A Class of Chaotic Neural Network with Morlet Wavelet Function Self-feedback

Yqoqun Xu[1] and Xueling Yang[2]

[1] Institute of System Engineering, Harbin University of Commerce, Harbin, 150028
[2] Department of mathematic, Harbin Engineering University, Harbin, 150001
xuyq@hrbcu.edu.cn

Abstract. A Chaotic neural network model with Morlet wavelet function self-feedback is proposed by introducing Morlet wavelet function into self-feedback of chaotic neural network. The analyses of the optimization mechanism of the networks suggest that Morlet wavelet function self-feedback affects the original Hopfield energy function in the manner of the sum of the multiplications of Morlet wavelet function to the state, avoiding the network being trapped into the local minima. The energy function is constructed, and the sufficient condition for the networks to achieve asymptotical stability is analyzed and is used to instruct the parameter set of the networks for solving traveling salesman problem (TSP). Simulation researches on 10-city TSP indicate that the proposed networks can find the optimal solutions of combinatorial optimization problems.

Keywords: Self-feedback, Chaotic neural network, Energy function, Morlet wavelet function.

1 Introduction

Chaotic neural networks (CNNs) can acquire the ability to escape from the local minima of the energy function by introducing chaotic search mechanism into the original Hopfield neural network (HNN) [1-7]. Different from external chaos, chaotic search mechanism in CNN is generated by the self-feedback item of CNN. Besides, CNN possesses abundant dynamics characteristics and can traverse every point of the system by chaotic search. However, CNN cannot be easy to converge to a point steadily. Hence, Chen and Aihara have proposed a transient chaotic neural network (TCNN) with chaotic simulated annealing (CSA) by introducing a linear self-feedback into the original HNN and reducing the self-feedback connection weight exponentially. It is ensured that the TCNN has transient chaotic search behavior and can converge to a point steadily. In addition, it overcomes the limitation of HNN which is not enough for the network to escape from the local minima. In the theory, Chen and Aihara have proven that the TCNN is asymptotical stability. This paper proposes a novel TCNN model with Morlet wavelet function self-feedback which has new characteristics different from linear self-feedback network. This paper analyzes the effect of energy

D.-S. Huang et al. (Eds.): ICIC 2011, LNBI 6840, pp. 32–40, 2012.

modifier item by Kwok unified framework theory, constructs energy function of the network, and further analyzes stability of the proposed network. The analysis of the energy function suggests that the proposed network model is asymptotical stability under the given conditions. The simulations of 10-city TSP suggest that the proposed chaotic neural network has a good optimal performance.

2 TCNN With Morlet Wavelet Function Self-feedback

The proposed TCNN with Morlet wavelet function self-feedback can be described as follows.

$$x_i(t) = \frac{1}{1+\exp(-y_i(t)/\varepsilon)} \tag{1}$$

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

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

$$u_i(t+1) = 4u_i(t)(1-u_i(t)) \tag{4}$$

$$z_i(t+1) = (1-\beta)z_i(t) \tag{5}$$

$$\phi(s,u,x) = \exp\{-[(x-u)/s]^2/2\}\cdot\cos[5(x-u)/s] \quad (s>0, -1<u<1) \tag{6}$$

where x_i is the output of neuron i ; y_i is the internal state of neuron i ; w_{ij} is the connection weight from neuron j to neuron i , $w_{ij} = w_{ji}$; I_i is an input bias of neuron i ; I_0 is a positive parameter; k is a damping factor of nerve membrane ($0<k<1$); z_i is the self-feedback connection weight; β is the damping factor of z_i .

The single neuron of the proposed TCNN model can be described as follows:

$$x(t) = \frac{1}{1+\exp(-y(t)/\varepsilon)} \tag{7}$$

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

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

$$u(t+1) = 4u(t)(1-u(t)) \tag{10}$$

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

$$\phi(s,u,x) = \exp\{-[(x-u)/s]^2/2\}\cdot\cos[5(x-u)/s] \quad (s>0, -1<u<1) \tag{12}$$

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

$\varepsilon = =1/11$, $y(1) = 0.283$, $k = 0.3$, $I_0 = 0.55$, $\beta = 0.0005$, $s(1) = 1.5$, $u(1) = 0.5$, $z(1) = 0.35$

The state bifurcation figures and the time evolution figures of the maximal Lyapunov exponent are respectively shown as Fig.1.

