 presents the non-monotonous degree of the activation function. Seen from the equations (4) and (5), it is concluded that the equation (4) is similar to the function of Sigmoid alone in form in the circumstance of the value of ω_1, ω_2 being between 0 and 1 without consideration of the monotonous nature. So the parameter ω_1, ω_2 presents a local non-monotonous phenomenon of the activation function. In other words, if the parameter ω_1, ω_2 borders on 1, the non-monotonous phenomenon of the activation function is very apparent; otherwise, if the parameter ω_1, ω_2 borders on 0, the non-monotonous phenomenon of the activation function is very weak.

3 Research on Single Neural Unit

In this section, we make an analysis of the neural unit of the Fourier series chaotic neural networks.

The single neural unit can be described as (7) ~ (9) together with (4) ~ (6):

$$x(t) = f(y(t)) \tag{7}$$

$$y(t+1) = ky(t) - z(t)(x(t) - I_0) \tag{8}$$

$$z(t+1) = (1-\beta)z(t) \tag{9}$$

In order to make the neuron behave transient chaotic behavior, the parameters are set as follows:

$$\varepsilon_0 = 0.02, \varepsilon_1 = 2, \varepsilon_2 = 2, \omega_1 = 1/3, \omega_2 = 1/3, y(1) = 0.283, z(1) = 0.4, k = 1, I_0 = 0.65$$

The state bifurcation figures and the time evolution figures of the maximal Lyapunov exponent are respectively shown as Fig.1~Fig.4 when β =0.004 and β =0.002.

Seen from the above state bifurcation figures, the neuron behaves a transient chaotic dynamic behavior. The single neural unit first behaves the global chaotic search, and with the decrease of the value of $z(0,0)$, the reversed bifurcation gradually converges to a stable equilibrium state. After the chaotic dynamic behavior disappears, the dynamic behavior of the single neural unit is controlled by the gradient descent dynamics. When the behavior of the single neural unit is similar to that of Hopfield, the network tends to converge to a stable equilibrium point. The simulated annealing parameter β affects the length of the reversed bifurcation, that is, the smaller value of β prolongs the reversed bifurcation.

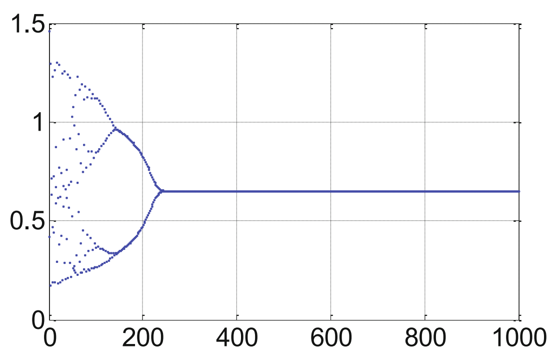

Fig. 1. State bifurcation figure of the neuron when β =0.004

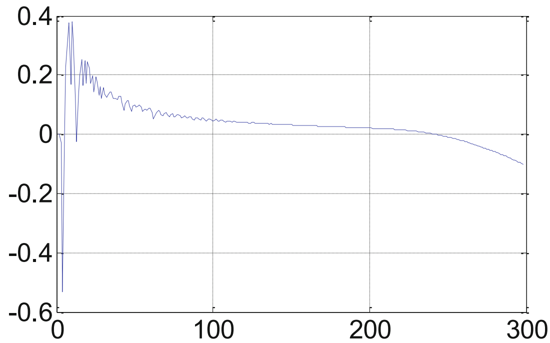

Fig. 2. Time evolution figure of the maximal Lyapunov exponent of the neuron when β=0.004

