**Article ID:** 1000- 8152(2000)06- 0847- 06

# Chaotic Neural Networks and Their Application to Optimization Problems

HE Yuyao
(College of Marine Engineering, Northwestern Polytechnical University· Xi'an, 710072, P. R. China)
WANG Lipo
(School of Electrical and Electronic Engineering, Nanyang Technological University· Singapore, 639798)

**Abstract:** In this paper, we first review three chaotic neural network models, and then propose a novel approach to chaotic simulated annealing. Second, we apply all of them to 10-city travelling salesman problem (TSP), respectively. The time evolutions of energy functions and outputs of neurons for each model are given. The features and effectiveness of four methods are discussed and evaluated according to the simulation results. We conclude that proposed neural network with simulated annealing have more powerful ability to obtain global minima than any other chaotic neural network model.

**Key words:** neural networks; transient chaos; simulated annealing; travelling salesman problem

**Document code:** A

# 混沌神经网络及其在最优化问题中的应用

贺昱曜 王力波
(西北工业大学航海工程学院· 西安, 710072) (南洋理工大学电气与电子工程学院· 新加坡, 639798)

摘要: 首先评述了三种混沌神经网络模型, 然后提出了一种新的混沌模拟退火算法. 其次将四种方法分别应用于 10 个城市的旅行推销商问题. 文中给出了每一模型神经元输出和能量函数随时间演变过程曲线. 根据仿真结果, 讨论了四种方法的特性与效果. 其结论为: 提出的模拟退火神经网络比其它网络模型更能获得全局最小解.

关键词: 神经网络; 暂态混沌; 模拟退火; 旅行推销商问题

## 1 Introduction

Many combinatorial optimization problems arising from science and technology are often difficult to solve entirely. Hence a great number of approximate algorithms for solving them have been developed[1,2]. Hopfield and Tank first applied the continuous-time, continuous-output Hopfield neural network (CTCO-HNN) to solving TSP[3], thereby initiating a new approach to optimization problems[4,5]. However, using the CTCO-HNN to solve TSP suffers from several shortcomings. First, the network is often trapped at a local minimum in the complex energy terrain because of its gradient descent property. Second, a CTCO-HNN may converge to an infeasible solution. Third, sometimes, a CTCO-HNN does not converge at all within presribed iteration[6].

In a number of methods proposed to overcome the shortcomings of CTCO-HNN, chaotic neural networks are shown to be powerful tools for escaping from local mininma. Based on chaotic property of biological neurons[7], Chen and Aihara proposed chaotic simulated annealing (CSA) to illustrate the features and effectiveness of a transiently chaotic neural network (TCNN) in solving optimization problems[8]. By adding a negative self-coupling to a chaotic neural network and gradually removing it, they used the transient chaos for searching and self-organizing. Chen and Aihara's approach significantly increased the probability of finding near-optimal solutions.

Wang and Smith presented an alternate approach to chaotic simulated annealing[9]. After analyzing the dynamics of the Euler approximation of CTCO-HNN as a function of the discretization time step, they suggested that starting from the Euler approximation of the CTCO-HNN with a large initial time step $\Delta t$, where the dynami-

Received date: 1999- 12- 30; Revised date: 2000- 07- 19.

cs are chaotic, and then gradually reducing the time step $\Delta t$, the system be assured to converge and to minimize the CTCO-HNN energy function. This approach eliminates the need for difficult choice of any other system paramerter.

Hayakawa et al proposed another approach for TSP by adding chaotic noise to the discretized CTCO-HNN[10]. The purpose is to help the network escaping from local minima more efficiently than mere white noise. The effect of the structure and distribution of chaotic noise on the performance of the neural network system was also discussed.