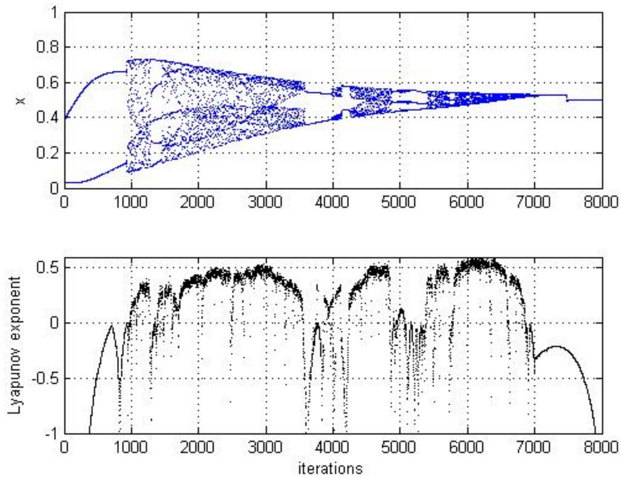

Fig. 1. State bifurcation figure and the maximal Lyapunov exponents of the single neuron

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_i, 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.

3 Energy Function

Analysis of the energy function is very important for the study of nonlinear dynamics systems, and through such an analysis, we can easily observe how the sigmoid function self-feedback affects the optimization performance of the proposed CNN. In this section, One is the unified framework theory that allows the construction and comparison of various models from the basic HNN by the introduction of an energy modifier. The other is Lyapunov stability analysis that we use to investigate the asymptotical stability of the proposed CNN model. Moreover, the unified framework theory is applied to construct the energy function for the Lyapunov stability analysis in this section.

3.1 Energy Modifier of CNN

Based on the unified framework, the energy function can be described as follows:

$$E = E_{Hop} + H \tag{13}$$

$$E_{Hop} = -\frac{1}{2}\sum_{i,j}^{n} w_{i,j}x_i x_j - \sum_{i}^{n} I_i x_i + \frac{\varepsilon_1}{\tau}\sum_{i}^{n}\int_0^{x_i} \ln\frac{x}{1-x}dx \tag{14}$$

Where E_{Hop} is the energy function of HNN, H is the energy modifier, the connection weight matrix W is symmetric, $W = W^T$ and $w_{ii} = 0$. For the proposed CNN model, the energy modifier can be described as

$$H = \lambda\sum_{i}^{n}\int_0^{x_i} z_i\phi(s_i, u_i, x - I_0)dx \tag{15}$$

In the following, we apply the unified framework to verify (15) is the energy modifier of the proposed CNN:

$$\begin{aligned}\frac{dy_i}{dt} &= -\frac{\partial E}{\partial x_i} = -\frac{\partial (E_{Hop} + H)}{\partial x_i} \\ &= -\frac{y_i}{\tau} + \left[\sum_{j\neq i}^{n} w_{ij}x_j + I_i\right] - \lambda z_i\phi(s_i, u_i, x_i - I_0)\end{aligned} \tag{16}$$

Where $-\frac{\partial E_{Hop}}{\partial x_i} = -\frac{y_i}{\tau} + \sum_{j}^{n} w_{ij}x_j + I_i$ and $\frac{dy_i}{dt} = -\frac{\partial E}{\partial x_i}$, the relationship are given by Hopfied. Applying Euler discretization, (12) can be rewritten as

$$y_i(t+1) = (1-\frac{\Delta t}{\tau})y_i(t) + \Delta t\left[\sum_{j\neq i}^{n} w_{ij}x_j(t) + I_i\right] - \Delta t\lambda z_i(t)\phi(s_i(t), u_i(t), x_i(t) - I_0) \tag{17}$$

where Δt is the time step. If the following relationships are satisfied:

$$k = 1-\frac{\Delta t}{\tau}, \quad \alpha = \Delta t, \quad \lambda = \frac{1}{\Delta t} \tag{18}$$

Then (17) is equal to (2). Therefore, (17) is a reasonable energy modifier of the proposed CNN.

Using the mean value theorem, (11) can be rewritten as

$$H = \lambda\sum_{i}^{n}\int_0^{x_i} z_i\phi(s_i, u_i, x - I_0)dx = \lambda\sum_{i}^{n} z_i x_i\phi(s_i, u_i, \overline{x}_i - I_0) \tag{19}$$

where $0 < \overline{x}_i < x_i(t)$. As seen from (19), the equation form of the energy modifier is a linear combination of sigmoid function and states.