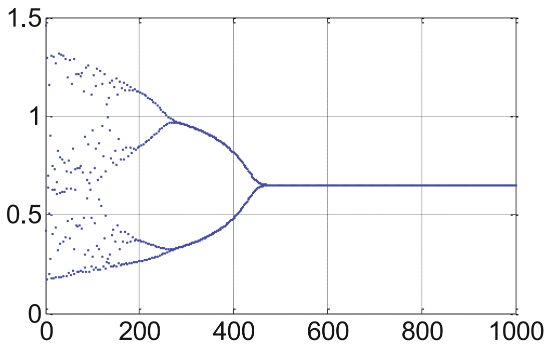

Fig. 3. State bifurcation figure of the neuron when β =0.002

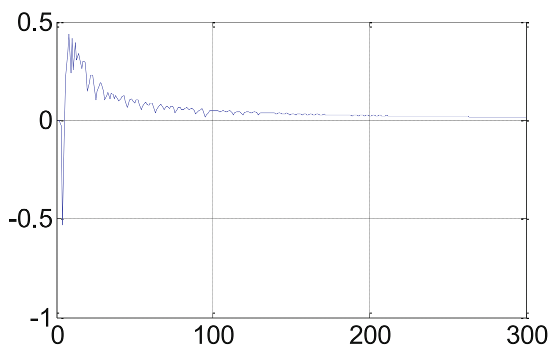

Fig. 4. Time evolution figure of the maximal Lyapunov exponent of the neuron when β=0.002

4 Application to Continuous Function Optimization Problems

In this section, we apply the Fourier series chaotic neural network to search global minima of the following function.

The function is described as follows [6]:

$$f_2(x_1, x_2) = (x_1 - 0.7)^2[(x_2 + 0.6)^2 + 0.1] + (x_2 - 0.5)^2[(x_1 + 0.4)^2 + 0.15] \quad (10)$$

The minimum value of (10) is 0 and its responding point is (0.7, 0.5).

The parameters are set as follows:

ε_0 =2.5, ε_1 =20, ε_2 =10, k =1, ω_1 =0.1, ω_2 =0.05, $y_1(1) = y_2(1)$ =0.283, β =0.05, α =0.4, $z_1(1) = z_2(1)$ =0.3, I_0 =0.65.

The time evolution figure of the energy function of FSCNN in solving the function is shown as Fig.5.

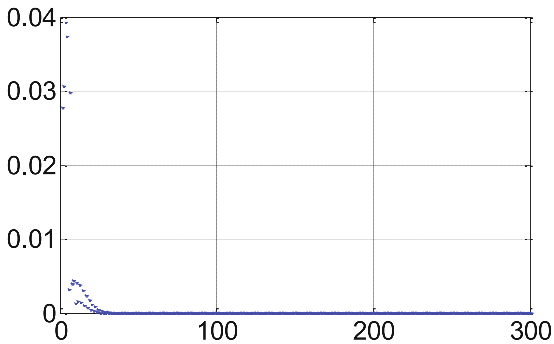

Fig. 5. Time evolution figure of energy function

The global minimum and its responding point of the simulation are respectively 1.3236e-016 and (0.7, 0.5).

This section indicates that FSCNN has a good performance to solve function optimization problems. In order to testify the performance of FSCNN, the new model is applied to solve 10-city traveling salesman problems.

5 Application to 10-City TSP

A solution of TSP with N cities is represented by N×N-permutation matrix, where each entry corresponds to output of a neuron in a network with N×N lattice structure. Assume V_{xi} to be the neuron output which represents city x in visiting order i. A computational energy function which is to minimize the total tour length while simultaneously satisfying all constrains takes the follow form:

$$E = \frac{A}{2}\sum_{x=1}^{n}(\sum_{i=1}^{n}V_{xi}-1)^2+\frac{B}{2}\sum_{i=1}^{n}(\sum_{x=1}^{n}V_{xi}-1)^2+\frac{D}{2}\sum_{x=1}^{n}\sum_{y=1}^{n}\sum_{i=1}^{n}d_{xy}V_{xi}V_{y,i+1} \quad (11)$$

A and B (A=B) are the coupling parameters corresponding to the constrains and the cost function of the tour length, respectively. d_{xy} is the distance between city x and city y.