In this paper, we first review the above three chaotic neural network models, and then we propose a novel approach to chaotic simulated annealing. By adding chaotic noise to each neuron of the discrete-time, continuous output Hopfield neural network (DTCO-HNN) and gradually decaying it, the proposed neural network is initially chaotic but eventaully convergent, therefore has richer and more flexible dynamics compared to the HNN. Second we apply all of them to 10-city TSP, respectively. Third the features and effectiveness of the four methods are discussed and evaluated according to simulation results. The results show that the proposed network has more powerful ability to obtain global minima than other networks when applied to TSP.

## 2 Chaotic neural network models

In this section, three chaotic neural network models are reviewed in detail, and then the proposed chaotic neural network is given.

### 2.1 Chaotic simulated annealing with decaying self-coupling

Chen and Aihara's transiently chaotic neural network is described as follows:

$$x_i(t) = f(y_i(t)) = \frac{1}{1+e^{-y_i(t)/\varepsilon}}, \quad (1)$$

$$y_i(t+1) = ky_i(t) + \alpha\left(\sum_j^n w_{ij}x_j(t) + I_i\right) - z_i(t)(x_i(t) - I_0), \quad (2)$$

$$z_i(t+1) = (1-\beta)z_i(t), \quad (i = 1, 2, \cdots, n), \quad (3)$$

where $x_i(t)$: output of neuron $i$; $y_i(t)$: internal state of neuron $i$; $w_{ij}$: connection weight from neuron $j$ to neuron $i$, $w_{ij} = w_{ji}$; $I_i$: input bias of neuron $i$; $\alpha$: positive scaling parameter for neuronal inputs; $k$: damping factor of nerve membrane, $0 \leqslant k \leqslant 1$; $z_i(t)$: self-feedback connection weight (refractory strength) $\geqslant 0$; $\beta$: damping factor of $z_i(t)$, $0 < \beta < 1$; $I_0$: positive parameter; $\varepsilon$: steepness parameter of the output function, $\varepsilon > 0$.

$z_i(0)$ should be carefully selected for the network to be chaotic. As $\beta \to 0$, the self-coupling term approaches to zero with time evolution in the form of $z_i(t) = z_i(0)e^{-\beta t}$, the system becomes a Hopfield-like network which converges to a fixed point, and the energy function is minimized. The speed of this annealing process is determined by $\beta$.

### 2.2 Chaotic simulated annealing with decreasing time-step

Considering the Euler approximation of the CTCO-HNN,

$$y_i(t+\Delta t) = \left(1-\frac{\Delta t}{\tau}\right)y_i(t) + \Delta t\left(\sum_j^n w_{ij}x_j(t) + I_i\right). \quad (4)$$

Wang and Smith suggested that one can start from Eq. (4) with a large initial time step $\Delta t(0)$, where the dynamics are chaotic, and then gradually reduce the time step $\Delta t$ as the network iterates, for example, by using the exponential decaying rule,

$$\Delta t(t+1) = (1-\beta)\Delta(t), \quad (5)$$

where $0 < \beta < 1$. This causes the network to go through a reverse bifurcation process as it starts with a chaotic state and ends with a stable convergent state. The system finally converges and minimizes the CHNN energy function.

### 2.3 Discretized continuous hopfield network with chaotic noise

Hayakawa added chaotic noise to the discretized CTCO-HNN to solve TSP. Their chaotic neural network model can be defined as Eq. (4), together with

$$x_i(t) = f(y_i(t) + A\eta_i(t)), \quad (6)$$

where A is a multiplier to the chaotic noise $\eta_i(t)$, and f is a sigmoidal activation function with $\varepsilon = 1$. Chaotic noise with different initial values is associated with each neuron, which can be generated from the logistic map

$$z_i(t+1) = az_i(t)(1-z_i(t)), \quad (7)$$

A normalized series can be obtained as follows,

$$\eta_i(t) = \frac{z_i(t) - \langle z \rangle_t}{\sigma_z}, \quad (8)$$

where $\sigma_z$ is the standard deviation of the series $z_i(t)$ over time and $\langle z \rangle$ is the average of $z_i$. To determine the feasibility of the solution at each of iteration a discretised state variable is introduced,

$$x_{ij}^d = \begin{cases} 1, & \text{iff } x_{ij} > \sum_i^n \sum_j^n x_{ij}(t)/n \times n, \\ 0, & \text{otherwise.} \end{cases} \quad (9)$$

Parameters to be adjusted are the logistic map parameter $a$ in Eq.(7), and the amplification factor $A$ in Eq.(6).