3.2 Asymptotical Stability

In the following, a computational energy function of the proposed CNN is constructed by applying the unified framework theory [11] and is proven to be asymptotically stable by applying Lyapunov stability theory.

Applying the unified framework discussed, the energy function can be described as

$$E(X)=-\frac{1}{2}\sum_{i,j}^{n} w_{i,j}x_i x_j-\sum_{i}^{n} I_i x_i-\frac{(k-1)}{\alpha}\varepsilon_1\sum_{i}^{n}\int_0^{x_i}\ln\frac{x}{1-x}dx+\frac{1}{\alpha}\sum_{i}^{n}\int_0^{x_i} z_i g(x-I_0)dx$$

then, the energy function can be described as

$$\alpha E(X)=-\frac{\alpha}{2}\sum_{i,j}^{n} w_{i,j}x_i x_j-\alpha\sum_{i}^{n} I_i x_i-(k-1)\varepsilon_1\sum_{i}^{n}\int_0^{x_i}\ln\frac{x}{1-x}dx+\sum_{i}^{n}\int_0^{x_i} z_i g(x-I_0)dx \tag{20}$$

where $W=W^T$, $w_{ii}=0$. Before analyzing the asymptotical stability of the discrete system containing (1) and (2), (1) and (2) are first combined into one equation as follows [1]:

$$\alpha\left[\sum_{j\neq i}^{n} w_{ij}x_j(t)+I_i\right]-z_i(t)g(x_i(t)-I_0)=\varepsilon_1\ln\frac{x_i(t+1)}{1-x_i(t+1)}-k\varepsilon_1\ln\frac{x_i(t)}{1-x_i(t)} \tag{21}$$

Applying the mean value theory, several equations can be described as follows:

$$\int_{x_i(t)}^{x_i(t+1)} z_i g(x-I_0)dx=\Delta x_i z_i g(\overline{x}_i-I_0) \tag{22}$$

where $\overline{x}_i(t)=x_i(t)+\overline{\theta}_i\Delta x_i$, $0\le\overline{\theta}_i\le 1$

$$\Delta x_i z_i\left[g(\overline{x}_i-I_0)-g(x_i-I_0)\right]\le\overline{\theta}_i z_i\left[\Delta x_i\right]^2/\varepsilon_2 \tag{23}$$

where $\xi_i=x_i(t)+\tilde{\theta}(\overline{\theta}_i\Delta x_i)$, $0\le\tilde{\theta}\le 1$ and the equality holds only for $\xi_i=I_0$

$$\int_{x_i(t)}^{x_i(t+1)}\ln\frac{x}{1-x}dx=\left[x_i(t+1)-x_i(t)\right]\ln\frac{\varphi_i}{1-\varphi_i}=\Delta x_i\ln\frac{\varphi_i}{1-\varphi_i} \tag{24}$$

where $\varphi_i=x_i(t)+\theta_i\Delta x_i$, $0\le\theta_i\le 1$

$$\Delta x_i\left(\ln\frac{x_i(t+1)}{1-x_i(t+1)}-\ln\frac{\varphi_i}{1-\varphi_i}\right)\ge 4(1-\theta_i)\left[\Delta x_i\right]^2 \tag{25}$$

where $\eta_{i1}=\varphi_i+\eta_1\left[(1-\theta_i)\Delta x_i\right]$, $0\le\eta_1\le 1$

$$\Delta x_i\left(\ln\frac{\varphi_i}{1-\varphi_i}-\ln\frac{x_i(t)}{1-x_i(t)}\right)\ge 4\theta_i\left[\Delta x_i\right]^2 \tag{26}$$

where $\eta_{i2}=x_i(t)+\eta_2\left[\theta_i\Delta x_i\right]$, $0\le\eta_2\le 1$。