This paper adopts the following 10-city unitary coordinates:

(0.4, 0.4439), (0.2439, 0.1463), (0.1707, 0.2293), (0.2293, 0.716), (0.5171, 0.9414), (0.8732, 0.6536), (0.6878, 0.5219), (0.8488, 0.3609),(0.6683, 0.2536), (0.6195, 0.2634). The shortest distance of the 10-city is 2.6776.

The reference [5] has presented that the effective activation function may adopt kinds of different forms, and should behave non-monotonous behavior. In this paper, ω_1, ω_2 that represents the non-monotonous degree is analyzed in order to simply ascertain the effect of the non-monotonous degree to FSCNN in solving 10-city TSP. Therefore, the models with different values of ω_1, ω_2 in solving 10-city TSP are analyzed as follows:

The parameters of the network are set as follows:

ε_1=20, ε_2=10, ε_0=1/30, k=1, α=0.6, $z(1)$=0.1, I_0=0.2,A=1.4, D=1.5,$\triangle$t=0.04.

2000 different initial conditions of y_{ij} are generated randomly in the region [0, 1] for different β. The results are summarized in Table1, the column 'NL', 'NG', 'LR' and 'GR' respectively represents the number of legal route, the number of global optimal route, the rate of legal route, the rate of global optimal route.

The lager value of the simulated annealing parameter β is regarded stronger if the network can all converge to the global minimum in 2000 different random initial conditions.

Seen from table 1, the follow observations can be drawn according to numerical simulation test:

First, the model with smaller ω_1, ω_2 s such as ω_1 =0.01, ω_2 =0.01; ω_1 =0.02, ω_2 =0.01; ω_1 =0.025, ω_2 =0.01; ω_1 =0.01, ω_2 =0.025; ω_1 =0.02, ω_2 =0.02 in solving 10-city TSP can all converge to the global minimum. But, it is not true that the smaller the parameter ω_1, ω_2 is, the more powerful the ability to solve 10-city. Because, for example, the parameter ω_1 =0.02, ω_2 =0.02 can all converge to the global minimum as β =0.001 while the parameter ω_1 =0.01, ω_2 =0.01 can almost converge to the global minimum as β =0.0008.

Second, with the decrease of the value of ω_1 and ω_2, the value of 'NG' becomes large gradually from 1893(ω_1 =0.02, ω_2 =0.02) to 1973(ω_1 =0.01, ω_2 =0.005) as β =0.01 .In other word, with the decrease of the value of ω_1 and ω_2, the ability to get global optimal route becomes strong.

Table 1. Results of 2000 different initial conditions for each value β on 10-city TSP

ω_1, ω_2	β	NL	NG	LR	GR
$\omega_1=0.001, \omega_2=0.001$	0.01	2000	1953	100%	97.65%
	0.008	2000	1983	100%	99.15%
	0.001	2000	1994	100%	99.7%
	0.0008	2000	1995	100%	99.75%
$\omega_1=0.005, \omega_2=0.01$	0.01	2000	1972	100%	98.6%
	0.008	2000	1994	100%	99.7%
	0.001	2000	1994	100%	99.7%
	0.0008	2000	1998	100%	99.9%
$\omega_1=0.01, \omega_2=0.01$	0.01	2000	1984	100%	99.2%
	0.008	2000	1998	100%	99.9%
	0.001	2000	1998	100%	99.9%
	0.0008	2000	2000	100%	100%
$\omega_1=0.02, \omega_2=0.01$	0.01	2000	1955	100%	97.75%
	0.008	2000	1945	100%	97.25%
	0.001	2000	1996	100%	99.8%
	0.0008	2000	2000	100%	100%
$\omega_1=0.025, \omega_2=0.01$	0.01	2000	1931	100%	96.65%
	0.008	2000	1931	100%	96.55%
	0.001	2000	2000	100%	100%
	0.0008	2000	2000	100%	100%
$\omega_1=0.01, \omega_2=0.005$	0.01	2000	1973	100%	98.65%
	0.008	2000	1994	100%	99.7%
	0.001	2000	1995	100%	99.75%
	0.0008	2000	1995	100%	99.75%
$\omega_1=0.01, \omega_2=0.02$	0.01	2000	1940	100%	97%
	0.008	2000	1928	100%	96.4%
	0.001	2000	1998	100%	99.9%
	0.0008	2000	1997	100%	99.85%
$\omega_1=0.01, \omega_2=0.025$	0.01	2000	1908	100%	95.4%
	0.008	2000	1923	100%	96.15%
	0.001	2000	2000	100%	100%
	0.0008	2000	2000	100%	100%
$\omega_1=0.02, \omega_2=0.02$	0.01	2000	1893	100%	94.65%
	0.008	2000	1927	100%	96.35%
	0.001	2000	2000	100%	100%
	0.0008	2000	2000	100%	100%