### 2.4 Simulated annealing with decaying chaotic noise

We propose a chaotic neural network model, which is described as

$$y_i(t+1) = \alpha\left(\sum_j^n w_{ij} x_j(t) + I_i\right) + \gamma\eta_i(t), \quad (10)$$

together with Eq.(1) $(i = 1, 2, \cdots, n)$,

where

$\eta_i(t)$: chaotic noise for input to $y_i(t)$ $(0 < \eta_i(t) < 1)$; $\gamma$: positive scaling parameter for chaotic noise. Other parameters have the same meaning as in Eqs.(1) and (2). The chaotic noise can be generated from the logistic map:

$$\eta_i(t) = z_i(t) - h, \quad (11)$$

$$z_i(t+1) = a z_i(t)(1 - z_i(t)). \quad (12)$$

We let $a(t)$ decay exponentially, so that $z_i(t)$ is initially chaotic and eventually settle to a fixed point $z^*$:

$$a(t+1) = (1-\beta)a(t) + \beta \cdot a_0 \quad (13)$$

and

$$h = z^* = 1 - 1/a_0, \quad (14)$$

where $1 \leqslant a_0 < 2.9$, $0 \leqslant \beta \leqslant 1$. With $\eta_i = 0$, the network described by Eqs.(10) ~ (14) reduces to the DTCO-HNN.

## 3 Application of the above methods to the travelling salesman problem

We adopt the formulation of Hopfield and Tank[3] for TSP. Namely, a solution of TSP with n cities is represented by $n \times n$-permutation matrix, where each entry corresponds to output of a neuron in a network with $n \times n$ lattice structure. Assume $x_{ik}$ to be the neuron output which represents city $i$ in visiting order $k$. A computational energy function which is to minimize the total tour length while simultaneously satisfying all the constraints takes the following form:

$$E = \frac{A}{2}\left\{\sum_{i=1}^n\left(\sum_{k=1}^n x_{ik} - 1\right)^2 + \sum_{k=1}^n\left(\sum_{i=1}^n x_{ik} - 1\right)^2\right\} + \frac{B}{2}\sum_{i=1}^n\sum_{j=1}^n\sum_{k=1}^n (x_{jk+1} + x_{jk-1}) x_{ik} d_{ij}, \quad (15)$$

where $x_{i0} = x_{in}$ and $x_{in+1} = x_{i1}$. $A$ and $B$ are the coupling parameters corresponding to the constraints and the cost function of the tour length, respectively. $d_{ij}$ is the distance between city $i$ and city $j$.

In our simulation, we will use the Hopfield-Tank original data on TSP with 10 cities.

### 3.1 Chaotic simulated annealing with decaying self-coupling

In Chen and Aihara's original paper[8], they applied transiently chaotic neural network to 10-city TSP. The parameters were set as follows:

$$k = 0.9, \quad \alpha = 0.015, \quad \varepsilon = 1/250,$$
$$I_0 = 0.65, \quad z(0) = 0.8,$$

5000 different initial conditions of $y_{ij}$ were generated randomly in the region [−1, 1] for each value of $\beta$ = 0.015, 0.010, 0.005, 0.003. The results are summarized in Table 1 (in Tables 1~4, NG = the number of global minimum, RG = the rate of global minimum, NI = the number of iteration). The time evolutions of the discrete energy function $E^d$, and the output of neuron $x_{23}$ with $\beta$ = 0.003 are shown in Fig.1 and Fig.2, respectively.