The detailed analysis of the asymptotical stability is shown as follows:

$$\alpha E(X(t+1))-\alpha E(X(t))$$

$$=-\frac{\alpha}{2}\sum_{i,j}^{n} w_{i,j}\Delta x_i\Delta x_j-\alpha\sum_{i,j}^{n} w_{i,j}x_j\Delta x_i-\alpha\sum_{i}^{n} I_i x_i-(k-1)\varepsilon_1\sum_{i}^{n}\int_{x_i(t)}^{x_i(t+1)}\ln\frac{x}{1-x}dx$$

$$+\sum_{i}^{n}\int_{x_i(t)}^{x_i(t+1)} z_i g(x-I_0)dx$$

where $\Delta x_i = x_i(t+1)-x_i(t)$, $\Delta x_j = x_j(t+1)-x_j(t)$. Substituting (26) into last formula, we have

$$\alpha E(X(t+1))-\alpha E(X(t))$$

$$=-\frac{\alpha}{2}\sum_{i,j}^{n} w_{i,j}\Delta x_i\Delta x_j-\sum_{i}^{n}\Delta x_i\left\{\alpha\left[\sum_{j}^{n} w_{i,j}x_j+I_i\right]-z_i g(x_i-I_0)\right\}$$

$$-(k-1)\varepsilon_1\sum_{i}^{n}\int_{x_i(t)}^{x_i(t+1)}\ln\frac{x}{1-x}dx+\sum_{i}^{n}\Delta x_i z_i\left[g(\overline{x}_i-I_0)-g(x_i-I_0)\right]$$

Substituting (17) and (19)-(22) into $\alpha E(X(t+1))-\alpha E(X(t))$, we have

$$\alpha E(X(t+1))-\alpha E(X(t))$$

$$=-\frac{\alpha}{2}\sum_{i,j}^{n} w_{i,j}\Delta x_i\Delta x_j-\sum_{i}^{n}\Delta x_i\left\{\varepsilon_1\ln\frac{x_i(t+1)}{1-x_i(t+1)}-k\varepsilon_1\ln\frac{x_i(t)}{1-x_i(t)}\right\}$$

$$+\ (k-1)\varepsilon_1\sum_{i}^{n}\Delta x_i\ln\frac{\varphi_i}{1-\varphi_i}+\sum_{i}^{n}\Delta x_i z_i\left[g(\overline{x}_i-I_0)-g(x_i-I_0)\right]$$

$$\le-\frac{\alpha}{2}\sum_{i,j}^{n} w_{ij}\Delta x_i\Delta x_j-\varepsilon_1\sum_{i}^{n}\left[\Delta x_i\right]^2\left[4(1-\theta_i)+4k\theta_i\right]+\sum_{i}^{n}\overline{\theta}_i z_i\left[\Delta x_i\right]^2/\varepsilon_2$$

$$=-\frac{1}{2}\Delta X(t)^T\left\{\alpha W'+2\varepsilon_1\left[4(1-\theta_i)+4k\theta_i\right]I_n\right\}\Delta X(t)$$

If the minimal eigenvalue λ' of the matrix W' satisfies

$$-\lambda' < \frac{2\varepsilon_1}{\alpha}[4(1-\theta_i)+4k\theta_i]=\frac{8\varepsilon_1}{\alpha}[1-(1-k)\theta_i] \quad (27)$$

Then the matrix $\alpha W'+2\varepsilon_1\left[4(1-\theta_i)+4k\theta_i\right]I_n$ is positive definite, then we can obtain $E(X(t+1))-E(X(t))\le 0$ and the proposed CNN model can achieve asymptotical stability.

4 Application to 10-City TSP

A solution of TSP with N cities is represented by $N \times N$ -permutation matrix, where each entry corresponds to the output of a neuron in a network with $N \times N$ lattice structure. Assume x_{ij} to be the neuron output that represents city i in visiting order j.

A computational energy function can be described as:

$$E = \frac{W_1}{2}\left\{\sum_{i=1}^{n}\left[\sum_{j=1}^{n} x_{ij} - 1\right]^2 + \sum_{j=1}^{n}\left[\sum_{i=1}^{n} x_{ij} - 1\right]^2\right\} + \frac{W_2}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}(x_{k,j+1} + x_{k,j-1})x_{i,j}d_{ik} \tag{28}$$

where $x_{i0} = x_{in}$ and $x_{i,n+1} = x_{i1}$. W_1 and W_2 are the coupling parameters corresponding to the constraints 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 parameters of the network are set as follows:

$$W_1 = 1,\ W_2 = 1,\ k = 1,\ I_0 = 0.75,\ z(1) = 0.1,\ \varepsilon = 0.2,\ \alpha = 0.8,\ u(1) = 0.5,\ s(1) = 1.8.$$

