Shannon Wavelet Chaotic Neural Networks

Yao-qun Xu[1,2], Ming Sun[1], and Ji-hong Shen[2]

[1] Institute of System Engineering, Harbin University of Commerce, 150028 Harbin, China
Xuyq@hrbcu.edu.cn, Snogisun@tom.com
[2] Department of Mathematics, Harbin Engineering University, 150001 Harbin, China
Shenjihong@hrbeu.edu.cn

Abstract. Chaotic neural networks have been proved to be strong tools to solve the optimization problems. In order to escape the local minima, a new chaotic neural network model called Shannon wavelet chaotic neural network was presented. The activation function of the new model is non-monotonous, which is composed of sigmoid and Shannon wavelet. First, the figures of the reversed bifurcation and the maximal Lyapunov exponents of single neural unit were given. Second, the new model is applied to solve several function optimizations. Finally, 10-city traveling salesman problem is given and the effects of the non-monotonous degree in the model on solving 10-city traveling salesman problem are discussed. The new model can solve the optimization problems more effectively because of the Shannon wavelet being a kind of basic function. Seen from the simulation results, the new model is powerful.

1 Introduction

Neural networks have been shown to be powerful tools for solving optimization problems. The Hopfield network, proposed by Hopfield and Tank [1, 2], has been extensively applied to many fields in the past years. The Hopfield neural network converges to a stable equilibrium point due to its gradient decent dynamics; however, it causes sever local-minimum problems whenever it is applied to optimization problems. 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 [3, 8~9]. The reference [4] has pointed out that the single neural unit can easily behave chaotic motion if its activation function is non-monotonous. And the reference [5] has presented that the effective activation function may adopt kinds of different forms, and should embody non-monotonous nature. In this paper, a new chaotic neural network model is presented to improve the ability to escape the local minima so that it can effectively solve optimization problems. 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 Shannon Wavelet, therefore the activation function is non-monotonous. And because Shannon wavelet 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

T.-D. Wang et al. (Eds.): SEAL 2006, LNCS 4247, pp. 244 – 251, 2006.

combinational optimizations and the effects of the non-monotonous degree in the model on solving 10-city TSP are discussed. The simulation results in solving 10-city TSP show that the new model is valid in solving optimization problems.

For any function $f(x) \in L_2(R)$ and any wavelet Ψ which is a basic function, the known formula can be described as follows:

$$f(x) = \sum_{j,k=-\infty}^{\infty} c_{j,k} \Psi_{j,k}(x) \tag{1}$$

2 Shannon Wavelet Chaotic Neural Network (SWCNN)

Shannon wavelet chaotic neural network is described as follows:

$$x_i(\mathrm{t}) = f(y_i(\mathrm{t})(1+\eta_i(\mathrm{t}))) \tag{2}$$

$$y_i(\mathrm{t}+1) = ky_i(\mathrm{t}) + \alpha\left[\sum_{\substack{j=1 \\ j\neq i}}^{n} w_{ij}x_j(\mathrm{t}) + I_i\right] - z_i(\mathrm{t})(x_i(\mathrm{t}) - I_0) \tag{3}$$

$$z_i(\mathrm{t}+1) = (1-\beta)z_i(\mathrm{t}) \tag{4}$$

$$\eta_i(\mathrm{t}+1) = \frac{\eta_i(\mathrm{t})}{\ln(\exp(1) + \lambda(1-\eta_i(\mathrm{t})))} \tag{5}$$

$$f(\mu) = \mathrm{Sigmoid}(\mu,\varepsilon_1) + coef \cdot \mathrm{Shannon}(\mu,\varepsilon_2) \tag{6}$$

$$\mathrm{Sigmoid}(\mu,\varepsilon_1) = \frac{1}{1+\exp(\mu/\varepsilon_1)} \tag{7}$$

$$\mathrm{Shannon}(\mu,\varepsilon_2) = \frac{\sin\pi(\mu/\varepsilon_2 - \frac{1}{2}) - \sin 2\pi(\mu/\varepsilon_2 - \frac{1}{2})}{\pi(\mu/\varepsilon_2 - \frac{1}{2})} \tag{8}$$

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, $\varepsilon_1, \varepsilon_2$ the steepness parameters of the activation function, β the simulated annealing parameter of the self-feedback connection weight $z_i(t)$, $\eta_i(t)$ the other simulated annealing parameter of the activation, I_0 a positive parameter and $coef$ the non-monotonous degree ($0 \le coef \le 1$).