Table 1 Results of 5000 different initial conditions for each value $\beta$ on 10-city TSP with Chen and Aihara's method

| $\beta$ | 0.015 | 0.010 | 0.005 | 0.003 |
|---|---|---|---|---|
| NG | 4946 | 4969 | 4998 | 5000 |
| RG | 98.9% | 99.4% | 99.9% | 100% |
| NI | 81 | 119 | 234 | 398 |

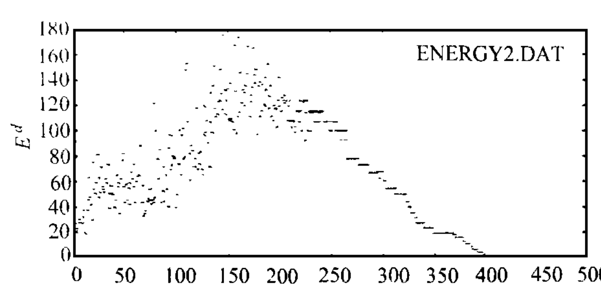


Fig.1 Time evolution of the discrete energy function for 10-city TSP with Chen and Aihara's method

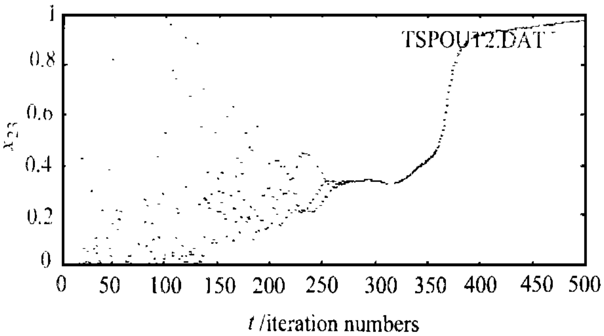


Fig. 2 Time evolution of the neuron output $x_{23}$ for 10-city TSP with Chen and Aihara's method

### 3.2 Chaotic simulated annealing with decreasing time-step

We apply Wang and Smith's CSA with decreasing time step to 10-city TSP. In order to decrease the influence of the gradient term $\partial E/\partial x_{ij}$ on chaotic neurodynamics at the early stage of the simulation, we multiply it by a coefficient $\alpha$. The parameters we choose are as follows:

$$\Delta t(0) = 2, \quad \tau = 0.5, \quad \alpha = 0.015.$$

We only adjust the parameter $\beta$. Fig. 3 and Fig. 4 show the time evolutions of the discrete energy function and one typical neuron state corresponding to an optimal solution for $\beta = 0.001$. Table 2 shows the feasibility and performance variation with different annealing factors.

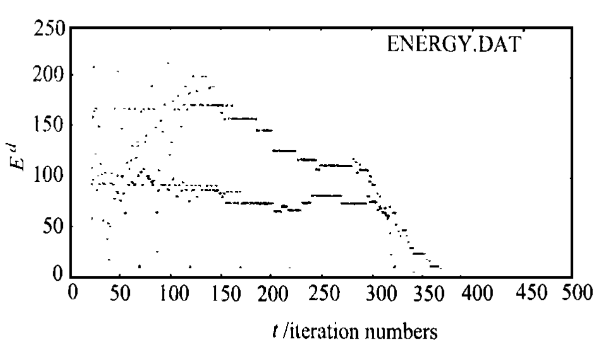


Fig. 3 Time evolution of the discrete energy function for 10-city TSP with Wang and Smith's method

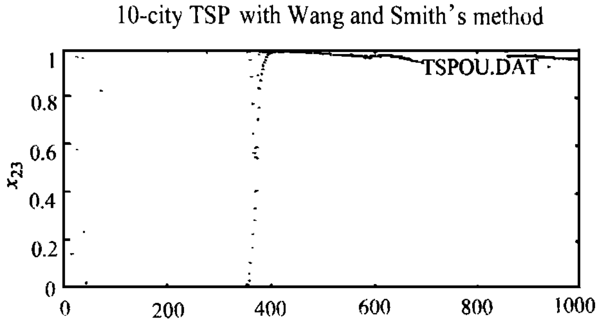


Fig. 4 Time evolution of the neuron output $x_{23}$ for 10-city TSP with Wang and Smith's method