Third, when the parameter $\omega_1=0.01, \omega_2=0.025$and $\omega_1=0.025, \omega_2=0.01$, the ability to all converge to the global minimum is more powerful, that is, the non-monotonous degree of the activation function has a effective on the solution of 10-city TSP.

However, as is analyzed in second, the ability in reaching 'NG' when the parameter $\omega_1=0.01, \omega_2=0.025$and $\omega_1=0.025, \omega_2=0.01$is weaker than that of $\omega_1=0.01, \omega_2=0.01$ when $\beta=0.01$and $\beta=0.008$. So, which model is needed is connected with the concrete request. However, in order to get the tradeoff effect, the value of $\omega_1=0.01, \omega_2=0.01$ may be chose.

6 Conclusion

The presented chaotic neural network called FSCNN is proved to be effective in solving optimization problems, and in the section of application to 10-city TSP, the model with different ω_1 , ω_2 is analyzed and made a comparison. As a result, the simple rule of the model is disclosed. However, there are a lot of questions in the model needed to research. For example, whether does the model in solving 30 or 60 cities' TSP accord with the same rule or so? And different networks have different parameters in solving TSP effectively, is there any way to weigh the different networks in the same parameters?

References

1. Hopfield, J.: Neural Networks and Physical Systems with Emergent Collective Computational Abilities. Proc. Natl. Acad. Sci. 79, 2554–2558 (1982)
2. Xu, Y.-q., Sun, M., Duan, G.-r.: Wavelet Chaotic Neural Networks and Their Application to Optimization Problems. LNCS, vol. 3791, pp. 379–384. Springer, Heidelberg (2006)
3. Potapove, A., Kali, M.: Robust chaos in neural networks. Physics Letters A 277(6), 310–322 (2000)
4. Shuai, J.W., Chen, Z.X., Liu, R.T., et al.: Self-evolution Neural Model. Physics Letters A 221(5), 311–316 (1996)
5. Chen, L., Aihara, K.: Chaotic Simulated Annealing by a Neural Network Model with Transient Chaos. Neural Networks 8(6), 915–930 (1995)
6. Ling, W.: Intelligence optimization algorithm and its application. Press of TUP (2001)
7. Xu, Y.-q., Sun, M.: Gauss-Morlet-Sigmoid Chaotic Neural Networks. In: Huang, D.-S., Li, K., Irwin, G.W. (eds.) ICIC 2006. LNCS, vol. 4113, pp. 115–125. Springer, Heidelberg (2006)
8. Xu, Y.-q., Sun, M., Shen, J.-h.: Gauss Chaotic Neural Networks. In: Yang, Q., Webb, G. (eds.) PRICAI 2006. LNCS (LNAI), vol. 4099, pp. 319–328. Springer, Heidelberg (2006)