In this model, the variable $z_i(t)$ corresponds to the temperature in the usual stochastic annealing process and the equation (4) is an exponential cooling schedule for the annealing as well as the equation (5). 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 [6] 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 *coef* presents the non-monotonous degree of the activation function. Seen from the equation (6), it is concluded that the equation (6) is similar to the function of Sigmoid alone in form in the circumstance of the value of *coef* being between 0 and 1 without consideration of the monotonous nature. So the parameter *coef* presents a local non-monotonous phenomenon of the activation function. In other words, if the parameter *coef* borders on 1, the non-monotonous phenomenon of the activation function is very apparent; otherwise, if the parameter *coef* borders on 0, the non-monotonous phenomenon of the activation function is very weak.

In order to gain insight into the evolution progress of the single neural unit, the research was made as follows.

3 Research on Single Neural Unit

In this section, we make an analysis of the neural unit of the Shannon Wavelet chaotic neural networks.

The single neural unit can be described as (9) ~ (12) together with (6) ~ (8):

$$x(t) = f(y(t)) \tag{9}$$

$$y(t+1) = ky(t) - z(t)(x(t) - I_0) \tag{10}$$

$$z(t+1) = (1-\beta)z(t) \tag{11}$$

$$\eta(t+1) = \frac{\eta(t)}{\ln(\exp(1) + \lambda(1-\eta(t)))} \tag{12}$$

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

$\varepsilon_1 = 0.004$, $\varepsilon_2 = 1.25$, $y(1) = 0.283$, $z(1) = 0.1$, $k = 1$, $\eta(1) = 0.8$, $\lambda = 0.5$, $I_0 = 0.5$

The state bifurcation figures and the time evolution figures of the maximal Lyapunov exponent are respectively shown as Fig.1~Fig.3 when $\beta = 0.004$ and $\beta = 0.002$.

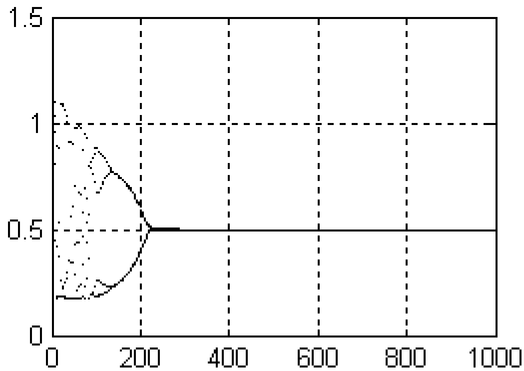

Fig. 1. State bifurcation figure of the neuron when $\beta = 0.004$

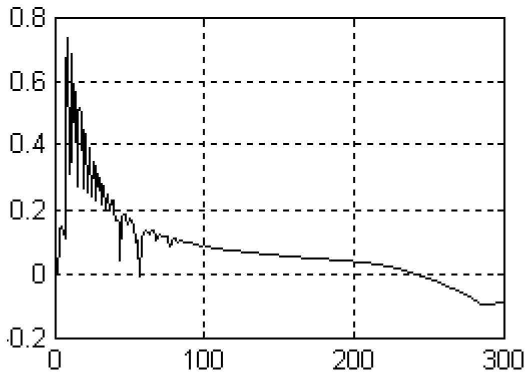

Fig. 2. Time evolution figure of the maximal Lyapunov exponent of the neuron when $\beta = 0.004$

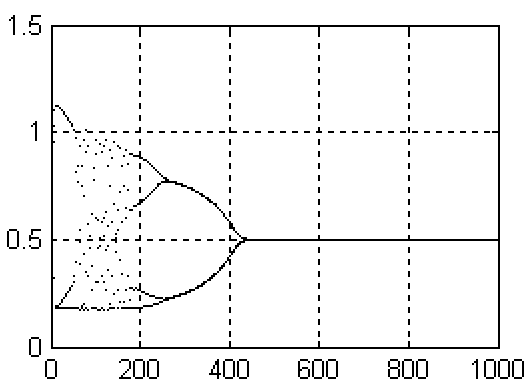

Fig. 3. State bifurcation figure of the neuron when $\beta = 0.002$

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 lager value of β prolongs the reversed bifurcation.