Table 2 Results of 100 different initial conditions for each value $\beta$ on 10-city TSP with Wang and Smith's method

| $\beta$ | 0.010 | 0.005 | 0.001 |
|---|---|---|---|
| NG | 65 | 81 | 92 |
| RG | 65% | 81% | 92% |
| NI | 130 | 225 | 965 |

### 3.3 Discretized continuous hopfield network with chaotic noise

In this method, chaotic noise is added to the discretized CTCO-HNN to solve the 10-city TSP. Since the network is no longer convergent, a statistical average over a period of time should be employed to determine the state of a neuron. Iterations less than 1000 correspond to the transient period. 2000 is chosen to be the maximum number of iterations. In order to determine the feasibility of the solution the visiting frequency to the optimal solutions is tested:

$$P_{os} = (\text{number of steps staying at the optimal solution})/(\text{total steps}). \quad (16)$$

$P_{os}$ is calculated for the last 1000 steps of 2000 iterative steps of computer runs. Because Hayakawa et al did not give the time evolutions of the discretized energy function and the output of neuron $x_{ij}$, we produce them in Fig. 5 and Fig. 6, the feasibility measure in Table 3. The logistic map parameter is chosen to be 3.81 and 3.93 since they belong to the chaotic region of the logistic map and are found to have good optimization ability on TSP.

Table 3 Results of 100 different initial conditions for each value $A$ on 10-city TSP with Hayakawa et al's method

| $A$ | 0.05 | 0.5 | 5.0 |
|---|---|---|---|
| NG | 10 | 70 | 32 |
| RG | 10% | 70% | 32% |
| NI | 2000 | 2000 | 2000 |

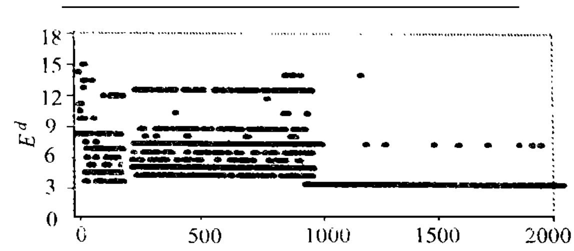


Fig. 5 Time evolution of the discrete energy function for 10-city TSP with Hayakawa et al's method

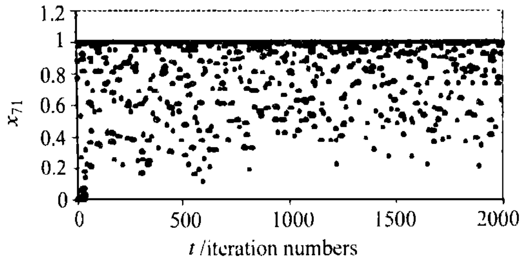


Fig. 6 Time evolution of the neuron output $x_{71}$ for 10-city TSP with Hayakawa et al's method

### 3.4 Simulated annealing with decaying chaotic noise

We analyze 10-city TSP using our approach described in Section 2.4. The parameters are set as follows:

$\alpha = 0.015, \quad \gamma = 0.1, \quad \varepsilon = 1/250,$

$a(0) = 3.9, \quad a_0 = 2.5, \quad h = 1 - 1/a_0 = 0.6.$

We vary only $\beta$ to investigate the dynamics of our chaotic neural network model. The results obtained are summarized in Table 4, Fig. 7 and Fig. 8 show the time evolutions of energy function $E^d$, and the output of neuron $x_{23}$, respectively.