Table 1. Results of 1000 different initial conditions for each value β on 10-city TSP

β	RGM(%)	RLM(%)	RIS(%)	AIC
0.01	68.9	31.1	0.0	86
0.005	76.9	23.1	0.0	126
0.004	80.0	20.0	0.0	147
0.003	84.9	15.1	0.0	172
0.002	89.8	10.2	0.0	212
0.001	100	0.0	0.0	252
0.0009	96.9	3.1	0.0	282
0.0008	96.4	3.6	0.0	299
0.0007	95.9	4.1	0.0	317
0.0006	95.5	4.5	0.0	352
0.0005	94.5	5.5	0.0	420
0.0004	89.7	10.3	0.0	564

1000 different initial conditions of y_i are generated randomly in the region [0, 1] for different β . The results are summarized in Table1, the column 'RGM', 'RLM', 'RIS' and 'AIC' respectively represents the rate of global minima, the rate of local minima, the rate of infeasible solutions and average iterations for convergence.

As seen from Table1, the CNN with sigmoid function self-feedback can solve the 10-city TSP effectively and find the global optimal solution with 1000 different initial conditions when β is 0.008. With the increase of the value of β when $\beta > 0.008$, the value of 'RGM' becomes small although it has a better constringency speed. At the same time, with the decrease of the value of β when $\beta < 0.008$, the value of ' AIC' becomes large although it has a better optimization effect.

5 Conclusion

The proposed CNN with sigmoid function self-feedback can acquire the ability to enhance the optimization and avoid trapping into local minimum. The analysis of the energy function of the proposed CNN with sigmoid function self-feedback indicates that there exists an energy modifier affecting the localized characterization ability of the proposed CNN. In addition, the stability analysis suggests that the proposed CNN with sigmoid function self-feedback can achieve asymptotical stability. The simulations on continuous function optimization problems and TSP indicate that the sigmoid function self-feedback has a better optimization ability.

References

1. Chen, L.N., Aihara, K.: Chaotic Simulated Annealing by a Neural Network Model with Transient Chaos. Neural Networks 8(6), 915–930 (1995)
2. Wang, L., Smith, K.: On Chaotic Simulated Annealing. IEEE Trans. on Neural Networks 9(4), 716–718 (1998)
3. Xu, Y.Q., Sun, M., Duan, G.R.: Wavelet Chaotic Neural Networks and Their Application to Optimization Problems. In: Adi, A., Stoutenburg, S., Tabet, S. (eds.) RuleML 2005. LNCS, vol. 3791, pp. 379–384. Springer, Heidelberg (2005)
4. Xu, Y.Q., Sun, M., Zhang, J.H.: A Model of Wavelet Chaotic Neural Network with Applications in Optimization. In: Proc. 6th World Congr. Intell. Control Autom., China, vol. 1, pp. 2901–2905 (2006)
5. Wang, L.: Intelligence Optimization Algorithm and its Application. Press of TUP (2001)
6. Xu, Y.-q., Sun, M., Guo, M.-s.: Improved Transiently Chaotic Neural Network and Its Application to Optimization. In: King, I., Wang, J., Chan, L.-W., Wang, D. (eds.) ICONIP 2006. LNCS, vol. 4233, pp. 1032–1041. Springer, Heidelberg (2006)
7. Lin, J.S.: Annealed Chaotic Neural Network with nonlinear Self-feedback and its Application to Clustering Problem. Pattern Recognition 34, 1093–1104 (2001)
8. Chen, L., Aihara, K.: Chaos and Asymptotical Stability in Discrete-time Neural Networks. Physica D 104, 286–325 (1998)

9. Kwok, T., Smith, K.: Experimental Analysis of Chaotic Neural Network Models for Combinatorial Optimization under a Unifying Framework. Neural Networks 13, 731–744 (2002)
10. Zhao, L., Sun, M., Cheng, J.H., Xu, Y.Q.: A Novel Chaotic Neural Network with the Ability to Characterize Local Features and its Application. IEEE Trans. On Neural Networks 20(4), 737–738 (2009)
11. Kwok, T., Smith, K.: A Unified Framework for Chaotic Neural-network Approaches to Combinatorial Optimization. IEEE Trans. Neural Network 10(4), 978–981 (1999)