4 Application to Continuous Function Optimization Problems

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

The function is described as follows [7]:

$$f_2(x_1,x_2) = (x_1 - 0.7)^2[(x_2 + 0.6)^2 + 0.1] + (x_2 - 0.5)^2[(x_1 + 0.4)^2 + 0.15] \quad (13)$$

The minimum value of (13) is 0 and its responding point is (0.7, 0.5).

The parameters are set as follows:

$\varepsilon_1 = 0.05$, $\varepsilon_2 = 10$, $\alpha = 0.08$, $k = 1$, $I_0 = 0.5$, $coef = 1/4$, β =0.002, $z(0,0)$ =[0.8, 0.8], $y(0,0)$ =[0.283, 0.283], $\eta(0,0)$ =[0.8, 0.8], $\lambda(0,0)$ =[0.01, 0.01].

The time evolution figure of the energy function of SWCNN in solving the function is shown as Fig.4.

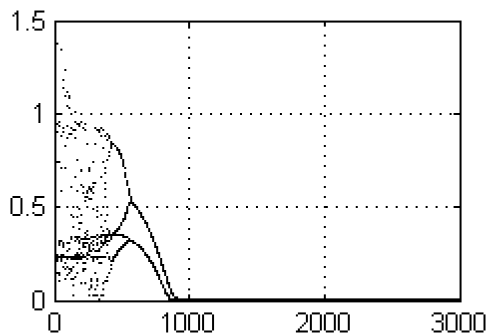

Fig. 4. Time evolution figure of energy function

The global minimum and its responding point of the simulation are respectively 2.1448e-015 and (0.7, 0.5).

This section indicates that SWCNN has a good performance to solve function optimization problems. In order to testify the performance of SWCNN, 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 [1]:

$$E = \frac{W_1}{2}\left\{\sum_{i=1}^{n}\left[\sum_{j=1}^{n} x_{ij} - 1\right]^2 + \sum_{j=1}^{n}\left[\sum_{i=1}^{n} x_{ij} - 1\right]^2\right\} + \frac{W_2}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}(x_{k,j+1} + x_{k,j-1})x_{ij}d_{ik} \qquad (14)$$

Where $x_{i0} = x_{in}$ and $x_{i,n+1} = x_{i1}$. W_1 and W_2 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 [6] has presented that the effective activation function may adopt kinds of different forms, and should behave non-monotonous behavior. In this paper, *coef* that represents the non-monotonous degree is analyzed in order to simply ascertain the effect of the non-monotonous degree to SWCNN in solving 10-city TSP. Therefore, the models with different values of *coef* in solving 10-city TSP are analyzed as follows:

The parameters of the network are set as follows:

$W_1 = 1$, $W_2 = 0.8$, $z_i(1) = 0.2$, $\alpha = 0.5$, $k = 1$, $\eta_i(1) = 0.8$, $I_0 = 0.5$, $\lambda = 0.008$, $\varepsilon_1 = 0.004, \varepsilon_2 = 2.5$.

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 *coef* s such as 0, 1/4, 1/5, 1/6, 1/7, 1/8, 1/9 and 1/10 in solving 10-city TSP can all converge to the global minimum. But, it is not true that the smaller the parameter *coef* is, the more powerful the ability to solve 10-city is. Because, for example, the parameter $coef = 1/10$ can all converge to the global minimum as $\beta = 0.0005$ while the parameter $coef = 1/9$ can all converge to the global minimum as $\beta = 0.0007$.