Table 4 Results of 5000 different initial conditions for each value $\beta$ on 10-city TSP with He and Wang's method

| $\beta$ | 0.1 | 0.015 | 0.003 | 0.001 |
|---|---|---|---|---|
| NG | 5000 | 5000 | 5000 | 5000 |
| RG | 100% | 100% | 100% | 100% |
| NI | 19 | 82 | 350 | 1022 |

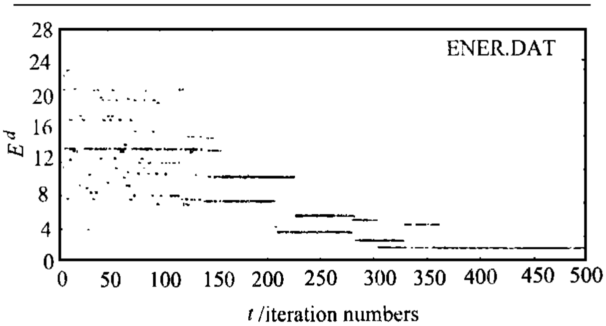


Fig. 7 Time evolution of the discrete energy function for 10-city TSP with He and Wang's method

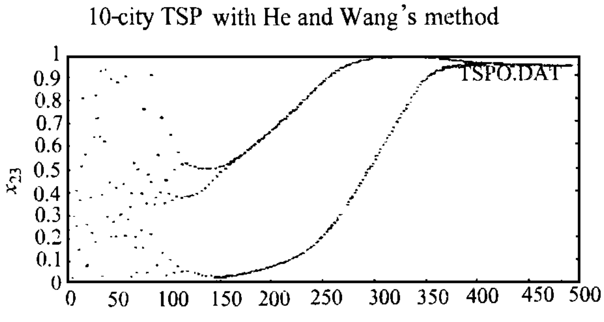


Fig. 8 Time evolution of the neuron output $x_{23}$ for 10-city TSP with He and Wang's method

## 4 Discussion

In this section, we illustrate the properties and optimization ability of chaos in neural networks by analyzing the results obtained above.

With Chen and Aihara's method applied to 10-city TSP, one can see from Fig. 2 that chaotic behavior of state sustains during the first 260 iterations. After the unstable wandering of the state ceases, the state has a little fluctuation and then tends to converge to a stable and optimal solution. The state transition from the inherent instability to order has the effect of chaotic wandering in search of the global minima. As shown in Table 1 the rate of global optimization is 98.8% for $\beta = 0.015$, 100% of the rate of global optimzation is also realized by simulations with a smaller $\beta = 0.003$, which corresponds to a slower annealing schedule. This can be interpreted as a prolonged chaotic search due to a lengthened region of instability. However, the tradeoff is the increased number of iterations required for convergence.

Fig. 4 shows how a neuron state follows its path contributing to an optimal solution by using Wang and Smith's CSA with decreasing time-step. Initially it oscillates between 0 and 1, but starts to visit intermediate values when iteration is between 200 and 350. The two branches finally merge together towards 1, which corresponds to a global minimum. The result suggests a search of the global minimum through chaotic dynamics. The corresponding wandering of the energy among local minima can be observed in Fig. 3. Finally the discrete energy function is globally minimized. In general, this method requires fewer parameters to be adjusted, but the difficulty arises when choosing an effective annealing rate $\beta$.

Since the discretized Hopfield network with chaotic noise is not convergent, no single stable state is attained by the neuron states. A typical neuron state iteration corresponding to an optimal solution is illustrated in Fig. 6. Although noisy, the state tends to have a denser distribution around the value of 1, except that at the beginning 0 is often visited. The final state of this neuron is 1 according to Eq. (16). Because of the noisy nature of the network a discretized energy is used. In Fig. 5, a transient region exists for iteration < 1000 where the energy attains non-optimal values. When iteration > 1000, the energy dramatically drops to a global minimum and stays around that value with occasional jumps to a local minimum. A maximum of 7/10 feasibility is obtained as shown in Table 3.