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

coef	β	NL	NG	LR	GR
$coef = 0$ (common network)	0.0003	2000	1923	100%	95.65%
	0.001	1998	1998	99.9%	99.9%
	0.0008	2000	2000	100%	100%
$coef = 1$	0.003	1837	620	91.85%	31%
	0.001	1891	1146	94.55%	57.3%
	0.0008	1904	1075	95.2%	53.75%
$coef = 1/2$	0.003	1962	1791	98.1%	89.55%
	0.0008	1925	1858	96.25%	92.9%
	0.0005	1842	1672	92.1%	83.6%
$coef = 1/4$	0.003	1975	1811	98.75%	90.55%
	0.0008	2000	1997	100%	99.85%
	0.00046	2000	2000	100%	100%
$coef = 1/5$	0.003	1979	1797	98.95%	89.85%
	0.001	2000	1999	100%	99.95%
	0.0009	2000	2000	100%	100%
$coef = 1/6$	0.003	1987	1819	99.35	90.95%
	0.001	2000	2000	100%	100%
$coef = 1/7$	0.003	1989	1806	99.45%	90.3%
	0.001	1999	1999	99.95%	99.95%
	0.0008	2000	2000	100%	100%
$coef = 1/8$	0.003	1990	1713	99.5%	85.65%
	0.0008	1999	1999	99.95%	99.95%
	0.0006	2000	2000	100%	100%
$coef = 1/9$	0.003	1993	1713	99.65%	85.65%
	0.0008	1999	1999	99.95%	99.95%
	0.0007	2000	2000	100%	100%
$coef = 1/10$	0.003	1998	1799	99.9%	89.95%
	0.0008	1999	1998	99.95	99.9%
	0.0005	2000	2000	100%	100%

Second, with the decrease of the value of $coef$, the value of 'NL' becomes large gradually from 1837 ($coef = 1$) to 2000 ($coef = 0$) as $\beta = 0.003$ .In other word, with the decrease of the value of $coef$, the ability to get legal route becomes strong.

Third, when the parameter $coef = 1/5$ and $coef = 1/6$, the ability to all converge to the global minimum is more powerful than that of $coef = 0$, that is, the non-monotonous degree of the activation function has a positive effect on the solution of 10-city TSP.

However, as is analyzed in second, the ability in reaching 'NL' when the parameter $coef = 1/5$ and $coef = 1/6$ is weaker than that of $coef = 0$. So, which model is needed is connected with the concrete request. However, in order to get the tradeoff effect, the value of $coef = 1/6$ may be chose. As the test result is not based on the theoretical analysis, the relationship between $coef$ and the performance need to be studied further.

6 Conclusion

The presented chaotic neural network called SWCNN is proved to be effective in solving optimization problems, and in the section of application to 10-city TSP, the model with different *coef* is analyzed and made a comparison. As a result, the simple rule of the model is disclosed. However, there are a lot of points in the model needed to be studied.

Acknowledgement

This work is supported by the Youth Science Foundation of Harbin 2005AFQXJ040.

References

1. Hopfield J J., Tank D W.: Neural Computation of Decision in Optimization Problems.Biol. Cybern.Vol.52. (1985)141-152
2. Hopfield, J.: Neural Networks and Physical Systems with Emergent Collective Computational Abilities. Vol. 79.In: Proc Natl Acad Sci, (1982) 2554-2558
3. Yao-qun Xu, Ming Sun, Guang-ren Duan. Wavelet Chaotic Neural Networks and Their Application to Optimization Problems. Lecture Notes in Computer Science, Vol. 3791. Springer (2006) 379-384
4. A Potapove, M Kali. Robust chaos in neural networks. Physics Letters A, vol277,no.6, (2000)310-322
5. Shuai J W, Chen Z X, Liu R T, et al. Self-evolution Neural Model. Physics Letters A, vol.221,no.5,(1996)311-316
6. Chen L, Aihara K. Chaotic Simulated Annealing by a Neural Network Model with Transient Chaos. Neural Networks. Vol.8,no.6,(1995)915-930.
7. Wang Ling. Intelligence optimization algorithm and its application. Press of TUP. (2001)
8. Y.-q. Xu and M. Sun. Gauss-Morlet-Sigmoid Chaotic Neural Networks. ICIC 2006, LNCS vol.4113,Springer-Verlag Berlin Heidelberg (2006) 115-125
9. Y.-q. Xu, M. Sun, and J.-h. Shen. Gauss Chaotic Neural Networks. PRICAI 2006, LNAI vol.4099,Springer-Verlag Berlin Heidelberg (2006) 319 – 328