We can see from Table 4 that $\beta$ can vary from a larger value of 0.1 to a smaller value of 0.001. When a larger $\beta (= 0.1)$ is used the neural network converges average with 19 iterations (for the discrete energy function). Among 5000 cases, 100 percent are to global minima (tour length = 2.690670). When $\beta$ gets smaller, the rate of global optimization minima is 100 percent, too. But the number of iterations is greatly increased for $\beta = 0.001$. The only difference for different $\beta$ is that the number of iterations is different.

## 5 Conclusions

We have introduced four methods of incorporation chaos into the Hopfield network for combinatorial optimizations. Two such methods, namely Chen and Aihara's CSA with a decaying self-coupling term and Wang and Smith's CSA with decaying time-steps, make use of chaotic annealing schemes analogous to the traditional simulated annealing. Both have features like chaotic wandering and transitions from instability to a stable state, which are found to have a novel ability to improve the optimization performance of the network. Existence of chaos and convergence to a stable solution in both methods are well established by the respective proofs. The Hopfield network with chaotic noise has the least number of adjustable parameters and is relatively efficient, but lacks convergence properties and is the least capable of reaching for an optimal solution.

Different from the first three methods, our method is to add chaotic noise to each neuron of DTCO-HNN and then gradually decay it. The network finally converges to optimal solution. The main feature of our method is that it is more efficient, the rate of optimization is higher than any other method and the parameter can be easily chosen when applied to 10-city TSP.

## References

[1] Reeves C R. Modern Heuristic Techniques for Combinatorial Problems [M]. Oxford: Blackwell, 1993

[2] Zanakis S H, Evans J R and Vazacopoulos A A. Heuristic methods and applications: a categorized survey [J]. Europ. J. Operat. Res., 1989, 43(1): 88-110

[3] Hopfield J J and Tank D W. "Neural" computation of decisions in optimization problems [J]. Biol. Cybern., 1985, 52(1): 141-152

[4] Hopfield J J. Neural networks and physical systems with emergent collective computational abilities [J]. Proc. Natl. Acad. Sci., 1982, 79(4): 2554-2558

[5] Hopfield J J. Neurons with graded response have collective computational properties like those of two-state neurons [J]. Proc. Natl. Acad. Sci., 1984, 81(5): 3088-3092

[6] Wilson G V and Pawley G S. On the stability of the traveling salesman problem algorthm of Hopfield and Tank [J]. Biol. Cybern., 1988, 58(1): 63-70

[7] Aihara K, Takabe T and Toyoda M. Chaotic neural networks [J]. Physics Letter A, 1990, 144(6, 7): 333-340

[8] Chen L and Aihara K. Chaotic simulated annealing by a neural network model with transient chaos [J]. Neural Networks, 1995, 8(6): 915-930

[9] Wang L and Smith K. On chaotic simulated annealing [J]. IEEE Transactions on Neural Networks, 1998, 9(4): 716-718

[10] Hayakawa Y, Marumoto A and Sawada Y. Effects of the chaotic noise on the performance of a neural network model for optimization problems [J]. Physical Review E, 1995, 51(4): 2693-2696

## 本文作者简介

**HE Yuyao** Received the B.S. degree from Beijing University of Aeronautics and Astronautics, the M.E. degree from Xi'an Jiaotong University, and the Ph.D. degree from Northwestern Poltechnical University in 1981, 1984, and 1999, respectively. From 1986 to 1987 he studied at the Department of Electrical and Computer Engineering, Clemson University, SC, U.S.A., mainly engaging in the research of computer controlled robotics. In 1998 he did collaborative research on neural networks at School of Computing & Mathematics, Deakin University, Victoria, Australia as a visiting professor. He is currently a professor. His research interests include power electronics, nonlinear system control and artificial neural networks.

**WANG Lipo** Received the Ph.D. degree in physics from Louisiana State University in 1988. In 1989 he worked at Stanford University, CA, as a postdoctoral Fellow. He is currently an associate professor. His research interests include theory and application of neural networks, pattern recognition and fuzzy neural network control.